How Working Memory Moderates Function Learning Behaviour: A Dual-Task Paradigm

by

Billal Ghadie

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of the requirements for the degree of

Master of Arts
in
Psychology

Carleton University

Ottawa, Ontario

Abstract

A breadth of research has demonstrated that many cognitive phenomena can be explained by a dual-processing account. However, little research has attempted to apply a dual-task paradigm to function learning. The present thesis aims to fill this gap in the literature by exploring the relationship between working memory and function learning behaviour. Eighty Carleton University students were randomly assigned to learn either a linear or bilinear function. Moreover, participants were randomly assigned to complete training and transfer under either single- or dual-task conditions. It was hypothesized that the secondary task would hinder performance resulting in a dependency on exemplar-based learning. Using a novel classification approach, the results showed that the secondary task reduced the stability of learning approach. However, the results remain inconclusive due to low power. Therefore, additional research is required to determine whether dual-task paradigms can be used to distinguish between rule- and exemplar-based processing in function learning.

Acknowledgments

To my colleagues, who have helped me through the highs and lows of my master's degree. To Guy, who's tutelage shaped the man I have become. To my mom Awatef, who was never short on love and support. Finally, to Fay for helping me keep my eyes on the stars and my feet on the ground. Thank you.

List of Contents

List of Tables

List of Figures

How Working Memory Moderates Function Learning Behaviour: A Dual-Task Paradigm

The ability to learn conceptual relationships is an important facet of human cognition. Primarily, cognitive psychologists have focused on the categorization aspect of concept learning (See Kruschke, 1992; Nosofsky & Kruske, 1992), which underlies humans' interpretation of the world. For instance, it allows botanists to classify plants as poisonous versus harmless, and students to perceive professors as engaging versus monotonous (Goldstone, Kersten, & Paulo, 2012). Research that examines how individuals learn conceptual relationships involving continuous stimuli has received somewhat less attention, however. This phenomenon is illustrated by runners' ability to predict the distance that they have travelled as a function of time, or bar patrons' attempts to estimate blood alcohol levels as a function of the number of alcoholic drinks consumed (Kalish, Lewandowsky, & Kruschke, 2004). Such relationships are typically described by a mathematical function and researchers assess them using a function learning tasks (Brown & Lacroix, 2018; Brown & Lacroix, 2017; Busemeyer, Byun, Delosh & McDaniel, 1997; Delosh, Busemeyer, & McDaniel, 1997; Kalish et al., 2004; Kwantes & Neal, 2006; Kwantes, Neal, & Kalish, 2012).

There is a growing body of research examining how individuals learn functional relationships. Traditionally, research has sought to describe the mechanisms that underlie function learning. In line with the categorization literature (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Maddox, Ashby, & Bohil, 2003; Waldron & Ashby, 2001), it has been assumed that learning functional relationships involves either a rule-based or an exemplar-based approach (McDaniel et al., 2014). The former requires individuals to develop a rule by becoming sensitive to the underlying regularities in stimuli-response magnitudes. During the transfer phase, the developed rule is used to extrapolate when presented with novel stimuli. In contrast, the exemplar-based approach postulates that the memorization of the stimulus-response pairs drives function learning. When novel stimuli are encountered, they are compared to existing exemplars and the most similar one determines the response (McDaniel et al., 2014). Intuitively, rule-based learning would seem more versatile because responses from exemplar-based learning are limited only to previously encountered values. Contemporary research, however, has shown that a hybrid approach seems to best explain function learning behaviour. Specifically, Delosh et al. (1997) created and validated the extrapolation associative model (EXAM) which posits that function learning involves both exemplar- and rule-based mechanisms. Extending this research, McDaniel et al. (2014) demonstrated that individuals vary in the degree to which they employ an exemplar-based or a rule-based approach. Furthermore, the preference for one approach over the other is moderated by working memory capacity. Those who display a larger working memory capacity tend to favour a rule-based approach whereas those with lower working memory capacity tend to favour an exemplar-based approach (McDaniel et al., 2014).

This dichotomy between rule- and exemplar-based learning bears many similarities to sequence learning (Curran & Keele, 1993; Keele, Ivry, Mayr, Hazeltine, & Heuer, 2003) and categorization research (Ashby et al., 1998; Maddox et al., 2003; Waldron & Ashby, 2001). Specifically, theories across both these learning domains have posited the existence of an explicit rule-based and implicit exemplar-based learning mechanisms. Nonetheless, very little research has been conducted using dual-task paradigms to better understand how these mechanisms contribute to function learning. In fact, only one set of experiments has attempted to examine function learning in this context and found ambiguous results (Brown & Lacroix, 2018).

Thus, the goal of the present thesis it to fill this gap in the literature by extending research that explores the cognitive mechanisms involved in function learning using a dual-task paradigm

(Brown & Lacroix, 2018). First, a theoretical background will be presented outlining the prevalent models that best describe function learning as well as typical function learning behaviour. Then, parallel research in sequence and category learning will be used to reinforce the validity of using dual-task paradigm in this line of research. To dissociate between the rule-based and exemplar-based memory systems in function learning, I adapted Curran and Keele's (1993) methodology. Specifically, levels of distraction during the training and transfer phases of a function learning task were manipulated through the presence or absence of a concurrent recall task (Brown & Lacroix, 2018). I then present the hypotheses of the proposed study followed by a detailed description of the employed methodological approach. Finally, results will be reported and interpreted within the context of the outlined theoretical framework.

Theories of Function Learning Behaviour

Individuals' ability to learn the relationship between two continuous X-Y variables across a set of training items is typically studied using the function learning paradigm (Delosh, Busemeyer, & McDaniel, 1997). First, participants are given a cover story instructing them to learn the relationship between two variables (e.g., a given quantity of fertilizer and plant growth). Then, for each trial, an X-value is presented (e.g., fertilizer) and individuals must estimate the corresponding Y-value (e.g., plant growth). Participants must guess at first, but feedback is given after each trial so that they can eventually infer the correct answer. Researchers have used a variety of functions that include linear, quadratic, or sinusoidal patterns (Brown & Lacroix, 2018; Delosh et al., 1997; McDaniel et al., 2014).

Once training is completed, participants proceed to the transfer phase. Their knowledge of the function is evaluated with a new set of X-values. Some lie within the training region and are called interpolation items while others go beyond the training region and are called extrapolation items (see Figure 1). To assess learning, researchers compare how closely the estimated Y-values match the defined function using absolute deviation (Brown & Lacroix, 2018; Delosh et al., 1997; McDaniel et al., 2014). Therefore, the relationship between accuracy during training and performance during transfer allows researchers to make inferences about how individuals learn functional relationships.

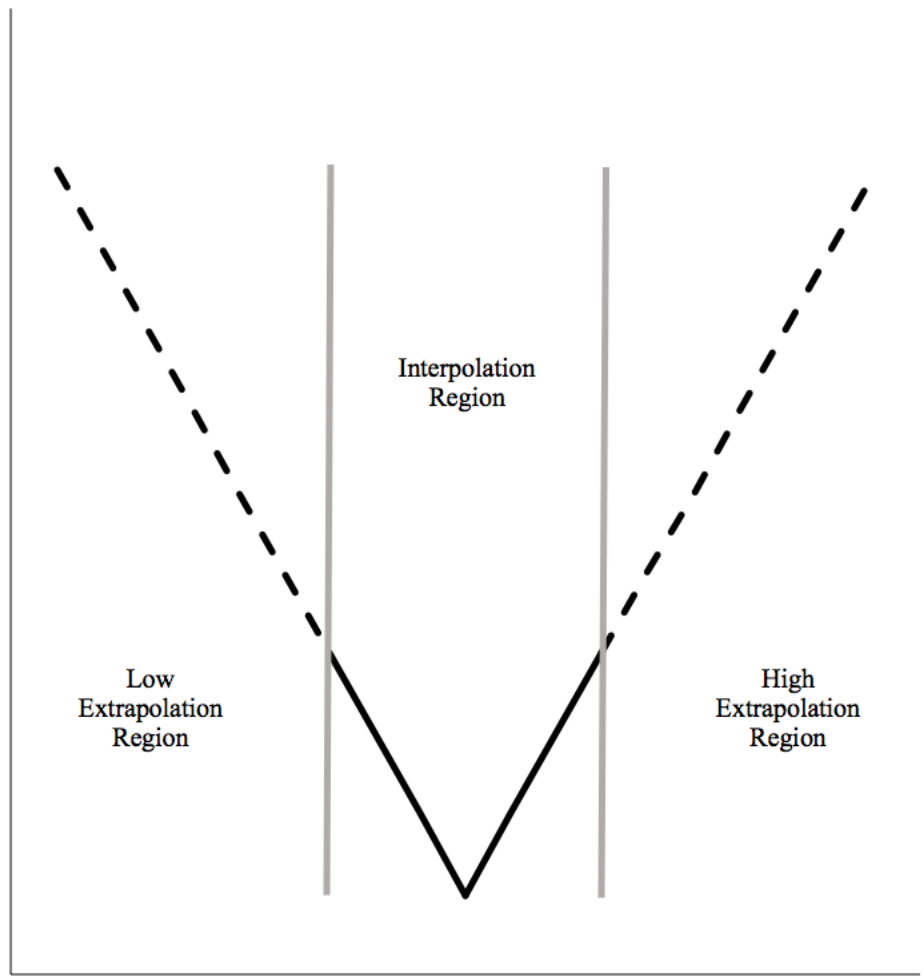

Figure 1. A sample V shaped function used during the transfer phase of a function learning task partitioned by region.

Function learning research has revealed a variety of robust phenomena. Most notably, individuals appear to learn positive linear functions more easily than any other function

(Brehmer 1974; Carroll 1963; Delosh et al., 1994). Initial research suggested that this predisposition resulted from individuals storing cue-criterion pairs in memory and fitting them to the most appropriate function. As simple functions will be learned more effectively than complex ones, it is unsurprising that individuals consistently attempt to fit a positive linear function to cue-criterion pairs (Carroll, 1963). However, Brehmer (1974) argued that this does not take into account the fact individuals learn positive linear functions better than negative linear functions. Thus, he proposed a two-stage model called the adaptive regression model wherein participants first discover the rule that defines the cue-criterion pairs and then subsequently learn to apply the rule to produce accurate responses. Brehmer (1974) assumed individuals are limited in the rules that they apply and that each one differs in strength. As such, he proposed that the order in which they apply each rule occurs hierarchically. To test this model, he conducted a function learning experiment in which participants were tasked with assigning numbers to twenty-line segments. Participants were instructed to use any rule they would like for labelling the presented lines. However, once chosen, participants could not change their rule. Following 20 classifications, participants were asked to draw the rule they had chosen. Participants completed 10 blocks of this task while being instructed to utilize a new rule at the onset of each block. The results demonstrated a predisposition to apply a positive linear rule first, followed by a negative linear rule, then an equal probability of quadratic and inverse-quadratic rule (Brehmer, 1974). Therefore, when learning functional relationships individuals are biased to think of linear functions first.

Despite the ease with which individuals can learn positive linear functions, they nonetheless demonstrate a consistent and systematic error in extrapolating positive linear functions (Brown & Lacroix, 2017; Delosh et al., 1997; Kwantes & Neal, 2006). Specifically, individuals underestimate Y-values during transfer in the low extrapolation region. Kwantes and Neal (2006) outlined two competing hypotheses to explain this phenomenon. First, individuals may be anchoring their responses to zero in the lower extrapolation region. Thus, as the X-values decrease, individuals' responses are reduced, reflecting the assumption that the functional relationship passes through the origin of the Cartesian plane. The second hypothesis posits that individuals assume the Y-value is similar to the presented X-value when they are unsure of response values (Kwantes & Neal, 2006). To differentiate which of the two hypotheses best explains why individuals underestimate Y-values in the low extrapolation region, Brown and Lacroix (2017) conducted a function learning task wherein the Y-intercept was manipulated. Specifically, participants learned a functional relationship with a moderate positive intercept, a large positive intercept, a control in which the intercept was zero, and a negative intercept. If the X-Y similarity theory holds, then consistent underestimation in the lower extrapolation should occur across all conditions except the control. In contrast, if individuals anchor their response values to zero, then participants in the positive intercept conditions should exhibit underestimation in the low extrapolation region, participants in the control condition should demonstrate no underestimation, and participants in the negative intercept condition should overestimate their response variables in the low extrapolation region. Brown and Lacroix's (2017) results were consistent with the anchoring hypothesis. Participants appeared to overestimate response values in the negative intercept condition and underestimate them in the two positive intercept conditions. Therefore, when learning a positive linear function, individuals are predisposed to anchor their responses to zero in the low extrapolation region. In turn, responses in the low extrapolation region will consistently be underestimated for linear functions with a positive intercept.

Along with general function learning behaviour, researchers have also begun to explore the specific mechanisms that allow for learning functional relationships. Because individuals can often extrapolate accurately, for example, researchers have argued that function learning must involve a rule-based approach (Carroll, 1963). However, Koh and Meyer (1991) proposed that models involving exemplar-based mechanisms with more sophisticated parameters for combining and storing stimuli can also account for accurate extrapolation. This fact made it more difficult to conclude which model is correct. Therefore, Koh and Meyer (1991) conducted a function learning task to evaluate three rule-based and four exemplar-based models. To compare them, participants were assigned to complete a function learning task defined by either a linear, logarithmic, or power function. Using extrapolation performance, Koh and Meyer assessed the predictions of the seven function learning models. Across three experiments, the most accurate predictor of extrapolation performance was a rule-based model according to which individuals develop an algorithmic rule during training that they use to accurately extrapolate. Thus, in line with previous research (Brehmer, 1974), function learning was postulated to rely on the use of a rule-based mechanism. However, prevailing cognitive theories have demonstrated that humans can engage in more than one kind of learning (Smith & Church, 2018). Therefore, rather than rely on an exclusive rule- or exemplar-based perspective, it is likely that function learning may be more plausibly described by a hybrid rule- and exemplar-based approach.

The first proponents of a hybrid approach to function learning can be traced back to Delosh et al. (1997). They claimed that prior function learning research (see Brehmer, 1974) had focused primarily on the learning rate of different function types. In fact, very few had up to that point attempted to discern function learning behaviour by testing extrapolation behaviour. Hence, due to the limited research, findings that ruled out the possibility of exemplar-based approach had to be viewed as preliminary. Therefore, Delosh et al. (1997) sought to evaluate the validity of four learning models. The first two models were strictly rule-based models which assume individuals can extrapolate accurately (Carroll, 1963; Brehmer, 1974; Koh & Meyer. 1991). The third model, the associative learning model (ALM), proposed that individuals rely on an exemplar-based approach to learn functional relationships (Busemeyer et al., 1997). As such, extrapolation beyond the interpolation region is impossible. The fourth model, extrapolation-association model (EXAM) is a hybrid model that incorporates the associative learning assumption of the ALM with a rule-based mechanism to allow for accurate extrapolation (Delosh et al., 1997). To assess these models, Delosh et al. (1997) conducted a function learning task using a quadratic, linear, and exponential functions to measure extrapolation performance. The results indicated that participants learned the linear function faster than the quadratic and exponential function. Furthermore, accuracy was consistently higher across all function conditions in the interpolation region compared to either extrapolation regions. Finally, the best predictor of extrapolation performance was EXAM. Therefore, Delosh et al. concluded that people's function learning performance may reflect a combination of exemplar- and rule-based approaches.

Nonetheless, while exploring individual differences in function learning performance, an interesting pattern emerged for a small subset of participants in the quadratic condition, as shown in Figure 2. Specifically, some participants did not extrapolate within the high extrapolation region, some extrapolated in accordance with a purely exemplar-based learning style as described by the ALM, and some extrapolated nearly perfectly as described by rule-based models. Therefore, there appears to be variability in extrapolation behaviour when learning functional relationships.

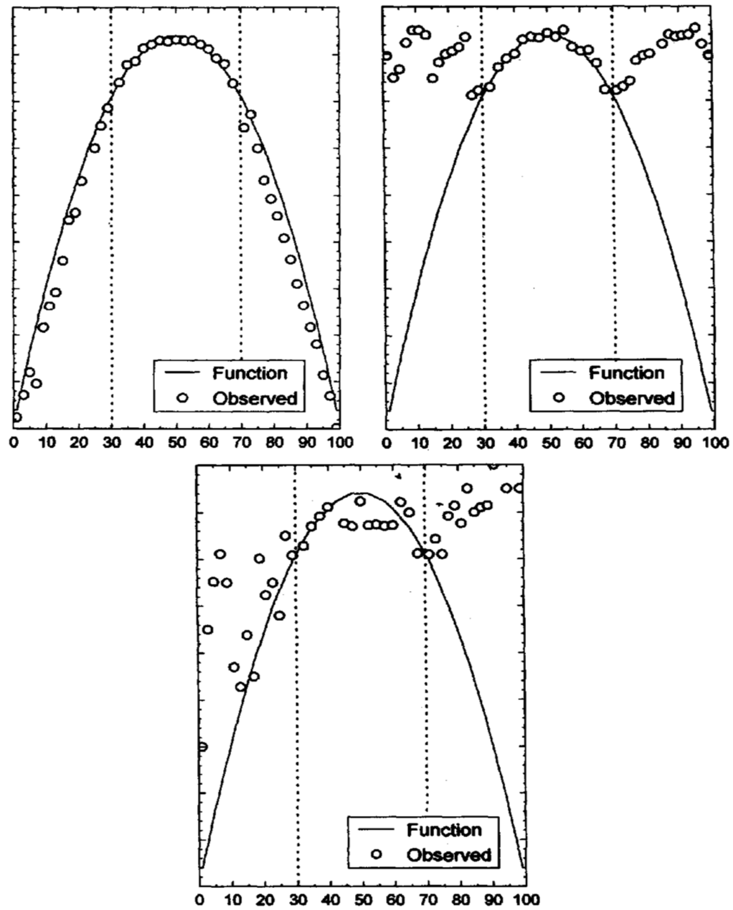

Figure 2. Individual differences in extrapolation performance revealed three distinct learning approaches: rule-based learners (top left), exemplar-based leaners (top right), and non-learners (bottom) (Delosh et al., 1997).

Currently, only one study by McDaniel et al. (2014) has explored how individuals differ in the degree to which they rely on exemplar- or rule-based learning during function learning. Borrowing from research in concept learning, they proposed that individuals tend to rely on either exemplar- or rule-based learning unless a task strongly favours one approach. Moreover, they contended that this predisposition is stable within individuals. In their experiment, participants completed a function learning task, wherein they learned a bilinear, V-shaped function. Through the assessment of mean absolute error (MAE) during the final block of training, participants were categorized as learners or non-learners. Then, those classified as learners were evaluated by comparing their MAE during transfer to the expected MAE produced by an exemplar learning model (McDaniel et al., 2014). Participants who performed significantly different from this MAE threshold set by the exemplar model were classified as rule learners. Otherwise, they were classified as exemplar learners. To corroborate this classification method, the MAE during transfer was correlated with the MAE during the last block of training. McDaniel et al. (2014) posited that rule-based learners would display a positive correlation between the last block of training and MAE whereas exemplar-based learners would not. For the latter group, they argued that knowledge of the cue-response pairings would not be related to extrapolation behaviour. The results confirmed their hypotheses. Thus, once again it appears that individuals have a tendency to utilize a particular approach when learning functional relationships.

Although the preference for a particular learning approach appears stable, little research has been conducted to examine what moderates these tendencies. To assess the cognitive capacities associated with rule- and exemplar-based learning, participants were asked to return for a second experimental session. This session involved replicating the results of McDaniel et al.'s initial study, completing the Raven's Advanced Progressive Matrices (RAMP) to assess fluid intelligence, and the Operation Span task (OSPAN) to measure working memory capacity. The goal was to determine if established measures of individual differences might predict participants' tendency to use a rule- or exemplary-based learning approach. As expected, during the function learning component of the replication, participants tended to favour either a rule-based or exemplar-based approach. Moreover, both fluid intelligence and working memory were significantly correlated with participants' preferred learning approach. However, fluid intelligence accounted only for a small portion of the variance found in function learning behaviour. In contrast, higher working memory appears to be more strongly associated with the predisposition to utilize a rule-based approach and not an exemplar-based approach. McDaniel et al. (2014) proposed that the larger working memory capacity facilitates the rule-based learning process. Specifically, these additional resources enable participants to extract patterns across trials, to partition complex functions into simple components, and to ignore initial biases. Therefore, when working memory capacity is insufficient to allow for such processes to operate, participants must rely on an exemplar-based approach to learn functional relationships.

Thus, it has become more apparent that function learning involves a hybrid approach. In fact, the only viable alternative to EXAM presently under consideration is a different hybrid model, the population of linear experts (POLE) model (Kalish et al., 2004). POLE assumes that individuals learn functional relationships through partitioning complex training items into multiple simpler components. Then, a linear expert function is developed that matches the partitioned training stimuli and produces the correct associated response values (Kwantes et al., 2012). When presented with novel stimuli, the closest matching stimulus in memory is activated, along with its respective expert, to produce an extrapolation response. Therefore, only one linear expert is required to accurately extrapolate linear functions. When a relationship is nonlinear, however, responses are generated from multiple experts with distinct slopes that maximize accuracy. The mixture-of-experts method prescribed by POLE incorporates a connectionist network known as ATRIUM (Erickson & Kruschke, 1998). ATRIUM proposes that categorization involves both a rule-based and exemplar-based mechanisms and comprises of a rule module, exemplar module, and a competitive gating mechanism that links the two (Erickson & Kruschke, 1998; Kruschke & Erickson, 1994). Consequently, it can be viewed as a hybrid model.

To validate POLE, a variant on the traditional function learning task was conducted wherein participants were trained to learn two separate line segments with different y-intercepts but similar slopes. Moreover, each segment was bound by a different region of the Cartesian plane: lower and upper regions. The transfer region composed of stimuli between both line segments that were not displayed during training. Thus, participants could extrapolate by connecting the two-line segments to form a quasi-sinusoidal function as predicted by the EXAM model or they could extrapolate in accordance with one of the line segments, which could be chosen randomly or due to a bias (Kalish et al., 2004). The results revealed that POLE described extrapolation behaviour better than EXAM. This suggests that individuals may not average stimuli-response pairs held in memory when presented with novel stimuli as proposed by EXAM. Rather, individuals may utilize a different rule-based approach in which they apply series of linear experts of varying slopes until one is found to maximize accuracy. Therefore, although there is dispute over how individuals generate responses, a hybrid rule- exemplar-based model appears to best describe function learning behaviour.

Finally, McDaniel, Dimperio, Griego, and Busemeyer (2009) would later conduct a function learning task to determine whether the EXAM or POLE model best describe function learning behaviour. They pointed out that POLE model has six free parameters whereas EXAM only has two. This gives POLE a large advantage in fitting data and makes direct comparison between the models difficult. Moreover, to avoid overfitting, researchers typically favour more parsimonious models (Neter, Kutner, & Wasserman, 1985). Therefore, McDaniel et al. argued that a priori predictions about the models' behaviour would be a stronger test of their adequacy. Thus, they conducted a function learning task in which they manipulated the density of transfer regions and item order. Then, they compared how well each model fit the training data and how well they accurately predicted participants transfer performance. The results indicated that the POLE model fit the training performance across all participants. However, the EXAM was an overall better predictor of transfer performance. The success of EXAM over POLE in predicting transfer behaviour is attributed to how each model is affected by stimuli density during training. Specifically, POLE becomes disproportionally influenced by the dense training regions resulting in transfer predictions based solely on the linear experts developed during training. As such, POLE fails to accurately predict performance when it must account for sparsely trained stimuli in the transfer region. In contrast, EXAM generates transfer responses based on a linear rule associated with the most similar cue-criterion pair encountered during training. Therefore, EXAM is able to make predictions, uninfluenced by the dense training region, that better captures participants performance (McDaniel et al., 2009). Thus, it appears that, in terms of predicting function learning behaviour, EXAM is superior to POLE.

Dual-Task Paradigms

Function learning behaviour is best explained by a hybrid, rule-based and exemplar-based, approach (Delosh et al., 1997; Kalish et al., 2004; Kwantes et al., 2012; McDaniel et al., 2009). Moreover, the preference for one approach over the other varies across individuals (McDaniel et al., 2004). Specifically, individuals who enjoy greater working memory capacity tend to utilize a rule-based approach. In contrast, lower working memory is associated with the preference for an exemplar-based approach. Rather than a single learning mechanism, however, the tendency to prefer one learning approach over another reflects the existence of two distinct learning mechanisms. The application of a dual-processing framework would further support the stable tendency for individuals to prefer either a rule- or exemplar-based strategy to learn functional relationships (McDaniel et al., 2014). In fact, a breadth of research has demonstrated that many cognitive phenomena can be explained by a dual-processing approach (Ashby, Alfonso-Reese, & Waldron, 1998; Curran & Keele, 1993; Knowlton & Squire, 1996; Maddox, Ashby, & Bohil, 2003; Smith & Church, 2017; Waldron & Ashby, 2001).

For instance, Curran and Keele (1993) proposed that sequence learning involves multiple learning systems. Specifically, they described an attentional system that relies on declarative memory and a nonattentional system that relies on procedural memory. To test their theory, they conducted a series of serial reaction time (SRT) tasks. Across a number of blocked training trials, participants were shown stimuli in one of four horizontal positions in a repeating pattern (Curran & Keele, 1993). For each presentation, participants were tasked with learning the keypresses that corresponded to the stimuli's position. Reaction times were used to measure how well participants learn the underlying sequential relationship between them. Participant learning was tested in two ways. The first occurred when participants completed an unexpected randomly sequenced block during training. The randomized block is typically followed by a block that restores the original structured sequence (Soetens, Melis, & Notebaert, 2004). This manipulation is key in determining if a pattern was learned. Specifically, researchers can infer learning if participants respond slower when presented with a randomly sequenced block compared to the preceding and subsequent sequenced blocks. The second sequential effect occurs after participants have been adequately trained to learn the sequential pattern. Following training, they complete four more blocks of trials, two randomly sequenced blocks, one block that follows the original sequence structure, then a final randomly sequenced block (Curran & Keele, 1993). The sequential knowledge participants gained during training is reflected in the decrease in response time when presented with a sequenced block compared to the flanking randomized blocks. Therefore, by manipulating the degree to which the attentional system can effectively function, these sequential effects can be used to differentiate between the attentional and nonattentional learning systems.

To differentiate between the two learning systems, Curran and Keele (1993) asked their participants to perform an SRT task concurrently with a tone-counting task. They posited that the addition of a concurrent task would tax the attentional resource used to learn the stimulus sequences. As such, the attentional learning system would be suppressed and any learning that occurred would result from the nonattentional learning system. Thus, participants were randomly assigned to a single- or dual-task condition. In the single-task condition, participants were trained to learn the sequence without the secondary task, then subsequently tested with the concurrent secondary task. In contrast, participants in the dual-task condition completed training and testing with the secondary task present. During training, the participants in the single-task condition performed significantly better than the dual-task condition. When tested under dual-task conditions, both groups demonstrated a similar sequential learning effect. Namely, participants in both the single- and dual-task conditions expressed the same sequential knowledge. Therefore, when left to operate freely, the attentional mechanism can effectively learn a sequence. However, when attention is suppressed through the addition of a dual-task, residual learning still occurs using the same mechanism utilized when participants were trained under dual-task conditions, the nonattentional system. Curran and Keele (1993) concluded that this demonstrates the two systems are distinct and operate in parallel.

Extending this research, Curran and Keele (1993) conducted a second experiment in which participants were trained to learn a sequence in an SRT task under dual-task conditions then tested without the distracting secondary task. They predicted that the sequential effects should be similar under single- and dual-task conditions. This would be consistent with their posited dual-systems model as attentional learning cannot occur when participants are trained under dual-task conditions. When the concurrent task is removed during testing, any learning displayed should be based exclusively on the nonattentional system, because it has not been given adequate practice for learning to occur. The results confirmed their hypothesis. During training, learning was demonstrated by the increase in reaction time when presented with a randomly sequenced block of trials. Moreover, when tested under single-task conditions, participants demonstrated a sequential learning effect, thus reflecting learning stemming from the nonattentional system (Curran & Keele, 1993). Therefore, these results provided further support that sequence learning involves multiple distinct learning systems.

More recently, Keele et al. (2003) presented a model supported by neuroimaging and behavioural research which put forward that the two sequence learning systems are distinct and operate in parallel. Keele et al. refer to these systems as the multidimensional and unidimensional systems. The multidimensional system uses attentional resources to learn the relation between events across different modalities (e.g., visual and auditory stimulus sequences). In contrast, the unidimensional system, which does not require attentional resources, allows for learning along a single dimension such as auditory, visual, or tactile sensation.

The behavioural evidence for these two systems is based on the aforementioned series of sequence learning studies conducted by Curran and Keele (1993). The neuroimaging evidence stems from research examining the neural correlates of sequence learning under single- and dual-task conditions (See Grafton, Hazeltine, & Ivry, 1995). Specifically, under single-task conditions, sequence learning was associated with activation in brain regions that mediate interdimensional association and explicit learning such as the inferior parietal-occipital lobe, inferior prefrontal cortex, and lateral premotor cortex. In contrast, sequence learning under dual-task conditions resulted in activation in entirely different brain regions responsible for implicit learning such as the left occipital cortex, left hemispheric motor cortex, and the junction between the parietal and occipital lobe. Therefore, these findings demonstrate conclusively that sequence learning does involve two distinct learning systems.

Of methodological importance, the sequence learning literature provides evidence that the application of a concurrent secondary task is effective in dissociating between multiple learning systems. By interfering with one system's ability to operate effectively, the behaviour of the second system can be inferred. In fact, research within the domain of category learning has favoured the use of a concurrent secondary task to test for the existence of multiple memory systems (Waldron & Ashby, 2001; Xing & Sun, 2017). In a typical categorization task, participants learn to classify stimuli as one of two distinct categories. Variations of categorization tasks include learning to categorize Gabor patches (Maddox et al., 2003; Waldron & Ashby, 2001; Xing & Sun, 2017), fictional animals (McDaniel et al., 2014), or making dichotomic predictions about the weather (Gluck, Shohamy, & Myers, 2002). Regardless of their variant, the application of a secondary task has aided researchers in understanding the mechanisms that allow for categorization to occur.

The most compelling evidence for multiple learning systems stems from a model proposed by Ashby et al. (1998). In their seminal paper, they proposed the competition between verbal and implicit systems (COVIS) model to explain categorization behaviour. The COVIS model describes two mechanisms that compete to produce the strongest response. The first system, the verbal system, relies on explicit declarative memory and is under conscious control (Ashby et al., 1998). The verbal system utilizes rule-based learning to learn easily verbalizable category rules, typically in the form of unidimensional rules (e.g., Respond A if the value on dimension x > c, for some constant c). In contrast, the implicit system is associated with procedural memory and operates outside of conscious control. Unlike the verbal system, the implicit system uses an exemplar-based approach to learn category structures in which the underlying category rule is impossible to verbalize. Throughout learning, the more successful of the two systems eventually dominates resulting in either verbal or implicit response behaviours.

As COVIS presents a compelling framework for category learning, researchers sought to validate it from a behavioural perspective. Accordingly, Waldron and Ashby (2001) conducted a categorization task wherein participants completed binary classification under single- or dual-task condition. The dual-task condition involved completing a numeric Stroop task concurrently with the classification task. Moreover, participants were randomly assigned to classify stimuli based on an explicit, unidimensional or an implicit, multidimensional category structure. Waldron and Ashby hypothesised that the concurrent Stroop task would interfere with working memory and attentional capacities and, as a result, impact participants' ability to learn the explicit rule. Because the implicit system outlined by COVIS does not require attentional processes to function, the concurrent task would have no impact on learning the implicit rule. As expected, the results revealed that participants in the dual-task condition required significantly more training to learn the category structures than participants in the single-task conditions. Moreover, the concurrent task impacted the explicit rule structure more than the implicit rule structure. Thus, this lends support for the existence of multiple learning systems. If categorization involved a single mechanism, then the addition of a secondary task would have impacted performance irrespective of category structure. However, because performance in the different category structures were significantly different, clearly, the two structures rely on distinct processes to function.

To further dissociate the verbal and implicit systems as well as lend further support to the COVIS model, Maddox et al. (2003) conducted a categorization task in which individuals had to classify sine wave gratings (i.e., Gabor patches). Participants were randomly assigned to complete either a rule-based or an information-integration categorization task. The former task involved learning to classify Gabor patches based on a single, easily verbalized rule associated with the explicit system. In contrast, the latter task involved classifying Gabor patches based on a complex rule that could not be easily verbalized and was associated with the implicit system. Furthermore, within each categorization task condition, participants received either immediate or delayed feedback upon making a response. Maddox et al. (2003) argued that in order to perform accurately in the information integration condition, participants had to receive immediate feedback following a response, as is true of all types of associative learning. However, because the explicit system can utilize working memory to store stimuli-response parings for an extended period of time, they posited that delayed feedback would have little impact on performance in the rule-based condition. Therefore, it was hypothesized that participants who received feedback immediately in the information integration condition would perform more accurately than participants who received delayed feedback. Moreover, participants in the rule-based condition would perform similarly regardless of when feedback was given. The results demonstrated that delayed feedback had minimal impact on rule-based categorization performance. In contrast, feedback delay greatly hindered information integration performance. Taken together, categorization appears to be yet another cognitive phenomenon that can be best explained through a dual-processing framework.

As discussed previously, function learning behaviour also appears to rely on a hybrid, rule- and exemplar-based system approach (Delosh et al., 1997; McDaniel et al., 2014).

Moreover, the tendency for an individual to use a rule-based approach over an exemplar-based approach is moderated by their working memory capacity (McDaniel et al., 2014). Parallel research in category and sequence learning have demonstrated a similar association between working memory and rule-based learning and have outlined extensive dual-processing theories to explain this phenomenon (Ashby et al., 1998; Curran & Keele, 1993; Keele et al., 2003; Maddox et al., 2003; Waldron & Ashby, 2001). This congruency between tasks implies a dual-processing framework might adequately explain function learning behaviour.

Presently, however, only one series of experiments has attempted to examine function learning in a dual-processing context. Specifically, across a series of experiments, Brown and Lacroix (2018) applied a concurrent task to a function learning task. In their first experiment, participants learned either a linear or a quadratic function with the absence or presence of a secondary task during training. The secondary task involved a memory scanning task wherein participants were shown a string of four letters and had to report whether a given letter was included in the presented string or not following each function learning trial. Brown and Lacroix (2018) hypothesised that the addition of the secondary task would tax working memory capacity, inhibiting the rule-based learning mechanism. Therefore, participants in the dual-task condition would exhibit an exemplar-based learning strategy. Moreover, due to the bias participants have for learning linear functions (Brehmer, 1974), participants would learn the linear function better than the quadratic function and as such, the prevalence of exemplar-based learning would be more apparent in the quadratic shape condition. Interestingly, a novel approach was applied to measure performance during training. Specifically, across all conditions, a response-criterion correlation was calculated to determine how well responses matched the correct criterion. Brown and Lacroix (2018) argued that the correlation, often referred to as the achievement index, would reduce the impact of response variability across trials that resulted from the addition of the secondary task. Using this achievement index, the results revealed that during training, participants in the linear dual-task condition performed better than participants in the quadratic dual-task condition. During transfer, when the function was linear, the addition of a concurrent task during training did not reduce participants' ability to learn the functional relationship. Moreover, with respect to the quadratic condition, there was no apparent difference in transfer performance when participants were trained under single- or dual-task conditions for the quadratic function. This null affect stemmed from the exceedingly low performance in the quadratic condition across both the single- and dual-task conditions.

Although the results with respects to the linear condition were to be expected, the absence of an effect in the quadratic condition was not. Brown and Lacroix (2018) attributed their findings to the difficulty of learning the quadratic shape function. Specifically, regardless of task condition, participants failed to perform adequately for this function. As research has demonstrated individuals can learn quadratic shape relationships fairly well (Delosh et al., 1997; McDaniel et al., 2014), this presented a problem. Thus, Brown and Lacroix (2018) conducted a second experiment with a few key modifications. First, rather than a linear or quadratic function, all participants learned a bilinear, V-shaped function under single- or dual-task conditions. Moreover, the secondary task consisted of a recall task. Participants were presented with four letters prior to each training trial and asked to recall the letters in their correct order following each trial. This adjustment in the concurrent task increased its difficulty, therefore increasing the strength of the manipulation. Similar to their first experiment, Brown and Lacroix (2018) hypothesized that the addition of a concurrent task would result in a decrease in training performance compared to participants who learned the relationship without the recall task. The difference in training performance would translate directly to transfer performance, that is, participants in the dual-task condition would perform worse than participants in the single-task condition. Finally, and of key importance, the concurrent task would inhibit working memory and as a result, interfere with the rule-based learning mechanism. Therefore, there would be a higher proportion of exemplar-based learners in the dual-task condition compared to the single-task condition.

Consistent with their first hypothesis, participants in the dual-task condition performed worse across the training blocks than those in the single-task condition. However, contrary to their second hypothesis, transfer performance across task conditions differed only in the interpolation region. Moreover, there was no difference in the proportion of rule-based and exemplar-based learning across task conditions. Brown and Lacroix (2018) concluded that the concurrent task did not impact participants' ability to learn the shape of the function. Rather, it impaired their ability to learn training exemplars resulting in poorer interpolation performance. Finally, participants in both the single- and dual-task conditions demonstrated a distinct dichotomous relationship; they either learned the function shape or they did not. Taken together, these results suggest that a concurrent recall task adequately interferes with working memory capacity, as shown by the poorer training performance under dual-task conditions. However, during transfer, when the concurrent task is removed, the rule-based mechanism appears to be able to operate freely resulting in similar extrapolation performance regardless of training condition. Speculatively then, retaining the secondary task during transfer might be the key in teasing apart the two learning systems.

The Present Thesis

The presented review has highlighted several key consistencies across learning paradigms. Namely, learning a sequence, a category, or a function involves two mechanisms, a rule-based and exemplar-based system (Ashby et al., 1998; Curran & Keele, 1993; Delosh et al., 1997; Kalish et al., 2004). The rule-based system, accessible by conscious thought, relies on attention and working memory to facilitate learning. In contrast, the exemplar-system operates without conscious thought and does not require attention and working memory to function. The preeminent theories in both category and sequential learning have demonstrated, both behaviourally (Curran & Keele, 1993; Knowlton et al., 1996; Waldron & Ashby, 2001) and neurologically (Ashby et al., 1998; Keele et al., 2003), that a dual-processing framework can explain the data. Currently, only one set of experiments has attempted to apply a dual-processing framework to function learning by differentiating between the two learning mechanisms (Brown & Lacroix, 2018). Although the secondary task impacted training performance in a way that was consistent with a dual-processing paradigm, transfer performance did not.

Therefore, the goal of the present thesis aimed to extend research involving the dual-processing paradigm in relation to function learning. Rather than adding a secondary task only during training, as previous research in categorization and function learning have done (Brown & Lacroix, 2018; Waldron & Ashby, 2001), the following experiment extended this methodology by also including a secondary task during transfer. Participants were randomly assigned to complete a function learning task in which they learned either a linear or bilinear V-shaped function. Across function conditions, the working memory load was manipulated for both the training and transfer phases. Therefore, participants were randomly assigned to complete training and transfer under either single- or dual-task conditions. This yielded four distinct conditions for each of the function types. First, a control condition wherein participants were trained and tested without the secondary task (i.e., SS condition). Secondly, a condition in which participants were trained and tested under dual-task conditions (i.e., DD condition). Third, a condition where participants were trained with the secondary task present then tested without the secondary task (i.e., DS condition). Finally, one in which participants were trained under single-task conditions then subsequently tested under dual-task conditions (i.e., SD condition).

The most crucial predictions targeted how performance differs when participants are trained and tested under different working memory loads. As suggested by Brown and Lacroix (2018), these differences were expected to remain consistent with both the linear and V-shaped conditions. In terms of the SD condition, when participants were trained without a secondary task, accuracy was hypothesized to be comparable to the SS group. However, when the secondary task was added during the testing phase, accuracy was hypothesized to decrease significantly. This prediction stems from Curran and Keele (1993) who argued that the addition of a secondary task inhibits the explicit system, and thus, performance reflects a reliance on the remaining implicit system. Therefore, the decrease in performance was predicted to reflect participants dependency on exemplar-based learning. Along with accuracy scores, the trend was hypothesized to emerge in the proportion of rule-based and exemplar-based learners. That is, compared to the SS group, there would be no difference in characterization of each type of learner during training. During transfer, however, more participants were expected to be characterized as exemplar-based learners than in the SS.

With respects to the DS condition, the predictions in the present thesis reflected past research. Specifically, the research by Brown and Lacroix (2018) found that the degree to which an individual learned a functional relationship was dichotomous, they either learned the relationship or not. To this effect, participants who were trained under dual-task conditions were expected to perform less accurately than the SS group. However, this difference would not be reflected in the transfer phase when the secondary task was removed. Because working memory was able to operate effortlessly, I hypothesized no difference in transfer performance between the DS group and the SS group (Brown & Lacroix, 2018). Moreover, consistent with previous research, the degree to which a participant favoured one learning approach over the other is predicted to remain consistent from training to transfer (Brown & Lacroix, 2018, McDaniel et al., 2014).

The addition of the SS and DD conditions allowed for an examination of function learning in relation to the availability of working memory resources. As such, within each condition, the tendency to utilize rule-based or exemplar-based learning was predicted to be stable (Brown & Lacroix, 2018, McDaniel et al., 2014). However, participants in the SS condition were hypothesized to perform better than any other condition. In contrast, participants in the DD condition were expected to perform the worst of all condition types. This significant decrease in performance was predicted to result from a predominance of exemplar-learners across both training and transfer. Finally, consistent with previous research, participants in the linear function condition were hypothesized to outperform participants in the V-shaped function condition (Delosh et al., 1997, McDaniel et al., 2014) regardless of level of distraction. Moreover, across all conditions, participants were expected to perform better in the interpolation region compared to either extrapolation regions (Brown & Lacroix, 2018; Delosh et al., 1997, McDaniel et al., 2014).

Method

Participants

Eighty undergraduate students enrolled in either a first- or second-year psychology course were recruited for the present experiment. Each participant was randomly assigned to one of eight conditions as shown in Table 1. All students were recruited using the Carleton University SONA system. Moreover, they received a course credit of 1% for their participation as compensation.

Table 1

Total sample size broken down across function type and working memory condition.

	Linear	Bilinear
Dual-Dual	15	6
Dual-Single	4	16
Single-Dual	16	3
Single-Single	3	17

Stimuli

Training stimuli consisted of seventeen evenly spaced integer values ranging from 60-140. The bilinear relationship (V-shaped) was composed of two linear functions with an inflection point at X = 100. Following McDaniel et al. (2014), when X was less than or equal to 100, the relationship was defined by the function: $y = 229.2 - 2.197x$. When X was greater than 100, the relationship was defined by the function: $y = 2.197x - 210$. Finally, following the procedure of Brown and Lacroix (2018), the linear relationship was defined by the function: $y = 0.72x + 5.5$. The training stimuli and their corresponding Y-value for both the linear and V-shaped functions are shown in Table 2. All Y-values were rounded to the nearest integer. Transfer stimuli was generated following a similar procedure. The low and high extrapolation regions consisted of 16 values ranging from 2-57 and 143-198, respectively. The interpolation region consisted of sixteen unique integer values ranging from 62-138. All transfer stimuli and corresponding Y-values is shown in Table 3.

The experiment was conducted using a desktop computer with a 14.8-inch monitor running the E-prime 2.0 Software (Schneider, Eschman, & Zuccolotto, 2002). Each function learning trial was displayed on a screen with a white background. The X-values were presented as the length of a red bar on a marked scale with 20 tick mark as shown in Figure 3. The scale ranged from 0 to 200. Therefore, each tick mark indicated a unit increase in X by 10. The stimuli were presented on the vertically centered top third of the display screen. Participants' responses were displayed in a text box one-third from the bottom and left of the display screen. The text box was labeled *Estimated Beros*.

Table 2

Stimuli and corresponding Y-values used during the training phase.

X-values	**Y-value**		**X-values**	**Y-values**	
	Quadratic	**Linear**		**Quadratic**	**Linear**
60	97	49	105	21	81
65	86	52	110	32	85
70	75	56	115	43	88
75	64	60	120	54	92
80	53	63	125	65	96
85	43	67	130	76	99
90	32	70	135	87	103
95	21	74	140	98	106
100	10	78			

Table 3

Stimuli and corresponding Y-values for items in the low extrapolation, interpolation, and high extrapolation regions during the transfer phase.

X-values	Low Extrapolation Y-values		X-values	Interpolation Y-values		X-values	High Extrapolation Y-values	
	Linear	**Bilinear**		**Linear**	**Bilinear**		**Linear**	**Bilinear**
2	7	225	62	51	93	143	108	104
6	20	216	68	54	80	147	111	113
10	13	207	72	57	71	151	114	122
14	16	198	78	62	58	155	117	131
18	19	190	82	65	49	158	119	137
21	21	183	88	69	36	162	122	146
24	23	176	92	72	27	165	124	153
28	26	168	98	76	14	168	126	159
32	29	159	102	79	14	172	129	168
35	31	152	108	83	27	176	132	177
38	33	146	112	86	36	179	134	183
42	36	137	118	91	49	182	137	190
45	38	130	122	93	58	186	139	199
49	41	122	128	98	71	190	142	207
53	44	113	132	101	80	194	145	216
57	47	104	138	105	93	198	148	225

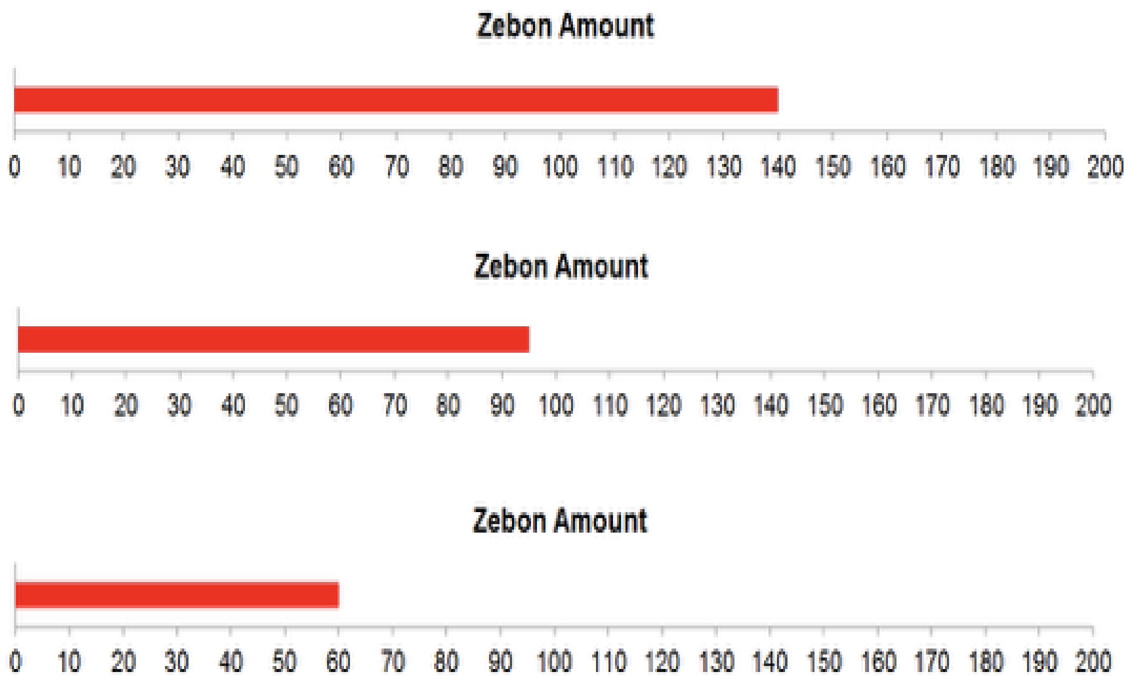

Figure 3. Visual depictions of presented stimuli across three separate trials.

Feedback was displayed in a text box one-third from the bottom and right of the display screen. The feedback text box was labeled *+/- Difference in Beros Estimation*. Both the participants' responses and feedback were displayed numerically in size 18, Times New Roman, black font. See Figure 4 for a sample function learning trial.

Experimental Design

Participants were randomly assigned to complete a function learning task wherein the relationship was defined by either linear or V-shaped function. Furthermore, the presence or absence of a secondary task during the training and transfer phases varied across participants. This task manipulation resulted in four task conditions: completing both training and transfer under a single-task condition, completing training and transfer under a dual-task condition, completing training under a dual-task condition and transfer under a single-task condition, and finally completing training under a single-task condition and transfer under a dual-task condition.

Thus, the experiment comprised of a 2 (function: linear vs. bilinear) x 4 (Task: SS vs. SD vs. DS vs. DD) between-subjects design.

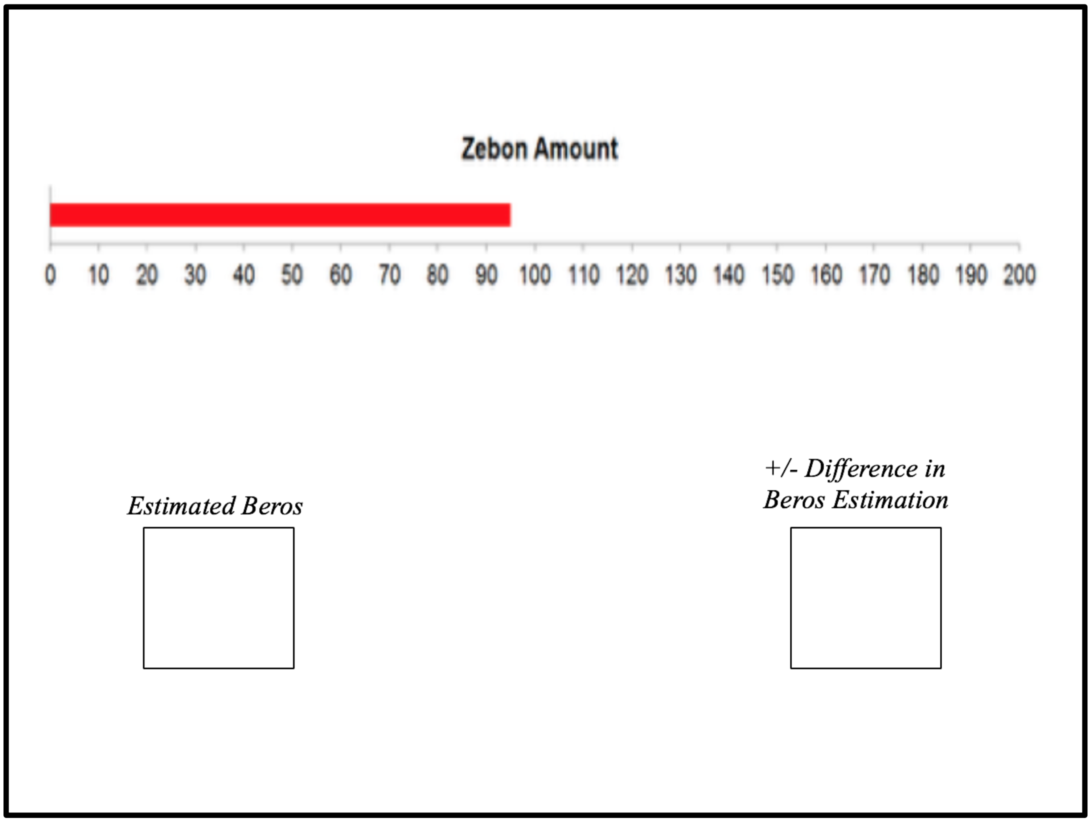

Figure 4. A sample function learning trial

Procedure

The experiment took place in small and quiet testing rooms. After informed consent was obtained, participants were seated at roughly 75 cm from the computer monitor. Moreover, each stimulus was displayed at approximately 5 degrees of visual angel. Under the supervision of an experimenter, the participants were given the following cover story (see Brown & Lacroix, 2018). Participants were told to assume the role of a scientist working for NASA who just discovered a new organism. This organism absorbs a fictional chemical called *Zebos* and emits a different fictional chemical called *Beros*. Participants read that their job, as a NASA scientist, was to predict the amount of *Beros* this new organism emits, based on the amount of *Zebos* it absorbed. After participants read the cover story, they completed three practice function learning trials that mimicked the training phase of the experiment. Furthermore, following training, participants completed three more practice trials that follow the transfer phase procedure. Moreover, the presence of the secondary task during the first or second half of the practice trials mirrored the condition the participant was assigned too. For example, participants in the SS condition did not practice the secondary task during the practice trials. In contrast, participants in the SD condition only practiced the secondary task during the final three practice trials. These practice trials ensured that the participants understood the experimental procedure. Participants were told to be as accurate as possible during the function learning task, however, those in the dual-task condition were told to achieve a minimum threshold of 90% accuracy on the secondary task. Following practice, participants moved to the training phase followed by the transfer phase. Upon completing the transfer phase, the experiment terminated, and participants were debriefed by the experimenter.

Function Learning Task

The function learning task was similar across both single- and dual-task conditions. Moreover, the procedure was also consistent across function type, the only difference being the shape of the function (e.g., linear vs. bilinear). The experiment included a training and transfer phase.

The training phase consisted of twelve blocks of seventeen trials in which each X-value was presented once per trial. Furthermore, the presentation of each X-value was randomized across each block. During each training trial, a red bar on a marked scale was displayed representing the amount of *Zebos* the organism observed. Then, using the numeric keypad, participants made their *Beros* estimate and used the “enter” key to input their response. There was no time constraint for participants to enter their estimation during this phase. Following the procedure outlined by McDaniel et al. (2014), feedback consisted of the signed deviation between the participants estimated *Beros* emitted and the actual *Beros* emitted. Feedback was displayed for 1500ms.

The transfer phase consisted of one block of forty-eight trials. For each transfer trial, the amount of *Zebos* absorbed was presented and participants entered the amount of *Beros* emitted, similar to the training phase. However, participants had 10 seconds to enter their estimations for each trial. Moreover, feedback was not displayed for the function learning trials during this phase. Across all participants, transfer stimuli were presented in ascending order, starting with the low extrapolation region.

Secondary Task

The secondary task was adopted from Brown and Lacroix (2018). The experiment proceeded as follows for participants assigned to a dual-task condition. First, they saw a fixation stimulus (i.e., the addition symbol) in the center of the screen for 500ms in size 25, Times New Roman font. Next, four consonants were randomly selected without replacement and presented horizontally in the center of the screen for 2000ms, in size 25, Times New Roman font. Then, the participants completed a function learning trial. Using the keyboard, they were given five seconds to enter the letters that they saw prior to the function learning trial. Participants entered their response by pressing the “enter” key. Finally, feedback for the secondary task was displayed visually and auditorily for 500ms. Specifically, when participants were correct, the word “Correct!” appeared in the center of the screen in green text, size 25 Times New Roman font followed by a sinewave, pure tone (500 Hz). In contrast, when they participants were incorrect, the word “Incorrect!” appeared in the center of the screen in red text, size 25 Times New Roman font followed by a sawtooth tone (115 Hz).

The single-task condition followed the exact same structure to ensure the two conditions are comparable temporally. Rather than consonants, however, participants saw a string of four asterisks prior to each function learning trial. Moreover, a single asterisk was displayed in the center of the screen for 500ms in place of the secondary task feedback following each function learning trial.

Results

Data Trimming

Three participants were removed from the dataset because they produced a high frequency of extreme responses that were more than five standard deviations above the mean of the training range (*M* = 80.2%). Following their removal, some data were still five standard deviations beyond the mean of the training (*M* = 0.36%, *SD* = 1.02%) and transfer (*M* = 1.79%, *SD* = 2.49%) range. A winsorizing procedure was thus employed to reduce the impact of these extreme values on further analyses (Ghosh & Vogt, 2012). Therefore, across the training blocks and transfer regions, these remaining outliers where transformed to values three standard deviations away from the means for each participant.

In addition, five participants were removed because they did not perform the secondary task according to the given instructions. Finally, two more participants were removed from the analyses because they left a significantly large proportion of transfer responses blank (*M* = 51.04%). Of the remaining participants, the proportion of blank responses per participant during training (*M* = 0.04%, *SD* = 0.27%) and transfer (*M* = 0.34%, *SD* = 1.23%) was negligible.

Working Memory Task

To measure accuracy on the working memory task, participants were given 0.25 points for every letter reported in the correct position within each trial for a maximum of 1. Throughout the experiment, participants maintained high accuracy on the working memory task (M = 91.98%, SD = 5.14%). To ensure that accuracy did not differ across function and Task Condition, a 2 (linear vs. bilinear) x 3 (SD vs. DS vs. DD) between-subjects ANOVA was conducted. The dependent variable was working memory accuracy. Neither of the main effects nor the two-way interaction reached significance ($ps > .2$). Therefore, working memory task accuracy was consistently high across the experimental conditions.

Training Performance

To assess how performance differed across function type and task condition, a three-way 2 x 4 x 12 Mixed-Design ANOVA was conducted. The between-subjects factors were Function Type (linear vs. bilinear) and Task Condition (DD vs. DS vs. SD vs. SS), and the within-subject factor was Blocks (1-12). Mean Absolute Error (MAE) was the dependent variable. It was calculated by taking the absolute difference of the participants response magnitudes and the correct response for each trial then averaging the differences for each block. Due to a violation in the Mauchly's test of sphericity and a moderate sized epsilon, the Greenhouse-Geisser correction was used to assess significance, χ^2 (67) = 369.3, $p < .001$, $\varepsilon = .360$. The results are shown in Figure 5.

As hypothesized, the analysis revealed a significant main effects of Block and Function Type, $F(3.96, 249.29) = 18.91$, $p < .001$, $\eta^2 = .23$ and $F(1, 67) = 47.5$, $p < .001$, $\eta^2 = .43$,

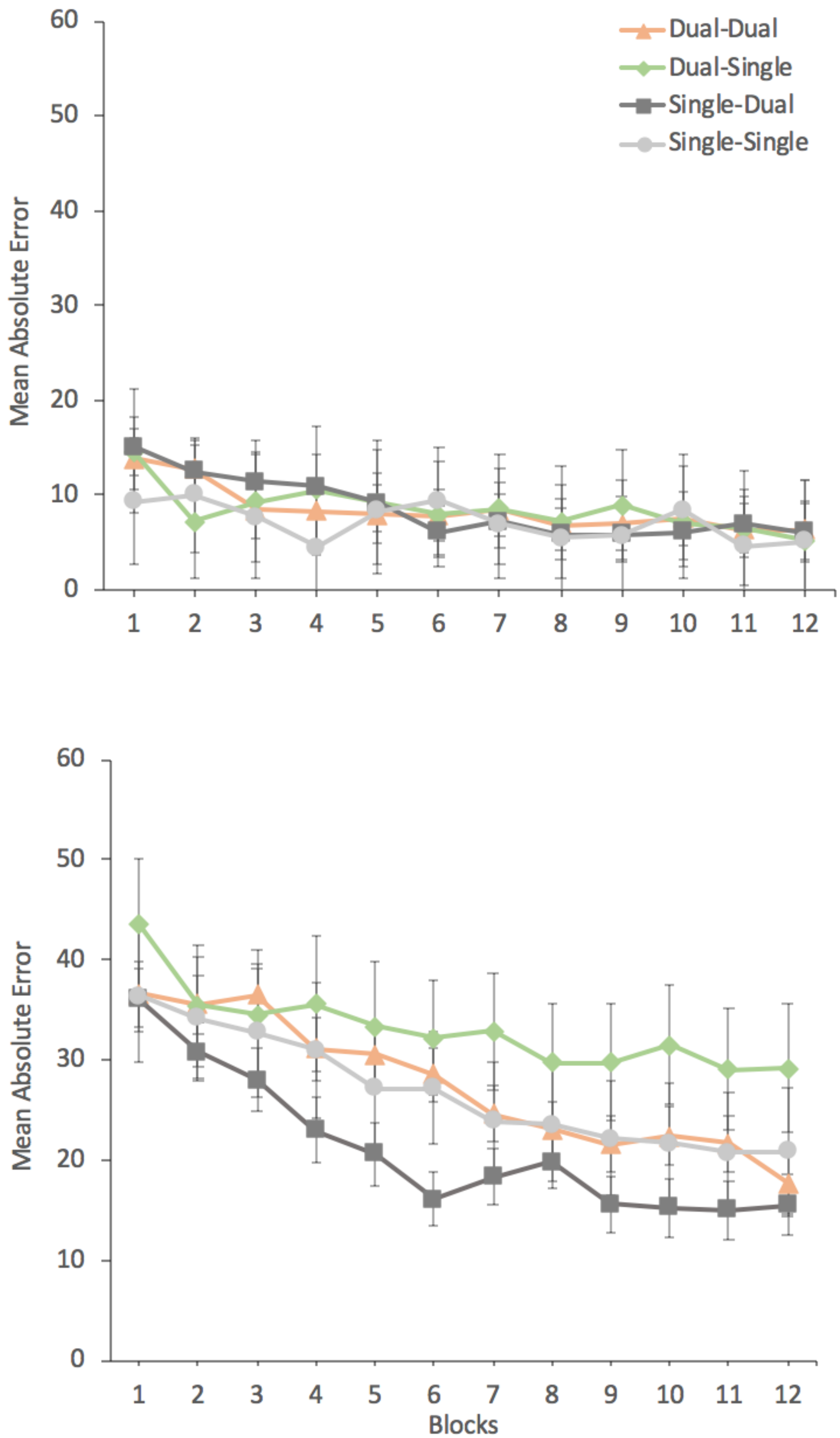

Figure 5. Training accuracy for the linear (top) and bilinear (bottom) conditions across working memory condition and training blocks.

respectively. This suggests that accuracy increased throughout training and that participants learning the linear function consistently outperformed those learning the bilinear function. However, the main effect of Task Condition failed to reach significant, $F(3, 67) = 0.73$, $p = .537$. The main effects were qualified by a single two-way interaction of Block by Function Type, $F(3.96, 249.29) = 4.27$, $p = .002$, $\eta^2 = .063$. Both the Block by Task Condition and Function Type by Task Condition failed to reach significance, $F(11.87, 249.29) = 0.67$, $p = .991$ and $F(3, 67) = 0.713$, $p = .548$, respectively. Finally, the two-way interactions were not qualified by significant three-way Function Type by Blocks by Task Condition interaction, $F(11.99, 249.29) = 0.51$, $p = .907$.

Thus, contrary to the hypothesis, accuracy across the training blocks differed depending on the function type participants learned irrespective of working memory condition. To decompose the significant two-way Block by Function Type interaction, the simple main effects were analyzed. Performance in both the linear $F(11, 341) = 10.632$, $p < .001$, and bilinear condition improved across training, $F(11, 418) = 23.184$, $p < .001$. Finally, a comparison of participants' performance on the last block of training showed that those in the linear condition ($M = 5.99$, $SD = 5.94$) were more accurate than those in the bilinear one ($M = 23.28$, $SD = 13.82$), $t(67) = 6.58$, $p < .001$, $d = 1.63$.

Transfer Performance

Next, a three-way 2 x 4 x 3 Mixed-Design ANOVA was conducted to evaluate transfer performance. The between-subject factors were Function Type (linear vs. bilinear) and Task Condition (DD vs. DS vs. SD vs. SS), and the within-subjects factor was Transfer Region (low extrapolation vs. interpolation vs. high extrapolation). Once again, the dependent variable was

MAE and was calculated following the same procedure as the training performance. The results are shown in Figure 6 and 7.

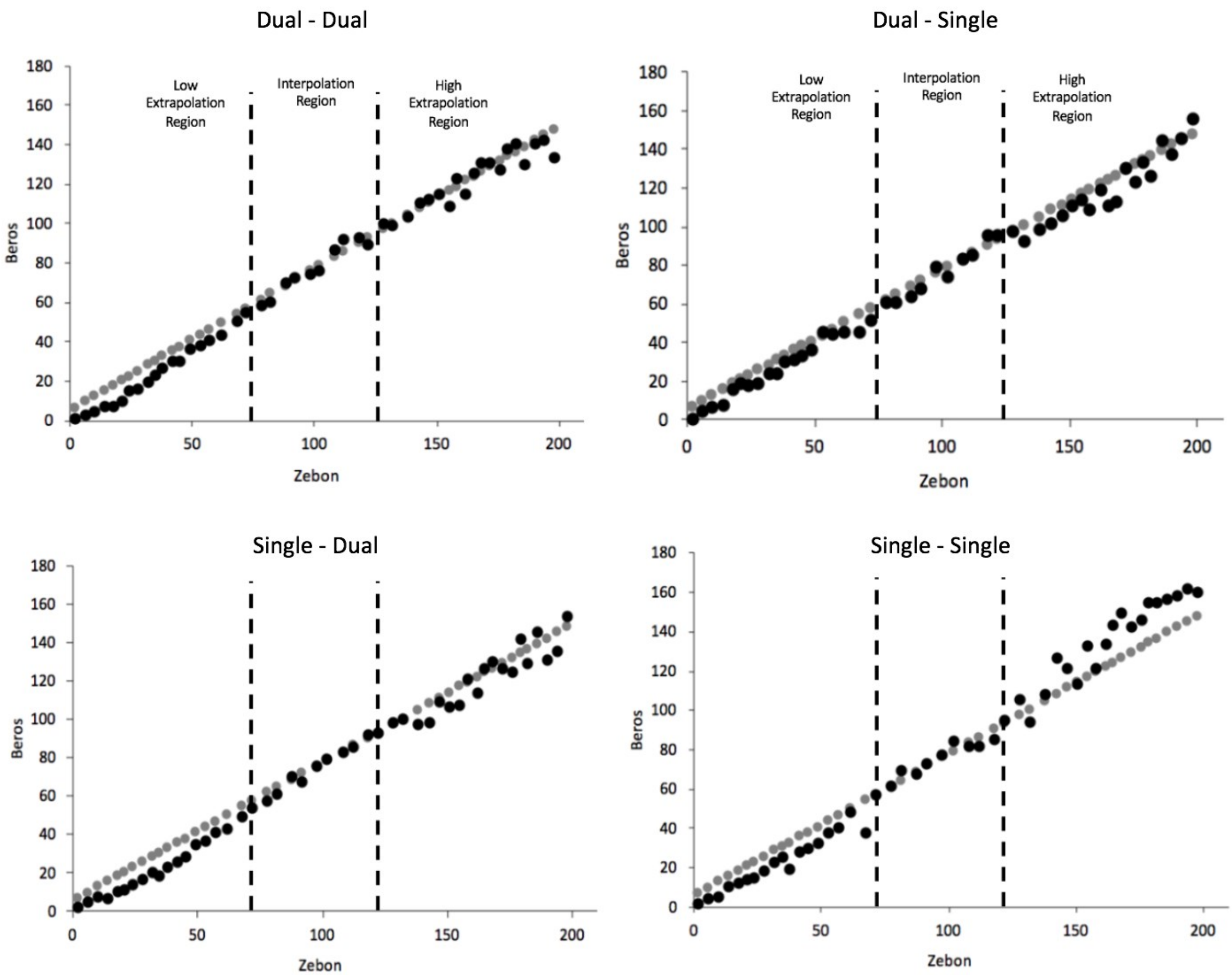

Figure 6. Transfer performance of participants who learned the linear function, partitioned by region, across Task conditions.

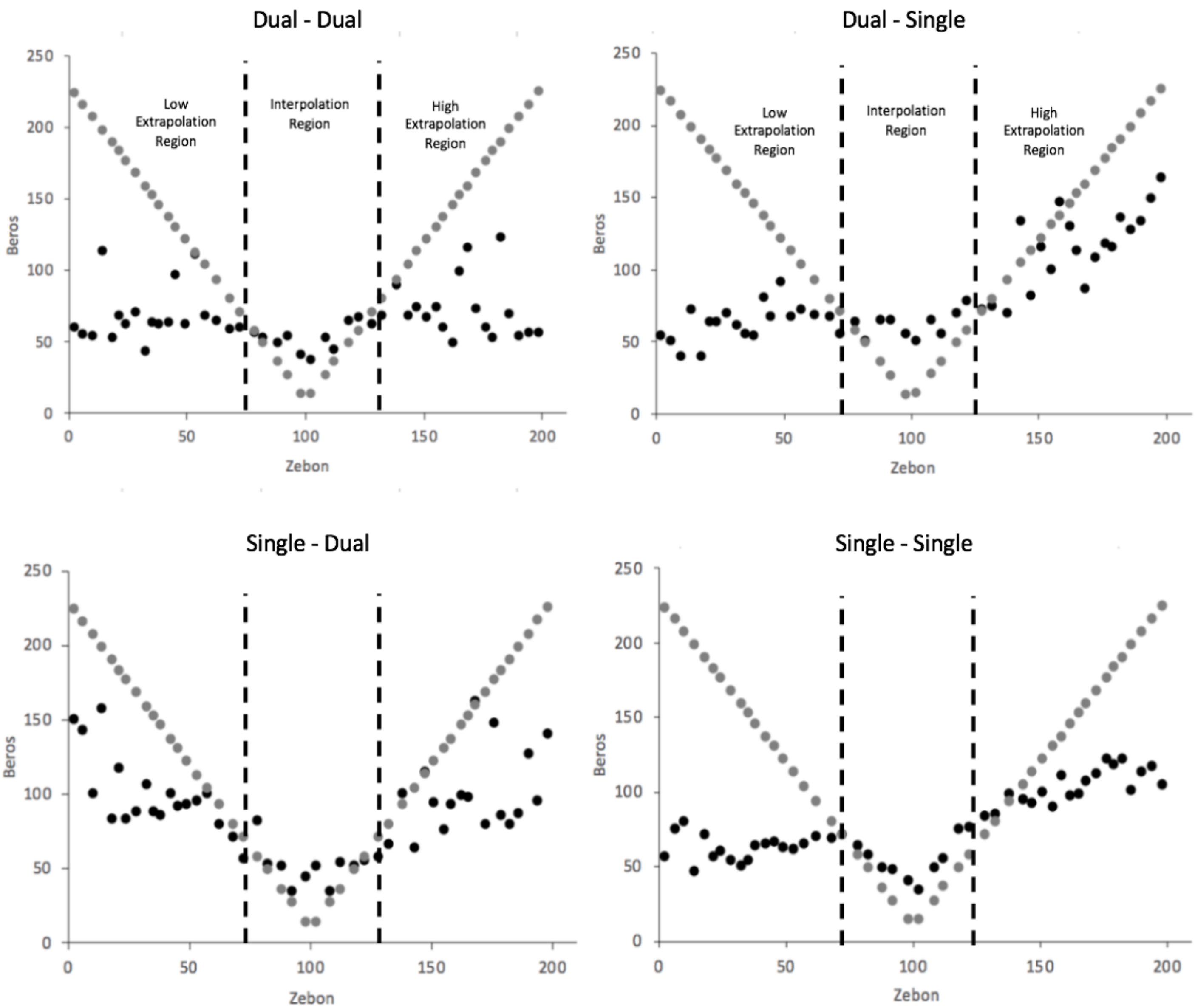

Figure 7. Transfer performance of participants who learned the bilinear function, partitioned by region, across Task conditions

Due to violations of sphericity and a large epsilon, the following analyses are presented with a Huynh-Feldt correction, χ^2 (2) = 15.63, p < .001, ε = .930. As predicted, and in line with previous research (Brown & Lacroix, 2018; Delosh et al., 1997; McDaniel et al., 2014), the analysis revealed significant main effects for the Transfer Region and Function Type, $F(1.86, 117.17) = 29.37$, $p < .001$, $\eta^2 = .32$ and $F(1, 67) = 135.55$, $p < .001$, $\eta^2 = .678$, respectively. In contrast, participants' accuracy did not significantly differ across Task Condition, $F(3, 67) = 1.02$, $p = .392$. The significant main effects were qualified by a significant two-way interaction of Transfer Region and Function Type, $F(1.76, 117.17) = 47.5$, $p < .001$, $\eta^2 = .43$. However, Task Condition did not significantly interact with Function Type or Transfer Region, $F(5.58, 117.17) = 0.91$, $p = .481$ and $F(3, 67) = 1.02$, $p = .392$. Finally, the three-way Transfer Region by Task Condition by Function Type interaction was not significant, $F(5.579, 117.17) = 1$, $p = .428$.

To further understand how function type and region affected transfer performance, a simple main effects analysis was conducted. It revealed a significant effect for the linear function type, $F(2, 65.51) = 11.44$, $p < .001$. Post-hoc analyses using a Bonferroni correction then showed that participants performed worse in the high extrapolation region ($M = 16.05$, $SD = 11.22$) than the low extrapolation ($M = 8.11$, $SD = 3.71$) and interpolation ($M = 8.41$, $SD = 8.96$) regions. The simple main effects analysis was also significant for the bilinear function type, $F(2, 64.91) = 76.44$, $p < .001$. As hypothesized, participants performed significantly better in the interpolation ($M = 24.39$, $SD = 15.15$) region compared to either the low ($M = 104.96$, $SD = 39.02$) or high ($M = 69.99$, $SD = 37.62$) extrapolation regions. Moreover, participants performed better in the high extrapolation region compared to the low extrapolation region. Hence, learning did occur across both function conditions, but there was little evidence that the secondary task had an impact on the results.

Assessment of Learners

To further examine how the addition of the secondary task affected the participants' ability to learn either function, an exploratory analysis was conducted. First, participants were categorized as either learners or non-learners through the assessment of MAE in the final block of training (McDaniel et al., 2014). Specifically, participants who achieved a MAE of 20 or less were classified as learners while the remaining participants were classified as non-learners. As shown in Table 4, 45 participants were classified as learners. Due to the ease at which individuals learn linear functions, it was unsurprising that a much higher proportion of participants in the linear condition exhibited learning (75%) compared to those in the bilinear condition (25%).

Table 4

The percentage of participants classified as learners broken down across function and working memory task conditions.

To further explore whether participants exhibited rule-based or exemplar-based learning, a novel approach was employed using the angle of inclination of the defined linear and bilinear functions. This was calculated by taking the arctangent of the slope formed by the x-axis and the

	Linear	Bilinear
Dual-Dual	(13) 92.31%	(5) 40%
Dual-Single	(3) 100%	(13) 23.08%
Single-Dual	(13) 92.31%	(3) 66.67%
Single-Single	(3) 100%	(17) 47.06%

*Brackets indicate total sample size of the respective cell

lines specified by the functions. For the linear function, the angle of inclination was 35.8º for both the low and high extrapolation regions. For the bilinear function, the angle of inclination was -65.5º for the low extrapolation region and 65.5º for the high extrapolation region. Participants were characterized as exhibiting rule-based learning if the angle of inclination of their applied function fell with within a ± 5º bound of the defined function. Otherwise, they were characterized as exemplar-based learners.

The results for the linear condition are shown in Table 5 and those for the bilinear condition in Table 6. The column labels represent learning approaches during transfer by region, specifically low extrapolation-high extrapolation pairs. The data suggests a trend such that a large proportion of participants favoured a mixed approach (68.97%), compared to a rule-based approach (27.59%) and an exemplar-based approach (3.44%) in the linear condition. In contrast, participants assigned to the bilinear condition trended towards a more stable strategy. Specifically, participants appeared to favour an exemplar approach (66.67%), over a mixed (20%) or rule-based approach (13.33%). Taken together, participant's strategies appeared to be stable across the low- and high-extrapolation regions, irrespective of working memory condition when participants learned the bilinear function. However, when learning the linear function, the addition of the secondary task seemed to hinder the participants ability to consistently use a rule-based approach throughout transfer.

Discussion

The goal of the present thesis was to determine the extent to which function learning can be explained by a dual-processing framework. Findings from analogous learning domains such as category and sequence learning (Ashby et al., 1998; Current & Keele, 1993) suggested that learning a functional relationship may also rely on exemplar- and rule-based approaches. Furthermore, the predisposition to prefer a rule-based approach is positively associated with an individual's working memory capacity (McDaniel et al., 2014). However, little research has attempted to determine the extent to which disrupting working memory processes interferes with

Table 5

Proportion of learning approaches in the linear condition by Task. The pairs represent approaches used in the low extrapolation – high extrapolation regions.

	Rule-Rule	Rule-Exemplar	Exemplar-Rule	Exemplar-Exemplar
Dual-Dual	(4) 33.33%	(8) 66.67%	(0) 0%	(0) 0%
Dual-Single	(1) 33.33%	(2) 66.67%	(0) 0%	(0) 0%
Single-Dual	(2) 18.18%	(8) 72.73%	(0) 0%	(0) 0%
Single-Single	(1) 33.33%	(2) 66.67%	(0) 0%	(0) 0%

Note. Brackets indicate the number of participants who exhibited each learning approach across working memory conditions

Table 6

Frequency count of participants learning approach in the bilinear condition by Task. The pairs represent approaches used in the low extrapolation – high extrapolation regions.

Bilinear	Rule-Rule	Rule-Exemplar	Exemplar-Rule	Exemplar-Exemplar
Dual-Dual	(0) 0%	(1) 50%	(0) 0%	(1) 50%
Dual-Single	(0) 0%	(1) 33.33%	(0) 0%	(2) 66.67%
Single-Dual	(1) 50%	(1) 50%	(0) 0%	(0) 0%
Single-Single	(1) 12.5%	(0) 0%	(0) 0%	(7) 87.5%

Note. Brackets indicate the number of participants who exhibited each learning approach across working memory conditions

function learning. Thus, participants completed an experiment for which function type and the presence of a secondary task varied to assess how impeding working memory affected their learning approach. As suggested by previous research, participants in linear condition were expected to outperform those in the bilinear condition (Delosh et al., 1997; McDaniel et al., 2014). Furthermore, during transfer, it was hypothesized that participants would perform better in the interpolation regions compared to either extrapolation regions. Finally, the central hypothesis for this thesis was that the addition of a secondary task would consistently increase the proportion of participants characterized as exemplar-based learners for both the linear and bilinear function conditions (Brown & Lacroix, 2018). As such, this increase would result in lower accuracy across training and transfer when the secondary task is present.

As predicted, participants performed significantly better when learning a linear function than a bilinear one. The ease with which individuals learned the linear function may reflect the hypothesized hierarchical nature of function learning. Specifically, research has consistently shown that individuals will first suppose functions to be positively linear and if this assumption does not work, they move on to negatively linear and subsequently non-linear assumptions (Brehmer, 1974; Delosh et al. 1994; McDaniel et al., 2014). Moreover, Byun (1996) demonstrated that participants make fewer errors when extrapolating a positive linear function than a logarithmic, negative power, and a positive power function. The most salient example of this phenomenon emerged in the dual-single and single-single conditions as shown in Figure 7. Specifically, participants appeared to extrapolate closer to the defined function in the high extrapolation region, represented by a positive linear function, compared to the low extrapolation region. The analysis also confirmed this thesis's second hypothesis. Participants were more accurate in the interpolation region than either the lower or higher extrapolation regions. This finding supports the notion that learning did occur for both functions. In fact, MAE error in both interpolation regions approximated the MAE during the final block of training, despite the fact that they involved novel stimuli. In sum, the current experiment replicated key benchmark function learning findings (Brehmer, 1974; Delosh et al., 1997; McDaniel et al., 2014).

Unfortunately, the addition of a secondary task failed to yield any conclusive effect on function learning performance. That is, MAE was similar across working memory conditions throughout training and transfer irrespective of whether participants learned a linear or bilinear function. Furthermore, the secondary task did not result in an increase in the proportion of exemplar-based learners following participants' categorization using their angle of inclination. Tables 5 and 6 suggests that the propensity to favour a rule-based or exemplar-based approach was dependent solely on the learned function. Specifically, participants who learned the linear function favoured a strictly rule-based approach considerably more than an exemplar-based approach. Conversely, participants who learned the bilinear function favoured an exemplar-based approach over a rule-based one. Taken together, it appears that the stability of individuals' learning approach superseded the interfering effect of inhibiting working memory. Speculatively, however, Figure 7 suggests the addition of a secondary task during the training and transfer phase may have impeded the participants' ability to employ a rule-based strategy in the ascending portion of the function. In contrast, when the secondary task was removed during either of these phases, participants seemed to have been able to extrapolate beyond the training range in the high extrapolation region. Nevertheless, none of the findings were statistically significant.

Despite its findings, the present these is not without limitations. Most prominently, the experiment lacks an adequate sample size and thus was severely unpowered. The effects of a low powered study are known and well documented, which is why psychologists strive to reach a high level of power (~80%) when conducting research (Fraley & Vazire, 2014; Funder & Ozer, 2019; Gignac & Szodorai, 2016). If the present experiment had been more adequately powered, then it might have been revealed that a secondary task can interfere with the mechanism associated with rule-based learning while leaving the exemplar-based system intact. In turn, this would parallel the findings in sequence learning and categorization that allowed researchers to conclude that they operate within a dual-processing framework. For example, Nissen and Bullemer (1987) concluded that multiple mechanisms must be involved in sequence learning when participants who suffered from Korsakoff's syndrome were able to exhibit sequential knowledge in an SRT task despite an explicit lack of awareness of the repeating pattern. Furthermore, Maddox and Ashby (2001) demonstrated that category learning is mediated by multiple learning systems by pairing a categorization task with a concurrent Stroop task resulting in an inability for participants to use rule-based learning. Therefore, it could then be concluded that function learning is yet another domain that may be explained by multiple learning systems.

Beyond its limitations, the present thesis contributes meaningfully to the growing function learning literature. Specifically, by using a slope analysis approach, I was able to capture specific differences in learning characteristics during the transfer phase in the low and high extrapolation region. This diverges from previous research that has typically averaged MAE across both regions (McDaniel et al., 2014), an approach which may not be representative of performance in individual regions as crucial information may be lost through averaging (Delosh et al., 1997). Indeed, a trend emerged when participants were characterized using the slope analysis approach. Those who learned the linear function showed variations in their learning style across both extrapolation regions. A considerable proportion of learners appeared to favour a blended approach, particularly a rule-based approach in the low extrapolation region and an exemplar-based approach in the high extrapolation region. In contrast, a similar effect did not emerge for participants who learned the bilinear function. This divergence may have reflected the tendency individuals have to apply a positive linear shape to functional relationships compounded by their predisposition to anchor responses to zero (Brehmer, 1974; Brown & Lacroix, 2018). In fact, as shown in Figure 6, across all four linear conditions participants appeared to consistently underestimate the defined function. As the defined function's intercept approximates zero, anchoring responses to zero would be an acceptable strategy to extrapolate accurately. Therefore, by using these heuristics, participants were able to reduce the impact of the secondary task, resulting in rule-based extrapolation.

This then begs the question that if the predisposition to anchor responses and assume a positive linear function resulted in rule-based learning processes, then why did some participants in the linear condition extrapolate following an exemplar-based approach in the high extrapolation region? Speculatively, this may be indicative of the secondary task having the indented effect of inhibiting rule-based learning. However, because participants find it easier to learn positive linear functions, this condition may not have led to lower accuracy. Recent work by Fischer and Holt (2017) examining the relationship between working memory capacity and function learning support this notion. Specifically, they conducted a function learning task in which participants learned either an asymptotic or an exponential function. Furthermore, they took an index of working memory capacity using two tasks, the digit span backward and letter-number sequencing task. Fischer and Holt argued that to extrapolate accurately in the asymptotic condition, participants need only to apply a simple positive linear rule. Therefore, exemplar-based abstraction would result in comparable accuracy to rule-based learning. In contrast, the exponential function requires complex rules to extrapolate accurately and as such, rule-based learners would outperform their exemplar-based counterparts.

As expected, Fischer and Holt's (2017) results showed that participants with higher working memory capacity tended to favour a rule-based approach whereas low working memory capacity was related to exemplar-based learning. Furthermore, rule-based learning only afforded an advantage to participants learning the exponential function as it required the application of more complex rules than the asymptotic function. As such, distinguishing between rule-based and exemplar-based learners using accuracy appears to be more challenging when the defined function is relatively simple.

Notably, recent research has demonstrated that participant classification using extrapolation accuracy may underestimate the proportion of rule-based learners. Specifically, Said and Fischer (2019) argued that while accurate extrapolation necessitates rule-learning, inaccurate extrapolation is not necessarily indicative of exemplar-learning. Therefore, using accuracy to characterize learning approaches may result in a misidentification of rule-learners as exemplar-learners. To test this idea, they asked participants to complete a standard function learning task, in which a negative exponential function was learned. Then, following the task's completion, they asked participants to draw the function that best represented the underlying functional relationship. Participants' learning approach was categorized through the standard approach using MAE (McDaniel et al., 2014) and a novel summary approach, which identified rule-based learning if the slope from the first and each subsequent extrapolation point was monotonically decreasing. The results were thought provoking. 25.5% of participants were characterized as rule-learners using the standard approach whereas 53.5% of participants characterized as rule-learners using the novel approach. Thus, MAE struggles to adequately identify participants' learning approach. As rule-based learners may achieve identical accuracy scores as exemplar-learners, a more robust measure is required. In particular, the use of a participants' angle of inclination may hold many benefits over other methods. Most notably, it incorporates the logic of the summary approach in the context of traditional transfer phases typically seen in function learning paradigms (Delosh et al., 1997; McDaniel et al., 2014; Said & Fischer, 2019). Moreover, it allows for individual classification when the defined function is non-linear or polynomial.

Therefore, future research should strive to validate the slope analysis approach by applying it to a high-powered function learning paradigm. For example, researchers could use a dual-paradigm procedure wherein participants' learning approach is characterized in both a function learning and categorization task. If the slope analysis is a valid method to classify function learning behaviour, then there should be agreement between preferred learning approach across both tasks (McDaniel et al., 2014). Alternatively, future research could replicate the present experimental design but also take indices of working memory. As such, working memory capacity should correlate positively with a preference for rule-based learning, as assessed by the slope analysis (Fischer & Holt, 2017; McDaniel et al., 2014).

A second area for future research to explore pertains to how a secondary task is integrated into a function learning task. Namely, the present experimental design displayed the four-letter target and participants only received feedback after a delay from the integrated function learning trial. This process is viewed as embedding a function learning task into a visual scanning task. Reversing this structure by embedding the secondary task in the function learning task might result in a more prominent impact on rule-based learning. Indeed, Xing and Sun (2017) attempted just this within the context of category learning to great success. Specifically, across three experiments, they had participants complete a categorization task, wherein they classified sine-wave gratings following either a rule-based (RB) or information integration (II) structure, with a concurrent secondary visuospatial memory task. The structure to which the tasks were presented varied across the three experiments. In the first experiment, the categorization and secondary tasked occurred sequentially. During the second experiment, the category learning task was embedded in the secondary task such that participants began a trial for the visuospatial task but entered a response and received feedback after completing a categorization trial. Finally, in the third experiment, the secondary task was embedded in the category learning task. As such, participants saw a categorization stimulus, completed the secondary task, then entered their response for the categorization task and received feedback. The results were clear. Participants learning the RB structure performed worse when the secondary task was embedded into the category learning task than when the opposite was done (Xing & Sun, 2017). This finding highlights the significant impact the structure of an experimental design can have on the relationship between working memory and learning. Therefore, future research should extend a similar manipulation in design to function learning. For example, an experiment in which a function learning stimulus is presented, and then, participants make an estimation and receive feedback following a delay from completing a secondary task. Such a manipulation may be key in successfully inhibiting working memory and thus, limiting rule-based learning.

In conclusion, theories across different domains have consistently shown that learning operates within a dual-processing framework consisting of an explicit rule-based and implicit exemplar-based mechanism (Ashby et al., 1998; Curran & Keele, 1993). Furthermore, the present thesis reported general trends wherein the addition of a secondary task resulted in a decrease in the stability of characterized learning approaches. Therefore, a dual-processing framework remains a viable explanation for variations found in function learning behaviour. Conducting an adequately powered function learning task would allow researchers to further our understanding of the mechanisms underlying how humans learn functional relationships.

References

Ashby, F. G., Alfonso-Reese, L., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105*(3), 442-481. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0033-295X.105.3.442

Björkman, M. (1965). Studies in predictive behavior: Explorations into predictive judgments based on functional learning and defined by estimation, categorization, and choice. *Scandinavian Journal of Psychology, 6*(3), 129-156. http://dx.doi.org/10.1111/j.1467-9450.1965.tb01030.x

Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior & Human Performance, 11*(1), 1-27. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1016/0030-5073(74)90002-6

Brown, M., & Lacroix, G. (2018). Working Memory Capacity and Individual Differences in Function Learning Behaviour (Unpublished doctoral dissertation). Carleton University, Ottawa, Canada.

Brown, M. A., & Lacroix, G. (2017). Underestimation in linear function learning: Anchoring to zero or x-y similarity? *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale, 71*(4), 274-282. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/cep0000129

Busemeyer, J. R., Byun, E., Delosh, E. L., & McDaniel, M. A. (1997). Learning functional relations based on experience with input-output pairs by humans and artificial neural networks. In K. Lamberts, & D. R. Shanks (Eds.), *Knowledge, concepts and categories; knowledge, concepts and categories* (pp. 408-437, Chapter xiii, 464 Pages) The MIT

Press, Cambridge, MA. Retrieved from http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/619165489?accountid=9894

Byun, E. (1996). *Interaction between prior knowledge and type of nonlinear relationship on function learning* (Doctoral dissertation, ProQuest Information & Learning).

Carroll, J. D. (1963). *Functional learning: The learning of continuous functional mappings relating stimulus and response continua* Available from PsycINFO. (615447650; 1965-11337-001). Retrieved from http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/615447650?accountid=9894

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(1), 189202. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0278-7393.19.1.189

Delosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 968-986. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0278-7393.23.4.968

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127*(2), 107-140. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0096-3445.127.2.107

Fraley, R. C., & Vazire, S. (2014). The n-pact factor: Evaluating the quality of empirical journals with respect to sample size and statistical power. *PLoS ONE*, *9*(10), e109019.

Fischer, H., & Holt, D. V. (2017). When high working memory capacity is and is not beneficial for predicting nonlinear processes. *Memory & Cognition, 45*(3), 404-412. doi:http://dx.doi.org.proxy.library.carleton.ca/10.3758/s13421-016-0665-0

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Ghosh, D., & Vogt, A. (2012, July). Outliers: An evaluation of methodologies. In *Joint statistical meetings* (Vol. 2012).

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78.

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning & Memory, 9*(6), 408-418. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1101/lm.45202

Goldstone, R. L., Kersten, A., & Carvalho, P. F. (2013). Concepts and categorization. In A. F. Healy, R. W. Proctor & I. B. Weiner (Eds.), *2nd ed.; handbook of psychology: Experimental psychology (vol. 4, 2nd ed.)* (2nd ed. ed., pp. 607-630, Chapter xix, 763 Pages) John Wiley & Sons Inc, Hoboken, NJ. Retrieved from http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/1240218640?accountid=9894

Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of Cognitive Neuroscience, 7*(4), 497-510. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1162/jocn.1995.7.4.497

Jitsumori, M. (1993). Category discrimination of artificial polymorphous stimuli based on feature learning. *Journal of Experimental Psychology: Animal Behavior Processes, 19*(3), 244-254. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0097-7403.19.3.244

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: Knowledge partitioning and function learning. *Psychological Review, 111*(4), 1072-1099. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0033-295X.111.4.1072

Keele, S. W., Ivry, R., Mayr, U., Hazeltine, E., & Heuer, H. (2003). The cognitive and neural architecture of sequence representation. *Psychological Review, 110*(2), 316-339. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0033-295X.110.2.316

Knowlton, B. J., Squire, L. R., Paulsen, J. S., Swerdlow, N. R., & Swenson, M. (1996). Dissociations within nondeclarative memory in Huntington's disease. *Neuropsychology, 10*(4), 538-548. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0894-4105.10.4.538

Koh, K., & Meyer, D. E. (1991). Function learning: Induction of continuous stimulus-response relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(5), 811-836. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0278-7393.17.5.811

Kruschke, J. K., & Erickson, M. A. (1994). Learning of rules that have high-frequency exceptions: New empirical data and a hybrid connectionist model. In *Proceedings of the sixteenth annual conference of the Cognitive Science Society* (pp. 514-519). Lawrence Erlbaum Associates.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*(1), 22-44. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0033-295X.99.1.22

Kwantes, P. J., & Neal, A. (2006). Why people underestimate y when extrapolating in linear functions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 32, 1019-1030. http://dx.doi.org/10.1037/0278-7393.32.5.1019

Kwantes, P. J., Neal, A., & Kalish, M. (2012). Item order matters in a function learning task. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale, 66*(2), 90-97. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/a0026639

Lum, J. A. G., Conti-Ramsden, G., & Lindell, A. K. (2007). The attentional blink reveals sluggish attentional shifting in adolescents with specific language impairment. *Brain & Cognition, 63,* 287–295

Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 650-662. http://dx.doi.org.proxy.library.carleton.ca/10.1037/0278-7393.29.4.650 Retrieved from http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/620119067?accountid=9894

McDaniel, M. A., Cahill, M. J., Robbins, M., & Wiener, C. (2014). Individual differences in learning and transfer: Stable tendencies for learning exemplars versus abstracting rules. *Journal of Experimental Psychology: General, 143*(2), 668-693. doi:http:// dx.doi.org.proxy.library.carleton.ca/10.1037/a0032963

McDaniel, M. A., Dimperio, E., Griego, J. A., & Busemeyer, J. R. (2009). Predicting transfer performance: A comparison of competing function learning models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(1), 173-195. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/a0013982

Muthén, B.,O. (1992). Latent variable modeling in epidemiology. *Alcohol Health and Research World, 16*(4), 286. Retrieved from https://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/1474321710?accountid=9894

Neter, J., Kutner, M.H., & Wasserman, W. (1985). *Applied linear statistical models*. Homewood, IL: Irwin.

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

Nosofsky, R. M., & Kruschke, J. K. (1992). Investigations of an exemplar-based connectionist model of category learning. In *Psychology of learning and motivation* (Vol. 28, pp. 207-250). Academic Press.

Said, N., & Fischer, H. (2019). Extrapolation accuracy underestimates rule learning: Evidence from the function-learning paradigm. Unpublished manuscript.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime *2.0*. Pittsburgh, PA: Psychological Software Tools.

Smith, J. D., & Church, B. A. (2017). Dissociable learning processes in comparative psychology. *Psychonomic Bulletin & Review,* http://dx.doi.org.proxy.library.carleton.ca/10.3758/s13423-017-1353-1 Retrieved from http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/1929446259?accountid=9894

Smith, J. D., Tracy, J. I., & Murray, M. J. (1993). Depression and category learning.*Journal of Experimental Psychology: General, 122*(3), 331-346. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1037/0096-3445.122.3.331

Soetens, E., Melis, A., & Notebaert, W. (2004). Sequence learning and sequential effects. *Psychological Research, 69*(1-2), 124-137. doi:http://dx.doi.org.proxy.library.carleton.ca/10.1007/s00426-003-0163-4

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review, 8*(1), 168-176. doi:http://dx.doi.org.proxy.library.carleton.ca/10.3758/BF03196154

Xing, Q., & Sun, H. (2017). Differential impact of visuospatial working memory on rule-based and information-integration category learning. *Frontiers in Psychology, 8*, 13. Retrieved from http://proxy.library.carleton.ca/login?url=https://search-proquest-com.proxy.library.carleton.ca/docview/1915374057?accountid=9894