Characterizing Colour-Word Contingency Learning

By

Ying-Hsin Lin

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Psychology

Waterloo, Ontario, Canada, 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis,

Including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Contingencies are constantly found in everyday situations and humans are extraordinarily adept at learning them, whether implicitly or explicitly. The learning of contingencies can benefit behaviour in many ways, some subtle and some much more apparent. But the question of how we can improve the efficiency of our performance and what factors influence a person's learning of contingencies remains relatively limited in exploration. In this dissertation, I use a simple contingency learning paradigm, the colour-word contingency paradigm (Schmidt, Crump, Cheesman, & Besner, 2007) to address three issues regarding the learning of contingencies: (1) Is there a cost in performance—in addition to a benefit—when learning contingencies?; (2) Does instance frequency play a role in the learning of contingencies?; and (3) Can people use higher-order associations such as semantic associations in learning contingencies? In the first series of experiments (Experiments 1-4), I found that although there is benefit in responding to events with high contingencies, there also is a cost to events with low contingencies. In the second series (Experiments 5-7), I showed that event frequency does play a role in contingency learning but that role is dependent on factors such as whether there is perfect contingency between events and how balanced the frequencies are in the high- and low-contingency conditions. In the third series (Experiments 8-10), I observed that people are capable of using higher-order associations—in this case, semantics—to guide the efficiency of their responding, however the use of such associations is dependent on whether the task encourages their use. This dissertation thereby addresses three issues where research has been quite tentative and, in some cases, the present findings are in contrast to the current literature. The overall goal was for this research to contribute to the growing literature on fundamental factors that can influence how we acquire contingencies, and to help us understand how we optimize our efficiency in learning, even when that learning occurs without awareness.

Acknowledgements

Thank you to all of my friends (you know who you are) and family who have supported me throughout the years that I spent completing my dissertation. Thank you to my family who have given me their continued support and provided me with all that they can to see that I can pursue my goals in my life. I cherish the every moment that we share and love you all very much.

Thank you to Dr. Colin MacLeod for being a great supervisor. I have learned an immense amount from you. Not only have you been supportive but you also have guided me and advised me to become a better researcher. Thank you to Drs. Jonathan Fugelsang and Daniel Smilek for serving on my committee and giving me invaluable advice and input.

Thank you to my friends, fellow and former lab members, colleagues and professors that I have met in Waterloo. Thank you for all your positive support, your advice and critiques, and all of your contributions. A special thank you to the research assistants who assisted with data collection.

Thank you to my friends in the UW Kendo Club. I appreciate all the team work, hard work, success, failures, laughter and all those (little) frustrations that we endured. I always had lots of fun whether it was during practice, tournaments, or when we just hung out. Thank you also for the invaluable experience I have gained about the management of a club. I hope the fun extends to the future generations.

And to Eddie, although I have known you only during the last phase of my dissertation, thank you for being my light. Your love and support is always encouraging and no words can describe my appreciation. Thank you for always finding ways to help me.

Table of Contents

Introduction

People are very proficient at learning associations between two or more events. Not surprisingly, then, the concept of association has been the focus of theory at least since the Greek philosophers (see, e.g., Sorabji, 2004) and was a major early emphasis in experimental psychology (see Boring, 1950). The learning of associations often assists our performance regardless of how that learning is accomplished—whether implicitly (without awareness) or explicitly (with awareness). As examples of explicit associative learning, we can strategically learn the Fibonacci sequence of numbers by associating them with patterns, or we can learn that the colour red is often associated with caution, danger, or stop, such that we exercise more caution and alertness when we encounter a red traffic light. Other associations are learned more implicitly. For example, athletes learn body cues from teammates and opponents that affect their playing of the sport. The learning of different types of associations has been explored by researchers examining implicit learning, statistical learning, probability learning, and contingency learning, among other forms of learning (see Shanks, 2007; De Houwer & Beckers, 2002).

This dissertation is an exploration of one form of human associative learning—contingency learning. *Contingency learning* can be defined as the acquisition of knowledge about co-variations or correlations between stimuli, between responses, or between stimuli and responses. Such learning influences performance in numerous ways, among them the speed and accuracy of responding. For example, if we have learned, whether implicitly or explicitly, that some stimulus—say, a heart—typically appears in the colour red and not in other colours, then we may well process that stimulus faster and/or more accurately when it appears in red than

when it appears in another colour. There is considerable evidence that this kind of learning occurs very rapidly, sometimes even within a single trial (e.g., Lewicki, 1985).

The efficiency of our performance while learning probabilities and contingencies is doubtless influenced by many factors. In past research, emphasis has often been placed on how the learning of associations can influence our behaviour in a positive way (Schmidt, Crump, Cheesman, & Besner, 2007). However, learning of probabilities and contingencies can also become so routine that any deviation from what has been learned can hurt performance (Wolfe et al., 2007; Hon & Tan, 2001). For example, in cities like Los Angeles where snowfall occurs rarely, average driving performance decreases dramatically and traffic becomes intolerable during a snowfall. This leads to more car accidents during that time, in part because people are accustomed to driving in a particular way and do not sufficiently adjust their behaviour under unfamiliar circumstances.

In this dissertation, I examine three main questions about how we learn contingencies. The first is whether contingency learning involves a cost as well as a benefit in terms of the efficiency of performance. The second addresses whether frequency of stimuli or of stimulus co-occurrence plays a role in learning contingencies. In the third, I examine whether people use higher-order semantic information while learning associations. These questions are important to consider because humans are able to grasp contingencies very quickly, such that research on fundamental factors that influence how we acquire contingencies can help us to understand how we maximize our efficiency.

The psychological literature on learning of associations can be traced back to animal research examining classical and operant conditioning (Thorndike, 1898; Pavlov, 1902). A dog salivates (conditioned response) to the sound of a bell (conditioned stimulus) after multiple

pairings of the sound with food (unconditioned stimulus). According to the Rescorla-Wagner (1972) model, conditioned responses are based not only on the statistical contingency between the conditioned stimulus (CS) and the unconditioned stimulus (US) but also on other conditioned stimuli that are associated with the unconditioned stimulus. Each time a conditioned stimulus is presented, it updates the association between the conditioned stimulus and the unconditioned stimulus. This model provides a basis for other models of human contingency learning (De Houwer & Beckers, 2002) and illustrates that learning of contingencies depends on multiple factors. Human contingency learning is often rightly compared to the ability of animals to learn associations after multiple pairings (See De Houwer & Beckers, 2002; Shanks, 2007).

In humans, extensive research has explored how quickly and effectively people can pick up information about and learn associations (Estes, 1964), possibly even effortlessly (Hasher & Zacks, 1984). Of course, sometimes humans use strategies to help with learning of associations (e.g., remembering new vocabulary for a test or connecting acronyms to their referents). But often, associative learning is beneath the level of awareness (Shanks, 2007). A now classic study by Nissen and Bullemer (1987) illustrates this point very well. In their serial reaction time task, participants responded by pressing the key corresponding to the location of a stimulus. Compared to people given a random sequence, people given a repeating sequence showed evidence of learning by their gradual reduction in response time over trials. Although the sequence was in fact irrelevant to responding to location, people nonetheless unintentionally learned sequence information and used it. Remarkably, even patients with Korsakoff's syndrome were able to benefit from the repeating sequences. This suggests that the learning was done implicitly, which is also consistent with unimpaired subjects not being aware of their learning.

Related research on artificial grammar learning and on visual statistical learning shows that humans can learn quite complex associations without being aware that they are actually learning (Reber, 1967). This type of implicit learning can also occur at quite an early age, as studies show that infants and children are also able to pick up complex associations simply by repeatedly processing them (Gomez & Gerken, 1999, Kirkham, Slemmer, & Johnson, 2002). Indeed, it is quite likely that language learning rests on this kind of statistical learning (Saffran, 2003).

Although the literature on contingency learning is vast, how and what people are actually learning in such situations is still not widely understood. In this dissertation, the aim is to find out how people learn contingencies and to explore factors that influence a specific type of contingency learning called *colour-word contingency learning* (Schmidt, Crump, Cheesman, & Besner, 2007). The colour-word contingency paradigm was selected because it provides a simple way to examine response time (RT) and accuracy in contingency learning by combining two dimensions that are very familiar to people and that frequently appear in everyday life— words and colours. This paradigm has recently been extensively used to study human contingency learning (Schmidt, 2013; Schmidt et al., 2007; Schmidt & Besner, 2008; Schmidt & De Houwer, 2012 a,b,c; Schmidt, De Houwer, & Besner, 2010). In what follows, I will first review the literature on colour-word contingency learning. I will then review more specific evidence that examines (1) the cost to performance in probability learning, (2) the role of frequency in such learning, and (3) the benefit of meaning in such learning.

The colour-word contingency paradigm

The colour-word contingency learning task refers to the learning of associations between non-colour words and the colours in which they are printed. Participants make identification

responses to the colours in which words are presented, for example, pressing a key for the colour

red when any word (e.g., TABLE or HORSE) is presented in red. In fact, the words themselves

could simply be ignored and successful performance would still be possible. This paradigm

examines the response time to events with different contingencies, determined by varying the

proportion of times that a given word appears in a given colour. In a typical contingency

learning study, there are three words and three colours. The probability of each word being

presented in each colour is varied, often to be either high contingency (80%) or low contingency

(20%). On high-contingency trials, a word is most often presented in a certain colour (e.g.,

PLATE in green) whereas on low-contingency trials, that word is presented less often in the

other two colours (e.g., PLATE in yellow or red).

In recent years, Schmidt, De Houwer, and their colleagues (2012c; Schmidt, 2013;

Schmidt & Besner, 2008; Schmidt et al., 2010) have explored this contingency task in some

fundamental ways, finding at the most basic level that response times are faster and errors are

less frequent on high-contingency trials than on low-contingency trials. This performance

difference between high-contingency and low-contingency trials has been labeled *the*

contingency effect.

The contingency effect has proved to be useful, for instance in explaining the item-

specific proportion congruent (ISPC) effect (see Lindsay & Jacoby, 1994; Jacoby, Lindsay, &

Hessels, 2003). The ISPC effect refers to the observation that there is an increase in the Stroop

effect—the interference in naming the print colour of an incompatible colour word (e.g., the

word RED printed in blue, where the response is "blue" to the print colour)—as the proportion of

congruent trials relative to incongruent trials increases. Schmidt and colleagues (Schmidt et al.,

2007) challenged the dominant view that the Stroop effect is due to participants shifting their

attention toward a congruent word and away from an incongruent word. Instead of this attentional account, Schmidt and colleagues argued that "participants implicitly learn contingencies (i.e., correlations) between words and responses and then use these contingencies to predict the specific response associated with each distracting word" (Schmidt & Besner, 2008, p. 515).

In an illustrative study (Schmidt et al., 2007), participants were shown one of four words presented in one of four colours on each trial. As usual, there were high contingency items (e.g., MOVE in blue) and low contingency items (e.g., MOVE in orange, yellow, or green). However, there were only two response keys, with two colours associated with each key (e.g., blue and green for the left key and yellow and orange for the right key). This permitted three response conditions. On stimulus match trials, the word that was usually presented in a certain colour was coupled with the expected response key (e.g., MOVE in blue). On response match trials, a word usually presented in a different colour was nevertheless paired with its expected response key (e.g., MOVE in green). On response mismatch trials, a word was presented in a colour in which it rarely appeared and the response was not the expected response (e.g., MOVE in yellow or in orange). Response mismatch trials were the slowest, but critically there was no response time difference between response match trials and stimulus match trials. This was taken as support for the idea that people are learning the contingencies between the words and the responses and not between the words and the colours. Thus, the colour-word contingency paradigm provides a way to examine the learning of contingencies between the nominally response-irrelevant word and the response to the relevant colour.

There are several factors that Schmidt and colleagues have found to influence colour-word contingency learning. Clearly, the paradigm allows people to learn colour-word

contingencies implicitly and quickly (Musen & Squire, 1993; Schmidt et al., 2007). But people can also unlearn contingencies very quickly. In one study (Schmidt et al., 2010), there were three learning blocks consisting of 18 trials each with the word-colour association occurring with 67% probability for high contingency and 33% probability for low contingency. Afterward, there were nine extinction blocks of 18 trials each where all words were presented in each of the three colours with equal probability (33% probability). Participants learned the contingencies quickly within the first learning block and there was a small carryover effect of that learning during the first extinction block, although that quickly disappeared. Although the small carryover onto the first extinction block does suggest that people can retain the information that is learned, it is also clear that they are primarily sensitive to the actual contingencies—or to the absence of contingencies—during the current block.

Schmidt and Besner (2008, p. 522) have argued very clearly that "contingency effects are solely facilitative in response times." In their Experiment 1, participants were shown trials forming three conditions: high contingency, low contingency, and no contingency (baseline). There were three words and two colours. Two of the words were presented in one colour 75% of the time (high contingency) and in the other colour 25% of the time (low contingency). The third word, which acted as a baseline, appeared equally often in each of two colours (50% of the time). Schmidt and Besner (2008) found that responses for high-contingency items were faster than those for baseline items and low-contingency items, which did not differ. Thus, they observed a response time benefit in responding to high-contingency items but no cost in responding to low-contingency items.

The picture was, however, more complicated. Although only a benefit was evident in response time in their Experiment 1, in their Experiment 2, Schmidt and Besner (2008) found

that the errors demonstrated a cost in performance. Because the threshold for the high contingency trials is lowered, "one of the incorrect responses is lowered when it is predictable (i.e., on low contingency trials) and it is therefore more likely that participants will make this particular error" (p. 522). Thus, there appeared to be a trade-off when learning contingencies such that there could be a benefit in response time coupled with a cost in response accuracy—a classic speed-accuracy trade-off (see Pachella, 1974). Schmidt and Besner (2008, p. 523) argue that the contingency learning effect is "consistent with the *(their)* proposed response mechanism where participants decrease the threshold for the expected (high contingency) response."

The facilitation of responding to high contingencies between words and colours as a result of contingency learning is seen as being beneath the level of awareness (see Experiment 3 in Schmidt et al., 2007). After participants were asked (1) whether they were aware of the contingencies and (2) to guess in which colour the word was most often presented, responses were grouped into three categories: aware, unaware, and chance. The magnitude of the contingency effect was unrelated to the level of awareness. In a different study, when participants were made aware of the contingencies by explicit instructions concerning what the contingencies were, the explicit instructions to learn the contingencies did increase the contingency effect (Schmidt & De Houwer, 2012b), but there still was a contingency effect even without explicit directions. Thus, although contingency effects can be obtained below the level of awareness, awareness apparently does enhance the contingency effect.

Schmidt and De Houwer (2012a) also found that contingency learning was not limited to just the learning of the simple association between a word and a colour. The contingencies that are learned can extend beyond the specific stimulus information that participants have experienced. For example, in a study in which participants were shown a nonword followed by a

target word in colour, the different contingencies of the nonword and the target word in colour influenced how likable the participants rated the target words on a subsequent task, with high nonword-target word contingency resulting in greater likability. This study so far is the only one to show that the colour-word contingency paradigm can generalize beyond specific stimulus-response pairings in the learning task, showing that other associations can influence the responses.

Given the research on how different factors influence colour-word contingency, Schmidt (2013) proposed a Parallel Episodic Processing (PEP) model to explain the contingency effect. In this model, derived from Logan's instance model of automaticity (Logan, 1988), people encode information about the stimuli, including the colour and the response, into episodic memory. At retrieval, people recover past episodes more fluently when they have been experienced more often, resulting in more instances in play, as is the case with high-contingency items. Thus, responses for high-contingency items are facilitated by the retrieval of faster stored instances. To retain the instances of contingencies in storage and to retrieve them at a subsequent time, Schmidt argued that some kind of limited capacity memory resources are required. In support of this requirement, when participants were placed under high load and given a secondary task, the contingency effect was reduced or even eliminated (Schmidt et al., 2010).

To explain how the contingency effect is due to facilitation of responding to high-contingency items, Schmidt (2013) argued for a threshold account. Initially, the response threshold is set to be similar for each item, whether high contingency or low contingency, principally because the contingency manipulation has not yet been influential. As trials progress, when people encounter a high-contingency trial, responses are facilitated because of the lowering

of the threshold of the expected response due to more and faster instances being available in memory. Consequently, it is easier to respond for the high-contingency items, leading to a faster response and, thus, facilitation. Significantly, Schmidt maintains that whereas the thresholds for high-contingency items are lowered with experience, the thresholds for low-contingency items are unaltered (see also Schmidt & Besner, 2008). As one illustration of this claim, Schmidt and Besner (p. 516) assert that "On a low contingency trial (e.g., BLUE in green), the distracting word is predictive of a specific response (i.e., BLUE predicts a blue key response). Thus, the threshold for this predicted response will be lowered. However, the response threshold for the remaining colours (green, yellow, and orange) will be unaltered. Because the correct response (green) is not the predicted response (blue), correct response latencies will not be speeded." Thus, their model predicts only a response time benefit for high-contingency items but not a response time cost for low-contingency items.

The colour-word contingency paradigm provides a simple paradigm to study how changes in probabilities and contingencies can alter performance without the person being aware of what they are learning—or even that they are learning. The comparison between high contingencies and low contingencies in the colour-word contingency paradigm is also generalizable to many studies in the attention literature that examine efficiency in performance when detecting or searching for high-probability versus low-probability targets (Hon & Tan, 2013; Hon, Yap, & Jabar, 2013; Wolfe & Van Wert, 2010). In this dissertation, then, to further understand the factors influencing contingency learning, this paradigm seemed ideal. Below, I review the relevant literature related to the three primary topics of my dissertation, the goal of which is to more fully characterize contingency learning with regard to: (1) evidence of a cost for low probability items, (2) the involvement of frequency, and (3) the contribution of higher-

order (semantic) associations. These questions are addressed to provide a basic understanding of how people learn contingencies (i.e., whether there are cost and benefits in performance) and whether people can use higher-order associations to benefit their performance.

Are there costs for low probability items in contingency learning?

In everyday situations, we encounter events that always occur (e.g., the sun comes up in the east every morning), events that sometimes occur (e.g., snow falls occasionally in the winter), and events that rarely occur (e.g., an earthquake hardly ever takes place in eastern Canada). Slower and more error-prone responding to low probability events, termed *the prevalence effect*, has been studied at least since Kundel (1982). Concepts from signal detection theory and probability matching have become the focus of many attention studies in domains such as detection tasks, vigilance tasks, and search tasks (Craig, 1978; 1987; Hon & Tan, 2013; Wolfe et al., 2007). In more applied settings, research has examined errors that occur in health care and military settings. For example, in the military, target detection during sentry duty decreased over time as the duration of the sentry duty increased (Johnson & McMenemy, 1989; see the classic clock test mimicking two-hour watch duties, Mackworth, 1948). In health care, tumors are often missed in radiology scans due to their relatively low probability of occurrence (Drew, Evans, Vo, Jacobson, & Wolfe, 2012). In situations where people encounter both high and low probability, there is a cost in responding to the low probability information (Craig, 1978; Laberge & Tweedy, 1964; Hon & Tan, 2013; Hon, Yap, Jabar, 2013; Wolfe et al., 2007). Often, these costs are evident even when the accuracy of target detection is very high. Responses to low probability targets have also been shown to be associated with greater dorsal prefrontal activation as opposed to greater ventral prefrontal activation for high probability targets (Casey et al., 2001; Hon, Ong, Tan & Yang, 2012). The relation between target probability and activation of

different regions is consistent with a biological index for the differences in activation associated with high and low probability events.

What happens when people encounter different event probabilities? When they experience both high and low probability information, it may be that the threshold of responding is altered to correspond to the different level of probability. Hon and Tan (2013) claim that there is an attentional locus and that response to target probability is affected by the amount of attentional resources available. In their task, participants detected the occurrence of two letters in two blocks of serially presented single letter targets. In both blocks, one target occurred 10% of the time and the other occurred 40% of the time; distracters occurred the remaining 50% of the time. The first block was a dual task block where participants also performed a counting task; the second block involved only the detection task. The dual task affected detection of the low-probability targets more than of the high-probability targets, suggesting that low-probability targets require more attentional resources. When attentional resources were made more available, the detection of low-probability targets in particular improved. Thus, Hon and Tan (2013) found that the size of the target probability effect was influenced by how available attentional resources were.

The findings from Hon and Tan (2013) suggest that low-probability targets require more attentional processing compared to high-probability targets. One might imagine that responding to high-probability targets would speed up over time. However, this assumption is in contrast to what work on vigilance and detection tasks has found (Craig, 1978; Craig, 1987). Moreover, Hon, Yap, and Jabar (2013) found that responding to the low-probability targets became slower over time whereas responding to the high-probability targets stayed the same across blocks. In their study, participants detected one of two targets. Over trials, one target occurred with low

probability (10% in Experiment 1 and 5% in Experiment 2) whereas the other occurred with high probability (40% in Experiment1 and 45% in Experiment 2). Results from both experiments showed that performance on low-probability targets slowed gradually until it reached a stable level whereas performance on high-probability targets remained constant over time. This study supports the notion that the response difference between high-probability and low-probability targets is driven by slowed responding to the low-probability targets, not by speeded responding to the high-probability targets. Hon et al. (2013) argue that, initially, perceptual templates for low-probability and high-probability targets are at the same level of accessibility in memory. Due to the low frequency of low-probability targets, however, the accessibility of their templates begins to fall and thus low-probability templates gradually need increased time to reach the same threshold of activation.

Similarly, Wolfe, Horowitz, Van Wert, Kenner, Place, and Kibbi (2007) have argued that the errors that are made during low-prevalence target searches (e.g., looking for weapons in a security scanning of baggage) are due not merely to careless lapses in attention but to a criterion shift with respect to responding to high-probability and low-probability targets. Low target-prevalent searches elicit more misses than high target-prevalent searches and it is argued that people attempt to equate the number of misses and false alarms, leading to a more liberal approach where a higher proportion of rare targets is being missed. In a baggage-scanning task where participants are looking for weapons, miss rates increased from an average of 0.20 at a 50% prevalence to 0.46 at a 2% prevalence, however, this increase of errors was not due to loss of sensitivity for low-prevalence conditions (*d'* for low-prevalence conditions did not decrease). They argued that participants shifted their criterion to say "no" in low-prevalence conditions.

Interestingly, when participants were given more time, the extra time did not influence their accuracy or sensitivity in detecting low-prevalence targets.

These studies indicate that the target prevalence effect is present across different types of attention tasks such that responding in low-probability conditions seems to be hindered in speed, in accuracy, or in both (see Craig, 1978). These studies also suggest that there is a threshold-like shift when participants encounter low-probability targets. Literature on signal-detection theory and vigilance tasks suggest that there are at least two types of accounts (Craig, 1978, 1987). People can fail to detect a low probability target because the expectancy of the target occurring is already low. As a result, the response criterion is altered such that there is a reduction in the likelihood of detection of the next target. A probability matching account suggests that people make an active effort to satisfy some criterion (e.g., to minimize errors) to stabilize response and signal rates (Craig, 1987).

Although Schmidt and colleagues (e.g., Schmidt & Besner, 2008) have emphasized only the benefit of responding to high-contingency trials during colour-word contingency learning, the prevalence effect in the attention literature suggests that there could be an analogous cost to performance for low-contingency trials that is similar to the findings from other attentional tasks. In the first part of my dissertation, I address whether a cost is evident when people encounter low-prevalence targets (10%) in addition to the benefit that they experience for high-prevalence targets (80%).

Does stimulus frequency play a role in contingency learning?

Contingency is the degree of association between two or more events whereas frequency refers to how often those events occur. In the colour-word contingency learning paradigm, high-contingency trials necessarily are also presented more frequently than low-contingency trials.

Thus, it is difficult to determine what the contribution of frequency is in colour-word contingency learning. Is it the total number of presentations of low-probability events and of high-probability events that matters, or is it the specific learned contingencies between the stimuli and between the stimuli and the responses that matters? The second chapter of the dissertation explores this issue, considering several aspects: (1) Could it be that frequency, not contingency, underlies the different responding patterns for nominally high-contingency and low-contingency items? and (2) Do the roles of frequency and contingency change when there is perfect contingency? The second part of this dissertation represents, then, an initial exploration of the relation between frequency and contingency. To determine the role of frequency, I pit frequency against contingency with the goal of determining whether frequency plays an independent role in contingency learning. In addition, I set out to determine how the role of frequency differs when there is a perfect—as opposed to partial—contingency.

The current literature on the colour-word contingency paradigm has only employed the typical paradigm wherein participants are presented with different words in colours and they respond to the colours of the words. In most cases, the contingency is 80% for high-contingency items and 20% for low-contingency items. Typically, all of the words are always presented in all three colours and, thus, responses are not mapped exclusively onto one word (Schmidt et al., 2007). For example, MOVE is presented in blue (high contingency) but (less often) also in green and in yellow (low contingency). Thus, although there is some evidence that the role of frequency cannot explain the colour-word contingency learning, it is unclear whether this is true when there is perfect contingency, that is, when a word (e.g., TABLE) is always associated with a certain response (e.g., "red"), as opposed to just usually (i.e., 80% of the time) associated with a given response.

There is evidence from previous studies that as the prevalence of targets increases there is a shift in response criterion. For example, in a simulated baggage task, very high prevalence rates (98%) elicit an increase in false alarms and a decrease in misses compared to a lower prevalence rate of 50% (Wolfe & Van Wert, 2010), suggesting a trade-off between false alarms and misses depending on the prevalence rate. What was also interesting is that they found that the very frequent targets did not elicit a faster response time but instead, the rare targets elicited a slower response time. Thus, the response criterion and the difference in responding to frequent targets vs. less frequent targets is dependent on whether there are more or fewer target-prevalent trials. This study suggests that a direct mapping of the response to a single word (100% contingency or prevalence) will alter the criterion and influence performance to an even greater extent.

Although there are by now quite a few studies that address colour-word contingency learning, there is only one study that has addressed frequency of colour-word contingency items, and then only to a limited extent. As described earlier, Schmidt et al. (2007, Experiment 4) have argued that in the colour-word contingency paradigm it is not the frequency of stimulus dimension pairings that is learned (i.e., the word being paired with the print colour); rather, it is the connection of the irrelevant dimension (the word) to the response (the key press for the colour).

Although Schmidt et al. (2007) maintained that contingency and not frequency explains colour-word contingency learning, I ask whether this holds true when people are also learning perfect contingencies. As discussed earlier, people are able to shift their criterion depending on the number of high-probability targets that they encounter. If this is the case, it is important to

consider the shift in criterion when there is a perfect contingency, that is, when they do not ever encounter a corresponding low-probability target.

It is in fact possible to equate frequency for high-contingency and low-contingency items, although this requires a rather unique set of conditions—in particular the perfect 100% contingency condition just described. A key question concerning frequency, then, is how much of the contingency effect would remain if, unlike the usual situation where high-contingency items are much higher in frequency than low-contingency items, the frequencies of the two conditions were balanced. In Chapter 2 (Experiments 5 to 7), I aim to examine whether colour-word contingency learning is partially attributable to confounded frequency.

Does meaning play a role in contingency learning?

Colour-word contingency learning happens very quickly. Despite learning being achieved very quickly, it may not all be beneficial: There may also be some costs involved with this type of contingency learning, as investigated in the first set of experiments of this dissertation. If indeed there are costs involved with responding to low prevalence targets, would it be possible for us to compensate for them and to improve our performance, such that the errors are minimized and our performance becomes more efficient?

In real life situations, the failure to find a weapon during a security scan of baggage obviously has potentially devastating consequences. However, because the prevalence of weapons in baggage is very low in frequency, there often are failures to detect these weapons, or misses. Research has examined how different training can improve performance, especially in visual search tasks (Wolfe et al., 2007; Wang, Lin, & Drury, 1997). For example, a promising training strategy is to provide people with search tasks with bursts of high-prevalence trials accompanied by feedback dispersed among low-prevalence trials. Wolfe et al. (2007) have

found that such intermittent bursts can decrease the likelihood of errors on low-prevalence trials. Other studies showed that prior exposure to high probability or low probability targets can influence later responding. The ability to detect low probability targets after encountering high probability targets elicits a greater decrement over time. In addition, false report rates to low probability targets are greater after responding to a set of high probability targets (Colquhoun & Baddeley, 1967).

Training can help improve the efficiency and accuracy in performance. Yet training requires time and is itself an external strategy to compensate for a performance detriment in some domain. What if people could "naturally" compensate for the performance cost of low-contingency items? Humans have an impressive ability to use higher-order associations to help with performance. Indeed, this has been the cornerstone of the priming literature. For example, in the domain of memory (Neely, 1977), when people are being tested on their response time to identify words, responding to a word preceded by a semantically related word (e.g., NURSE preceded by DOCTOR) is facilitated compared to when successive words are semantically unrelated.

Past research has also found that people can learn semantic information that is beyond what is directly presented. Goschke and Bolte (2007) found that people are able to learn abstract sequential structures and that they can do so implicitly. In their study, participants named line drawings of exemplar objects which belonged to one of four semantic categories (animals, clothing, body parts, furniture). Although the objects were presented in random order, there was a repeating semantic category sequence. Participants were able to acquire the implicit knowledge of the category sequence order, and that led to anticipation of naming the upcoming exemplar objects belonging to the next category, and hence to speeded responding. When a

repeating category sequence switched to a random sequence, participants' response times increased, providing evidence that they were retaining implicit knowledge about the category sequence. Moreover, this learning was done implicitly: When given a subsequent explicit recognition test, participants were not able to identify the category sequences. These results provide evidence that we are able to retain sequential relations between categories of stimuli without reference to the features of the exemplars. People can learn higher order associations on tasks that look like simple stimulus-response conditioning: In this case, perceiving and naming an object can implicitly activate a representation of its semantic category, despite the task not requiring processing at the level of categories.

Similarly, Schmidt and De Houwer (2012a) found evidence that the learning of contingencies in a colour-word contingency learning paradigm can extend beyond the simple associative learning of words and colours. In their study, on each trial, participants were presented with a nonword for 250 ms, then a target that appeared for 2000 ms or until a response was made. Participants pressed one key when the target was negative in valence and another key when the target was positive in valence. Target words were either of a positive valence (HUG, FLOWER) or of a negative valence (GUN, CRIME). There were four nonwords and each was presented most often with one target word (high contingency) and equally often with the other three target words (low contingency). On stimulus match trials, a nonword was presented with the most frequent target word (e.g., NIJARON-HUG). On valence match trials, a nonword was presented with the target word of the same valence as its most frequent target word (e.g., NIJARON-FLOWER). On valence mismatch trials, a nonword was presented with the target word in the valence opposite to its most frequent target (e.g., NIJARON-GUN).

Schmidt and De Houwer (2012a) found that valence-mismatch trials were responded to slower than valence match and stimulus match trials. Because there was no difference found between valence match and stimulus match trials, participants evidently did not learn the specific association between the nonword and that target word. Instead, participants learned the semantic association between the nonword and the valence of the word. Furthermore, on a subsequent task when participants explicitly rated how likable each nonword was, Schmidt and De Houwer observed that the ratings of nonwords paired with positive valence targets were themselves more positive than was the case for nonwords associated with negative valence targets. Thus, this study provides some evidence that contingency learning is not limited to specific colour to word associations but can be extended to learning of meaningful relations, in this instance of valence associations.

The storing of information concerning responses, stimuli, and valences in episodic traces that influence performance at retrieval provides some evidence that people are able to use semantic associations and that the use of these higher order associations can influence performance in contingency learning. People seem to be able to extract information from simple associative learning tasks, information that goes beyond the given stimuli. In the last part of the dissertation, I address whether we can observe the use of semantic associations to help with learning of contingencies without any explicit instructions to do so.

The Research Rationale

The principal aims of this dissertation are, first, to examine whether the learning of contingencies is accompanied by costs as well as benefits, second, to explore whether the influences of contingency and frequency can be separated, and third, to determine whether people use higher order associations to facilitate their learning. The colour-word contingency

20

paradigm is an ideal tool to address these questions because learning of high-contingency items and low-contingency items can be compared in a single task. The learning of different contingencies also occurs across trials, which is very different from visual search tasks or detection tasks which examine the contingencies within a single trial. Furthermore, as Schmidt and his collaborators have argued (e.g., Schmidt & De Houwer, 2012c), the learning of contingencies in this paradigm is attained very quickly and below the level of awareness, which can be generalized to everyday learning.

This dissertation consists of three chapters. In the first chapter, I examine whether there are costs involved in the colour-word contingency paradigm in addition to the benefits emphasized in the previous literature. Importantly, the literature on the prevalence effect in visual search reviewed in the introduction suggests that focusing only on the benefit of learning contingencies may provide only half of the picture and thereby fail to capture fully how people learn contingencies. In the studies in this dissertation, based in part on the prevalence effect, it is hypothesized that there is a cost as well as a benefit involved in colour-word contingency learning. The problem is that comparison of a high-contingency condition to a low-contingency condition does not provide a baseline against which to measure costs and benefits separately (see Jonides & Mack, 1984, for a review of the issues involved in cost/benefit analysis). I add a baseline condition to examine whether a cost is present.

In the second chapter, the central question concerns the relative roles of frequency and contingency in the contingency learning task. The type of contingency trial ordinarily is confounded with frequency: High-contingency trials are presented more often than low-contingency trials. As a result, it is difficult to discern the difference between contingency and frequency. Also, past literature on colour-word contingencies have focused on examining the

differences in contingencies with a range of 10% to 80%. All words are always presented in every colour. As a result, there are no situations where a word is only directly associated with one colour. This part of the dissertation is an exploratory one in part to examine whether direct 100% mapping of word and response influences the contingency effect. Of course in this situation there will be a strong, consistent contingency effect. But will frequency also play a role? Intuition certainly suggests that it will. One way to determine this is to pit frequency against contingency, examining the case where high-contingency items have frequencies equal to low-contingency items: Will the size of the contingency effect be reduced or even eliminated when the frequency confound is removed?

Lastly, the broad cognitive literature has emphasized the benefit of semantic associations in facilitating human performance. The question in the third chapter of the dissertation is whether without explicit instructions we can and do use semantic associations to help with the implicit learning of colour-word contingencies. If we are able to use semantic associations, then making meaningful connections available should contribute to learning the contingencies and potentially also to transfer of one learned contingency to a semantically related situation. This chapter is intended to explore the role of higher-order versus stimulus level factors in the learning of contingencies.

General Method

The basic procedures used across studies are sufficiently similar to warrant this section on general method features. The basis of the methodology comes from Schmidt et al. (2010). In their study, when participants performed their first contingency block (with 80% high contingency and 20% low contingency), a contingency effect developed very quickly. Moreover, a small transfer effect was found when participants then switched to a non-contingency block (with all words associated with each colour with equal probability). To implement the basic procedure in our laboratory, and to confirm that contingency learning can carry over to a situation involving no contingencies, Experiment 1 replicated Schmidt et al. (2010). The nine subsequent experiments examined whether (1) there is both a cost and a benefit to contingency learning (Experiments 2, 3, and 4), (2) the benefit in contingency learning is due to contingency, frequency, or both (Experiments 5, 6, and 7), and (3) semantic representations beyond the actual stimuli can influence contingency learning (Experiments 8, 9, and 10).

Participants. All participants were undergraduate students recruited from the University of Waterloo in exchange for credit or for financial remuneration.

Stimuli. In the typical experiment in the dissertation, unless otherwise noted, the irrelevant word dimension consisted of three common 5-letter words, and the three display colours were red, green, and yellow. All stimulus words were presented in lower case at the centre of the screen in Courier New font 16 on a black background.

Apparatus. All experiments were programed using E-prime v2.0 (Psychology Software Tools, Pittsburgh, PA, USA), with all stimuli presented and responses recorded through the program. All displays other than trial stimuli—such as fixation and feedback—were presented in white on a black background in Courier New font size 18 unless otherwise noted.

Procedure. On each trial, participants saw a central fixation cross for 250 ms, followed by a word in lower case in the same location until response or until 2000 ms had elapsed. On trials that were incorrect or too long, feedback consisting of three x's (XXX) was shown in white at the centre of the screen for 1000 ms. Participants used three fingers on their right hand to respond to the colour of the word by pressing keys "j", "k", or "l," one key corresponding to each colour.

Experiment 1: Transfer

To import the colour-word contingency learning paradigm into the lab, and to provide a basis for later studies that examine carryover effects, the first step was to replicate Schmidt et al. (2010). Based on Schmidt's work, I anticipated a quickly established contingency effect when the contingencies were in place, and a diminishing carryover effect of the contingencies learned in the initial blocks when an unannounced switch was made to blocks with no contingencies. Despite cancellation of the contingencies, participants were expected to show an initial difference in responding to high-contingency items versus low-contingency items, but this difference should gradually dissipate without the contingencies being re-established.

Method

Participants. Thirty students from the University of Waterloo completed one session in exchange for course credit. No descriptive data were collected regarding these participants.

Design. Participants completed two phases: a Contingency Phase and a No Contingency Phase. There were 8 blocks of 30 trials each for the Contingency Phase and 8 blocks of 27 trials each for the No Contingency Phase. First, however, participants completed 30 practice trials to familiarize them with the procedure. During practice, three other words (PEDAL, GLOBE, JUDGE) appeared in red, yellow, and green with equal probability.

Contingency Phase. Participants were shown one of three words (i.e., MONTH, UNDER, PLATE) on each trial, the same words as used by Schmidt and De Houwer, (2012c) but in English rather than Dutch. Every word appeared in one colour 80% of the time (e.g., MONTH was presented in yellow 80% of the time, UNDER in green 80% of the time, and PLATE in red 80% of the time). This created the high-contingency condition. Every word also appeared in each of the other two colours 10% of the time, constituting the low-contingency condition. See Table 1 for a description of the stimuli and trials. Word and colour combinations were counterbalanced across participants.

No Contingency Phase. In this second phase, every word appeared in each of the three colours equally often (33%), eliminating the previous contingency. For example, MONTH now appeared in red 3 times, in green 3 times, and in yellow 3 times (see Table 1). All pairings of colours and words were counterbalanced.

Results

Errors and Outliers. First, incorrect responses were removed from the analysis of response times (RTs) (Overall: 4.61% errors; Contingency Phase: 3.25% of high-contingency, 7.50% of low-contingency; No Contingency Phase: 3.19% of high-contingency, 6.16% of low-contingency). A 2 (phase) x 2 (contingency) repeated measures ANOVA was conducted on the error rates. Participants made significantly fewer errors on high-contingency items ($M = 0.032$) than on low-contingency items ($M = 0.068$), $F(1,29) = 30.09$, $MSE = 0.001$, $p < 0.001$, $\eta^2_p = 0.509$. Neither the main effect of phase nor the interaction was significant, $F(1,29) = 2.40$, $MSE = 0.001$, $p = 0.132$, $\eta^2_p = 0.077$, and $F(1,29) = 2.11$, $MSE = 0.001$, $p = 0.157$, $\eta^2_p = 0.068$, respectively.

Next, all response times less than 200 ms were removed as anticipation responses.

Finally, RTs were trimmed using a 2.5 standard deviation criterion (see Van Selst & Jolicoeur,

1994) by subject, phase, and type of contingency. Overall, 2.90% of the correct data were

trimmed (Contingency Phase: 2.80% for high-contingency items, 3.0% for low-contingency

items; No Contingency Phase: 3.06% for high-contingency item, 2.94% for low-contingency

items). These same data handling procedures will be followed in all subsequent experiments: (1)

removal of errors, (2) trimming of very short anticipation responses, and (3) conservative

removal of outliers.

Response Times. The data of principal interest were the RTs, with 2 standard error bars

shown in Figure 1.   A 2 X 2 repeated-measures ANOVA examined the effects of phase

(Contingency, No Contingency) and contingency (high, low) on RTs for colour identification.

The analysis revealed that overall participants responded significantly faster during the

Contingency phase than during the No Contingency phase, $F(1,29) = 15.58$, $MSE = 1093.83$, $p <$

0.001, $\eta^2_p = 0.349$. Furthermore, participants responded reliably more quickly to high-

contingency items than to low-contingency items, $F(1,29) = 41.04$, $MSE = 684.87$, $p < 0.001$, η^2_p

$= 0.586$, an overall contingency effect of 30 ms. Importantly, the interaction between phase and

contingency was also significant, $F(1,29) = 7.58$, $MSE = 299.40$, $p = 0.01$, $\eta^2_p = 0.207$. Using a

paired samples t-test, the contingency effect was found to be larger for the Contingency phase (M

$= 39$ ms) than for the No Contingency phase ($M = 21$ ms), $t(29) = 2.75$, $SE = 6.32$, $p = 0.01$.

Block Analysis. To examine the effect of contingency across blocks, separate 2

(contingency) x 8 (blocks) ANOVAs were conducted for the two phases.

Contingency phase. There was a significant main effect of contingency, with participants

40 ms faster on high-contingency items ($M = 495$ ms) than on low-contingency items ($M = 535$

ms), $F(1,29) = 45.27$, $MSE = 4328.84$, $p < 0.001$, $\eta^2_p = 0.610$. There was also a significant main

effect of blocks, $F(7,203) = 3.80$, $MSE = 3257.95$, $p = 0.001$, $\eta^2_p = 0.116$. There was no reliable

contingency by block interaction, $F(7,203) = 1.47$, $MSE = 1659.73$, $p = 0.179$, $\eta^2_p = 0.048$,

indicating a stable contingency effect across blocks, consistent with rapid learning of the

contingency. Trend analysis revealed that participants showed a linear trend across blocks,

responding overall slower as the blocks progressed, $F(1,29) = 14.24$, $MSE = 4597.36$, $p = 0.001$,

$\eta^2_p = 0.329$. The contingency x linear trend across block interaction was significant, $F(1, 29) =$

3.45, $MSE = 2136.25$, $p = 0.074$, $\eta^2_p = .106$.

   No Contingency phase. There was a significant main effect of contingency, with

participants responding 21 ms faster for high-contingency items ($M = 528$ ms) than for low-

contingency items ($M = 549$ ms), $F(1,29) = 15.57$, $MSE = 3360.65$, $p < 0.001$, $\eta^2_p = 0.349$. The

main effect of blocks was not significant, $F(7,203) = 0.47$, $MSE = 2976.31$, $p = 0.852$, $\eta^2_p =$

0.016. Surprisingly, the interaction of contingency with blocks also was not significant, $F(7,203)$

$= 1.72$, $MSE = 1347.51$, $p = 0.11$, $\eta^2_p = 0.056$, but the contingency x linear trend across block

interaction was significant, $F(1, 29) = 6.54$, $MSE = 1274.93$, $p = 0.016$, $\eta^2_p = .184$, declining

from an initial 32 ms to a final 10 ms effect.

Discussion

   Participants responded faster to high-contingency items than to low-contingency items in

the first phase, when the contingency was present. Moreover, they learned the contingency

information very quickly—it appeared full blown in the very first block and did not change over

the 8 blocks. This pattern replicates that reported by Schmidt et al. (2010). In the second phase,

where the contingency was cancelled and all items occurred with equal probability, there was

clear evidence of transfer: Participants still responded to high-contingency items faster than they

did to low-contingency items. This pattern is largely consistent with the findings of Schmidt et al. (2010), supporting the conclusions that these contingencies are quickly acquired and do transfer to sequences without contingencies, indicating that these contingencies are gradually unlearned once they have been cancelled. However, how quick versus gradual that unlearning is may not be entirely consistent over studies, comparing the present data to those of Schmidt et al. (2010).

Chapter I: Costs and Benefits in Contingency Learning

Experiment 2: Baseline

The standard assumption is that the contingency effect is due to speeding of the high-contingency items with practice. As Schmidt and Besner (2008, p. 522) put it "contingency effects are solely facilitative in response times." Of course, it is in fact difficult to determine whether the effect is due to a benefit, such that responses to high-contingency items are faster, or to a cost, such that responses to low-contingency items are slower—or both. To address whether there is a benefit or a cost or both within the contingency effect, a baseline (neutral, no-contingency control condition) is necessary against which to measure cost and benefit deviations (cf. Jonides & Mack, 1984). Of all the prior studies, only that of Schmidt and Besner has included a no contingency condition, and that study was unique in only using two response colours. The fundamental question of costs and benefits therefore deserves further attention.

The purpose of Experiment 2 was to investigate relative costs and benefits by incorporating a baseline. To accomplish this, a word that appeared in each of the three colours with equal probability (i.e., there was no contingency) was included. Participants should as usual be able to learn the contingencies, responding faster to words presented in their high-contingency colour and slower to words presented in one of their low-contingency colours. Then, by comparison to the baseline (no contingency) item as a control, deviations that represent speeding up or slowing down can be differentiated. The expectation is that the intuitive benefit will be observed, so the key question relates to whether there will also be a cost. Based on previous findings that rare occurrences of targets can hurt performance (Hon et al., 2013; Wolfe, Horowitz, Van Wert, Kenner, Place, & Kibbi (2007), it is reasonable to anticipate a cost where responding is slower to low-contingency items than to the baseline item.

Method

      Participants. Thirty-one participants from the University of Waterloo (*M* age = 19.43

years, 25 female, 6 male) completed one session in exchange for course credit. All had normal

or corrected-to-normal vision.

      Materials and Procedure. There were 30 practice trials and 8 blocks of 48 experimental

trials. On each practice trial, a row of asterisks (*****) was presented with an equal probability

of occurring in each of the three colours (yellow, green, red). In the experimental phase, four

words (MONTH, PLATE, UNDER, CLOCK) were presented in three colours: red, yellow, and

green. The proportions of three of the words were manipulated such that each word appeared in

one colour 83.33% of the time (high contingency) and in the each of the other two colours 8.33%

of the time (low contingency). For example, in a block of 48 trials, MONTH appeared in the

colour red 10 times (high contingency) and in green once and in yellow once (both low

contingency). In addition, one word was presented in each of the colours with equal probability

(33.33%). For example, CLOCK appeared in yellow, green, and red 4 times each in a block of

48 trials. This equal-probability word acted as the baseline control (no contingency). All of the

stimuli and trial types are displayed in Table 2. All pairings of colours and words were

counterbalanced across participants. The contingencies stayed the same across all 8

experimental blocks. Otherwise, the procedures were identical to those of Experiment 1, except

that there was no transfer No Contingency phase.

Results

      Errors and Outliers. First, incorrect responses were removed from the analysis of RTs.

Overall, 2.97% of responses were errors. Mean error proportions were 2.38% for high

contingency, 4.10% for low contingency, and 3.90% for no contingency. A repeated-measures

ANOVA conducted on the three types of contingency revealed a significant main effect, $F(1.59,$ $47.54) = 6.20$, $MSE = 0.001$, $p = 0.007$, $\eta^2_p = 0.171$, with a Greenhouse-Geisser correction. Three pairwise t-tests were conducted using the Bonferroni correction with an adjusted alpha level of 0.017 per test (0.05/3). Errors were fewer for high-contingency items than for low-contingency items, $t(30) = 2.95$, $SE = 0.01$, $p = 0.006$, or for baseline items, $t(30) = 4.06$, $SE = 0.004$, $p < 0.001$. Errors were not significantly different for low-contingency items and baseline items, $t(30) = 0.33$, $SE = 0.01$, $p = 0.746$.

Next, all response times less than 200 ms were removed as anticipation responses. Finally, RTs were trimmed using a 2.5 SD criterion by subject and type of contingency. Overall, 3.01% of the correct data were trimmed (2.56% of high contingency items, 4.56% of low contingency items, and 3.38% of baseline items).

Response Times. Figure 2 presents the RT data. A repeated measures ANOVA was conducted with three levels of contingency (high, low, baseline). There was a significant main effect of contingency, $F(2,60) = 21.85$, $MSE = 297.60$, $p < 0.001$, $\eta^2_p = 0.421$. Three follow-up t-tests were conducted using the Bonferroni adjusted alpha levels of 0.017 per test (0.05/3). High-contingency items ($M = 517$ ms) were responded to 29 ms faster than low-contingency items ($M = 546$ ms), $t(30) = 6.19$, $SE = 4.60$, $p < 0.001$, replicating the typical contingency effect. There was also a benefit of high contingency, such that participants were 19 ms faster responding to high-contingency items than to no-contingency baseline items ($M = 536$ ms), $t(30) = 5.63$, $SE = 3.36$, $p < 0.001$. Although the difference between responses to low-contingency items and no-contingency baseline items was only marginally significant, $t(30) = 1.90$, $SE = 5.02$, $p = 0.067$, the low-contingency items were 10 ms slower than the no-contingency baseline items, consistent with a cost.

Block Analysis. To address whether there were any trends across blocks for the different contingency conditions, a 3 (contingency) x 8 (blocks) repeated measures ANOVA was conducted. Of course, given the preceding analysis, there was a main effect of contingency, $F(1.68,50.42) = 23.00$, $MSE = 2965.66$, $p < 0.001$, $\eta^2_p = 0.434$, with a Greenhouse-Geisser correction. But the main effect of block was not significant, $F(7,210) = 1.86$, $MSE = 3669.31$, $p = 0.078$, $\eta^2_p = 0.058$, and most important there was no suggestion of a contingency X block interaction, $F < 1$. Put simply, the pattern across blocks was very consistent, as Figure 2 shows.

Discussion

Participants responded faster to high-contingency items relative to baseline no-contingency items, reflecting a benefit of high contingency. This is consistent with the findings of Schmidt and Besner (2008). But particularly in later blocks there also was a divergence of the low-contingency items from the baseline no-contingency items, indicative of a cost. Although not significant overall in this experiment, this cost consistently recurs in subsequent experiments, so I take it to be real. One possibility for why a cost is evident primarily on later blocks is that both low-contingency and no-contingency stimuli are quite infrequent per block. Thus, it may take longer for participants to differentiate the low-contingency items from the baseline stimuli whereas the much more frequent high-contingency stimuli are readily differentiated from both early on.

Experiment 3: Baseline Transfer

To explore further the contribution of cost to the contingency effect, Experiment 3 examined how the learning in an initial contingency phase affected performance when there was a switch to a no-contingency phase in which all items appeared with equal probability. The existence of a no-contingency baseline item in the learning phase provided a bridge between the two phases. I expected there to be some carryover of the benefit and that this would decline over blocks, but I was particularly interested in the potential cost and how it would behave as a function of transfer from contingency to no contingency and back again.

Method

Participants. Thirty participants (mean age of 20.03 years, excluding one participant who failed to report their age; 26 females, 4 males) from the University of Waterloo completed one session in exchange for course credit. All had normal or corrected-to-normal vision.

Apparatus. All stimuli were presented in 18-point Courier New font on a black background.

Design. In the initial practice phase, there were 48 practice trials during which participants were shown a row of asterisks (*****) in red, yellow, or green, with equal probability. Following practice, there were three main phases: Contingency Phase I, No Contingency Phase, and Contingency Phase II. There were 6 blocks of 48 trials in each of these phases.

Contingency Phases I and II: Participants were shown one of four words (i.e., MONTH, UNDER, PLATE, CLOCK) on each trial; these were the same words as in Experiment 2. Three of the words appeared in one colour 83.33% of the time (e.g., MONTH was presented in yellow 83.33% of the time, UNDER in green 83.33% of the time, and PLATE in red 83.33% of the

33

time). This created the high-contingency condition. The same three words also appeared in the

other two colours a total of 16.67% of the time, constituting the low-contingency condition. The

last word (e.g., CLOCK) served as a no-contingency baseline, appearing in each of the three

colours with a probability of 0.33. Table 3 presents the stimuli and trials. All words and colours

were counterbalanced across participants.

No Contingency Phase: In the No Contingency Phase, all four words appeared in each of

the three colours equally often (33%), eliminating the previous contingencies; Table 3 displays

the stimuli. All pairings of colours and words were counterbalanced.

Results

Errors and Outliers. First, incorrect responses were removed from the analysis of RTs.

Overall, there were 4.76% errors. To address whether there was a difference in errors in the

phases, three *t*-tests were conducted using the Bonferroni correction with an adjusted alpha level

of 0.017 per test (0.05/3) for each phase. For Contingency Phase I, high-contingency items ($M =$

3.50%) had fewer errors than low-contingency items ($M = 5.65\%$), $t(29) = 3.26$, $SE = 0.01$, $p =$

0.003. There was no difference in errors for high-contingency items and baseline items ($M =$

4.40%), $t(29) = 1.49$, $SE = 0.01$, $p = 0.147$, nor for low-contingency items and baseline items,

$t(29) = 1.72$, $SE = 0.01$, $p = 0.10$. For the No Contingency Phase, there were no differences in

errors over the 3 conditions (high: 4.63%; low: 5.70%; baseline: 5.74%, $t(29) = 3.42$); high vs

low $t(29) = 1.65$, $SE = 0.01$, $p = 0.109$, high vs baseline, $t(29) = 1.86$, $SE = 0.004$, p = 0.073; and

low vs baseline, $t(29) = 0.08$, $SE = 0.01$, $p = 0.940$. For Contingency Phase II, high-contingency

items ($M = 4.22\%$) had fewer errors than low-contingency items ($M = 6.48\%$), $t(29) = 3.42$, $SE =$

0.01, $p = 0.002$, and more errors than baseline items ($M = 1.40\%$), $t(29) = 7.60$, $SE = 0.004$, $p <$

0.001. In addition, low-contingency items had more errors than baseline items, $t(29) = 6.31$, *SE* = 0.01, $p < 0.001$.

Next, all response times less than 200 ms were removed as anticipation responses. Finally, RTs were trimmed using a 2.5 SD criterion by subject and type of contingency. For Contingency Phase I, trimming resulted in removing 3.07% of high-contingency items, 2.65% of low-contingency items, and 2.90% of baseline items; for the No-Contingency Phase, trimming resulted in removal of 3.01% of high-contingency items, 3.10% of low-contingency items, and 3.54% of baseline items; for Contingency Phase II, trimming resulted in removal of 3.17% of high-contingency items, 2.68% of low-contingency items, and 2.80% of baseline items.

Response Times. Figure 3 presents the RT data. To examine the contingency effect, a separate repeated measures ANOVA on the three types of contingency (high, low, and baseline) was conducted for each of the phases.

Contingency Phase I. There was a main effect of contingency, $F(2,58) = 11.00$, *MSE* = 556.11, $p < 0.001$, $\eta^2_p = 0.275$. Three *t*-tests were conducted using the Bonferroni correction with an adjusted alpha level of 0.017 per test (0.05/3). High-contingency items (*M* = 539 ms) were responded to 29 ms faster than low-contingency items (M=568 ms), $t(29) = 4.03$, *SE* = 7.08, $p < 0.001$, and 15 ms faster than baseline items (M=554 ms), $t(29) = 3.18$, *SE* = 4.76, $p = 0.003$. Although low-contingency items were slower than baseline items, this 14 ms difference was not significant given the Bonferroni correction, $t(29) = 2.16$, *SE* = 6.20, $p = 0.039$, although entirely consistent with the significant costs otherwise observed in Experiment 2-4.

No Contingency Phase. There was no main effect of contingency, $F < 1$. Surprisingly, given the results of Experiment 1 and of Schmidt et al. (2010), there was no evidence of transfer of the contingencies learned in Contingency Phase 1 to the No Contingency Phase. There was no

contingency x linear trend across block interaction, $F(1,29) = 0.012$, $MSE = 1300.24$, $p = 0.915$, $\eta^2_p < 0.001$.

Contingency Phase II. The contingencies learned in Contingency Phase I did return immediately, though, once they were re-established in Contingency Phase II. There was a main effect of contingency, $F(1.51, 43.89) = 18.37$, $MSE = 647.03$, $p < 0.001$, $\eta^2_p = 0.388$, with Greenhouse-Geisser correction. Three further *t*-tests were conducted using the Bonferroni correction with an adjusted alpha level of 0.017 per test (0.05/3). High-contingency items ($M = 555$ ms) were 35 ms faster than low-contingency items ($M = 590$ ms), $t(29) = 4.82$, $SE = 7.13$, $p < 0.001$, and were also 14 ms faster than baseline items ($M = 569$ ms), $t(29) = 3.01$, $SE = 4.57$, $p = 0.005$. Low-contingency items were 21 ms slower than baseline items, $t(29) = 4.04$, $SE = 5.11$, $p < 0.001$.

Block Analysis. To address whether there were any trends over blocks for each type of contingency, separate 3 (contingency) X 6 (blocks) repeated measures ANOVAs were conducted for each of the three phases.

Contingency Phase I. There was a main effect of contingency, $F(1.66, 48.09) = 10.44$, $MSE = 4387.50$, $p < 0.001$, $\eta^2_p = 0.265$. There was no effect of block, $F(5,145) = 1.41$, $MSE = 3904.51$, $p = 0.224$, $\eta^2_p = 0.046$, and there was no interaction between contingency and block, $F(5.42, 157.19) = 0.93$, $MSE = 3746.44$, $p = 0.470$, $\eta^2_p = 0.031$. The *t*-tests examining the main effect of contingency were reported above.

No Contingency Phase. There was a main effect of block, $F(5,145) = 2.29$, $MSE = 2908.60$, $p = 0.05$, $\eta^2_p = 0.073$, and a linear trend over blocks, $F(1,29) = 8.11$, $MSE = 3772.73$, $p = 0.008$, $\eta^2_p = 0.219$. There was no main effect of contingency, $F < 1$, and no interaction between contingency and block, $F(10, 290) = 1.06$, $MSE = 1253.32$, $p = 0.397$, $\eta^2_p = 0.035$.

Contingency Phase II. There was a main effect of contingency, $F(1.48, 42.88) = 16.49$, $MSE = 4348.40$, $p < 0.001$, $\eta^2_p = 0.363$. The *t*-tests for type of contingency were reported above. There was no main effect of block, $F < 1$, nor an interaction of contingency with block, $F(5.74, 166.57) = 1.77$, $MSE = 4220.48$, $p = 0.112$, $\eta^2_p = 0.057$.

Discussion

Once again, the contingency effect was found to include both a benefit, such that responding was faster to high-contingency items than to baseline items, and a cost, such that responding was slower to low-contingency items than to baseline items. These effects both were apparent in Contingency Phase I and in Contingency Phase II. To strengthen the conclusion that the contingency effect contains both components—costs as well as benefits—I conducted a further analysis treating experiment (Experiment 2 vs 3) as a factor, combining the two contingency phases in Experiment 3. These all were basically identical in procedure.

A 3 (Contingency: high, low, baseline) X 2 (Experiment 2, Experiment 3) ANOVA was conducted. There was a main effect of contingency, $F(1.72, 101.47) = 29.53$, $MSE = 493.84$, $p < 0.001$, $\eta^2_p = 0.334$. Three *t*-tests were then conducted using the Bonferroni adjusted alpha levels of 0.017 per test (0.05/3). Responses to high-contingency items ($M = 528$ ms) were 29 ms faster than to low-contingency items ($M = 557$ ms), $t(60) = 6.86$, $SE = 4.16$, $p < 0.001$, demonstrating the basic contingency effect. There was a benefit such that responses were 17 ms faster to high-contingency items than to baseline items ($M = 545$ ms), $t(60) = 5.92$, $SE = 2.88$, $p < 0.001$. Critically, there was a reliable cost such that responses were 12 ms slower to low-contingency items than to baseline items, $t(60) = 2.90$, $SE = 3.95$, $p = 0.005$.

It is important to note that there was no speed-accuracy trade-off: More errors were produced on low-contingency items than on baseline items, indicating that there is a consistent

cost—not only in response time but also in accuracy. The overall findings suggest that, just as in visual search tasks, learning of rare occurrences of targets can suffer (Wolfe et al., 2007). The high-contingency items are handled more efficiently, but at the same time the low-contingency items are handled less efficiently.

There was one puzzling result in Experiment 3, in contrast to Experiment 1. For reasons that are not clear, this time there was no evidence of any carryover effect in the No Contingency Phase: The contingency effect simply disappeared immediately in the first block of transfer. This certainly highlights the idea that we can unlearn contingencies very quickly (Schmidt et al., 2007), although it remains to be determined why these effects sometimes briefly persist and sometimes do not when the contingencies have been cancelled. One possibility (admittedly not tested here) is that the inclusion of a no-contingency item among the high-contingency and low-contingency items in Experiment 3—the sole difference between Experiments 3 and 1—is somehow crucial, providing a bridge that was missing in Experiment 1. This idea would be worthy of empirical test in the future. For the present, the primary purpose of Experiment 3 has been achieved in unambiguously demonstrating a cost component as well as a benefit component to the contingency effect.

Experiment 4: Baseline Nonwords

In Experiments 2 and 3, there was a benefit in responding to high-contingency items and a cost in responding to low-contingency items. In Experiment 4, I set out to determine whether familiar, meaningful stimuli are necessary for the contingency effect to develop, and whether both costs and benefits would still be observed for unfamiliar, non-meaningful stimuli. Past literature on the Stroop effect has shown that nonwords cause less interference than words (Klein, 1964; see MacLeod, 1991, for a review). No research in the colour-word contingency learning paradigm has examined the influence of nonwords. If nonwords show both a benefit and a cost, then semantic processing of the stimuli is not necessary and the contingency effect would appear to rely on more primitive learning. This would also generalize the materials for which human contingency learning has been demonstrated.

Method

Participants. Thirty-one participants (*M* age = 20.77 years, 25 female, 6 male) from the University of Waterloo completed one session in exchange for course credit. All had normal or corrected-to-normal vision.

Materials and Procedure. The method and procedures were identical to those of Experiment 2 except that pronounceable nonwords were substituted for words. The four nonwords used were: FLABE, THROG, DWIPS, and BRASK. Table 4 displays the stimuli and trials.

Results

Errors and Outliers. First, incorrect responses were removed from the analysis of RTs. Overall, 3.17% of responses were errors. Mean error proportions were 2.31% for high contingency, 6.25% for low contingency, and 3.76% for no contingency (baseline). A repeated-

measures ANOVA conducted on the three types of contingency revealed a significant main effect, $F(1.27, 38.07) = 14.96$, $MSE = 0.001$, $p < 0.001$, $\eta^2_p = 0.333$, with a Greenhouse-Geisser correction. Three pairwise t-tests were conducted using the Bonferroni correction with an adjusted alpha level of 0.017 per test (0.05/3). There were fewer errors for high-contingency items than for low-contingency items, $t(30) = 4.54$, $SE = 0.09$, $p < 0.001$, or for baseline items, $t(30) = 4.05$, $SE = 0.004$, $p < 0.001$. There also were fewer errors for baseline items than for low-contingency items, $t(30) = 2.95$, $SE = 0.008$, $p = 0.006$.

Next, all response times less than 200 ms were removed as anticipation responses. Finally, RTs were trimmed using a 2.5 SD criterion by subject and type of contingency. Overall, 2.81% of the correct data were trimmed (2.70% of high-contingency items, 2.87% of low-contingency items, and 3.07% of baseline items).

Response Times. Figure 4 presents the RT data. A repeated measures ANOVA was conducted with three levels of contingency (high, low, and baseline). There was a significant main effect of contingency, $F(1.66, 49.64) = 32.08$, $MSE = 519.95$, $p < 0.001$, $\eta^2_p = 0.517$, with Greenhouse-Geisser correction. Three follow-up t-tests were conducted using the Bonferroni adjusted alpha levels of 0.017 per test (0.05/3). High-contingency items were responded to faster than low-contingency items, $t(30) = 6.91$, $SE = 6.05$, $p < 0.001$, replicating the typical contingency effect. There was also a benefit of high contingency, such that participants were faster responding to high-contingency items than to no-contingency baseline items, $t(30) = 6.57$, $SE = 3.93$, $p < 0.001$. And again low-contingency items were slower than no-contingency baseline items, $t(30) = 2.86$, $SE = 5.58$, $p = 0.008$, consistent with the cost observed in the preceding experiments in this chapter.

Block Analysis. To address whether there were any trends across blocks for the different contingency conditions, a 3 (contingency) x 8 (blocks) repeated measures ANOVA was conducted. Of course, given the preceding analysis, there was a main effect of contingency, $F(1.66, 49.93) = 32.49$, $MSE = 4233.97$, $p < 0.001$, $\eta^2_p = 0.520$, with a Greenhouse-Geisser correction. There was also a main effect of block, $F(7,210) = 6.95$, $MSE = 3744.64$, $p < 0.001$, $\eta^2_p = 0.188$, and a contingency by block interaction, $F(6.74, 202.15) = 2.14$, $MSE = 4180.41$, $p = 0.043$, $\eta^2_p = 0.067$. The source of the interaction likely is the "bumpiness" over blocks in the low-contingency condition, which is not readily interpretable. It is worth noting that, in every condition, there was a main effect of block accompanied by a linear trend across blocks; I report just the linear trend analyses here: for high-contingency items, $F(1,30) = 14.05$, $MSE = 1373.84$, $p = 0.001$, $\eta^2_p = 0.319$; for low-contingency items, $F(1,30) = 17.53$, $MSE = 3461.09$, $p < 0.001$, $\eta^2_p = 0.369$; and for baseline items, $F(1,30) = 12.41$, $MSE = 2447.78$, $p = 0.001$, $\eta^2_p = 0.293$. As we have seen in many of the contingency learning experiments carried out in our lab, participants characteristically slow down across blocks, perhaps tiring of the task.

Discussion

As in all of these experiments, participants responded faster to high-contingency items than to baseline items, reflecting a benefit of high contingency. But there also was a significant cost for responding to low-contingency items relative to no-contingency items. The use of nonwords in this experiment confirms that familiar stimuli such as words are not necessary for contingency learning. People can learn contingencies with non-meaningful units and the contingency learning appears to be as quick and robust as it is with meaningful units. This finding is also relevant to the issue of using meaning in contingency learning, to which I will

return in the third chapter. For the present, it solidifies that both costs and benefits contribute to the contingency learning effect, generalizing this conclusion across materials.

Chapter I Discussion

Humans can acquire contingency information very quickly (Gehring et al., 1992; Schmidt et al., 2007). In the past, this learning typically has been attributed to one component—a benefit due to speeded processing of high-contingency items (e.g., Schmidt et al., 2007). But in fact isolating benefits vs costs has not been possible in most previous studies (apart from Schmidt & Besner, 2008) because there was no no-contingency baseline against which to measure cost/benefit deviations. In three experiments in this chapter (Experiments 2, 3, and 4) in which a baseline no-contingency condition was inserted, I have shown that, in addition to the benefit for high-contingency items, there is also a reliable cost for low contingency items.

The only previous study to address this issue by including a no-contingency baseline was that of Schmidt and Besner (2008), and their data supported only a benefit. What might explain this discrepancy? The only possibility that is readily apparent is that they used only two responses (two colours) whereas the studies reported in chapter 1 all used three responses/colours. Exactly why this procedural difference would produce the different pattern is not clear, however, and another possibility is simply that their result was a Type II error. The cost has been consistently present over a series of three experiments.

Although not entirely consistent, one observation in the present experiments is that the divergence between the low-contingency items and the baseline items seems to emerge somewhat slowly over the first couple of blocks. It is possible that the slowing of performance on low-contingency items is due to the small number of presentations of these items in each block (only 16.67% of all trials), making it hard for people to learn them—and to differentiate them from the baseline trials, which are also relatively infrequent (only 33.33% of all trials) compared to the much more frequent high-contingency items (83.33% of all trials). Also, the

low-contingency items may be increasingly surprising as trials progress. Consistent with this analysis, it is interesting to note that in Contingency Phase II of Experiment 3 this pattern was not apparent, possibly because Contingency Phase I provided sufficient experience.

There are studies in the search literature that illustrate this difference in attending to low-probability vs high-probability targets. Hon et al. (2013) found that people are slower at detecting low probability targets and, based on this, proposed that perceptual templates of low and high probability targets are initially activated at the same level. However, the activation of the templates for low probability targets falls due to their infrequent occurrence. Thus, there is greater perceptual evidence needed for the threshold to be reached for low probability targets, contributing to the overall increased time to detect them. They also proposed that this pattern could be due to a criterion shift from a more liberal criterion to a more conservative criterion for low probability targets. Notice that the Hon et al. idea is just the converse of the idea proposed by Schmidt and Besner—that the changes occur in the thresholds for the high-probability targets and not for the low-probability targets.

A threshold account similar to that of Hon et al. (2013) has been put forward in research on missing rare targets in visual search tasks (Wolfe, Horowitz, & Kenner, 2005). People slow down after making mistakes but speed up after successes, suggesting that they shift their decision criterion depending on whether they encounter high-prevalence or low-prevalence targets. Participants become more conservative in their response criterion when they encounter more low-prevalence targets. When targets are frequent, "no" response times are slower than "yes" responses because "no" responses will more likely lead to a mistake. Correspondingly, for low prevalence targets, "no" responses are faster and "yes" responses are slower. Wolfe et al. argued

that people are constantly adjusting their criterion and that the distribution of target probabilities affects how conservative the response criterion is.

The conclusion of this chapter is that low-contingency items are responded to more slowly than baseline items: There is a cost as well as a benefit in contingency learning. Thus the typical contingency effect, the difference between high contingency and low contingency, would appear to derive from the sum of two components—faster responding to high-contingency items coupled with slower responding to low-contingency items, something that can only be observed by incorporation of a no-contingency baseline condition.

Chapter II: Role of Frequency in Contingency Learning

Experiment 5: Frequency vs. Contingency I

A possible explanation for why the cost is harder to observe than the benefit is that the low-contingency items take some time to learn because of their relatively lower frequency compared to high-contingency items. As just discussed, according to Hon et al. (2013), in a task where accuracy remains very high, high-probability targets become invariant across training whereas low probability targets get slower. They argued that this may be due to the rare occurrence of the low-probability targets. The purpose of the experiments in this second chapter is to determine whether there are differences in performance for high-probability targets versus low-probability targets—in other words, to address the role that frequency of presentation may play in the contingency effect.

An additional question motivating the experiments of Chapter 2 concerned how the roles of frequency and contingency change when there is a direct mapping of the word and the colour response—when the contingency is 100%. In an event where a specific word only appears in a certain colour that elicits a certain response (e.g., MONTH appears only in red), does the size of the contingency effect change? To pursue the role of frequency, the 100% contingency case provides an important tool, so its own effect must also be considered.

Experiment 5 crossed frequency with contingency to examine two conditions: high contingency-low frequency (HiC-LoF) and low contingency-high frequency (LoC-HiF). In the HiC-LoF condition, a given word was associated with one particular colour but that word itself occurred with relatively low frequency. In the LoC-HiF condition, each word appeared in multiple colours but now the word occurred with relatively high frequency. If frequency plays a role in contingency learning, then increasing the frequency of low-contingency items and

reducing the frequency of high-contingency items should reduce the contingency effect. Indeed, as Table 5 shows, in the present case, individual high-contingency items appeared equally as often as individual low-contingency items (i.e., MONTH appeared only in red 6 times; UNDER appeared in yellow 6 times and in green 6 times). Would equating the frequency of high-contingency and low-contingency items diminish or eliminate the contingency effect? Put another way, does the contingency effect rely on a frequency confound?

Method

      Participants. Thirty-two students (*M* age = 19.81 years, 23 female, 9 male) from the University of Waterloo completed one session in exchange for course credit. All had normal or corrected-to-normal vision.

      Materials and Procedure. There were 30 practice trials and 8 blocks of 30 experimental trials. A row of asterisks (*****) was presented on each of the practice trials and occurred with equal probability in each of three colours (yellow, green, red). In the main experiment, three words (MONTH, PLATE, UNDER) were presented in three colours: red, yellow, and green. Instead of the typical high-contingency and low-contingency conditions, there were two conditions: HiC-LoF and LoC-HiF. In the HiC-LoF condition, one word was shown in one colour 100% of the time (e.g., MONTH appeared six times in a block, always in red). The words in the HiC-LoF condition were high in contingency because they always appeared only in one colour but they were also low in frequency because they appeared relatively few times per block (6 occurrences in a block of 30 trials). In contrast, for the LoC-HiF condition, the words were low in contingency because they occurred in the other two colours with 50% probability but high in frequency because they occurred more often (e.g., UNDER appeared in yellow 6 times and in green 6 times, for a total of 12 occurrences in a block of 30 trials). Ultimately, then, the

47

frequencies of all items, whether high-contingency or low-contingency, were the same. Table 5 displays the items and conditions. All pairings of colour and words were counterbalanced. The contingency stayed the same across all 8 blocks.

Results

      Errors and Outliers. First, incorrect responses were removed from the analysis of RTs. Overall, there were 4.13% errors (3.97% HiC-LoF; 4.17% LoC-HiF). A paired samples *t*-test demonstrated no significant difference in errors between the conditions, $t(31) = 0.37$, $SE = 0.01$, $p = 0.715$.

      Next, all response times less than 200 ms were removed as anticipation responses. Finally, RTs were trimmed using a 2.5 SD criterion by subject and type of contingency. Overall, 2.95% of responses were trimmed (2.64% of HiC-LoF items, 3.02% of LoC-HiF items).

      Response Times. Figure 5 displays the RT data. A 2 (contingency) x 8 (blocks) repeated measures ANOVA was conducted. For the first time in this series of experiments, when there was a contingency present, there was no main effect of contingency; that is, participants did not respond to high-contingency items (M = 525 ms) faster than to low-contingency items (M = 518 ms), $F(1,31) = 1.34$, $MSE = 3954.97$, $p = 0.257$, $\eta^2_p = 0.041$. Thus, there also was no effect of frequency apparent. There was also no main effect of blocks, $F(4.72, 146.26) = 1.13$, $MSE = 4067.11$, $p = 0.347$, $\eta^2_p = 0.035$, nor any interaction of contingency and blocks, $F(5.03,155.85) = 0.95$, $MSE = 2177.45$, $p = 0.451$, $\eta^2_p = 0.030$, both with the Greenhouse-Geisser correction, indicating that performance was consistent across blocks, showing no effect of either contingency or frequency.

Discussion

Impressively, the contingency effect was eliminated when the frequency of the colour-word association for high-contingency items was reduced and the frequency of the colour-word association for low-contingency items was correspondingly increased. It is noteworthy that the way this was accomplished in Experiment 5 resulted in any given low-contingency word occurring in a particular colour with the same frequency as any given high-contingency word, equating word-colour stimulus probability across the two conditions. This suggests that frequency does play a role in the contingency effect—that part of the reason for high-contingency items being responded to more quickly is that they simply are experienced more often across trials.

In this experiment, the words in the low-contingency conditions were never in the same colour as those in the high-contingency condition, thus participants could learn that high contingency words were only associated with one colour. This experiment indicates that equating the frequencies of the low-contingency word-colour pairings with those of the high-contingency word-colour pairings eliminated the effect of contingencies. This leads to the further suggestion that it is the individual stimuli that are being learned, not the conditions that they form, an idea that has been influential in the Stroop literature as well (Melara & Mounts, 1993; Sabri, Melara, & Algom, 2001; Melara & Algom, 2003). When a particular high-contingency stimulus occurs as often as a particular low-contingency stimulus, there is no contingency effect because, in keeping with instance theory (Logan, 1988) and therefore in line with Schmidt's (2013) PEP model, the number of instances of each of these stimuli in memory is identical.

Experiment 6: Frequency vs. Contingency II

Experiment 5 suggested that frequency can and does play a key role in the contingency effect. In Experiment 6, a converging method was used to determine whether frequency and contingency both influence what has previously been seen as purely a contingency effect. Here, the two types of contingency items (high, low) were divided into six conditions: HiC-MedF (high contingency-medium frequency), HiC-LoF (high contingency-low frequency), HiC-ExLoF (high contingency-very low frequency), LoC-HiF (low contingency-high frequency), LoC-MedF (low contingency-medium frequency), and LoC-LoF (low contingency-low frequency).

With this set of conditions, either contingency can be held constant to examine the effect of frequency or frequency can be held constant to examine the effect of contingency. For example, to determine whether there is a contingency effect without a frequency effect, response times from HiC-MedF and LoC-MedF can be compared; the same can be done for HiC-LoF and LoC-LoF. In essence, this holds the frequency constant to examine the contingency effect. Analogously, to determine whether there is a frequency effect without a contingency effect, response times from LoC-HiF and LoC-LoF can be compared; the same can be done for HiC-MedF and HiC-LoF. To determine whether there is a frequency effect, this holds the contingency constant.

Table 6 illustrates the stimuli and conditions. This method permits addressing two questions: (1) what is the role of contingency when frequency is held constant, and (2) what is the role of frequency when contingency is held constant? Based on the findings of Experiment 5, if the manipulation of frequency was the key to eliminating the contingency effect, then the expectation was that there would be separable effects of both frequency and contingency.

Method

      Participants. Forty students (*M* age = 19.93 years; 27 female, 13 male) from the University of Waterloo completed one session in exchange for course credit. All had normal or corrected-to-normal vision.

      Apparatus. Except for the critical stimulus words that were presented in yellow, green, or red, all other materials were presented in white at the centre of the screen on a black background in the Courier New font size 18.

      Materials and Procedure. There was one practice block of 30 trials, then 16 experimental blocks of 42 trials each. A row of asterisks (*****) presented on the practice trials had an equal probability (33%) of occurring in each of the three colours (yellow, green, red).

      For the experimental trials, six words (MONTH, PLATE, UNDER, CLOCK, TABLE, WORLD) were presented in three colours: red, yellow, and green. As usual, there were two possible contingencies: high and low. There also were three possible relative frequencies: high, medium, and low. This resulted in a total of 6 different possible conditions: HiC-MedF, HiC-LoF, HiC-ExLoF, LoC-LoF, LoC-HiF, and LoC-MedF. The four condition of interest were HiC-LoF, HiC-MedF, LoC-LoF, LoC-MedF. The HiC-ExLoF and the LoC-HiF were eliminated from analysis; these two conditions were only included to balance the occurrences of words in the overall experiment, as shown in Table 6.

      Frequency. In each study block, words presented with high frequency were presented 16 times, those presented with medium frequency were presented 8 times, those presented with low frequency were presented 4 times, and those presented with extra-low frequency were presented twice. Note that the extra-low frequency and the high frequency conditions were included

simply to balance overall total presentations of words/colours between conditions, and consequently are not incorporated in the data analysis.

Contingency. When a word appeared in high contingency, it was presented in one colour 100% of the time (e.g., MONTH appeared in red only, UNDER appeared in yellow only, PLATE appeared in green only). When words appeared in low contingency, they appeared in two colours with equal (50%) probability (e.g., CLOCK appeared equally often in red or green, TABLE in green or yellow, and WORLD in red or yellow).

Results

Errors and Outliers. First, incorrect responses were removed from the analysis of RTs. Overall, there were 3.80% errors; these broke down over the 6 conditions as follows: HiC-LoF: 3.40%; HiC-MedF: 2.66%; HiC-ExLoF: 4.22%; LoC-HiF: 4.40%; LoC-LoF: 3.67%; LoC-MedF: 3.91%. Next, all response times less than 200 ms were removed as anticipation responses. Finally, RTs were trimmed using a 2.5 SD criterion by subject and type of contingency. Overall, 3.09% of correct responses were trimmed.

Response Times. First, a 2 (high, low contingency) x 2 (low, medium frequency) x 16 (blocks) repeated measures ANOVA was conducted. There was a main effect of contingency, with participants responding 15 ms faster to high-contingency items ($M = 580$ ms) than to low-contingency items ($M = 595$ ms), $F(1,39) = 16.67$, $MSE = 8561.73$, $p < 0.001$, $\eta^2_p = 0.299$. There was a main effect of frequency, with low frequency items ($M = 595$ ms) responded to 15 ms slower than medium frequency items ($M = 580$ ms), $F(1,39) = 10.10$, $MSE = 13367.14$, $p = 0.003$, $\eta^2_p = 0.206$. There was also a significant main effect of block, $F(8.85, 345.06) = 4.23$, $MSE = 13499.88$, $p < 0.001$, $\eta^2_p = 0.098$, with Greenhouse Geisser correction. Critically, there was no interaction between contingency and frequency, $F < 1$, and no other interactions were

significant, all Fs < 1. This pattern suggests that frequency and contingency both contributed but that they made separate contributions to the overall effect, given the absence of any interaction between contingency and frequency.

Discussion

The findings of Experiment 6 suggest roles both for frequency and for contingency in the contingency learning paradigm, in agreement with the findings from Experiment 5. Indeed, in the present case, the contributions of frequency and contingency to the "contingency effect" were virtually identical and appeared to be independent, although this could very well be a function of the values of frequency and contingency chosen here. Thus, it seems that people are encoding two types of information. First, they are able to learn contingency information between the word and the response. For example, they learn that MONTH is mostly presented in red, requiring pressing of the red button. Second, people are sensitive to the relative number of presentations. More presentations of a stimulus (e.g., WORLD presented in yellow and red with an overall frequency of 8 times per block) leads to faster learning than does fewer presentations (e.g., WORLD presented in yellow and red with an overall frequency of 4 times per block). When contingency is equated, there is a role for frequency; when frequency is equated, there is a role for contingency.

Experiment 7: Double Frequency

In Experiments 5 and 6, words in the high contingency condition were 100% contingent with the associated colours. This perfect contingency, under which a given word occurs in only one colour, may be a special case, and in fact the rest of the literature has rarely if ever used a 100% condition. I next wanted to know what happens with frequency when high-contingency items do not involve 100% contingency but instead include variability similar to that in the experiments in the first chapter of this dissertation and most of the published literature. In Experiment 7, I returned to an 80%/20% contingency. This time, I contrasted two frequency conditions, one where all contingencies were presented with twice the frequency of the other. Would the two contingency conditions (high and low) interact with the two frequency conditions (single, double)? To examine this, I contrasted a set of items with the usual contingency and frequency—Hi (8x), Lo (1x) to another set of items with the same contingency but doubled frequency—Hi2 (16x), Lo2 (2x). Table 7 makes this comparison more concrete.

If there is an influence of frequency, the conditions with double frequency should make learning easier and, therefore, there should be faster response times for conditions with double frequency than for conditions with single frequency. Of course, there should also be a contingency effect for both frequency conditions. Would the contingency effect differ, though, as a function of item frequency?

Method

Participants. Thirty-one students (*M* age = 20 years; 20 female, 11 male) from the University of Waterloo completed one session in exchange for course credit. All had normal or corrected-to-normal vision.

Apparatus. Except for the stimulus words that were presented in yellow, green, or red, all other stimuli were presented in white at the centre of the screen on a black background in Courier New font size 18.

Materials and Procedure. There were 30 practice trials followed by 6 blocks of 90 experimental trials each. During the practice trials, a row of asterisks (*****) was presented with equal probability in each of the three colours (yellow, green, red).

In the experimental trials, six words (MONTH, PLATE, UNDER, CLOCK, TABLE, WORLD) were presented in three colours: red, yellow, and green. There were two possible contingencies, high vs. low, and two possible frequencies, single vs. double. The four conditions of principal interest were Hi (8x), Lo (1x) and Hi2 (16x), Lo2 (2x), where the number in parentheses represents that condition's frequency in each block. Table 7 provides details of the stimuli in each block. Basically, in each study block, the three single frequency words were presented 8 times for the high-contingency item and 1 time for each of the two low-contingency items. In the same block, the three double frequency words were presented 16 times for the high contingency item and 2 times for each of the two low-contingency items.

Results

Errors and Outliers. First, incorrect responses were removed from the analysis of RTs. Overall, there were 3.94% errors (3.41% of Hi, 3.60% of Hi2, 5.8% of Lo, 5.69% of Lo2). A 2 (high contingency, low contingency) x 2 (single frequency, double frequency) repeated measures analysis revealed a main effect of contingency such that responses to high-contingency items (M = 0.35) produced fewer errors than responses to low-contingency items (M = 0.55), $F(1,30)$ = 13.94, MSE = 0.001, p = 0.001, η^2_p = 0.317. There was no main effect of frequency, nor an interaction of contingency and frequency, both Fs < 1.

Next, all response times less than 200 ms were removed as anticipation responses. Finally, RTs were trimmed using a 2.5 SD criterion by subject and type of contingency. Overall, 3.05% of correct responses were trimmed (3.02% of Hi, 3.00% of Hi2, 3.41% of Lo, 3.18% of Lo2).

Response Times. A 2 (high, low) X 2 (single, double) x 6 (6 blocks) repeated measures ANOVA was conducted. There was a main effect of contingency, with participants responding 18 ms faster to high-contingency items ($M = 540$) than to low-contingency items ($M = 558$), $F(1,30) = 11.21$, $MSE = 5597.39$, $p = 0.002$, $\eta^2_p = 0.272$. There was a main effect of block, $F(5,150) = 2.44$, $MSE = 4220.03$, $p = 0.037$, $\eta^2_p = 0.075$. Critically, there was neither a main effect of frequency, $F(1,30) = 0.27$, $MSE = 4199.26$, $p = 0.611$, $\eta^2_p = 0.009$, nor an interaction of contingency with frequency, $F(1,30) = 0.13$, $MSE = 4295.68$, $p = 0.723$, $\eta^2_p = 0.004$.

There was, however, an interaction of frequency with block, $F(4.21,126.29) = 5.08$, $MSE = 2009.70$, $p = 0.001$, $\eta^2_p = 0.145$, with Greenhouse-Geisser correction. There was no interaction of contingency with block, $F(3.6, 108.01) = 1.35$, $MSE = 4117.29$, $p = 0.258$, $\eta^2_p = 0.043$, with the Greenhouse-Geisser correction. Finally, there was a three-way interaction of contingency with frequency with block, $F(5,150) = 2.95$, $MSE = 2038.48$, $p = 0.014$, $\eta^2_p = 0.089$, with the Greenhouse Geisser correction. To unpack the three-way interaction, I analyzed the interaction of frequency with block separately for the high-contingency and low-contingency conditions. For high-contingency items, there was a significant main effect of block, $F(3.57, 106.98)$ 4.14, $MSE = 1717.10$, $p = 0.002$, $\eta^2_p = 0.121$. There was no main effect of frequency nor was there an interaction of frequency with blocks, both $Fs < 1$. For low-contingency items, there was no main effect of frequency, $F < 1$. There was also no main effect of block, $F(5,150) = 1.55$, $MSE = 5960.13$, $p = 0.177$, $\eta^2_p = 0.049$. However, there was a significant interaction of

frequency with block, $F(5,150) = 4.37$, $p = 0.001$, $\eta^2_p = 0.127$. Analysis of low-contingency

items with original frequency and doubled frequency was also conducted separately. For low-

contingency items, there was a significant main effect of blocks for original frequency, $F(3.54,$

$106.06) = 3.06$, $MSE = 9118.14$, $p = 0.024$, $\eta^2_p = 0.093$ with Greenhouse-Geisser correction, but

there was no main effect of blocks for double frequency, $F(5,150) = 1.34$, $MSE = 2762.15$, $p =$

0.249, $\eta^2_p = 0.043$.

Discussion

There was a contingency effect for both the single frequency and the double frequency

items. However, there was no frequency effect nor did contingency and frequency interact.

Essentially, the contingency effect was unaltered by doubling frequency. This finding of a null

frequency effect might initially seem in contrast to the findings in Experiments 5 and 6. The fact

that Experiments 5 and 6 differed from Experiment 7 in that the first two had 100% high

contingencies and the present one had 80% high contingencies could have had an influence.

However, I suspect that the critical difference is that, in Experiment 7, frequency did not

compete with contingency: Both sets of items had the same relative frequencies, as is the case in

the previous studies in the literature and in Chapter 1. It is only when a frequency manipulation

"corrects" for or counteracts contingencies that a role for frequency will be observed, as was the

case in Experiments 5 and 6.

Chapter II Discussion

The experiments in Chapter I raised the possibility that there is a difference in response criterion depending on the target's prevalence. This in turn could suggest a role for target frequency in the contingency effect. Although contingency and probability/frequency are conceptually distinct, it is the case that in the existing contingency studies (e.g., Schmidt et al., 2007; Schmidt & Besner, 2008; Schmidt et al., 2010), high contingency items also have been presented more often than low contingency items. The goal of Chapter II, therefore, was to re-emphasize the distinction between contingency and frequency. Contingency is the association between two or more events; frequency refers to how often events occur. In our studies, when a word is most often associated with a certain colour, this is considered to be a high contingency item (e.g., PLATE is presented most often in green) whereas when that same word is infrequently paired with another colour, it is considered to be a low contingency item (e.g., PLATE in yellow or green). The question, then, is: Are people learning the overall number of presentations of two or more events, the actual contingencies of those events, or both?

In this chapter, the experiments showed that frequency can and does affect responding in the contingency task, however, this influence depends on how frequency is manipulated. In Experiment 5, increasing the number of occurrences of the low-contingency items and decreasing the number of occurrences of the high-contingency items—such that the frequencies wound up identical in the two conditions—resulted in the disappearance of the contingency effect. In Experiment 6, constructing items to allow holding frequency constant to examine contingency showed a reliable contingency effect, but the reverse was also true: Holding contingency constant and examining frequency showed a reliable frequency effect. In what we know as the contingency learning task, there clearly is a role for frequency. But Experiment 7 demonstrated that this role must involve frequency manipulations at the level of the individual

items: When the number of occurrences of each condition—high-contingency and low-contingency—was simply doubled, that did not produce a frequency effect.

Two features of the experiments in Chapter II are noteworthy, although their interpretation is not straightforward. First, frequency and contingency did not interact; when frequency exerted an influence, that influence was independent of the contingency effect. Second, the frequency effects were evident here only when contingencies were "perfect"—when high-contingency items were unique (e.g., UNDER appeared in red 100% of the time). How critical that is remains to be seen. For the present, the experiments in Chapter II stand in marked contrast to the argument of Schmidt et al. (2007) against a stimulus familiarity account—that participants are not simply learning frequencies of items but are sensitive principally to contingencies. In contrast, the experiments in Chapter II suggest that there is a role of frequency in the contingency learning paradigm.

Chapter III: Use of Semantics in Contingency Learning

Experiment 8: Semantic vs. Episodic Contingency

Chapter III addresses the third and final question under consideration in this dissertation: Do people go beyond the stimuli in contingency learning, to learn something about their meaning? The work of Schmidt and De Houwer (2012a) with evaluative stimuli suggests that this might be the case, but their procedure required them to graft meaning on with an extra stimulus, thereby substantially altering the standard paradigm. The focus here, instead, was on the learning of categorical information about the words within the standard contingency learning paradigm. Put simply, is the contingency effect influenced by whether words are semantically related versus unrelated?

Critically in this final series of experiments, there were no explicit instructions regarding what the categories were except in the last experiment. Probabilities for each set of words were manipulated such that it was more beneficial for people to learn that semantic categories—not just individual words—were associated with specific colours. Will the size of the contingency effect be greater for semantically associated words as opposed to episodically associated words? Do people learn anything about the items beyond the fact that specific words are most often connected to specific colours? If they can and do, then semantically related words should help participants bring to mind categories and doing so should facilitate their performance, resulting in a larger contingency effect for semantically related words (semantic contingency) relative to unrelated words (episodic contingency). If they cannot or do not go beyond the stimuli themselves, then performance under semantic contingency should not differ from that under episodic contingency.

Method

  Participants. Thirty-five students assigned to the semantic condition (*M* age = 20.47 years[1]) and 32 participants assigned to the episodic condition (*M* age = 21.84 years; 24 female, 8 male) completed one session in exchange for course credit. All were students from the University of Waterloo who had normal or corrected-to-normal vision.

  Materials and Procedure. Only the words differed across the semantic and episodic conditions.

  There were 30 practice trials followed by 8 blocks of 30 experimental trials. Six different words were presented on the practice trials, each having a constant probability of occurring in each of the three colours (yellow, green, red). In the semantic condition, the six words were selected from three semantic categories: instruments, animals, and occupations. For each category, there were two semantically related words: FLUTE, PIANO, HORSE, CAMEL, FARMER, NURSE. The two words of each category appeared in the same colour 40% of the time (high contingency) for an overall 80% contingency for the category. These words appeared in the other two colours 20% of the time (low contingency). For example, in a block of 30 trials, FLUTE and PIANO appeared in the colour red 4 times each (high contingency), but collectively as a category 8 times; they appeared in green or yellow one time each (low contingency). For the episodic condition, the same procedure was applied, but the pairings were arbitrary in that the words were not categorically related: BAKER, ELBOW, APPLE, PIANO, HORSE, TRAIN. All pairings of colours and words were counterbalanced. The contingency stayed the same across all 8 blocks. The stimuli are illustrated in Table 8.

---

[1] No gender information was collected

Results

Errors and Outliers. First, incorrect responses were removed from the analysis of RTs. Overall, there were 3.42% errors for the semantic condition (high: 2.89%, low: 5.54%) and 3.53% for the episodic condition (high: 3.16%, low: 5.01%). A 2 (contingency: high, low) X 2 (relatedness condition: semantic, episodic) repeated-measures ANOVA revealed a significant main effect of contingency, $F(1,64) = 9.32$, $MSE = 0.001$, $p = 0.003$, $\eta^2_p = 0.127$. The main effect of relatedness was not reliable, $F(1,64) = 2.86$, $MSE = .001$, $p = .096$, $\eta^2_p = 0.043$. There was, however, a contingency by relatedness interaction, $F(1,64) = 9.81$, $MSE = 0.001$, $p = 0.003$, $\eta^2_p = 0.133$. Two paired samples t-tests were conducted to further examine this interaction. For the semantic condition, low-contingency items ($M = 0.055$) were more prone to errors than high-contingency items ($M = 0.029$), $t(34) = 3.83$, $SE = 0.01$, $p = 0.001$. However, for the episodic condition, there was no difference in errors between high-contingency items ($M = 0.033$) and low-contingency items ($M = 0.032$), $t(30) = 0.07$, $SE = 0.01$, $p = 0.943$.

Next, all response times less than 200 ms were removed as anticipation responses. Finally, RTs were trimmed using a 2.5 SD criterion by subject and type of contingency. Overall, 3.06% of the RTs were trimmed for the semantic condition (high: 3.05%, low: 3.09%) and 2.96% were trimmed for the episodic condition (high: 2.94%, low: 3.02%).

Response Times. The RT data are displayed in Figure 8. A mixed 2 (contingency) x 2 (relatedness) x 8 (blocks) ANOVA was conducted, with the relatedness factor between subjects. There was a main effect of contingency, $F(1,65) = 72.78$, $MSE = 2874.92$, $p < 0.001$, $\eta^2_p = 0.528$. There also was a main effect of block, $F(5.69,370.01) = 2.21$, $MSE = 5917.79$, $p = 0.040$, $\eta^2_p = 0.033$. There was, however, no main effect of relatedness, $F(1,65) = 2.55$, $MSE = 60907.87$, $p = 0.115$, $\eta^2_p = 0.038$. And critically, there was no significant interaction between contingency and

relatedness, $F(1,65) = 1.71$, $MSE = 2874.92$, $p = 0.195$, $\eta^2_p = 0.026$. There was also no three-way interaction of contingency X blocks X experiments, $F(7,455) = 1.25$, $MSE = 2348.59$, $p = 0.274$, $\eta^2_p = 0.019$. In fact, no other interactions were significant, all Fs < 1.

Discussion

Semantic pairs were not responded to faster than episodic pairs, as shown by the non-significant main effect of relatedness and the non-significant interaction of relatedness with contingency. In other words, the size of the contingency effect was very similar both for semantically associated words and for episodically unrelated words, as Figure 8 portrays. So there was no evidence of participants using semantic representations to improve their efficiency in responding in this task. Instead, the results suggest that subjects are simply learning primitive stimulus-level pattern-colour contingencies, with little if any knowledge of the characteristics of the particular words beyond perhaps their physical appearance.

Experiment 9: Semantic Transfer

Experiment 8 showed that participants learn contingencies for semantically related words and episodically related words equally well. However, the between-subjects manipulation of the two types of words may have made it difficult for participants to determine the role of categorical information. An alternative and potentially more effective and powerful way to address this issue would be to use a within-subject design and look for transfer. Would there be any carryover effect of a contingency effect established for one set of words onto a different set of semantically related words where all words occurred in equal probability. Essentially, this would follow the procedure of Experiment 2, where transfer was observed in moving from a contingency phase to a no-contingency phase. Would the fact that words were taken from the same semantic category when a contingency was present carry over, for at least a little while, to other words from the same semantic category without the contingency?

Method

Participants. Thirty students from the University of Waterloo (*M* age = 19.70 years; 19 females, 11 males) completed one session in exchange for course credit. All had normal or corrected-to-normal vision.

Design. There were 30 practice trials. Asterisks in the practice trials had an equal probability of occurring in each of the three colours. There followed two experimental phases: the Contingency Phase and the No Contingency Phase. There were 8 blocks of 30 trials for the Contingency Phase and 7 blocks of 27 trials for the No Contingency Phase. The small difference in number of trials per block was necessitated by the composition of items in the two phases. Overall, there were six words in the Contingency Phase (FLUTE, PIANO, HORSE, CAMEL, FARMER, NURSE) and another six words in the No Contingency phase (BANJO, CELLO,

ZEBRA, TIGER, BAKER, LAWYER). Words were of high contingency or low contingency

with one of three colours: red, yellow, or green.   However, unknown to the participants, pairs of

words actually belonged to three categories, and the categories were also of high or low

contingency (see Table 9). There were three categories— animals, professions, and

instruments—with each containing four words, two of which appeared in the Contingency Phase

and two of which appeared in the No Contingency Phase.

Contingency Phase. Each category appeared in one colour 80% of the time with each of

the two words in that category appearing 40% of the time (high contingency). The same

category appeared in the other two colours 20% of the time (low contingency). For example, in

a block of 30 trials, HORSE and CAMEL appeared in the colour red 4 times each (high

contingency). In addition, either HORSE appeared in green once and CAMEL appeared in

yellow once or the converse (low contingency). Table 9 lays out the stimuli and conditions. All

pairings of colours and words were counterbalanced. The contingencies stayed the same across

all 8 blocks.

No Contingency Phase. Although the three categories remained the same across the two

phases, the two words in each category were different in the second phase from those in the first

phase. As well, the contingency was cancelled, with each word appearing equally often in every

colour. For example, in a block of 27 trials, given that HORSE and CAMEL had been used in

the Contingency Phase, ZEBRA and TIGER were used in the No Contingency Phase, each

appearing in red 3 times, green 3 times, and yellow 3 times, as illustrated in Table 9. All

pairings of colours and words were counterbalanced.

Results

Errors and Outliers. First, incorrect responses were removed from the analysis of RTs. Overall, there were 3.32% errors (Phase 1: high = 2.38%, low = 4.24%; Phase 2: high = 3.77%, low = 3.91%). A repeated-measures ANOVA was conducted on the two levels of contingency and the two different phases and revealed a significant main effect of contingency, $F(1,29) = 5.18$, $MSE = 0.001$, $p = 0.030$, $\eta^2_p = 0.151$. Responses to high-contingency items ($M = 0.031$) contained fewer errors than did responses to low-contingency items ($M = 0.041$).

The main effect of phase was not significant, $F(1,29) = 2.53$, $MSE < 0.001$, $p = 0.123$, $\eta^2_p = 0.080$. Overall errors were quite equivalent for the No Contingency Phase ($M = 0.039$) and the Contingency Phase ($M = 0.033$), $F(1,29) = 2.53$, $MSE < 0.001$, $p = 0.123$, $\eta^2_p = 0.080$. The phase by contingency interaction approached significance, $F(1, 29) = 4.07$, $MSE = 0.001$, $p = 0.053$, $\eta^2_p = 0.123$, consistent with most of the change in error rates between phases being due to the increase in high-contingency errors, with a small decrease in low-contingency errors.

Next, all response times less than 200 ms were removed as anticipation responses. Finally, RTs were trimmed using a 2.5 SD criterion by subject and level of contingency. Overall, 2.82% were trimmed (Phase 1: 2.92% of high-contingency items, 2.77% of low-contingency items; Phase 2: 2.89% of high-contingency items, 2.68% of low-contingency items).

Response Times. A repeated measures ANOVA was conducted with two levels of contingency (high, low) and two phases (Contingency, No Contingency). There was a significant main effect of contingency such that responses were 15 ms faster to high-contingency items ($M = 561$ ms) than to low-contingency items ($M = 576$ ms), $F(1,29) = 18.11$, $MSE = 362.10$, $p < 0.001$, $\eta^2_p = 0.384$, replicating the typical contingency effect.

There was a main effect of phase where the No Contingency Phase ($M = 580$ ms) was overall 22 ms slower than the Contingency Phase ($M = 558$ ms), $F(1,29) = 8.97$, $MSE = 1616.37$, $p = 0.006$, $\eta^2_p = 0.236$. More important, there was also an interaction of phase with contingency, $F(1,29) = 6.09$, $MSE = 397.06$, $p = 0.020$, $\eta^2_p = 0.174$. Two t-tests were conducted using the Bonferroni adjusted alpha levels of 0.025 per test (0.05/2) to probe the interaction. For the Contingency Phase, high-contingency items ($M = 546$ ms) were responded to 24 ms faster than low-contingency items ($M = 570$ ms), $t(29) = 3.70$, $SE = 6.42$, $p = 0.001$, replicating the typical contingency effect. For the No Contingency Phase, although it was in the right direction, the 6 ms difference between responses for high-contingency items ($M = 577$ ms) and low-contingency items ($M=583$ ms) did not reach significance, $t(29) = 1.90$, $SE = 3.06$, $p = 0.068$.

Block and Transfer Analysis. Because of the differing number of blocks for the Contingency Phase and the No Contingency Phase (Phase 1: 8 blocks; Phase 2: 7 blocks), separate contingency by blocks repeated measures ANOVAs were conducted for each phase.

Contingency Phase: There was a main effect of contingency, $F(1,29) = 15.41$, $MSE = 5124.75$, $p < 0.001$, $\eta^2_p = 0.347$. There was, however, no main effect of block, $F(4.37, 126.66) = 1.48$, $MSE = 7480.58$, $p = 0.209$, $\eta^2_p = 0.048$, and no interaction of contingency with block, $F < 1$.

No Contingency Phase: The main effect of contingency was only marginally significant, $F(1,29) = 3.55$, $MSE = 996.06$, $p = 0.069$, $\eta^2_p = 0.109$, indicative of very little in the way of a carryover effect. There was a main effect of block, $F(6,174) = 2.76$, $MSE = 2646.97$, $p = 0.014$, $\eta^2_p = 0.087$, but no interaction of contingency with block, $F < 1$. Indeed, a t-test over just the first three blocks of the No Contingency Phase showed no reliable contingency effect, $t(29) = 1.12$, $SE = 4.74$, $p = 0.273$.

Discussion

As previously, participants were able to learn contingency information as evidenced by the faster response times to high-contingency items than to low-contingency items here. However, in support of the findings from Experiment 8, there was no evidence that participants used categories to facilitate their learning of contingencies because no carryover contingency effect was observed from the Contingency Phase to the No Contingency Phase. Again, what subjects appear to be learning in the contingency learning task is quite primitive, possibly having to do only with the differing physical forms of the stimuli. However, the determination of a semantic effect is on the assumption that there is a carryover effect from the Contingency phase to the No Contingency phase: It also is possible that if there was any weakness of the carryover, then no semantic effect would be observed.

Experiment 10: Semantic Baseline

Experiments 8 and 9 showed no evidence of the use of semantic categories to influence contingency learning. It could be argued, however, that the designs of Experiments 8 and 9 did not encourage participants to use semantic categories. Although the categories were of high contingency with a particular colour, it was also the case that both of the exemplar words within a category were highly contingent with the colour. Thus, participants could simply have learned the contingencies of the individual words without using the semantic category because the category only provided redundant information.

In Experiment 10, I examined whether semantic effects extended from one word to another within a category during contingency learning itself. Here, there were four types of contingency items: high-contingency items, high-related items, low-related items, and low-contingency items, all occurring during the time when contingencies were in place (i.e., this was not a transfer study). A high-contingency item was, as usual, a word presented in a certain colour most times (e.g., FLUTE in red 7 times). A high-related item was a word both semantically related to and presented in the same colour as a high-contingency item, but with considerably lower presentation frequency (e.g., PIANO in red once). A low-related item was the same word as the high-related item but presented in a low-contingency colour (e.g., PIANO in green once). Finally, as usual, a low-contingency item was a word presented in a colour other than the one in which it was most often presented (e.g., FLUTE in yellow once). Table 10 displays the organization of items and conditions.

Participants in Experiment 10 were explicitly told that the words belonged to the three categories: animals, instruments, and professions. However, the individual words were not provided. The use of explicit instructions is modeled after Schmidt and De Houwer's (2012c)

69

explicit contingency instruction, which increased the contingency effect. If participants use the semantic categories in their learning, then responses to high-related items should benefit from being related to high-contingency items. If participants still do not use semantic categories, then responses to high-related items should show no benefit relative to responses to low-related items or to low-contingency items. A key feature of this experiment, then, is that the semantic relatedness manipulation occurs during the contingency phase, not in a no-contingency transfer phase, and that the category relations are explicitly noted in the instructions, making the semantic association more apparent and thereby encouraging its use.

Method

  Participants. Forty-four students (*M* age = 20.75 years; 31 female, 13 male) from the University of Waterloo completed one session in exchange for course credit or for $10. All had normal or corrected-to-normal vision.

  Apparatus, Materials and Procedure. All procedural details including the practice were the same as in Experiment 8 except for the frequency and the type of contingency items. There were six words (FLUTE, PIANO, HORSE, CAMEL, FARMER, NURSE) made up of two words from each of three categories—animals, professions, and instruments. Participants were informed of the three categories, but not of the individual words, before commencing.

  There were four types of contingency items: high, high-related, low-related, and low. For high-contingency items, words appeared 70% of the time contingent with a particular colour (FLUTE in red). For high-related item, same-category words appeared only 10% of the time but were still contingent with the same colour as the high-contingency item (PIANO in red). Low-related items were the same words as high-related items but presented in a different colour (PIANO in green). Low-contingency items were the same words as high-contingency items but

70

presented in a different colour (FLUTE in yellow). All pairings of colour and words were counterbalanced. The contingencies stayed the same across all 8 blocks.

Results

 Errors and Outliers. First, incorrect responses were removed from the analysis of RTs. Overall, there were 3.52% errors (high = 2.71%, high-related = 6.44%, low-related = 3.79%, low = 6.06%). A repeated-measures ANOVA conducted on the four types of contingency revealed a significant main effect of contingency, $F(2.39,100.24) = 6.74$, $MSE = 0.003$, $p = 0.001$, $\eta^2_p = 0.138$. Six t-tests were then conducted using the Bonferroni correction with an adjusted alpha level of 0.008 per test (0.05/6). High-contingency items elicited fewer errors than both high-related items, $t(42) = 4.19$, $SE = 0.01$, $p < 0.001$, and low-contingency items, $t(42) = 3.44$, $SE = 0.01$, $p = 0.001$. There were no other significant differences.

 Finally, RTs were trimmed using a 2.5 SD criterion by subject and type of contingency. Overall, 2.86% were trimmed (2.85% of high, 2.33% of high-related, 3.64% of low-related, 2.62% of low).

 Response Times. A 4 X 8 repeated measures ANOVA was conducted with the four levels of contingency (high, high-related, low-related, low) and blocks. With the Greenhouse-Geisser correction, there was a significant main effect of contingency, $F(3,126) = 26.17$, $MSE = 13328.36$, $p < 0.001$, $\eta^2_p = 0.384$. There also was a main effect of block, $F(3.54, 148.75) = 4.29$, $MSE = 33776.48$, $p = 0.004$, $\eta^2_p = 0.093$, but there was no interaction of contingency with block, $F(10.94, 459.60) = 1.68$, $MSE = 16660.21$, $p = 0.075$, $\eta^2_p = 0.038$.

 To compare the contingency conditions, six t-tests were conducted using the Bonferroni adjusted alpha levels of 0.008 (0.05/6). High-contingency items were the fastest, confirming a benefit: They were faster than low-related, $t(43) = 5.70$, $SE = 9.77$, $p < 0.001$, faster than high-

related, $t(43) = 3.99$, $SE = 6.36$, p < 0.001, and faster than low-contingency, $t(43) = 7.36$, $SE = 9.09$, $p < 0.001$. The critical analysis evaluated whether there was a difference between high-related and low-related items. The relevant t-tests demonstrated that high-related items were faster than both low-related items, $t(43) = 3.42$, $SE = 8.87$, $p < 0.001$, and low-contingency items, $t(43) = 4.97$, $SE = 8.34$, $p < 0.001$. There was, however, no difference between low-related items and low-contingency items, $t(43) = 1.27$, $SE = 8.81$, $p = 0.210$.

Discussion

This final experiment of Chapter 3 illustrates that participants can use semantic category information but that this use has to be evaluated appropriately. When participants see high-contingency items (e.g., FLUTE in red), the semantic category for instruments is brought to mind and thus high-related items (e.g., PIANO in red; same category, same colour as high-contingency items) are responded to faster even though they have the same low frequency as the low-contingency items. The finding that high-related items (e.g., PIANO in red) are faster than low-related items (e.g., PIANO-green), despite having the same frequency of presentation, provides evidence that participants learned that the category "instruments" is highly contingent with a certain colour (e.g., red).

The benefit of semantic relatedness is seen only when the task encourages participants to use semantic information and perhaps only when contingencies are in place (i.e., no benefit was apparent in examining semantic transfer without contingencies as in Experiment 9). The within-task manipulation of different frequencies for two of the high-contingency items in Experiment 10 allowed the observation of a semantic influence in contingency learning. Experiments 8 and 9 illustrate that participants do not use semantic categories to help with contingency learning when there is no encouragement to do so. In both of those experiments, pairs of words that were

of high contingency and belonged to the same category (e.g., FLUTE in red and PIANO in red) were still individually presented more often than their counterpart low-contingency items. Thus, there was no need to use the semantic information to facilitate the learning when there was enough information to learn the contingencies without the use of semantics.

Chapter III Discussion

Do we use higher-order associations as a matter of course? As seen in Experiment 1, people are very good at connecting two or more events, and these associations are learned very quickly, at least with the small stimulus sets used here. When we think of words, we routinely associate the meaning to the word and this meaning extends to other similar words (e.g., spreading activation; Collins & Loftus, 1975; Posner & Snyder, 1975). Semantic priming has been found to facilitate responding to words that are similar in meaning. For example, when we think about the word *doctor*, similar words come to mind very readily, such as the word *nurse*, and we can then process these related words more fluently. Similarly, in a Deese-Roediger false memory task (Roediger & McDermott, 1995), people create false memories of unpresented words that are similar to the meaning of the words that were actually presented.

Chapter III investigated the possible role of meaning in contingency learning. These three experiments suggest that people do not routinely use semantic association in contingency learning unless there is encouragement and an opportunity to apply semantics during responding. When the learning of associations is done very quickly and the task can be completed without attention to the semantic relations of the words, people show no evidence of the use of semantic categories. In Experiments 7 and 8, responses were not faster to semantically related words than to semantically unrelated words. The size of the contingency effect was very similar for both types of words in Experiment 8, and there was no evidence of any carryover from a Contingency Phase to a No Contingency Phase for words that were semantically related in Experiment 9. Participants seemed to be learning only the basic word-colour response association and not using any factors beyond the physical stimuli even though the use of categorization would potentially provide additional information to support processing and facilitate responding.

The failure to use semantic categories in the experiments is in contrast to studies in other domains suggesting that people process semantically related words relatively easily, with little intention. However, the present findings do show that, given the opportunity to use meaning (Experiment 10), semantics can play a role in contingency learning. Semantic priming in various tasks such as Stroop-related tasks and lexical decision tasks certainly has been found to facilitate responding (see MacLeod, 1991; see Logan, 1980).

Similarly, Schmidt and De Houwer (2012a) also found that learning of contingencies can go beyond the physical word and colour that are shown. Recall that in their study, there were three pairings between nonwords and target words. First, certain nonwords were presented most often with a frequent target (NIJARON-HUG). Second, some nonwords were presented with the same valence as the most frequent target (NIJARON-FLOWER) and third, some nonwords were presented with the valence opposite that of the most frequent target (NIJARON-GUNS). Participants responded to the valence of the target, pressing one key for positive valence responses and another key for negative valence responses. Schmidt and De Houwer found that responses for target words were faster if the distracting nonword was presented most often with the same valence as the target relative to a different valence. Apparently, participants were able to learn the valence associated with the distracting nonword. They were also able to learn the associations between the nonwords and the general valence of multiple words (Schmidt & De Houwer, 2012a, Experiment 2). Thus, participants were able to learn the contingencies between the valences of the words to the distracting nonwords, supporting the claim that under the right circumstances participants can learn contingencies beyond the physical stimuli. Of course, to observe this result, Schmidt and De Houwer substantially modified the paradigm.

One crucial difference between Experiments 8 and 9 and Experiment 10 is the opportunity for the expression of the semantic categories. In Experiments 8 and 9, in the high-contingency condition (e.g., animals in red 8 times out of 10), the individual words (e.g., CAMEL and HORSE) each were also of high contingency (4 times each). Thus, there was no real need to use the category "animal" because the individual exemplar words were already high contingency. Participants could accomplish contingency learning simply by learning the associations between the individual words and the colours without appeal to meaning.

Experiment 10 shows that learning of a word that is a high-contingency item (occurring with high frequency) influences another exemplar word of the same category (with low frequency). The critical comparison is between two conditions: high-related words vs. low-related words. In a high-related word condition, the word is an exemplar of the same category as the word in the high-contingency condition. In a low-related word condition, however, the word is the same word as the high-related word but in a low-contingency colour. The faster responding observed for the high-related words provides evidence that people can use semantic information acquired during contingency learning.

In Chapter III, the main finding is that most contingency learning appears to be quite primitive, relying largely on the physical stimuli. But Experiment 10 demonstrates that, at least under some circumstances, participants can use higher-order semantic information to facilitate the learning of contingencies. However, if the opportunity for expression of the use of semantics is not readily available, participants do not routinely use semantic relations to facilitate the learning. Despite extensive evidence that we are very good at making higher-order associations, if the task does not encourage their use, people will not take up this approach and will instead use the more efficient way which, in this case, is to rely on the individual stimuli.

General Discussion

The three chapters of this dissertation have addressed three key issues related to the learning of contingencies, reaching the following conclusions: (1) learning of colour-word contingencies involves not only a benefit to high-contingency items but also a cost to low-contingency items, (2) frequency of presentation of co-occurring events does have an influence on our learning of those contingencies, and (3) when the task allows opportunities to use higher-order information, such as semantics, such information can influence the learning of contingencies, although ordinarily the learning is more primitive, focusing on the stimuli themselves.

Recently, Schmidt (2013; Schmidt et al., 2007) has adapted the instance theory of automaticity (Logan, 1988) to explain colour-word contingency learning, the idea being that participants store the stimuli and responses in instances that then are retrieved during subsequent performance of the task. On each trial, a stimulus will evoke relevant instances. Those stimuli for which more instances—and especially more consistent instances—exist in memory likely will also cue faster retrieval. This readily explains the benefit in the high-contingency condition, although it is less clear how it explains the cost that was consistently observed in Chapter I.

I will now assess how my data from the 10 experiments reported here fit with Schmidt's (2013) PEP model, an instance-based explanation of the colour-word contingency effect. (1) *The theory explains the contingency effect in terms of faster responding to high contingency items only.*

As just noted, the version of instance theory laid out by Schmidt (2013) as an explanation of colour-word contingency learning easily handles the benefit due to high-contingency items; presumably, this is why it was proposed. But therein lies the problem: In contrast to the one

previous study that reported only a benefit (Schmidt & Besner, 2008), Experiments 2, 3, and 4 reported here replicated the benefit but also found a consistent cost in the response times for low-contingency items compared to baseline no-contingency items.

Schmidt and Besner (2008) argued that participants lower the threshold for responding to an expected item but do not alter the threshold of other items. This clearly would lead to faster responding to high-contingency items than to other items, whether low contingency or no contingency (baseline). Under their view, the response times for low-contingency items should not be slower than baseline because the thresholds for responding to low-contingency items and baseline no-contingency items would remain the same. There was one problem in their own data for their argument: more errors for low-contingency items than for baseline items. They explained this by arguing that, due to the response thresholds for high-contingency items being lowered, people are biased to respond to high-contingency items, resulting in errors for low-contingency items.

How could the cost observed here be explained in terms of the instance account? I agree that the number and consistency of the instances for high-contingency items is greater than is the case for no-contingency baseline controls, hence the benefit. Correspondingly, there are also more instances to retrieve for baseline items than for low-contingency items, baseline items having been more often encountered than low-contingency items. Given more baseline instances in memory, there is also a greater likelihood of faster instances for baseline items relative to low-contingency items, explaining the cost for the low-contingency items. This is also in keeping with the visual search for rare targets explanations discussed throughout this dissertation (Hon & Tan, 2013; Wolfe et al., 2007), which argue—in contrast to Schmidt and Besner—that the threshold for rare items does change across trials.

(2) *It is unclear what the theory predicts about the contingency effect as practice progresses.*

Past research has not directed much attention to how the colour-word contingency effect should change with extended practice. In particular, how does the effect behave over blocks? The only previous study to report a block analysis was that of Schmidt et al. (2007). They found that the contingency effect was consistent across blocks accompanied by an overall decrease in response time. Intriguingly, although the experiments reported in this dissertation also found the contingency effect to be remarkably consistent over blocks, these experiments have almost always shown performance slowing down over blocks. And most salient, there have been consistent costs as well as benefits over blocks.

It may be the case that RTs diverge for low-contingency items and baseline items as blocks progress, although this pattern is not entirely clear. Possibly, it takes some time to learn the low-contingency items due to their low frequency, which would seem to fit with the visual search for rare targets work just alluded to. If word and colour and response are stored into instances, the number of instances for low-contingency items will be relatively fewer than those of either high-contingency or baseline items. Based on past literature, people are constantly switching the response criterion across time depending on the different types of contingency trials that they receive (cf. Wolfe et al., 2007) so a difference in performance between high-contingency and low-contingency trials may be at least in part due to the generally slower learning of low-contingency items (Hon & Tan, 2013), again consistent with a cost.

(3) *The theory suggests that the colour-word contingency effect cannot be attributed to the role of frequency.*

Schmidt et al. (2007) maintained that the contingency effect cannot be explained as due to frequency. Recall that, in their study, two colours (blue, green) were responded to with the

left response key and two colours (yellow, orange) were responded to with the right response key. Thus, as usual, there were high contingency items (e.g., MOVE in blue) and low contingency items (MOVE in yellow, orange, or green). They found that response time to words that were more frequently presented in a particular colour (e.g., high-contingency items such as MOVE in blue) was the same as response time to words that were associated with the same response key as MOVE in blue but were of lesser frequency (e.g., a low-contingency item such as MOVE in green). Although their findings provide a clear basis for arguing that the contingency effect cannot be due simply to frequency effects, the experiments in this dissertation provide evidence that whether frequency plays a role is dependent on the context. For example, in Experiment 5, equating the frequency of high-contingency and low-contingency items eliminated the contingency effect. In contrast, in Experiment 7, when the frequencies of word-colour pairings were doubled, there was no evidence of faster responding and the contingency effect was unaltered.

Thus, Experiments 5 through 7 show that whether frequency plays a role in the contingency effect is dependent on the context in which the contingencies are presented.

(4) *The theory emphasizes the storing and retrieval of colour, word, and response in an instance but there is no role for higher-order associations such as semantics.*

Chapter III (Experiments 8, 9, and 10) extends the current research on colour-word contingency learning to whether people can use higher-order associations and not simply respond to stimulus events alone. Past literature has suggested that the colour-word contingency effect can be generalized beyond simple stimulus-response pairs (Schmidt & De Houwer, 2012a), at least when an additional item provides a kind of semantic context. In the present experiments, the use of higher-order associations is dependent on the context of the stimulus-response pairs.

80

In Experiments 8 and 9, there was no evidence that participants were using semantic associations to help with their learning of contingencies. This may be because these two studies provided situations where there was redundancy of information so there was no need to use semantic associations to assist learning: The stimuli themselves were adequate. However, in Experiment 10, where the opportunity to use semantic associations was enhanced, there was evidence that people are faster for even low-frequency items if they are associated with the high-contingency items. The present studies (especially Experiment 10) suggest that colour-word contingency learning cannot be explained simply by word-response pairing, in contrast to the view espoused in the previous literature (Schmidt & Besner, 2008; Schmidt et al., 2007; Schmidt, 2013).

The experiments in this dissertation suggest that the parallel episodic model of Schmidt (2013) is in need of modification to capture these new findings. Colour-word contingency learning does not seem to rest entirely on word-response instance retrieval at a primitive level, nor is it uncontaminated by frequency effects, nor is it entirely about the benefit of high contingency. As so often happens in cognition—consider the Stroop effect, for example (see MacLeod, 1991)—the story behind why colour-word contingency learning occurs seems to be more complex than initially thought.

Summary

The 10 experiments in this dissertation make the following contributions to the colour-word contingency literature:

(1) The learning of colour-word contingencies leads not only to faster responding to high-contingency items but also to slower responding to low-contingency items. This balance may be affected by the range of probabilities used, although more studies have to be conducted to confirm this hypothesis.

(2) The frequencies of high-contingency and low-contingency items can affect the size of the contingency effect. This seems to depend somewhat upon whether word-response mappings are unique, and also on the balance of frequencies in the high-contingency and low-contingency conditions.

(3) Higher-order associations, such as semantics, may be used when the task provides the opportunity to do so. Otherwise, learning seems to focus mainly on the individual stimuli and is rather primitive.

People learn that the colour red is often associated with caution, danger, or stop and we exercise more caution and alertness when we encounter a red traffic light. Many times, we hope that learning of associations can improve our performance in some way but, in reality, whether learning occurs depends on many factors. However, there are other circumstances where learning of associations can be detrimental to our performance and under which the use of higher-order relations is dependent on the task and the context in which we find ourselves. For example, the overlearned skill of driving a car on the left side in North America can have drastic consequences in another culture where the driver sits on the right side. Human learning is complex, and even a seemingly basic learning process such as that in the colour-word contingency paradigm is more complicated—but potentially also more revealing—than initially thought.

References

Abernethy, B. (1991). Visual search strategies and decision-making in sport. *International Journal of Sports Psychology, 22*, 189-210.

Boring, E.G. (1950). *A history of experimental psychology*. New York: Appleton-Century-Crofts.

Casey, B.J., Forman, S.D., Franzen, P., Berkowitz, A., Braver, T.S., Nystrom, L.E., Thomas, K.M., & Noll, D.C. (2001). Sensitivity of prefrontal cortex to changes in target probability: A functional MRI study. *Human Brain Mapping, 13*, 16-33. doi: 10.1002/hbm.1022

Collins, A.M., & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407-428. http://dx.doi.org/10.1037/0033-295X.82.6.407.

Colquhoun, W.P., & Baddeley, A.D. (1967). Influence of signal probability during pretraining on vigilance decrement. *Journal of Experimental Psychology, 73*, 153-155. http://dx.doi.org/10.1037/h0024087

Craig, A. (1978). Is the vigilance decrement simply a response adjustment towards probability matching? *Journal of Human Factors, 20*, 441-446. doi:10.1177/001872087802000408.

Craig, A. (1987). Signal detection theory and probability matching apply to vigilance. *Journal of Human Factors, 29*, 645-652. doi: 10.1177/001872088702900604.

Crump, M.J.C., Vaquero, J.M.M., & Milliken, B. (2008). Context-specific learning and control: The roles of awareness, task relevance, and relative salience. *Consciousness and Cognition, 17*, 22-36. doi:10.1016/j.concog.2007.01.004.

De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theories

    on human contingency learning. *The Quarterly Journal of Experimental Psychology,*

    55B, 289-310.

Drew, T., Evans, K., Vo, M.L.H., Jacobson, F.L., & Wolfe, J.M. (2013). Informatics in

    radiology: What can you see in a single glance and how might this guide visual search

    in medical images? *Radiographics, 33,* 263-274.

    http://dx.doi.org/10.1148/rg.331125023.

Estes, W.K. (1964). All-or-none processes in learning and retention. *American Psychologist,*

    19, 16-25. http://dx.doi.org/10.1037.

Gehring, W.J., Gratton, C., Coles, G.H, & Donchin, E. (1992). Probability effects on stimulus

    evaluation and response processes. *Journal of Experimental Psychology: Human*

    Perception and Performance, 18, 198-216. http://dx.doi.org/10.1037/0096-

    1523.18.1.198

Gomez, R.L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific

    and abstract knowledge. *Cognition, 70,* 109-135. doi:10.1016/S0010-0277(99)00003-7

Goschke, T., & Bolte, A. (2007). Implicit learning of semantic category sequences: Response-

    independent acquisition of abstract sequential regularities. *Journal of Experimental*

    Psychology: Learning, Memory, and Cognition, 33, 394-406.

    http://dx.doi.org/10.1037/0278-7393.33.2.394.

Gwinn, R., & Leber, A. (2014). *Features previously associated with a unique response attract*

    attention. Paper presented at the 56th annual meeting of the Psychonomic Society, Long

    Beach, LA.

Hasher, L., & Zacks, R.T. (1984). Automatic processing of fundamental information: The case

    of frequency of occurrence. *American Psychologist, 39*, 1372-1388.

    http://dx.doi.org/10.1037/0003-066X.39.12.1372.

Hon, N., & Tan, C-H. (2013). Why rare targets are slow: Evidence that the target probability

    effect has an attentional locus. *Attention, Perception & Psychophysics, 75*, 388-393.

    doi: 10.3758/s13414-013-0434-0

Hon, N., Ong, J., Tan, R., & Tang, H.Y. (2012). Different types of target probability have

    different prefrontal consequences. *NeuroImage, 59*, 655-662.

    doi:10.1016/j.neuroimage.2011.06.093.

Hon, N., Yap, M.J., & Jabar, S.B. (2013). The trajectory of the target probability effect.

    Attention, Perception, & Psychophysics, 75, 661-666. doi: 10.3758/s13414-013-0429-

    x.

Jacoby, L.L., Lindsay, D.S., & Hessels, S. (2003). Item-specific control of automatic processes:

    Stroop process dissociations. *Psychonomic Bulletin & Review, 10,* 639-644.

Johnson, R.F., & McMenemy, D.J. (1989). Target detection, rifle marksmanship, and mood

    during three hours of simulated sentry duty. *Proceedings of Human Factors and*

    Ergonomics Society Annual Meeting, 33, 1414-1418. doi:

    10.1177/154193128903302011.

Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological*

    Bulletin, 96, 29-44. http://dx.doi.org/10.1037/0033-2909.96.1.29

Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy:

    Evidence for a domain general learning mechanism. *Cognition, 83*, B35-B42.

    doi:10.1016/S0010-0277(02)00004-5.

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology, 77*, 576-588. http://dx.doi.org/10.2307/1420768

Kundel, H.L. (1982). Disease prevalence and radiological decision making. *Investigative Radiology, 17*, 107-109.

Kundel, H. L. (2000). Disease prevalence and the index of detectability: A survey of studies of lung cancer detection by chest radiology. In E.A. Krupinski (Ed.), *Medical imaging 2000: Image perception and performance*. Proceedings of the International Society of Optical Engineering, 3981, 135-144.

Laberge, D., & Tweedy, J.R. (1964). Presentation probability and choice time. *Journal of Experimental Psychology, 68*, 477-481. http://dx.doi.org/10.1037/h0048691

Lewicki, P. (1985). Nonconscious biasing effects of single instances on subsequent judgments. *Journal of Personality and Social Psychology, 48*, 563–574.

Lindsay, D.S., & Jacoby, L.L. (1994). Stroop process dissociations: The relationships between facilitation and interference. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 219–234.

Logan, G.D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology, 12*, 523-553. doi:10.1016/0010-0285(80)90019-5

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492-527. http://dx.doi.org/10.1037/0033-295X.95.4.492.

MacLeod, C.M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163-203. http://dx.doi.org/10.1037/0033-2909.109.2.163.

Mackworth, N.H. (1948). The breakdown of vigilance during prolonged visual search.

    Quarterly Journal of Experimental Psychology, 1, 6-21.

    http://dx.doi.org/10.1080/17470214808416738

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects.

    Psychological Review, 110, 422-471. http://dx.doi.org/10.1037/0033-295X.110.3.422

Melara, R. D., & Mounts, J. R. (1993). Selective attention to Stroop dimensions: Effects of

    baseline discriminability, response mode, and practice. *Memory & Cognition, 21*, 627-

    645. http://dx.doi.org/10.3758/BF03197195

Musen, G., & Squire, L.R. (1993). Implicit learning of colour-word associations using a Stroop

    paradigm. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*,

    789-798. http://dx.doi.org/10.1037/0278-7393.19.4.789

Neely, J.H. (1977). Semantic priming and retrieval from lexical memory: Roles of

    inhibitionless spreading activation and limited-capacity attention. *Journal of*

    Experimental Psychology: General, 106, 226-254. http://dx.doi.org/10.1037/0096-

    3445.106.3.226.

Nissen, M.J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from

    performance measures. *Cognitive Psychology, 19*, 1-32. doi:10.1016/0010-

    0285(87)90002-8

Pachella, R. (1074). The interpretation of reaction time in information-processing research. In

    B. H. Kantowitz (Ed.). *Human information processing: Tutorials in performance and*

    cognition (pp. 41-82). Hillsdale, NJ: Erlbaum.

Pavlov, I. P. (1902). *The work of the digestive glands: lectures*. C. Griffin.

Posner, M.I., & Snyder, C.R.R. (1975). Attention and cognition control. In R.L. Solso (Ed.),

> *Information processing and cognition: The Loyola symposium.* Hillsdale, NJ: Erlbaum.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and*

> *Verbal Behavior, 6,* 855-863. doi:10.1016/S0022-5371(67)80149-X.

Roediger, H.L., & McDermott, K.B. (1995). Creating false memories: Remembering words

> not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and*
>
> *Cognition, 21,* 803-814. http://dx.doi.org/10.1037/0278-7393.21.4.803.

Sabri, M., Melara, R. D., & Algom, D. (2001). A confluence of contexts: Asymmetric versus

> global failures of selective attention to Stroop dimensions. *Journal of Experimental*
>
> *Psychology: Human Perception and Performance, 27*(3), 515-537.
>
> http://dx.doi.org/10.1037/0096-1523.27.3.515

Saffran, J.R. (2003). Statistical language learning: Mechanisms and constraints. *Current*

> *Directions in Psychological Science*, *12*, 110–114. doi:10.1111/1467-8721.01243

Schmidt, J.R. (2013). The parallel episodic processing (PEP) model: Dissociating contingency

> and conflict adaptation in the item-specific proportion congruent paradigm. *Acta*
>
> *Psychologica, 142,* 119-126. doi:10.1016/j.actpsy.2012.11.004.

Schmidt, J.R., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing

> to do with congruency and everything to do with contingency. *Journal of Experimental*
>
> *Psychology: Learning, Memory and Cognition, 34*, 514-523.
>
> http://dx.doi.org/10.1037/0278-7393.34.3.514.

Schmidt, J.R., Cheesman, J., & Besner, D. (2013). You can't Stroop a lexical decision: Is

> semantic processing fundamentally facilitative? *Canadian Journal of Experimental*
>
> *Psychology, 67,* 130-139.   http://dx.doi.org/10.1037/a0030355.

Schmidt, J.R., Crump, M.J.C., Cheesman, J., & Besner, D. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition, 16*, 421-435. doi:10.1016/j.concog.2006.06.010

Schmidt, J.R., & De Houwer, J. (2012a). Contingency learning with evaluative stimuli: Testing the generality of contingency learning in a performance paradigm. *Experimental Psychology, 59*, 175-182. http://dx.doi.org/10.1027/1618-3169/a000141.

Schmidt, J.R., & De Houwer, J. (2012b). Does temporal contiguity moderate contingency learning in a speeded performance task? *Quarterly Journal of Experimental Psychology, 65*, 408-425. doi:10.1080/17470218.2011.632486.

Schmidt, J.R., & De Houwer, J. (2012c). Learning, awareness, and instruction: Subjective contingency awareness does matter in the colour-word contingency learning paradigm. *Consciousness and Cognition, 21,* 1754-1768. doi:10.1016/j.concog.2012.10.006.

Schmidt, J.R., De Houwer, J., & Besner, D. (2010). Contingency learning and unlearning in the blink of an eye: A resource dependent process. *Consciousness and Cognition, 19,* 235-250. doi:10.1016/j.concog.2009.12.016

Shanks, D.R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology, 60*, 291-309. doi:10.1080/17470210601000581.

Shim, J., Carlton, L.G., Chow, J.W., & Chae, W. (2005). The use of anticipatory visual cues by highly skilled tennis players. *Journal of Motor Behaviour, 2,* 164-175. doi:10.3200/JMBR.37.2.164-175

Sorabji, R. (2004). *Aristotle on memory* (2nd ed.). Chicago: University of Chicago Press.

Thorndike, R.L. (1898). Animal intelligence: An experimental study of the associative

    processes in animals. *Psychological Monographs: General and Applied, 2*, i-109. doi:

    10.1037/h0092987

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the

    effectiveness of reinforcement and nonreinforcement. In A.H. Black & W.F. Prokasy

    (Eds.), *Classical Conditioning II: Current research and theory* (pp. 64–99). New York:

    Appleton-Century-Crofts.

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier

    elimination. *Quarterly Journal of Experimental Psychology, 47A*, 631-650.

    http://dx.doi.org/10.1080/14640749408401131

Wang, M.J.J., Lin, S.C., & Drury, C.G (1997). Training for strategy in visual search.

    International Journal of Industrial Ergonomics, 20, 101-108. doi:10.1016/S0169-

    8141(96)00043-1.

Wolfe, J.M., Horowitz, T.S., & Kenner, N.M. (2005). Rare items often missed in visual

    searches. *Nature, 435,* 439-440. doi:10.1038/435439a

Wolfe, J.M., & Van Wert, M.J. (2010). Varying target prevalence reveals two dissociable

    decision criteria in visual search. *Current Biology, 20*, 121-124.

    doi:10.1016/j.cub.2009.11.066.

Wolfe, J.M., Horowitz, T.S., Van Wert, M.J., Kenner, N.M., Place, S.S., & Kibbi, N. (2007).

    Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of*

    Experimental Psychology: General, 136, 623-638. http://dx.doi.org/10.1037/0096-

    3445.136.4.623.

Appendix A

Tables: All tables depict the occurrences of each stimulus in each condition per block.

Table 1: Exp. 1 (Transfer). There were two phases (Contingency, No Contingency) with two contingency conditions (Hi, Lo).

Contingency Phase			No Contingency Phase		
Hi Contingency	Lo Contingency		Formerly Hi Contingency	Formerly Lo Contingency	
Month (8x)	Month (1x)	Month (1x)	Month (3x)	Month (3x)	Month (3x)
Under (8x)	Under (1x)	Under (1x)	Under (3x)	Under (3x)	Under (3x)
Plate (8x)	Plate (1x)	Plate (1x)	Plate (3x)	Plate (3x)	Plate (3x)

Table 2: Exp. 2 (Baseline). There were three conditions: high contingency, low contingency, and no contingency (baseline).

Hi Contingency	Lo Contingency	
Month (10x)	Month (1x)	Month (1x)
Under (10x)	Under (1x)	Under (1x)
Plate (10x)	Plate (1x)	Plate (1x)
No Contingency		
Clock (4x)	Clock (4x)	Clock (4x)

Table 3: Exp. 3 (Baseline Transfer). There were three conditions, high contingency, low contingency, and no contingency (baseline).

Hi Contingency	Lo Contingency	
Month (10x)	Month (1x)	Month (1x)
Under (10x)	Under (1x)	Under (1x)
Plate (10x)	Plate (1x)	Plate (1x)
No contingency		
Clock (4x)	Clock (4x)	Clock (4x)

Table 4: Exp. 4 (Baseline Nonwords). There were three conditions, high contingency, low contingency, and no contingency (baseline).

Hi Contingency	Lo Contingency	
throg (10x)	throg (1x)	throg (1x)
dwips (10x)	dwips (1x)	dwips (1x)
brask (10x)	brask (1x)	brask (1x)
No Contingency		
flabe (4x)	flabe (4x)	flabe (4x)

Table 5: Exp. 5 (Frequency vs Contingency I). There were two conditions (HiC-LoF and LoC-HiF).

Conditions	Stimuli	
HiC-LoF	Month (6x)	
LoC-HiF	Under (6x)	Under (6x)
	Plate (6x)	Plate (6x)

Table 6: Exp. 6 (Frequency vs Contingency II). There were 6 conditions in total but the bottom two conditions were included only to balance colours, and so were eliminated for analysis purposes (HiC-ExLoF and LoC-HiF).

Conditions	Stimuli		Frequency
HiC-MedF	Month (8x)		8
HiC-LoF	Under (4x)		4
LoC-MedF	World (4x)	World (4x)	8
LoC-LoF	Clock (2x)	Clock (2x)	4
HiC-ExLoF	Plate (2x)		2
LoC-HiF	Table (8x)	Table (8x)	16

Table 7: Exp. 7 (Double Frequency). There were six words presented in high or low contingency in two different frequency conditions, one with double the frequency of the other (Hi-Lo vs Hi2-Lo2).

Hi Contingency (Hi)		Lo Contingency (Lo)	
Month (8x)	Month (1x)	Month (1x)	
Under (8x)	Under (1x)	Under (1x)	
Plate (8x)	Plate (1x)	Plate (1x)	
Hi Contingency (Hi2)		Lo Contingency (Lo2)	
Clock (16x)	Clock (2x)	Clock (2x)	
Table (16x)	Table (2x)	Table (2x)	
World (16x)	World (2x)	World (2x)	

Table 8: Exp. 8 (Semantic vs. Episodic). High and low contingency items were either from the same category (semantic) or from different categories (episodic).

Condition: Between-Subjects	Hi Contingency		Lo Contingency
Semantic: Two related words	flute (4x)	piano (4x)	flute (1x) piano (1x)
	horse (4x)	camel (4x)	horse (1x) camel (1x)
	farmer (4x)	nurse (4x)	farmer (1x) nurse (1x)
Episodic: Two unrelated words	apple (4x)	piano (4x)	apple (1x) piano (1x)
	horse (4x)	train (4x)	horse (1x) train (1x)
	baker (4x)	elbow (4x)	baker (1x) elbow (1x)

Table 9: Exp. 9 (Semantic Transfer). There were two phases (Contingency and No Contingency) with categorized items of high or low contingency in the first phase.

Contingency Phase			No Contingency Phase		
Hi Contingency		Lo Contingency	Equal probability		
flute (4x)	piano (4x)	flute (1x) piano (1x)	banjo cello (3x)	banjo cello (3x)	banjo cello (3x)
horse (4x)	camel (4x)	horse (1x) camel (1x)	zebra tiger (3x)	zebra tiger (3x)	zebra tiger (3x)
farmer (4x)	nurse (4x)	farmer (1x) nurse (1x)	baker lawyer (3x)	baker lawyer (3x)	baker lawyer (3x)

Table 10: Exp. 10 (Semantic Baseline). There were two phases: a Contingency Phase with four contingency conditions and a No Contingency Phase.

Contingency Phase				No Contingency Phase		
Hi	Hi-Related	Lo-Related	Lo	Equal probability		
flute (7x)	piano (1x)	piano (1x)	flute (1x)	banjo cello (3x)	banjo cello (3x)	banjo cello (3x)
horse (7x)	camel (1x)	camel (1x)	horse (1x)	zebra tiger (3x)	zebra tiger (3x)	zebra tiger (3x)
farmer (7x)	nurse (1x)	nurse (1x)	farmer (1x)	baker lawyer (3x)	baker lawyer (3x)	baker lawyer (3x)

Experiment 1
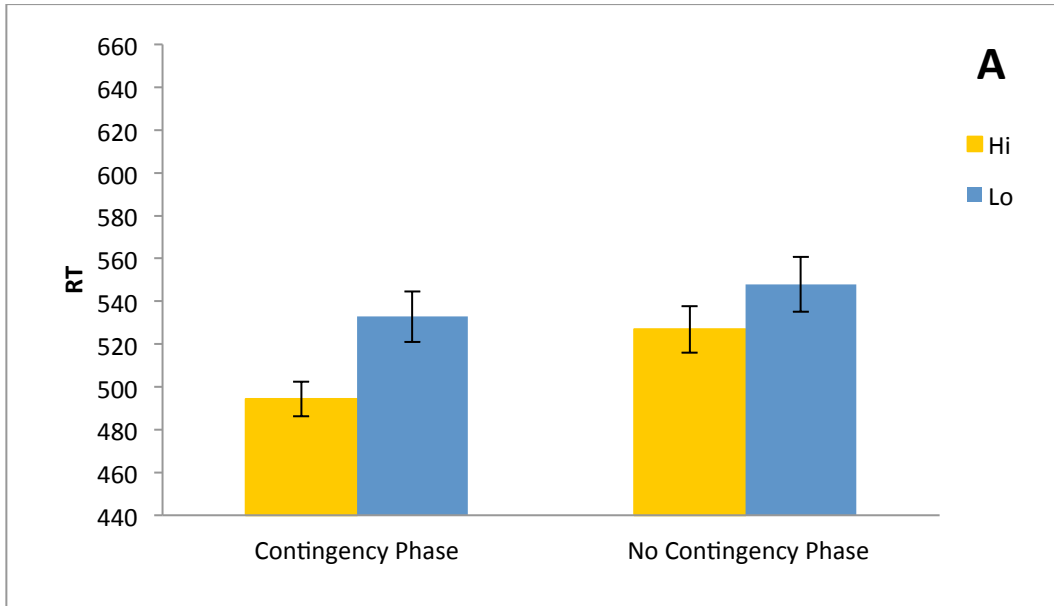
Figure 1A: Transfer. Response times for high vs. low contingency items for the Contingency Phase and the No Contingency Phase. Error bars depict two standard errors.
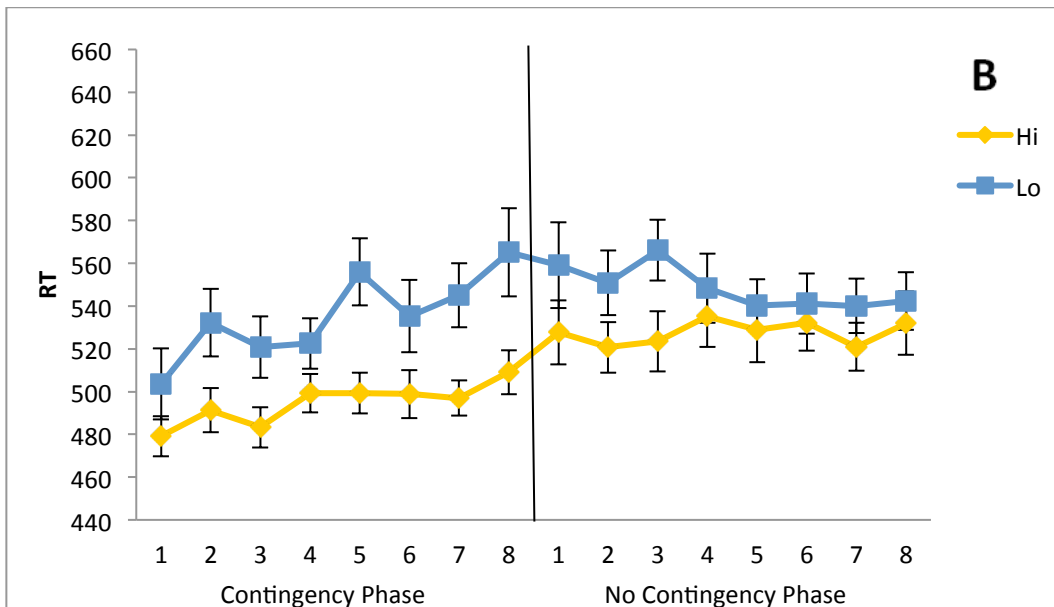
Figure 1B: Transfer. Response times by blocks for high vs. low contingency items for the Contingency Phase and the No Contingency Phase. Error bars depict two standard errors.

Experiment 2
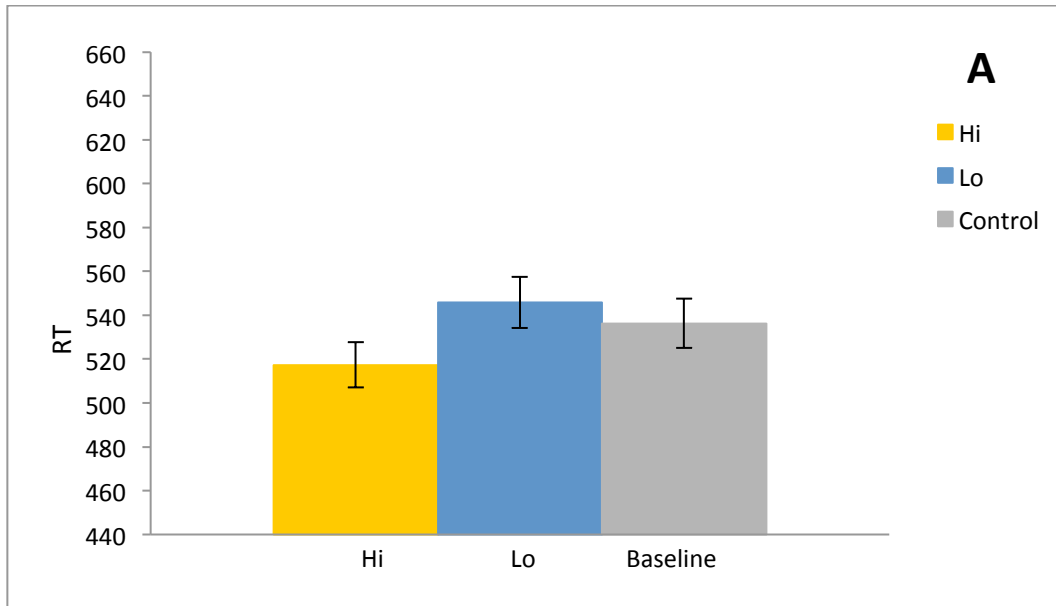
Figure 2A: Baseline. Response times for high contingency, low contingency, and no contingency (baseline) items. Error bars depict two standard errors.
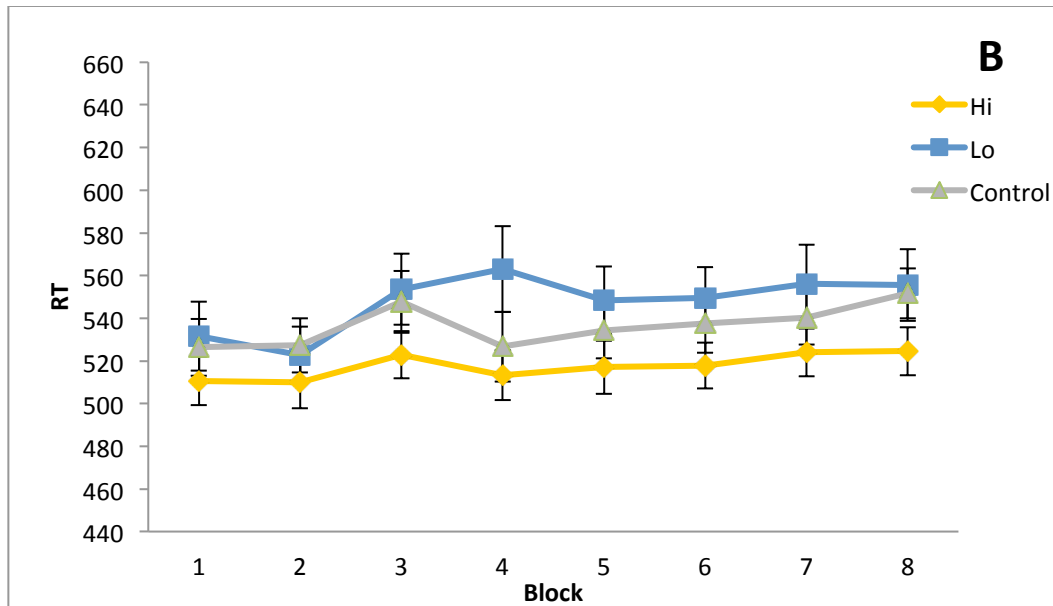
Figure 2B: Baseline. Response times by blocks for high contingency, low contingency, and no contingency (baseline) items. Error bars depict two standard errors.
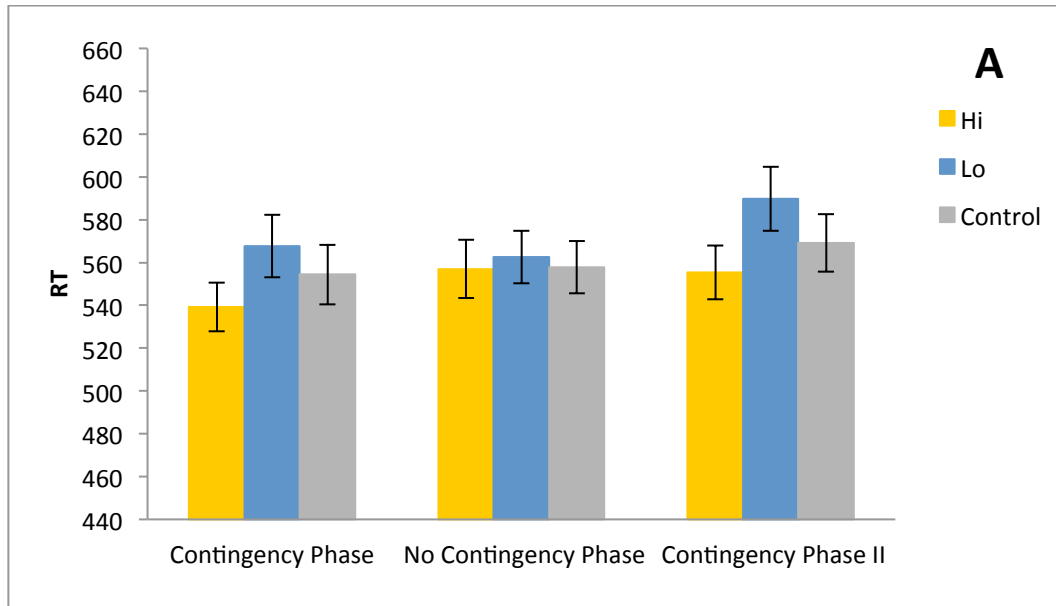
Experiment 3

Experiment 3A: Baseline Transfer. Response times for high, low, and no contingency items for Contingency Phase I, No Contingency Phase, and Contingency Phase II. Error bars depict two standard errors.
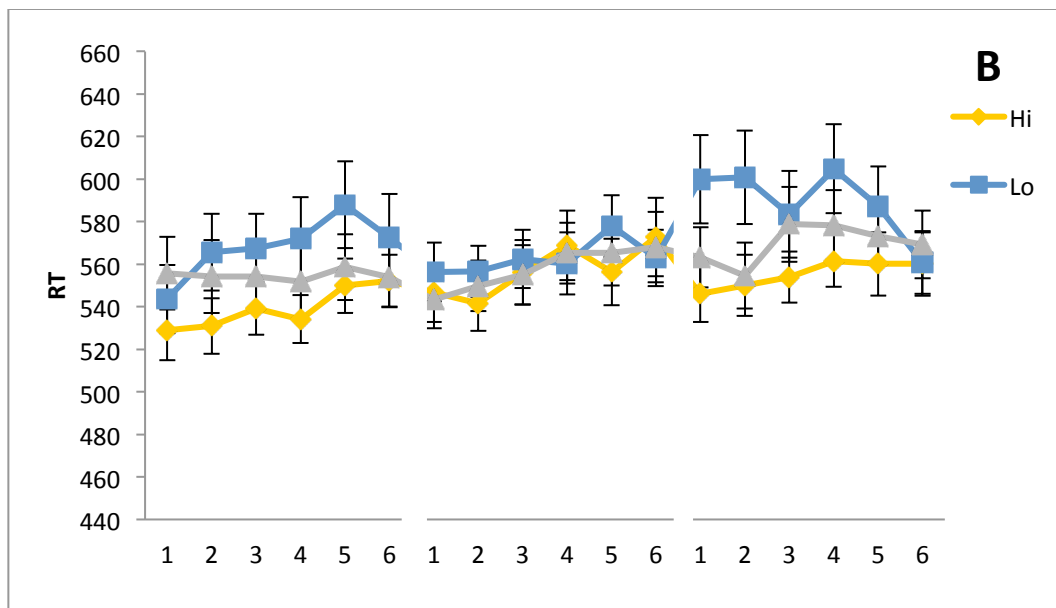
Experiment 3B: Baseline Transfer. Response times by blocks for high, low, and no contingency items for Contingency Phase I, No Contingency Phase, and Contingency Phase II. Error bars depict two standard errors.

Experiment 4
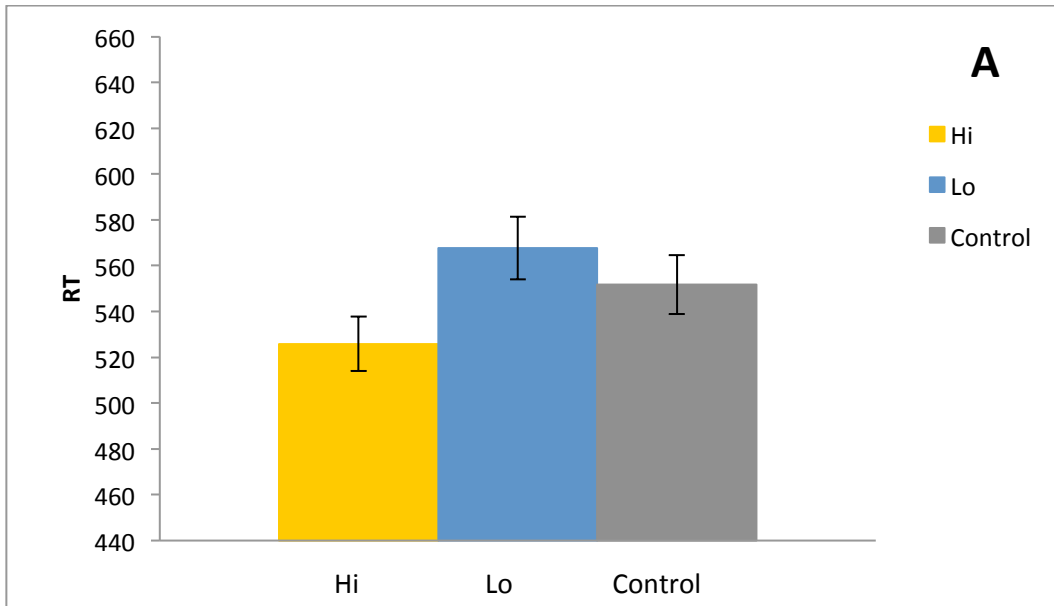
Figure 4A: Baseline Nonwords. Response times for high contingency, low contingency, and no contingency (baseline) items. Error bars depict two standard errors.
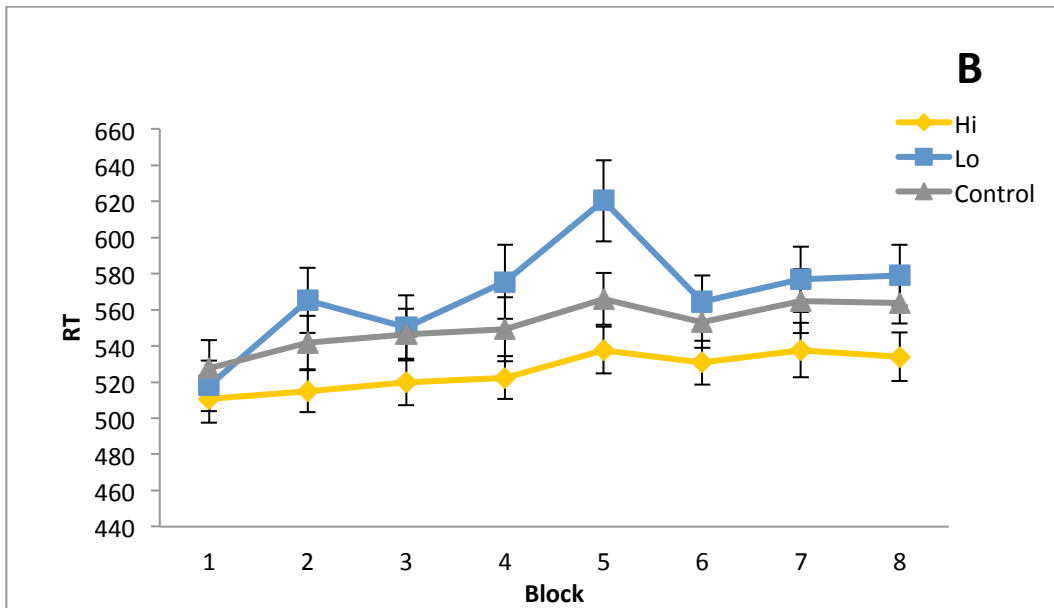
Figure 4B: Baseline Nonwords. Response times by blocks for high contingency, low contingency, and no contingency (baseline) items. Error bars depict two standard errors.

Experiment 5
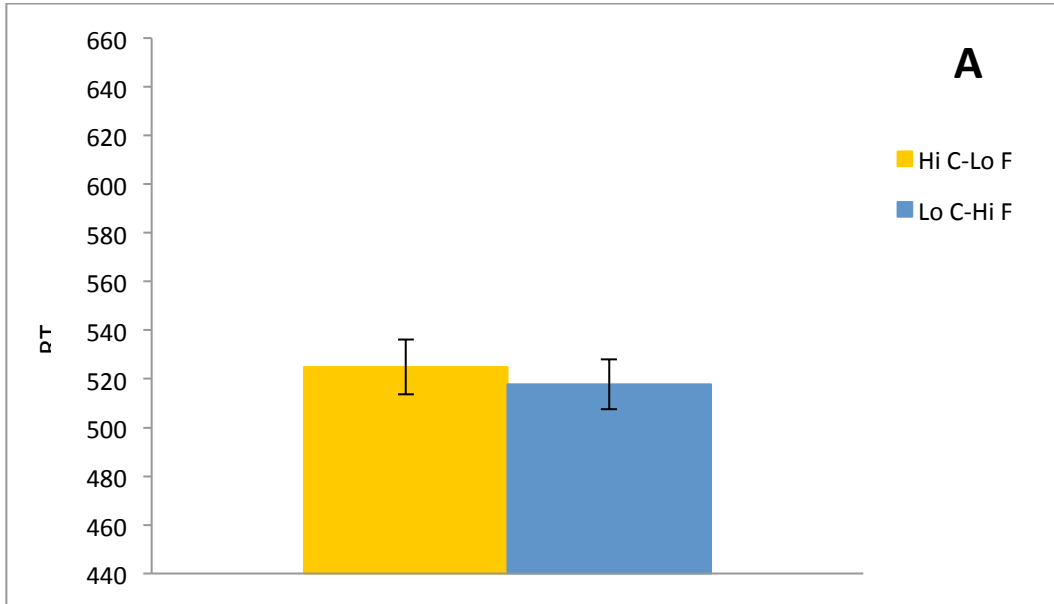
Figure 5A: Frequency vs. Contingency I. Response times for two conditions: HiC-LoF and LoC-HiF. Error bars depict two standard errors.
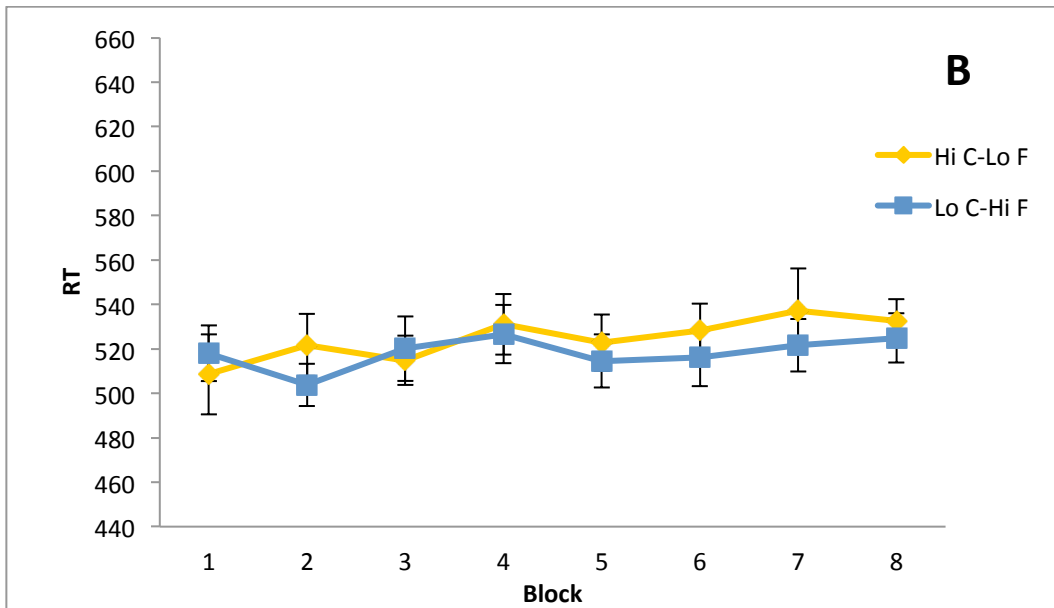
Figure 5B: Frequency vs. Contingency I. Response times by blocks for two conditions: HiC-LoF and LoC-HiF. Error bars depict two standard errors.

Experiment 6
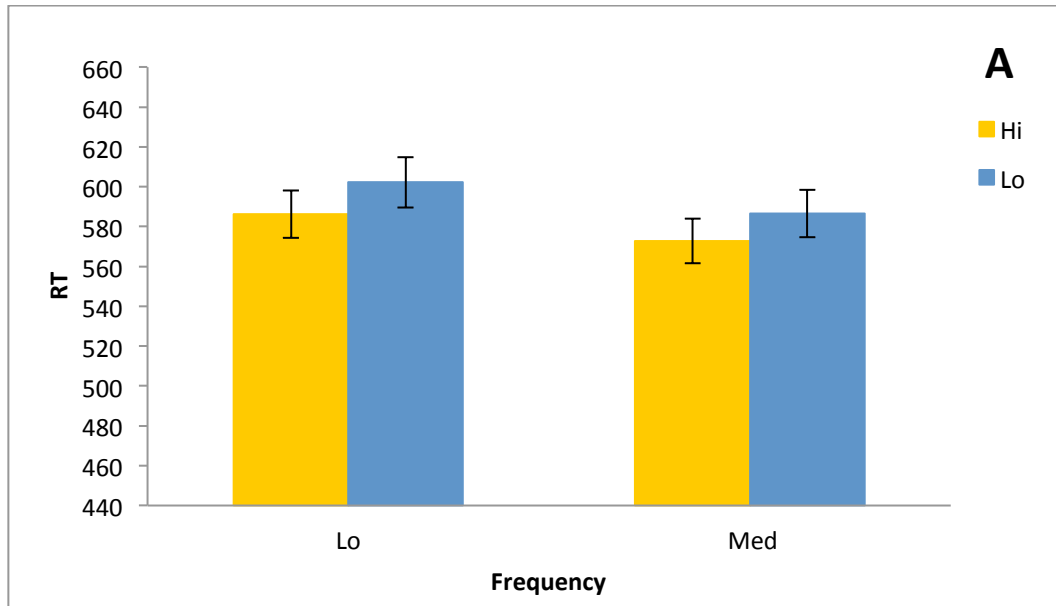
Figure 6A: Frequency vs. Contingency II. Response times as a function of high vs. low contingency items and Lo vs. Med Frequency. Error bars depict two standard errors.
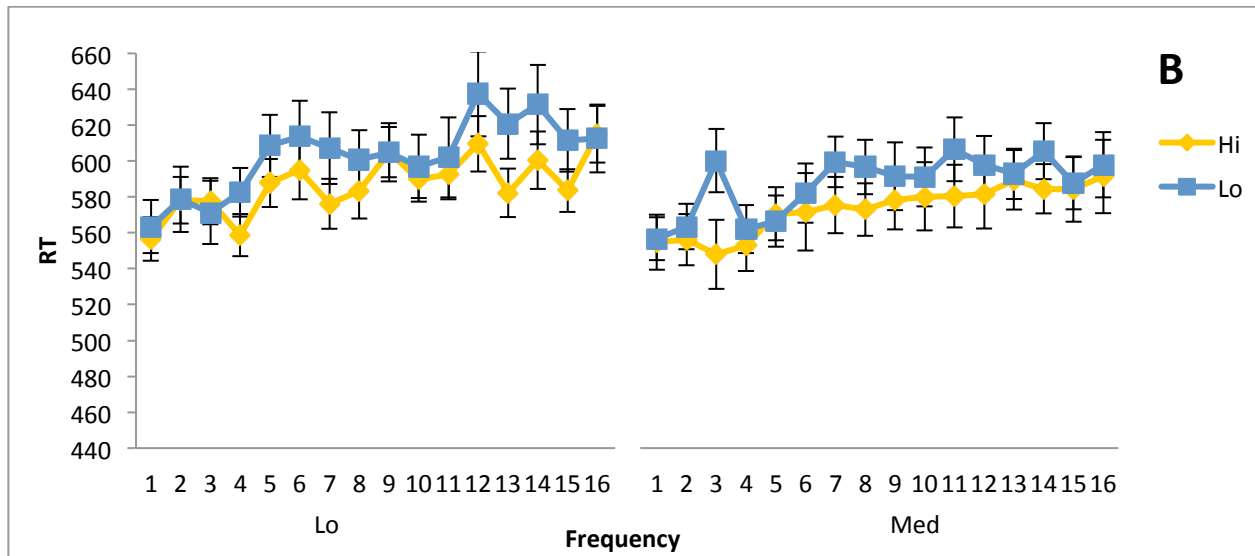
Figure 6B: Frequency vs. Contingency II. Response times by blocks as a function of high vs. low contingency items, and Lo vs. Med frequency. Error bars depict two standard errors.

Experiment 7
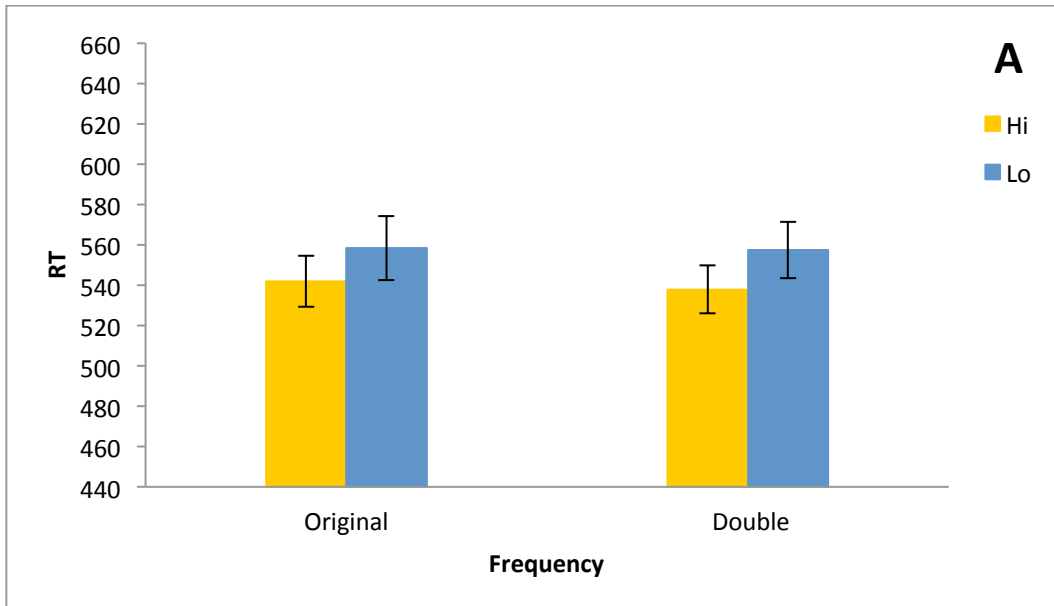
Figure 7A: Double Frequency. Response times for high vs. low contingency items for two frequency conditions (Original, Double). Error bars depict two standard errors.
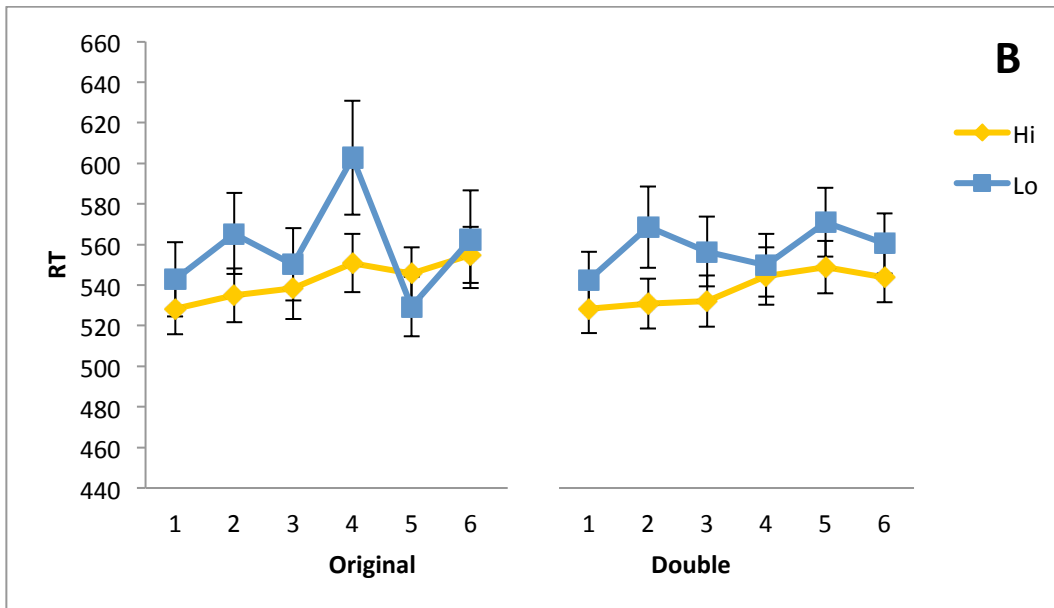
Figure 7B: Double Frequency. Response times by blocks for high vs. low contingency items for two frequency conditions (Original, Double). Error bars depict two standard errors.

Experiment 8
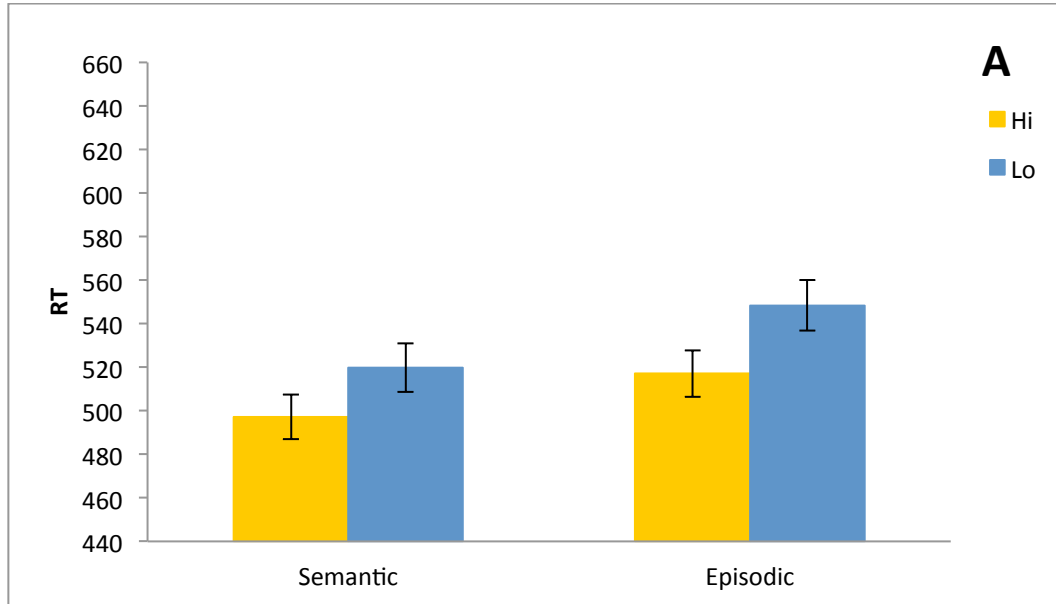
Figure 8A: Semantic vs. Episodic. Response times for high vs. low contingency items for two between-subjects conditions (Semantic, Episodic). Error bars depict two standard errors.
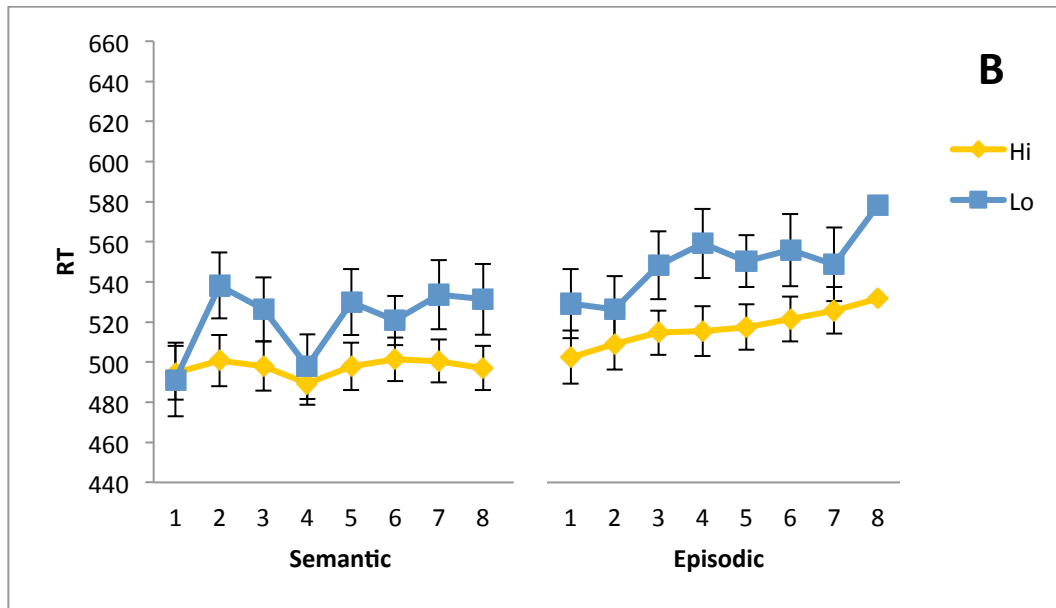
Figure 8B: Semantic vs. Episodic. Response times by blocks for high vs. low contingency items for two between-subjects conditions (Semantic, Episodic). Error bars depict two standard errors.
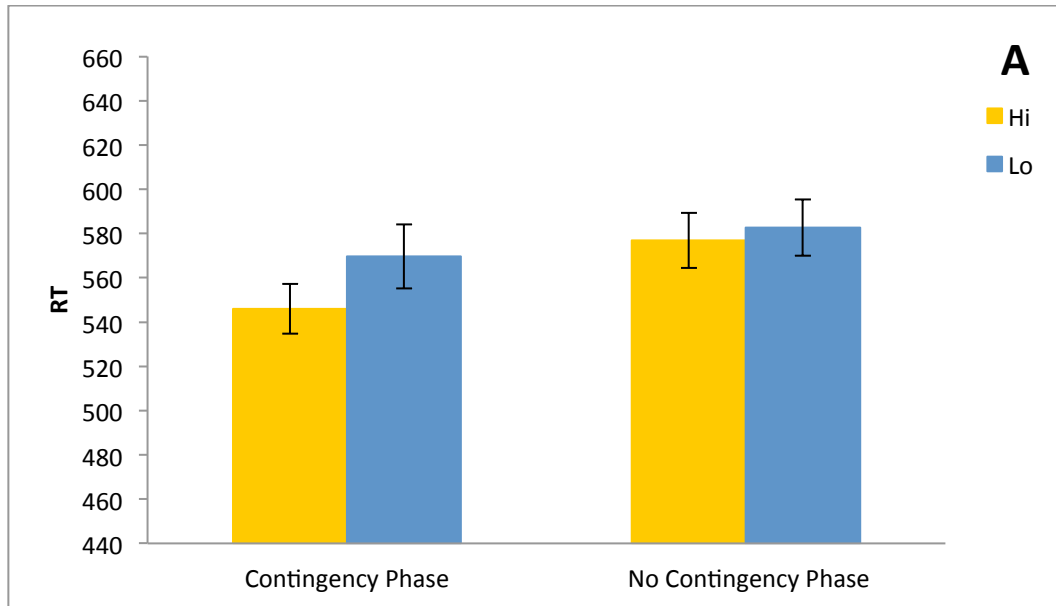
Experiment 9

Figure 9A: Semantic Transfer. Response times for high vs. low contingency items for the Contingency Phase and the No Contingency Phase. Error bars depict two standard errors.
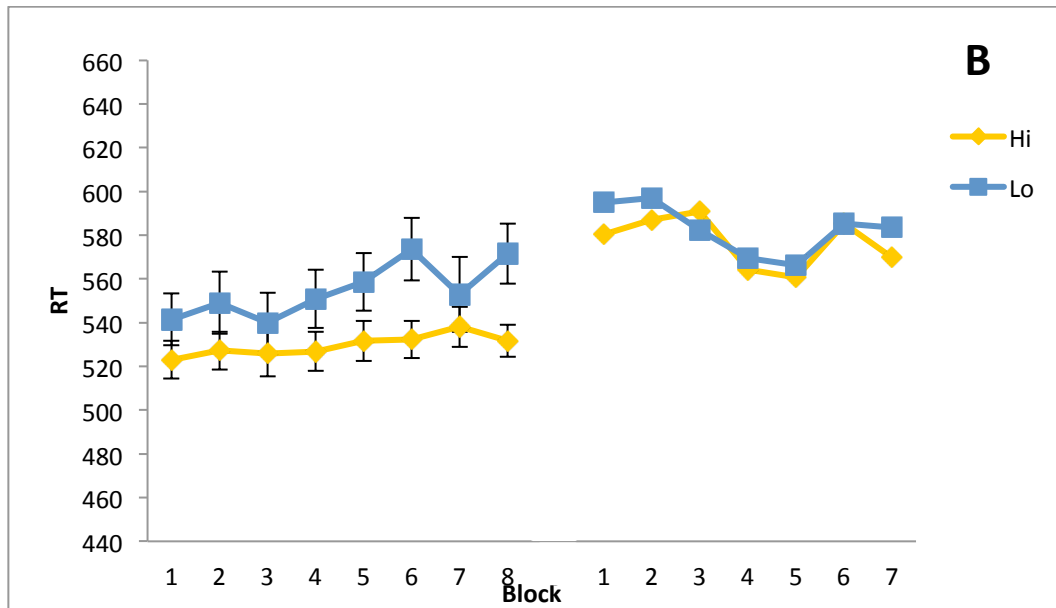

Figure 9B: Semantic Transfer. Response times by blocks for high vs. low contingency items for the Contingency Phase and the No Contingency Phase. Error bars depict two standard errors.
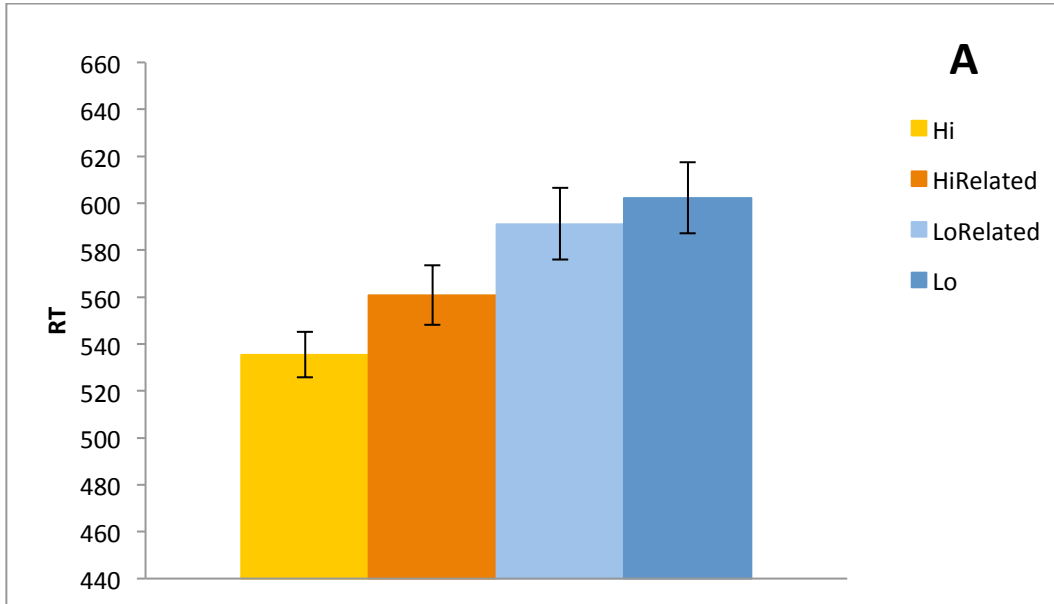
Experiment 10

Figure 10A: Semantic Baseline. Response times for Hi, Hi-Related, Lo-Related, and Lo contingency items. Error bars depict two standard errors.
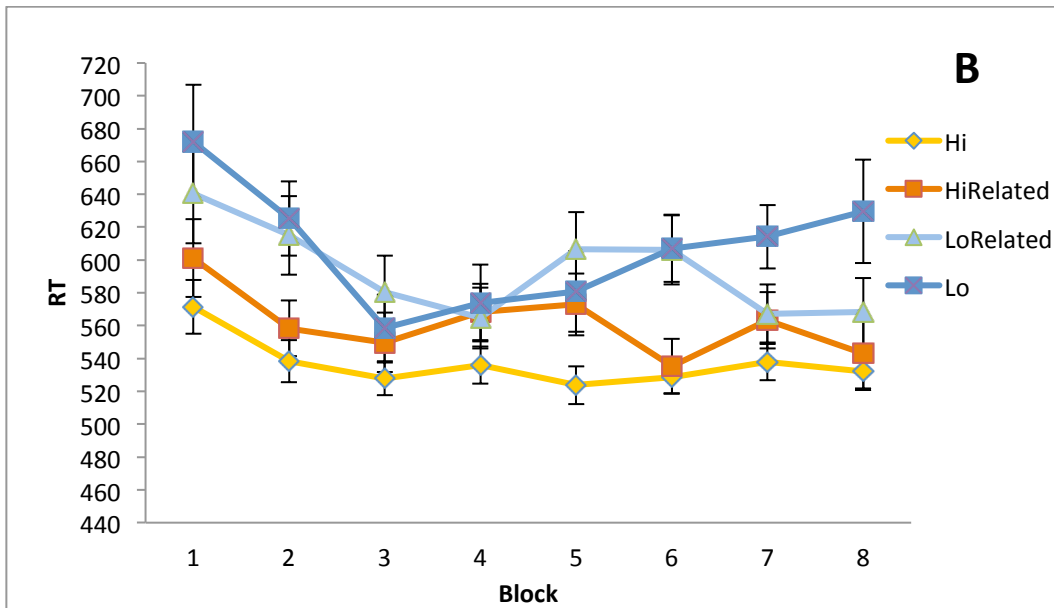

Figure 10B: Semantic Baseline. Response times by blocks for Hi, Hi-Related, Lo-Related, and Lo contingency items. Error bars depict two standard errors.