7-17-2020

All Together Now: Effects of Simultaneous Presentation and Stimulus Complexity on Categorization Performance and Strategy Preferences

Reshma Gouravajhala
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Psychological & Brain Sciences

Dissertation Examination Committee:
Mark McDaniel, Chair
Julie Bugg
Andrew Butler
Brett Hyde
Henry Roediger, III

All Together Now: Effects of Simultaneous Presentation and Stimulus Complexity on Categorization Performance and Strategy Preferences

by

Reshma Gouravajhala

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2020
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would first like to express my sincere gratitude to my advisor, Mark McDaniel, for his unwavering support and guidance not only during my dissertation process, but also over the last five years. Mark has treated me as an equal since I arrived at Washington University and provided me the scaffolding to grow as a student and as a scientist. While I am thankful for our many conversations about cognitive psychology and education, I am perhaps even more appreciative of Mark's enthusiasm for chatting about sports at every possible opportunity. I look forward to continuing all these conversations long into the future.

I would also like to thank Roddy Roediger, Andy Butler, Julie Bugg, and Brett Hyde for their willingness to serve on my dissertation committee and for their helpful feedback throughout the process. I am also thankful to Chris Wahlheim for his generosity in sharing his time, advice, and data visualization expertise during his career at Washington University and beyond.

I would like to next thank the present and former members of the Memory and Complex Learning Lab – especially Francis Anderson and Toshi Miyatsu – for being the best lab mates. They were my first source of support in St. Louis, and I am so grateful to have had their companionship over the years. I know we never got those matching tattoos, but I will cherish memories of all those hours in the lab – brainstorming about research methods, laughing about the most obscure facts and debate topics over lunch, walking each other through statistical analyses, and trading facts about exercise routines – for the rest of my life.

My graduate school experience would have been considerably worse were it not for the constant support and commiseration of friends and fellow sufferers. I am especially thankful for

Öykü Üner, Emily Streeper, Eylül Tekin, Sam Chung, Emily Hanson, Marina Gross, and Katie Bentley for the work parties, the "work" parties, and all the tea/dinner/ice cream dates in between. I am also grateful to Vidhi Shah, Emily Phillips, and Nancy Ding for not only for their constant support from afar, but also their willingness to be subjected to long-winded rambles about my research at any given moment.

I am particularly grateful to Eric Failes for being the best partner through everything. I deeply appreciate his insistence in celebrating even my most minor accomplishments, his valiant (though futile) efforts to curb my comma and parentheses use in all my writing since my Master's proposal, his excellent advice regarding research methods and statistical analyses, and his unyielding faith in my abilities.

Finally, I would like to express my deepest gratitude to my parents and my older brother for their words of encouragement, unconditional love, and ability to find humor and joy in every situation. My parents taught me to ask a million questions of the world and instilled in me the confidence and drive to seek out answers. And my brother gave me the best advice for retaining my happiness and sanity during grad school before I had even started: "If you feel two weeks behind on work, you are already a week ahead of everyone else, so take it a day at a time." I have repeated this advice often, and along with my brother's many other words of wisdom, it has helped me immensely. I would not be the person I am today without my family, and this dissertation is for them.

Reshma Gouravajhala

Washington University in St. Louis

August 2020

ABSTRACT OF THE DISSERTATION

All Together Now: Effects of Simultaneous Presentation and Stimulus Complexity on Categorization Performance and Strategy Preferences

by

Reshma Gouravajhala

Department of Psychological & Brain Sciences

Washington University in St. Louis, 2020

Professor Mark McDaniel, Chair

The present study investigated the effects of simultaneous (relative to sequential) presentation on participants' learning of both simple (Experiment 1) and complex (Experiment 2) categories. Previous research has studied the impact of presentation mode on categorization of novel transfer objects, but the present study is the first to date that also examined its influence on learners' strategy preferences. Participants were trained (using a blocked observational training procedure) on two categories of abstract shapes that were defined by a bi-dimensional disjunctive rule. Some participants were shown objects sequentially, while others were presented with an organized simultaneous display depicting all to-be-learned stimuli at once. During training, participants responded to block-by-block strategy probes that provided online insight into the extent to which they were utilizing rule-based or exemplar-based strategies. Following training, participants classified ambiguous, rule-favored, and memory-favored transfer objects, and also reported any rules they had developed during training. Measures of working memory were also obtained. Participants' categorization performance on the transfer tasks were conditionalized on degree of rule acquisition as well as their block-level strategy preferences. The findings revealed that, in contrast to existing literature, simultaneous presentation generally promoted an exemplar-based approach to category learning.

Chapter 1: Introduction

Categorization is a fundamental cognitive skill that individuals of all ages use to classify novel stimuli based on shared characteristics and associations with previously learned information. Quick and efficient categorization is integral to many aspects of daily life, from the mundane (whether or not tomatoes are fruit) to the critical (what diagnosis to give a patient presenting a certain set of symptoms).

In the typical category learning experiment conducted in the laboratory, participants are presented with to-be-learned stimuli during a training phase, and then tested on their ability to categorize novel stimuli. Importantly, over the course of the task, participants make decisions about how to learn a given category. Indeed, much of category learning research over the last 60 years has focused on two cognitive strategies – rule abstraction and exemplar memorization – that can be used to determine category membership of to-be-learned stimuli, and their psychological and neurobiological underpinnings (Ashby & Ell, 2001; Levine, 1975). Briefly, rule abstraction involves hypothesis testing to develop a rule that can be used to categorize novel stimuli into learned categories, whereas exemplar memorization involves storing presented stimuli in long term memory, and classifying novel stimuli based on similarity to previously seen exemplars.

Critically, most category learning experiments to date share a methodological feature that might not accurately reflect how individuals learn new categories in more educationally relevant settings. Namely, typical laboratory studies present to-be-learned stimuli to participants in a sequential fashion: each exemplar is presented (with or without its associated category label) one at a time throughout a task. However, a quick glance at a science textbook or lecture slide reveals that, in the classroom, students are often presented with to-be-learned stimuli in a more simultaneous fashion, where multiple exemplars from different categories are presented at once (Figure 1.1).

Figure 1.1 Page from the *Smithsonian Field Guide to the Birds of North America*, depicting simultaneous presentation of stimuli from bird categories.

The present study aimed to extend previous work by examining whether presentation mode (sequential versus simultaneous presentation) affected individuals' categorization strategy preferences as they learned to categorize rule-based stimuli of varying complexity. Prior to describing our experiments, we first review the small but relevant literature on presentation mode in category learning, provide an overview of key findings regarding learners' categorization strategy preferences in a standard sequential presentation paradigm, and then theorize how these patterns might change under simultaneous learning conditions in the present study.

1.1 Sequential versus Simultaneous Presentation of To-Be-Learned Stimuli

At a broad level, there are theoretical advantages of simultaneous presentation over sequential presentation. Specifically, simultaneous presentation of exemplars might benefit comparative processes due to attenuated memory demands relative to sequential presentation (Kang & Pashler, 2012; Meagher et al., 2017). Unlike in sequential presentation, participants attempting to abstract a rule during simultaneous presentation would not be required to hold items in working memory as they compare and contrast across to-be-learned stimuli to identify diagnostic features. However, because exemplar memorization revolves around storing instances in long term memory, lowered working memory demands resulting from simultaneous presentation might not benefit individuals using a memory-based strategy to the same degree.

Previous studies of simultaneous versus sequential presentation chiefly focused on blocking (presenting all stimuli from the same category before showing stimuli from another category) versus interleaving (intermixing stimuli from different categories) of to-be-learned stimuli during training. Typically, sequential presentation has been compared to two types of simultaneous presentation: exemplars from the same category (referred to as "simultaneous blocked") and exemplars from different categories (referred to as "simultaneous interleaved").

These two types of simultaneous presentation have been theorized to offer unique merits. For example, simultaneous blocked presentation might allow participants to observe the range of features within a category, while simultaneous interleaved presentation encourages learners to contrast features between categories and identify diagnostic criteria (Andrews et al., 2011).

However, a direct comparison of sequential versus simultaneous presentation has yielded mixed results thus far: some studies have shown no difference in novel categorization performance (often used interchangeably with “transfer” or “generalization”) between sequential and simultaneous presentation (Andrews et al., 2011; Wahlheim et al., 2011), while others have shown the predicted benefits of simultaneous over sequential presentation (Higgins & Ross, 2011; Kang & Pashler, 2012; Meagher et al., 2017).

When comparing within the two simultaneous presentation types, a clearer pattern has emerged: simultaneous blocked presentation has resulted in worse generalization than simultaneous interleaved presentation (Andrews et al., 2011; Kang & Pashler, 2012; but see Sana et al., 2017, for some indication that simultaneous blocked might be more beneficial than sequential blocking). These findings indicate that performance when stimuli are presented simultaneously is maximized when participants are able to compare and contrast between exemplars of different categories.

Extending upon this idea, Meagher et al. (2017) recently conducted a naturalistic category learning study in which they presented participants organized simultaneous displays of complex and highly variable rock categories, interspersed within more typical sequential presentation blocks. The displays included all to-be-learned stimuli, ordered by category (Figure 1.2), which the authors believed would allow participants to observe both the shared features within a category and the diagnostic features that distinguished the categories at the same time. When their rock category stimuli included highly variable exemplars, Meagher et al. found that the simultaneous-interspersed presentation resulted in better performance only in the categorization of previously trained objects; there were no conditional differences with respect to generalization (categorization of novel exemplars from the trained categories). However, when more atypical rocks were removed from the dataset, categorization performance of novel objects was higher following the simultaneous-interspersed condition, in comparison to the sequential presentation-only condition (Experiment 2). These findings highlight that, when category stimuli are well-structured (i.e., have high within-category similarity), having the opportunity to simultaneously compare within and across categories can be beneficial for category learning.

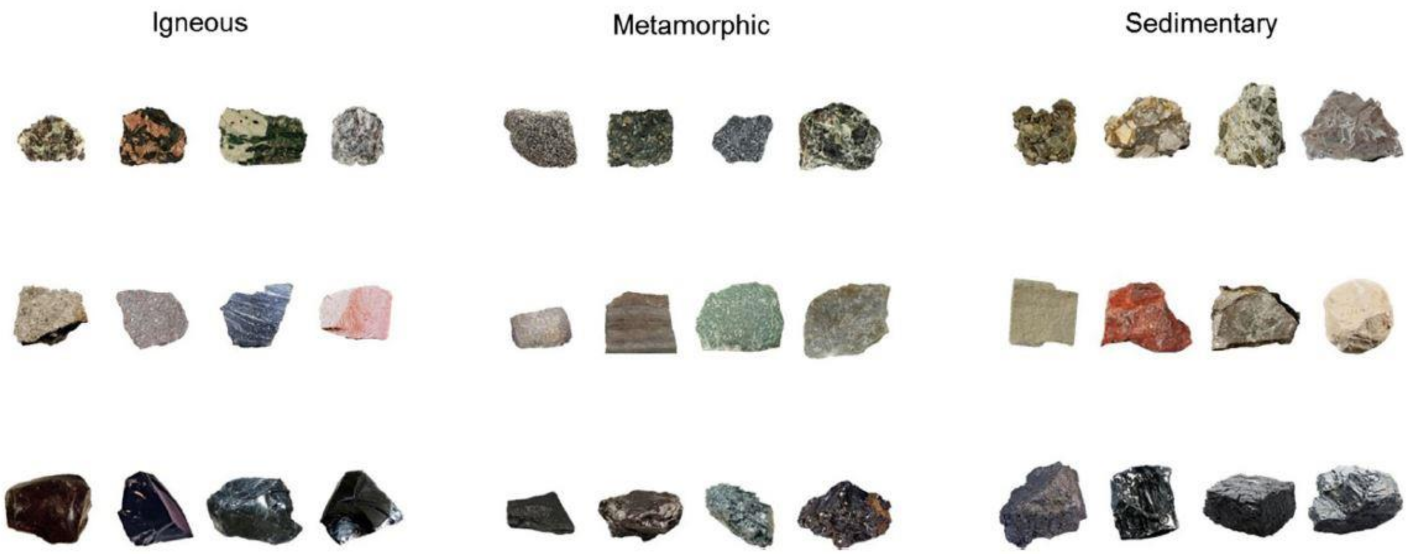

Figure 1.2. Organized simultaneous display of complex, multi-dimensional rock stimuli, organized according to their category membership. Taken from Meagher et al. (2017).

1.2 Categorization Strategy Preferences during Sequential Presentation

As previously mentioned, much of the category learning literature to date has focused on rule abstraction and exemplar memorization. For years, researchers argued that either all learners engaged in a rule-based approach to category learning (Bourne, 1974; Nosofsky et al., 1994; Trabasso & Bower, 1968) or that all learners employed a memory-based approach (Kruschke, 1992; Medin & Schaffer, 1978). In contrast to these views, some category learning researchers proposed hybrid models wherein rule abstraction and exemplar memorization could both be used to categorize stimuli (Anderson & Betz, 2001; Ashby et al., 1998; Erikson, 2008; Erikson & Kruschke, 1998).

Some task-level factors have been shown to influence learners' strategy preferences. For example, highly structured categories (i.e., ones that have high within-category similarity) tend to promote rule abstraction, while less well-structured categories appear to induce exemplar memorization (Craig & Lewandowsky, 2012). Furthermore, when Homa et al. (1981) increased the number of to-be-learned stimuli in their category learning task, participants tended to preferentially engage in rule abstraction, suggesting that participants adopted a rule abstraction strategy when the number of to-be-learned stimuli exceeded their capacity for memorization.

More relevant to the present study, there is also some evidence that increased stimulus complexity (as defined by number of dimensions along which stimuli vary) encourages participants to engage in more exemplar-based category learning. For example, Regehr and Brooks (1993) used artificial stimuli (imaginary animals whose features varied along five binary dimensions, three of which were relevant for categorization into two categories) to demonstrate that when participants were tasked with learning stimuli containing many idiosyncratic features (relative to those that followed a more systematic structure), they favored exemplar memorization relative to rule abstraction (but see Minda & Smith, 2001, for the opposite finding, perhaps owing to their modeling-based approach). Relatedly, recent work in naturalistic category learning has shown that participants learning to categorize highly multi-dimensional rocks might prefer (and benefit more from) memory-based strategies (Nosofsky et al., 2018).

In addition to task-level influences on categorization strategies, there is also evidence for learner-level factors that influence strategy preferences. In fact, McDaniel et al. (2014) identified individual differences in the degree to which learners adopted rule- and memory-based strategies on the same complex tasks (see also Medin et al., 1984). Critically, McDaniel et al. found these strategy preferences to be stable across multiple tasks, suggesting that some learners might exhibit a general tendency to prefer rule abstraction, while others typically engage in exemplar memorization. Indeed, Little and McDaniel (2015) asked younger adults to learn how to categorize eight abstract shapes defined by a bi-dimensional disjunctive relational rule. Using both an objective transfer measure and participants' strategy self-reports, the authors found that 52% of participants self-reported a global rule abstraction preference, while 37% of participants classified themselves as predominantly memorizers (and the remaining 11% reported roughly equal preference between the two strategies).

These previous studies highlight that, rather than assuming that all people approach a category learning task similarly, researchers and educators should be mindful of individual differences in strategy preferences. However, there is an additional complication to consider: participants might not persist with one strategy throughout a category learning task. In fact, there is considerable evidence of strategy shifting (from rule abstraction to exemplar memorization, or vice versa) in a variety of category learning tasks (Gouravajhala et al., 2019; Hoffmann et al., 2016; Johansen & Palmeri, 2002; Kalish et al., 2005). For example, in both artificial and natural language learning, this same pattern (transitioning from exemplar memorization to rule abstraction) is evident, as individuals initially categorize exemplars by memorizing specific bigrams and trigrams, but then, with practice, learn to abstract a grammar (Bourne et al., 1999).

To determine whether younger and older adults switched strategies during a feedback category learning task, Gouravajhala et al. (2019) utilized novel block-by-block strategy probes where participants self-reported strategy preferences throughout training, rather than simply relying on a single global strategy questionnaire, as is more typical (Wahlheim et al., 2016). Gouravajhala et al. found that a striking 93.3% of participants shifted between rule abstraction and exemplar memorization strategies at least once, and that 32.5% of participants shifted at least five times over the course of training. Furthermore, participants appeared to switch strategies during training as a direct consequence of high performance error on previous blocks.

Taken together, these findings suggest that strategy preferences during a category learning task, once believed to be the same across all learners, can be influenced by numerous task- and learner-level factors. To our knowledge, no category learning study to date has examined whether presentation mode is a task-level factor that influences strategy preferences (including the nuances of individual differences and strategy shifting behaviors that might emerge). Given the prevalence of simultaneous presentation in real-world educational settings, we believed it critical to address this issue.

1.3 The Present Study

We developed the present study with two primary objectives in mind: we aimed (1) to replicate and extend previous work by assessing the effects of presentation mode on rule-based category learning of simple (Experiment 1) and complex (Experiment 2) stimuli, and (2) to directly investigate whether simultaneous presentation, relative to sequential presentation, would differentially impact participants' categorization strategy preferences during our task.

To this end, younger adult participants were trained (over the course of several blocks) on the category membership of abstract shapes that adhere to a bi-dimensional disjunctive rule. Many of the procedural details of our experiments followed Gouravajhala et al. (2019), but we implemented three critical changes. Firstly, we added a simultaneous presentation condition modeled after the organized display layout used in Meagher et al. (2017). Secondly, we included measures of working memory to test theoretical benefits of simultaneous presentation on categorization. Finally, we utilized an observational training paradigm[1], such that all to-be-learned stimuli and their associated category labels were presented in temporal conjunction.

Participants' learning of the rule-based categories was measured in two ways. First, participants were asked at the end of training to report any rules they had developed. Second, for a more objective measure of participants' category learning, we measured participants' category learning by testing their ability to categorize ambiguous, rule-favored, and memory-favored transfer objects (Gouravajhala et al., 2019; Little & McDaniel, 2015; Wahlheim et al., 2016). Briefly, for each ambiguous item, there was a corresponding training object that was highly similar in form. However, when categorized according to the correct rule, ambiguous objects in fact belonged to the opposite category of their training object counterparts; thus, categorization according to perceptual similarity would yield an incorrect response. Rule- and memory-favored transfer objects assessed learners' acquisition of the correct rule and their memory for the trained objects, respectively. Together, these transfer objects aimed to assess how well participants had learned the rule-based categories.

To address the second main question of the present study, we also obtained measures of participants' strategy preferences across both presentation modes. Following Gouravajhala et al. (2019), participants were asked to provide strategy reports following each block of training. Importantly, we utilized this strategy probe methodology (in lieu of a global strategy preference questionnaire administered once at the end of training) in order to identify strategy preference dynamics during training. Next, we outline potential outcomes as they relate to our primary study objectives.

[1] The observational training paradigm was used to control for stimulus presentation time across conditions.

1.3.1 Presentation Mode and Categorization of Novel Transfer Objects

As previously described, though there have been mixed findings in comparisons of participants' categorization performance of novel items across sequential and simultaneous presentations, simultaneous presentation has been theorized to promote rule-based learning (Higgins & Ross, 2011; Kang & Pashler, 2012; Meagher et al., 2017). Thus, we would expect a greater number of participants in the simultaneous condition, relative to the sequential condition, to acquire and self-report the correct bi-dimensional rule in the present study. Moreover, if this were the case, we would also expect differences in their categorization of transfer objects, with participants in the simultaneous condition exhibiting higher rule-based accuracy than those in the sequential condition. Lastly, given that the benefit of simultaneous presentation is believed to result from an attenuation of working memory demands involved in hypothesis testing and rule abstraction, we would also expect that participants with low working memory capacities would especially benefit from simultaneous presentation on rule-based transfer performance.

However, a key methodological feature of the present study might lead to the emergence of the opposite patterns. Most prior studies utilized partially simultaneous displays (where only some of the to-be-learned stimuli were presented in temporal conjunction per trial), and the only study to date that utilized a fully simultaneous display (Meagher et al., 2017) interspersed these trials with more traditional sequential presentations of to-be-learned stimuli. The fully organized simultaneous displays in the present study might in fact improve memory for the trained stimuli, relative to those presented sequentially, for a couple of reasons. Firstly, whereas participants in the sequential condition were required to allocate the same amount of time and attention to every training object, those viewing the simultaneous display were able to allocate more or less time and attention to specific instances as needed over the course of training. Secondly, because the displays were organized by category (i.e., all category members were grouped together), the spatial proximity of objects within each category might have helped learners store more objects in memory. In fact, research in visual memory has found that structured displays (e.g., visual displays in which items are grouped together by proximity or similarity) improve memory for trained objects by reducing cognitive and neural loads, and effectively increase long-term memory capacity (Luria & Vogel, 2014; Gao et al., 2015; Magen & Emmanouil, 2019; Xu & Chun, 2007). If the same findings emerge in our category learning task, then fewer participants in the simultaneous presentation condition would be expected to acquire the correct rule (relative to those receiving sequential presentation), transfer performance would reflect improved memory for trained items, and working memory capacity would not be expected to impact categorization performance.

1.3.2 Presentation Mode and Strategy Preferences

Having detailed our predictions regarding the impact of presentation mode on categorization performance given previous findings, we now theorize how presentation mode might impact strategy preferences in the present study. We also make predictions below on the downstream effects these strategy preferences (with or without any individual differences that emerge) might have on strategy dynamics and categorization performance of novel transfer objects.

1.3.2.1 Possibility 1: No Individual Differences in Strategy Preferences

Owing to a few key methodological details in the present study, one possibility is that individual differences would not emerge in either the sequential or simultaneous presentation condition. Specifically, previous studies (all using a sequential presentation paradigm) showing individual differences in younger adults' strategy preferences have utilized feedback training, in which participants are presented with to-be-learned stimuli and tasked with categorizing each object prior to receiving corrective feedback (Gouravajhala et al., 2019; Little & McDaniel, 2015).

Without this regular feedback, participants might be less likely to engage in hypothesis testing and rule abstraction, especially when learning multi-dimensional categories (Alfonso-Reese, 1996; Ashby & Maddox, 2005; Ashby et al., 1998; Salatas & Bourne, 1974). Moreover, we used a limited set of to-be-learned stimuli that were repeated numerous times during training, which would also be expected to drive participants towards a memorization strategy (Homa et al., 1981; Kang & Pashler, 2012). Taken together, these findings suggest that all participants in the sequential condition might exhibit an exclusive preference for exemplar memorization.

There are reasons to believe that participants under simultaneous presentation conditions might also solely prefer one strategy, though there is less clarity on which one they would be directed towards. One possibility is that those in the simultaneous presentation condition would also prefer exemplar memorization, not only because of the reasons outlined above (observational training and repetition of a small set of stimuli), but also because learners might find a full organized display too overwhelming, and struggle to abstract a rule.

Alternatively, the opposite pattern might emerge for participants in the simultaneous condition. In the past, studies using a sequential presentation paradigm have found limited comparisons across training objects, perhaps owing to memory demands associated with maintaining all relevant objects in working memory. Indeed, participants have often only made comparisons between their current training object and those in a few preceding trials during training (Carvalho & Goldstone, 2014; Jones et al., 2006; Stewart & Brown, 2004). Simultaneous presentation of to-be-learned stimuli has been theorized to diminish working memory demands involved in hypothesis testing by allowing participants to have the opportunity to compare across training objects without needing to store them in memory (Higgins & Ross, 2011; Meagher et al., 2017). Thus, in a fully simultaneous display, there would be a relative ease of comparison across all objects (both within and between categories), which would be expected to encourage learners to engage in hypothesis testing more readily than in the sequential presentation condition. More specifically, Carvalho and Goldstone (2012) proposed that participants are motivated to look for similarities when comparing members of the same category, and differences when contrasting members of different categories. In doing so, participants become better equipped to identify diagnostic characteristics of categories in the task. Taken together, if this were the case, then we would expect all participants in the simultaneous condition to endorse a rule-based strategy.

It is important to note that some of the aforementioned task-level factors (e.g., stimulus repetition) would require multiple training blocks to take effect. Thus, we believed it necessary to assess participants' specific strategy preferences not only at the end of training (as was done in Gouravajhala et al., 2019), but also at the beginning of training and averaged across all training blocks. If no individual differences emerge in the present study, then we would expect all participants to show their predicted strategy preference on average or by the end of training.

1.3.2.1.1 Strategy Preference Dynamics during Training

We believed that the frequency of strategy shifts in the present study would be low for two reasons. Firstly, if participants in both sequential and simultaneous presentation conditions elect to endorse a single strategy (as a result of the factors described above), they would not be expected to switch back and forth between the two strategies. (A minor caveat: because task-level factors might require time to take effect, some participants might shift in the initial blocks before settling on the one strategy they endorse for the remainder of training. As a consequence, we will identify learners' strategy preferences at multiple points of training.)

Secondly, previous research on strategy shifts has revealed that participants' switching between rule abstraction and exemplar memorization strategies was directly precipitated by high performance error. In order for participants to be as reactive to any performance error in this study, they would have had to not only test themselves on each object and its associated category label (as they would be required to do in a feedback learning paradigm), but also monitor their learning. As this course of events seemed unlikely, we believed that there would be a decrease in strategy shifts in the present study, relative to Gouravajhala et al. (2019).

1.3.2.1.2 Categorization of Novel Transfer Objects and Acquisition of the Correct Rule

With respect to participants' categorization of ambiguous, rule-favored, and memory-favored transfer objects, we would expect performance to be driven by their preferred strategies. In other words, if all participants (across both presentation modes) exhibit a strong memorization preference, then we would expect transfer performance to reflect that tendency. Specifically, participants would be expected to categorize ambiguous objects according to perceptual similarity, show poor performance on rule-favored objects, and exhibit high accuracy on the memory-favored objects. However, if all participants in the simultaneous condition instead prefer a rule abstraction strategy by the end of training or on average, we would expect to find condition differences on novel transfer performance. Moving to the question of whether more participants in one condition would identify the correct rule, we would again expect participants' acquisition of the full rule to be directly impacted by their strategy preferences in a similar manner.

1.3.2.2 Possibility 2: Individual Differences in Strategy Preferences

Another possibility regarding strategy preferences in the present study is that individual differences do emerge. In an extension of McDaniel et al. (2014), it could be the case that participants' preferences for rule abstraction or exemplar memorization reflect stable tendencies (developed over a lifetime of category learning experiences) that neither observational training nor presentation mode could override. Thus, it is possible that the individual difference patterns observed in past studies (with roughly equivalent numbers of rule abstractors and exemplar memorizers) might emerge in the sequential condition (Little & McDaniel, 2015; Wahlheim et al., 2016).

However, if either of the previous hypotheses are affirmed for strategy preferences in the simultaneous condition (i.e., pushed towards memorization because of the training format, size of stimulus set, and feeling overwhelmed, or pushed towards rule abstraction because of lower working memory demands), we would expect a greater proportion of learners with that preference in that condition. In other words, we believed it possible that the frequencies of rule abstractors and exemplar memorizers in the simultaneous presentation condition would not be equal, unlike in the sequential condition.

If individual differences do arise in the present study, then we would be able to help further an existing debate about working memory's role in categorization strategy preferences. The findings thus far are mixed: Little and McDaniel (2015) found that individual differences in strategy preferences were not related to working memory capacity, but Wahlheim et al. (2016) found that younger adults with higher working memory capacity preferentially endorsed exemplar memorization over rule abstraction.

1.3.2.2.1 Strategy Preference Dynamics during Training

Even if individual differences in strategy preferences were to emerge, participants in the present study would not be expected to switch strategies very often due to the observational nature of training. In terms of potential condition differences, we would predict that, if anything, participants in the simultaneous presentation condition might exhibit more strategy shifts as the experiment progresses, simply as a result of boredom from repeatedly reviewing the same organized display. Specifically, participants in this condition might be motivated to test out another strategy if they feel the task is too repetitive with their current strategy. As described previously, participants' preferences will be calculated at the beginning and end of training, as well as on average.

1.3.2.2.2 Categorization of Novel Transfer Objects and Acquisition of the Correct Rule

As before, we would expect participants' categorization of novel ambiguous, rule-based, and memory-based transfer objects to reflect their strategy preferences. If individual difference patterns differ by presentation mode, we would expect there to be significant interactions between presentation mode and strategy preference when investigating performance differences on each of the three transfer tasks. Lastly, we would predict that participants' acquisition of the correct rule would again be influenced by their strategy preferences.

Chapter 2: Experiment 1

2.1 Method

Stimulus materials, programming scripts used for analysis, and raw anonymized data are available to the interested reader on the Open Science Framework (OSF; https://osf.io/he48n/). The experiment described below was approved by the Institutional Review Board of Washington University in St. Louis, and administered using Collector, a PHP-based software.

2.1.1 Participants and Design

The participants were 160 younger adults[2] at Washington University in St. Louis who received partial course credit for their participation in the experiment. Participants were divided into two between-subjects conditions: sequential presentation (*N* = 83) and simultaneous presentation (*N* = 77). Participants were recruited through the University's cloud-based participant management software system, SONA. The experiment was conducted online, and participants were instructed to complete the tasks individually, in one sitting, and without the use of electronic devices to aid them. Numerous probe questions were included to help ensure that participants paid attention to the task.

2.1.2 Procedure

The experiment lasted approximately 45 min in total. Over the course of the experiment, participants completed two working memory tasks (backward digit span and a shortened operation span), training (including block-by-block strategy probes), three transfer tasks (categorization of ambiguous, rule-favored, and memory-favored objects), and a brief global strategy questionnaire. Each task is described in further detail below:

[2] Sample size was determined by conducting an *a priori* power analysis on the interaction term for the 2 (presentation mode: sequential versus simultaneous) x 3 (rule acquisition / strategy group levels) χ^2 test. The results of the power analysis revealed that at least 76 participants were required in each condition.

2.1.2.1 Backward Digit Span

Following the consent process, participants completed a backward digit span task. In this task, a sequence of numbers (ranging from four to nine digits, increased incrementally during the task) was presented on a computer screen against a white background, with each digit on display for 1 s. At the end of each sequence, participants were prompted to report the numbers in that sequence in the reverse order of presentation. Participants first completed two practice trials (consisting of a three-digit sequence and a four-digit sequence) to gain familiarity with the task. Following these practice trials, participants completed 11 test trials, consisting of two trials per digit length (i.e., two sets of a four-digit sequence, then two sets of a five-digit sequence, etc.). This task lasted approximately five min.

2.1.2.2 Training

Following the backward digit span, participants were then provided the following brief instructional overview of the category learning component of the experiment:

> *"You will now complete a category learning experiment in which you will be presented with images of shapes, along with their associated category labels. You must learn how to categorize each shape into one of two categories: Blicket or Dax. You may choose to formulate a rule to learn the shapes, or memorize the shapes and associated category labels instead. Both strategies can be used to successfully complete this task. After you complete the learning phase, you will then be tested on how well you have learned the two categories."*

While it is not typical to alert participants to potential categorization strategies, these instructions aimed to help ensure that strategy probes presented throughout the proposed experiment did not encourage a shift to any particular strategy[3]. Participants were then asked about the instructions to check for comprehension and attentiveness. Next, all participants completed a training phase, which lasted approximately 15 min.

To help situate the present study in the existing body of literature and facilitate comparisons with past research, we chose to use previously created stimuli (Little & McDaniel, 2015). The training procedure was modeled largely on those used by Wahlheim et al. (2016) and Gouravajhala et al. (2019), with some important differences. In this experiment, training was conducted using an observational learning procedure during which participants learned to categorize objects into the “Blicket” or “Dax” category (see Figure 2.1 for examples of training stimuli). All category stimuli were made up of two colored shapes with one shape inside the other. These categories are defined by a disjunctive rule: if objects’ inner and outer shapes matched on either color or form, the object belonged to the “Blicket” category, whereas if the shapes differed in both color and form, the object belonged to the “Dax” category.

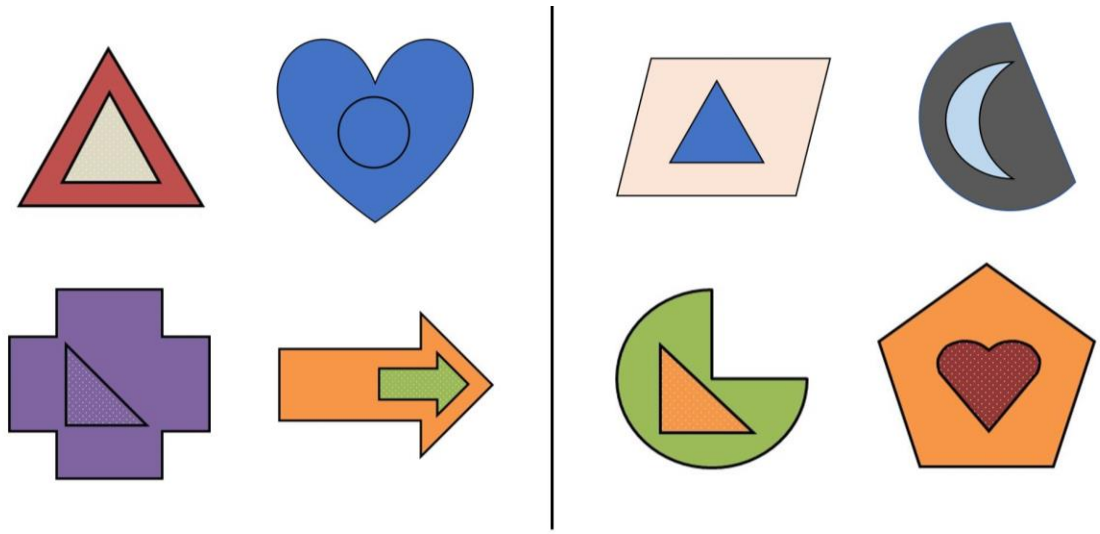

Figure 2.1 Eight of the 12 training stimuli used in Experiment 1. The four objects on the left belong to the “Blicket” category and include inner and outer shapes that share the same color or

[3] In the past, participants have rarely reported a preference for any strategy other than rule abstraction or exemplar memorization, and so there was limited concern that naming the two dominant strategies would shift participants away from another strategy they might otherwise have used (Gouravajhala et al., 2019).

form. The four objects on the right belong to the “Dax” category and include inner and outer shapes that share neither the same color nor form.

The training phase was divided into discrete learning blocks. As shown in Figure 2.2, younger adult participants in Wahlheim et al. (2016) and Gouravajhala et al. (2019) reached ceiling performance after approximately 10 blocks of training. However, because participants in the present study will be learning under observational learning conditions, they might experience slower learning rates (Ashby et al., 2002). To help offset this potential slower rate of learning, the present experiment included 12 learning blocks in the training phase.

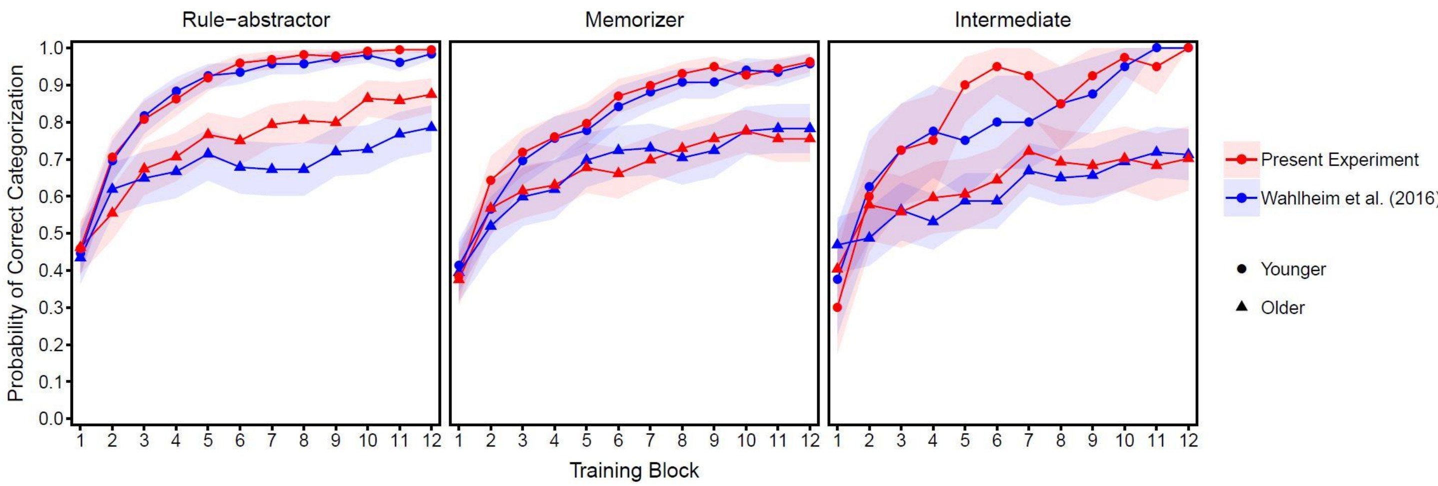

Figure 2.2. Probability of correct categorization across training blocks in Wahlheim et al. (2016, in blue) and Gouravajhala et al. (2019, in red). For the purposes of the present study, it is only important to note that younger adults’ performance (as depicted by round dots) reaches ceiling by approximately Block 10 in both studies. Taken from Gouravajhala et al. (2019).

In each block, 12 training items (six objects from each category) were presented on a computer screen against a white background. The associated category labels for each item were displayed directly beneath the stimulus for the duration of each trial. The same 12 items were repeated, either sequentially or simultaneously, across the 12 training blocks. For participants in the sequential presentation condition, each item was presented individually for 5 s in a predetermined random order, resulting in 12 separate learning trials. Training items were presented in a new random order across each sequential block. Participants in the simultaneous presentation condition were presented with an organized grid of all training objects (with left-right order on the screen counterbalanced across blocks) for 60 s during each block. The order of items within each half of the organized simultaneous display was randomized in each block. In both conditions, each block (whether containing multiple trials or a single simultaneous display) was followed by a self-report strategy probe, described in detail next.

2.1.2.2.1 Block-by-block Strategy Probes

After each training block, participants were probed about the degree to which they used rule-based or memorization-based strategies to categorize the objects in the previous two blocks. The following questions were presented (in a random order for each block) to participants: “How often did you apply a rule?”; “How often did you memorize objects and their associated category labels?”; “How often did you use a strategy other than rule abstraction or exemplar memorization?”. Participants rated their strategy use on a 5-point Likert scale, ranging from 1 (Never) to 5 (Always), and were instructed to use the full range of the scale and give an extreme rating only when they used a strategy exclusively during the preceding blocks. Not only were these probes utilized to reveal important information about individuals’ strategy preferences throughout training, but they also served to interject an interactive component to the observational learning paradigm.

2.1.2.3 Transfer

Following the 12 blocks of training, participants were then tested on their ability to categorize objects into “Blicket” and “Dax” categories on three different transfer tasks. The transfer phase lasted approximately 15 min.

2.1.2.3.1 Ambiguous Object Categorization

In this phase, participants categorized a set of 12 new ambiguous objects into either the "Blicket" or "Dax" category. As shown in Figure 2.3 (left panel), these transfer objects were created to be highly similar in form to training objects. However, when categorized according to the correct rule, each transfer object belongs to the opposite category than it did during training. As such, this task offers a more objective index of participants' strategy preferences, relative to their strategy self-reports. For example, a rule abstraction strategy preference would be revealed if ambiguous objects are categorized according to the rule, and an exemplar memorization preference would be revealed if ambiguous objects are instead categorized primarily according to memory for trained exemplars. Objects (without their associated category labels) were presented sequentially in a random order, and participants were provided labels ("Blicket" and "Dax") for their categorization selection. Participants were given 5 s to categorize each object. No feedback was provided during this task.

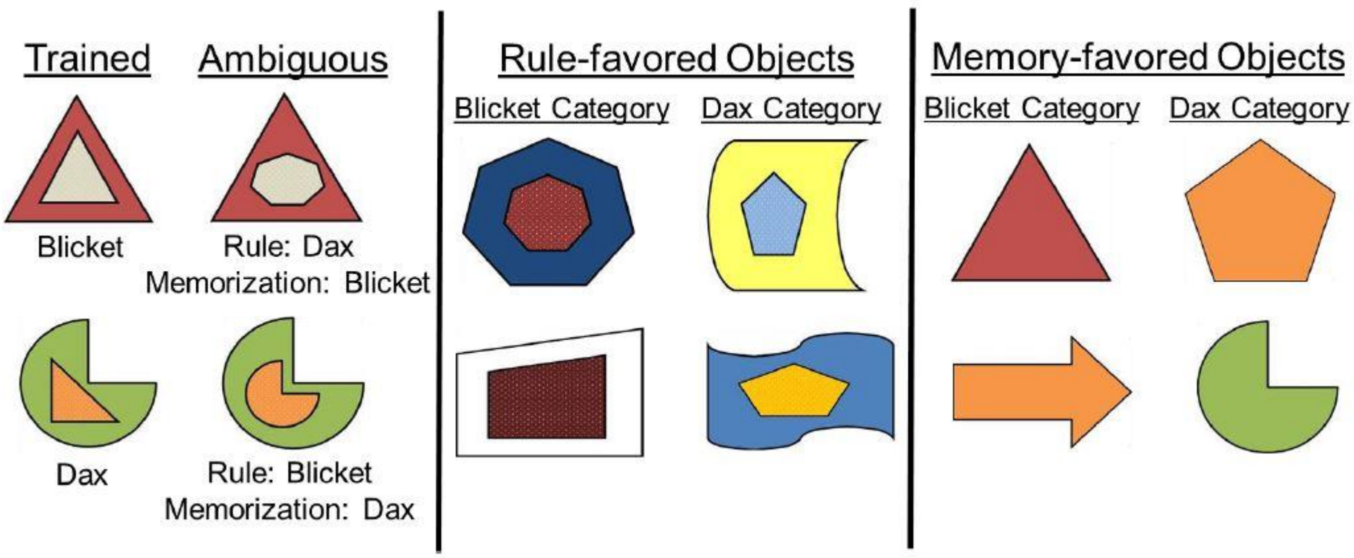

Figure 2.3. Left panel: Examples of training stimuli (left column) and their respective ambiguous transfer objects (right column). The ambiguous objects place rule-abstraction and memorization strategies in opposition to one another. Therefore, categorization of an ambiguous object according to the rule results in classification opposite that to which the perceptually similar training items belonged. Middle panel: Examples of novel objects used to assess rule-abstraction

for categories of training objects. Right panel: Examples of novel objects used to assess memory for categories of training objects. Taken from Gouravajhala et al. (2019).

2.1.2.3.2 Global Strategy Probes

After categorizing ambiguous transfer objects, participants completed a two-question strategy questionnaire, with questions presented in a random order for each participant. On the rule-oriented question, participants were instructed to verbalize any rule they used to classify objects (if they had developed one) so far in the experiment. There was no character limit imposed on participant responses. Participants who did not develop a rule were asked to type "No rule" or leave the answer field blank. While the rule question was of primary importance in the present experiment, a memory-oriented question was also presented to participant in order to not bias participants towards a rule-based strategy for the remainder of the category learning task. Thus, participants were also asked to report how many of the training objects they had memorized, and were provided the options "None", "Some", and "All" to choose from. The global strategy questionnaire was self-timed.

2.1.2.3.3 Rule-favored Object Categorization

In this phase, participants categorized 12 new objects that were perceptually dissimilar to any previously presented objects into the "Blicket" and "Dax" categories (see Figure 2.3, middle panel). Six of the objects followed the rule for the "Blicket" category, and six followed the rule for the "Dax" category. Rule-favored objects will be presented sequentially in a random order, without any associated category labels. As with ambiguous transfer objects, participants were given 5 s to select their response. No feedback was provided during this phase.

2.1.2.3.4 Memory-favored Object Categorization

In this final transfer phase, participants were presented with 12 new objects that comprised only of the outer shape (with the original color) of training objects (see Figure 2.3, right panel). As with the other transfer objects, memory-favored objects were presented sequentially in a random order, and no feedback was provided. The order of presentation between the rule-favored and memory-favored transfer was counterbalanced across participants, such that half of all participants completed the rule-favored transfer task following the ambiguous transfer task, and the other half completed the memory-favored task first.

2.1.2.4 Operation Span

In the final task, participants completed a shortened version of the Operation Span (Unsworth et al., 2005). In this task, participants were shown sets of letters (displayed for 1 s each) separated by arithmetic problems (displayed for a maximum of 15 s each). Each arithmetic problem was presented along with a solution, and participants were tasked with determining whether the solution was correct on each trial. After each set of letters and arithmetic problems, participants were then asked to report all the letters in the preceding series in the exact order of presentation. Sets ranged in size from three to seven letters, presented in a pre-determined random order for all participants. The task lasted approximately 10 min.

2.2 Results

We first present descriptive analyses for our two working memory tasks (backward digit span and operation span). Then, we address the question of whether participants in the sequential and simultaneous conditions differed in their accuracy of the rule, and present analyses with transfer performance conditionalized on learners' degree of rule acquisition. Here, we also present secondary analyses pertaining to the relationship between working memory capacity and rule-based transfer performance. Next, we address our second question about the relationship between presentation mode and categorization strategy preferences. We determined whether participants in the two conditions differed in their strategy preferences on the first training block, on the final training block, and on average across all training blocks. Then, we present analyses pertaining to participants' strategy preference dynamics during training. Lastly, we re-conditionalized participants according to the classification method that best captured their transfer performance (as determined by linear discriminant analyses; LDAs), and identified any differences in their categorization of ambiguous, rule-favored, and memory-favored transfer objects.

The level of significance was set at $\alpha = .05$. Logistic mixed effect regression models were created using the *glmer* function in the *R* package *lme4,* and subsequent linear combination tests of the fixed effects were conducted using the *glht* function in the *R* package *multcomp* (Bates et al., 2015). Subjects were entered as random effects in all models. Groups comparisons were reported as Wald *z* tests, and fixed effect estimates for all relevant analyses were converted to odds ratios (*OR*) for easier interpretation.

2.2.1 Working Memory

2.2.1.1 Backward Digit Span

In this task, participants were shown 12 sets of number sequences (each comprising four to nine digits), and then reported each series in the opposite order of presentation. For each set, participants were awarded credit (one point) if their response contained all presented numbers in the exact opposite order of presentation (i.e., reporting "5, 4, 3, 2, 1" following presentation of "1, 2, 3, 4, 5"). Thus, the total possible score on this task was 12 points. There were no differences in performance on this task between those in the sequential presentation condition ($M = 6.25$, $SD = 2.34$) and those in the simultaneous condition ($M = 6.01$, $SD = 2.43$), $t(156.02) = .63$, $p = .53$.

2.2.1.2 Operation Span

In this task, participants were shown 10 sets of three to seven letters intermixed with math problems (accompanied by potential solutions), and they were instructed to verify the accuracy of each solution while maintaining the letters in memory. Following each equation-letter set, participants were then asked to report the letters in the exact order of presentation. Participants' scores on this task were determined by calculating the sum of all perfectly recalled letter sets, for a total possible score of 75 points. As with the backward digit span, there were no differences in performance between those in the sequential presentation condition (M = 48.37, SD = 18.27) and those in the simultaneous condition (M = 46.33, SD = 17.50), $t(157.83) = .72$, $p = .48$.

2.2.1.3 Composite Working Memory Measure

Having established that there were no differences in working memory performance between conditions, we then examined the relationship between the two working memory measures collapsed across presentation modes, and found a significant correlation: $r = .50$, $t(158) = 7.23$, $p < .001$. Next, we standardized participants' performance in both tasks, and created an average working memory composite that was used in all future analyses involving working memory. Finally, as shown in Figure 2.4, we affirmed that there was a wide range of working memory composite scores in our sample.

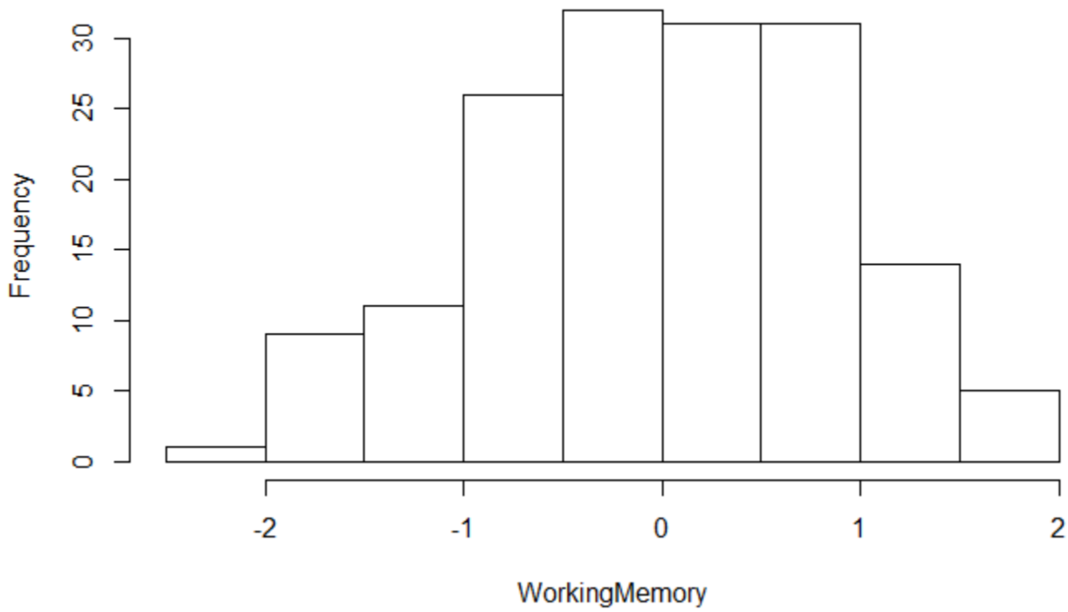

Figure 2.4. Histogram of the working memory composite scores in Experiment 1.

2.2.2 Effects of Presentation Mode on Rule Acquisition and Transfer

We first coded participants' free response reports on the rule question of the global strategy questionnaire. The author and a research assistant in the laboratory independently coded participants' self-reported rules into categories of "full rule" (responses mentioned both shape and color as diagnostic dimensions for classification of stimuli into "Blicket" and "Dax" categories), "partial rule" (responses mentioned only shape or color as a diagnostic dimension for classification) and "incorrect rule"[4] (responses mentioned using dimensions other than shape or color for classification, or participants stated that they had not used a rule and/or indicated reliance on memorization instead). The two raters exhibited excellent initial agreement (95.6% of responses), and they resolved disagreements through discussion. The frequencies of participants

[4] Following Gouravajhala et al. (2019), participants who reported use of an incorrect rule or no rule were grouped together. Furthermore, models that included four rule acquisition groups (full rule, partial rule, incorrect rule, no rule) failed to converge.

classified according to their degree of rule acquisition are displayed in Table 2.1. We first conducted a 2 (presentation mode: sequential vs. simultaneous) x 3 (rule acquisition: full vs. partial vs. incorrect) chi-square test of independence, which revealed a non-significant presentation mode x rule acquisition interaction, $\chi^2(2) = 4.03$, $p = .13$.

Table 2.1 Frequencies of Rule Acquisition in Experiment 1

	Degree of Rule Acquisition		
Condition	Full	Partial	Incorrect/None
Sequential	34	14	35
Simultaneous	20	17	40

We then conditionalized participants' transfer performance on their degree of rule acquisition (Table 2.2). It is important to note that categorization performance on our three tasks is measured differently. Owing to the nature of the ambiguous transfer task, performance closer to 1.00 reflects categorization according to the rule, whereas a score closer to 0.00 signifies categorization according to perceptual similarity. On the transfer task of rule-favored and memory-favored objects, however, performance is measured simply as a proportion of accurate responses.

Table 2.2 Probabilities of Correct Categorization of Transfer Objects as a Function of Rule Acquisition in Experiment 1

		Degree of Rule Acquisition		
Object Type	Presentation Mode	Full	Partial	Incorrect/None
Ambiguous	Sequential	.93 [.90, .95]	.35 [.28, .42]	.25 [.21, .30]
	Simultaneous	.84 [.80, .89]	.42 [.35, .49]	.20 [.16, .24]
Rule-favored	Sequential	.95 [.93, .97]	.71 [.64, .78]	.57 [.52, .62]
	Simultaneous	.95 [.93, .98]	.66 [.66, .72]	.56 [.51, .60]
Memory-favored	Sequential	.74 [.70, .78]	.76 [.70, .83]	.81 [.77, .85]
	Simultaneous	.89 [.85, .93]	.84 [.79, .89]	.83 [.80, .87]

Note: Performance on ambiguous and rule-favored objects reflect probabilities of categorization according to the correct rule. For ambiguous objects, performance of 1.00 would indicate perfect categorization according to the correct rule, and performance of 0.00 would indicate perfect categorization according to memory for perceptually similar training objects. Performance on Memory-favored objects reflects probabilities of correct categorization based on perceptually similar trained items. Margins of error for 95% confidence intervals are displayed in brackets.

To identify any performance differences, we conducted a no-intercept logistic mixed effect regression model that contained 19 fixed effects. The first 18 fixed effects each corresponded to a group (e.g., sequential presentation partial rule learner categorizing rule-favored objects) that was dummy coded (using 0s and 1s). The final fixed effect was working memory, which served as a covariate in the model. Overall, the model accounted for 39.53% of the variance (conditional $R^2 = .40$). We then conducted a series of theoretically motivated linear combination hypothesis tests to determine the significance of any effects (see Figure 2.5).

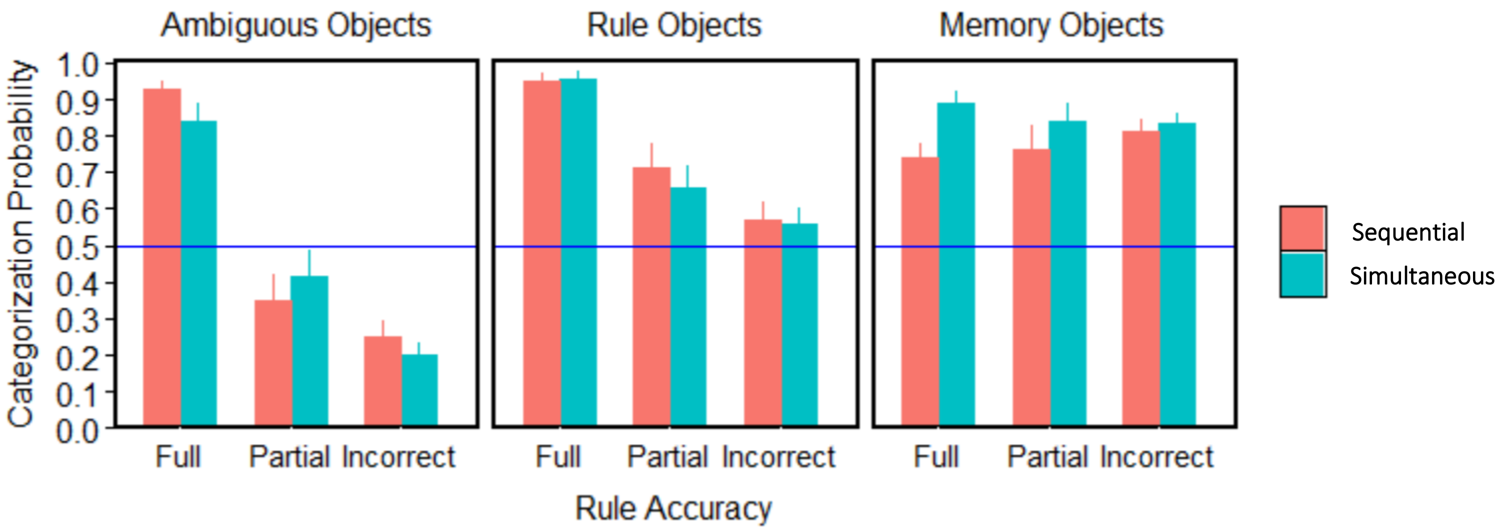

Figure 2.5. Probability of correct categorization of ambiguous, rule-favored, and memory-favored transfer objects in Experiment 1, conditionalized on the degree of participants' rule acquisition (full, partial, and incorrect) and presentation mode condition (sequential in red, simultaneous in teal).

Collapsing across rule acquisition levels, when going from sequential to simultaneous presentation, the odds of categorizing ambiguous objects according to the rule decreased by a factor of 3.03 (z = -2.21, p = .03), and the odds of an accurate response when categorizing memory-favored objects increased by a factor of 7.54 (z = 3.91, p < .001). There was no significant difference between presentation modes on categorization accuracy of rule-favored objects (OR = .90, z = -.19, p = .85).

Next, as a validity check for our scoring rubric, we compared whether participants in each of the three rule acquisition groups differed in their transfer performances, when collapsed across the two presentation conditions. In comparison to participants in the partial rule group, those in the full rule group categorized significantly more ambiguous objects according to the rule (OR = 236.81, z = 13.06, p < .001) and also had higher odds of a correct response when categorizing rule-favored objects (OR = 109.62, z = 9.20, p < .001). There was no difference between the two groups in categorization of memory-favored objects (*OR* = 1.48, *z* = .91, *p* = .36). When going from the incorrect rule to the full group, the odds of categorizing ambiguous objects according to the rule increased by a factor of 1889.35 (*z* = 18.47, *p* < .001) Moreover, the odds of a correct response on the rule-favored task increased by a factor of 374.98 (*z* = 12.40, *p* < .001), but there were no differences between the two groups in their performance on the memory-favored transfer task (*OR* = .91, *z* = -.24, *p* = .81).

Lastly, we tested the significance of relevant interactions. The presentation mode x full rule versus partial rule group interaction was significant only for ambiguous objects (*OR* = .34, *z* = -2.58, *p* = .01), but not for rule-favored or memory-favored objects (both *p*s > .05). Participants in the sequential condition who acquired the full rule had especially greater odds (in comparison to their simultaneous counterparts) of categorizing ambiguous objects according to the rule. The presentation mode x full rule versus incorrect rule group interactions were nonsignificant for ambiguous or rule-favored objects (both *p*s > .05) but were marginally significant for memory-favored objects (*OR* = 2.06, *z* = 1.83, *p* = .07). Specifically, participants in the simultaneous condition had greater odds of an accurate response when categorizing memory-favored objects if they had acquired the full rule. Lastly, the presentation mode x partial rule versus incorrect rule group interaction was significant only for ambiguous objects (*OR* = 2.00, *z* = 1.77, *p* = .08). As seen in Figure 2.5, the odds of categorizing ambiguous objects according to the rule were especially decreased for participants in the simultaneous condition who did not acquire a rule.

2.2.2.1 Effects of Working Memory Capacity on Rule-based Transfer

We next addressed a secondary question of whether participants' working memory capacities predicted rule-based transfer performance (i.e., categorization of ambiguous and rule-favored objects).

We first conducted a linear regression on mean ambiguous transfer performance, with presentation mode, working memory capacity, and presentation mode x working memory included as factors. The model accounted for 1.59% of the variance (adjusted $R^2 = .02$). After accounting for working memory capacity, there was a significant effect of presentation mode, $\beta = -.13$, $t(156) = -2.27$, $p = .02$, such that those in the simultaneous condition categorized more ambiguous objects according to perceptual similarity, relative to those in the sequential presentation condition. There was, however, neither a significant main effect of working memory capacity nor a significant presentation mode x working memory capacity interaction, both ps > .05.

We then conducted a linear regression on mean accuracy on the rule-favored objects, again with presentation mode, working memory capacity, and presentation mode x working memory included as factors in the model. This model accounted for .82% of the total variance (adjusted $R^2 = .008$). After accounting for working memory capacity, there was a marginally significant effect of presentation mode, $\beta = -.07$, $t(156) = -1.95$, $p = .05$, such that participants in the simultaneous condition showed poorer performance when categorizing rule-favored objects, relative to their counterparts in the sequential condition. Again, there were no other significant effects, both ps > .05.

2.2.3 Categorization Strategy Preferences during Training

Using the heuristic outlined in Gouravajhala et al. (2019), we compared participants' responses on the 5-point rule and memory probes[5] that followed each training block to classify each individual according to their responses. On each block, participants were assigned a strategy preference based on any numerical difference between their probe responses. For example, if on

[5] Participants gave consistently low ratings on the "Other" strategy probes, and so we did not include those strategy reports in determining their strategy preference on each block (consistent with Gouravajhala et al., 2019).

the third block, a participant responded with a “4” on the rule probe, and gave a “3” rating for the memory probe, that participant was classified as having a rule preference for that block. Importantly, because probes were presented in quick succession, we believed any numerical difference between them was meaningful. Individuals who provided the same numerical ratings for both the rule and memory probes were classified as having an equal preference for that block.

We used these block-level strategy preferences to identify whether individual differences in strategy preferences emerged at three points of training (following the first block, following the final block, and averaged across all blocks), and in either case, whether strategy preference patterns differed as a function of participants’ presentation mode condition.

2.2.3.1 First-block Strategy Preference

We first focused on participants’ strategy preferences following Block 1, which would reflect their initial choices after their first exposure to their assigned presentation mode. As a reminder, if there were no individual differences in strategy preferences at this point of training, then all participants within each condition would be expected to endorse the same strategy. However, as shown by the frequencies of participants classified according to this first-block method (Table 2.3), participants in the task endorsed different strategies.

Table 2.3 Frequencies of First-block Strategy Preferences in Experiment 1

	First-block Strategy Preference		
Condition	Rule	Memory	Equal
Sequential	41	23	19
Simultaneous	24	39	14

To determine whether these strategy preference differences differed by condition, we conducted a 2 (presentation mode: sequential vs. simultaneous) x 3 (first-block strategy: rule-first vs. memory-first vs. equal preference-first) chi-square test of independence, which revealed a significant presentation mode x first-block strategy interaction, $\chi^2(2) = 9.12$, $p = .01$.

Because of the 2 x 3 nature of the experimental design, the chi-square test was insufficient in revealing the underlying reasons for this interaction. Thus, we conducted a generalized linear model (using a Poisson error distribution) to predict frequency by presentation mode (dummy coded using 0s and 1s) and first-block strategy (with group membership dummy coded using 0s and 1s). In addition to these dummy-coded main effects, we included three presentation mode x first-block strategy interaction terms to create a fully saturated model that accounted for all variability in frequency counts.

We then performed a series of theoretically motivated linear combination tests of significance to identify any condition differences in each of the first-block strategy groups. The odds of endorsing a rule-first strategy decreased by a factor of 1.71 when going from sequential to simultaneous presentation ($z = 2.08$, $p = .04$). The odds of endorsing a memory-first strategy increased by a factor of 1.69 when going from sequential to simultaneous presentation ($z = -2.01$, $p = .04$). Participants in both presentation mode conditions, however, did not differ in their odds of endorsing rule and memory strategies equally following the first block of training ($OR = 1.36$, $z = .87$, $p = .39$).

Lastly, we conducted a linear regression on working memory capacity with rule-first, memory-first, and equal preference-first groups included as factors (dummy coded using 0s and 1s for group membership), collapsed across presentation mode condition. The results of our subsequent linear combination tests of significance revealed no significant effects, all *p*s > .05.

2.2.3.2 Final-block Strategy Preference

To determine if there were still individual differences in strategy preferences following the final block of training, we re-classified individuals according to their strategy preferences following the 12th training block. Again, as shown in Table 2.4, the frequencies of participants according to this final-block method revealed that participants in either condition did not settle on a single strategy preference by the end of training.

Table 2.4 Frequencies of Final-block Strategy Preferences in Experiment 1

	Final-block Strategy Preference		
Condition	Rule	Memory	Equal
Sequential	42	16	25
Simultaneous	34	24	19

To test for condition differences in strategy preferences, we conducted a 2 (presentation mode: sequential vs. simultaneous) x 3 (final-block strategy: rule-final vs. memory-final vs. equal preference-final) chi-square test of independence. Unlike with first-block strategy preference frequencies, there was no significant presentation mode x final-block strategy preference interaction, $\chi^2(2) = 3.04$, $p = .22$.

As before, we conducted a linear regression to examine the relationship between working memory capacity and each of the three final-block strategy groups (dummy coded using 0s and 1s for group membership), collapsed across presentation modes. There were no significant effects, all *p*s > .05.

2.2.3.3 Average Strategy Preference

Finally, we re-classified participants according to their average strategy preference over the 12 training blocks. In other words, for each individual, we determined which strategy preference they endorsed most often. For a specific example, if a participant exhibited a rule preference on 7 blocks, a memorization preference on 3 blocks, and an equal strategy preference on 2 blocks, they would then be classified as having an average rule preference[6]. The frequencies of participants classified according to this average strategy method are displayed in Table 2.5.

Table 2.5 Frequencies of Average Strategy Preference in Experiment 1

	Average Strategy Preference		
Condition	Rule	Memory	Equal
Sequential	46	17	20
Simultaneous	32	21	24

We conducted a 2 (presentation mode: sequential vs. simultaneous) x 3 (average strategy: rule vs. memory vs. equal) chi-square test of independence, which revealed a non-significant presentation mode x average strategy preference interaction, $\chi^2(2) = 3.08$, $p = .21$. Thus, there were no differences in average strategy preferences across the two presentation mode conditions.

Again, we conducted a linear regression on working memory capacity with average rule, average memory, and average equal strategy preference groups included as factors (dummy coded using 0s and 1s for group membership). There were no significant effects, all *p*s > .05.

[6] Two participants exhibited equal endorsement of rule and memorization strategies across blocks, and were classified as having an average equal preference.

2.2.4 Strategy Preference Dynamics during Training

Having explored strategy preferences at different points of training, we next addressed the issue of strategy preference dynamics in the present experiment. Using the same method described in Gouravajhala et al. (2019), we calculated whether individuals' strategy preferences changed – rule to memory, rule to equal, memory to equal, memory to rule, equal to rule, equal to memory – in consecutive training blocks[7]. Thus, each participant was able to switch for a total possible 11 times. For participants in both conditions, we calculated the total number of strategy switches or shifts (see Figure 2.6).

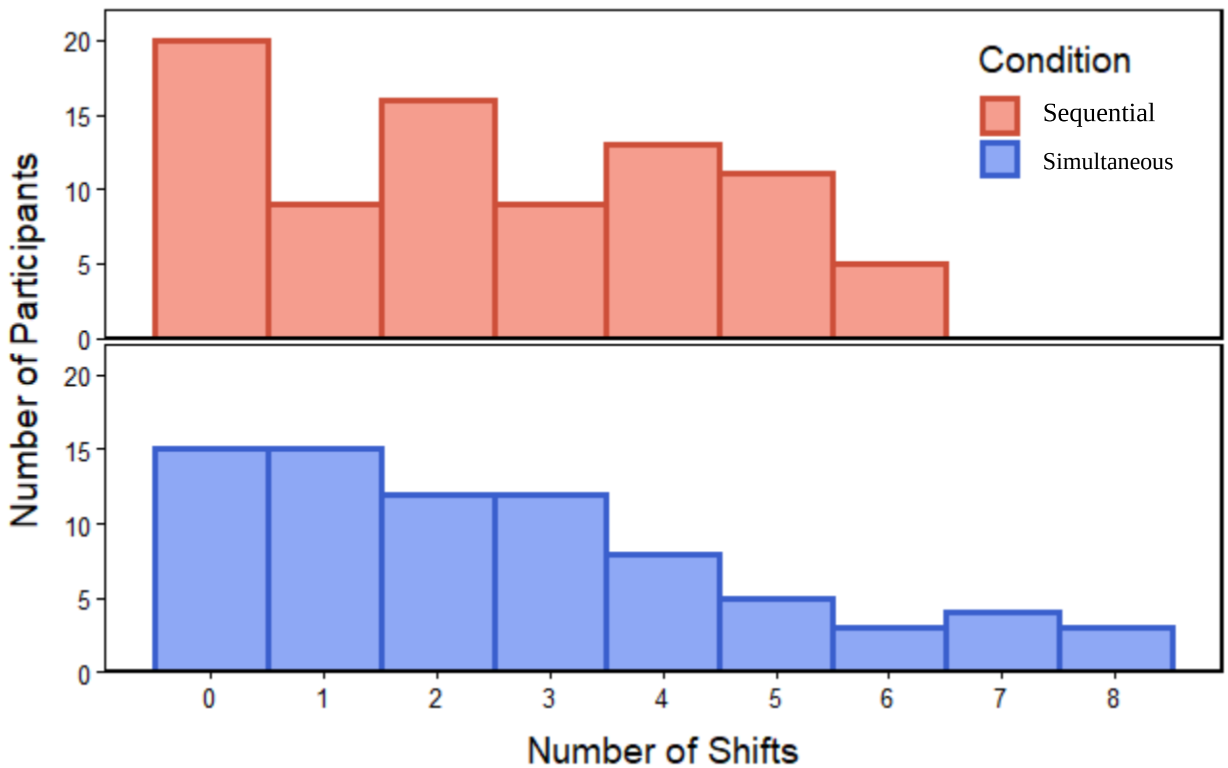

Figure 2.6. Histogram depicting the number of times, in Experiment 1, participants in the sequential presentation condition (top panel, in red) and the simultaneous condition (bottom panel, in blue) shifted from one strategy to another on consecutive blocks.

[7] All six patterns were considered to reflect a strategy switch, as they indicated a change in preference from the previous block.

While participants in both conditions exhibited shifting behavior, there was no difference in the total number of shifts by participants in the sequential condition (M = 2.47, SD = 1.93) and simultaneous condition (M = 2.62, SD = 2.25), t(150.37) = -.46, p = .65. Additionally, the correlation between working memory and total number of shifts was not significant, r = .04, t(158) = .48, p = .63. Finally, as predicted, there were significantly fewer shifts in the present experiment (collapsed across presentation mode condition; M = 2.54, SD = 2.10) relative to in Gouravajhala et al. (2019; M = 3.44, SD = 2.62), t(222.61) = 3.08, p = .002. As a reminder, the observational nature of the present study (where participants were not provided direct correct-answer feedback following each trial) was expected to result in fewer strategy shifts (Kalish et al., 2005).

2.2.4 Effects of Participants' Strategy Preferences on Transfer

Although we previously examined condition differences in participants' transfer performance (conditionalized on their degree of rule acquisition), those analyses neglected learners' strategy preferences. Thus, we were interested in whether the strategy preference patterns that emerged during training had a downstream impact on categorization performance.

Though we had developed three approaches to classify participants (first-block strategy preference, final-block strategy preference, and average strategy preference) that could feasibly be used to conditionalize transfer performance, we believed that learners' first-block strategy preferences would be least aligned with their categorization of transfer objects. After all, participants reported first-block preferences after the first block of training, and then categorized transfer objects at a minimum of eleven training blocks later. Relatedly, we believed the strategy shifting behavior exhibited in the present study (78% of participants shifted at least once, and

32.5% shifted at least four times) lent additional support against conditionalizing transfer performance on participants' first-block preferences.

With respect to the other two classification methods, we predicted that learners' final-block strategy preferences would more closely align with subsequent transfer performance, as their preferences might reflect their final representations before beginning the transfer phase (as was theorized by Gouravajhala et al., 2019). To test this hypothesis, we conditionalized participants' transfer performance both on final-block strategy preferences and on average strategy preferences. We then conducted separate linear (canonical) discriminant analyses to determine whether mean transfer performance was better able to separate participants into their final-block strategy groups or their average strategy groups. The results of these analyses indicated that, while both classification methods captured similar amounts of variance, conditionalizing transfer performance according to participants' final-block strategy preferences was slightly more effective (canonical R^2 = .325) relative to conditionalizing on average strategy preference (canonical R^2 = .289).

Thus, we chose to conditionalize categorization transfer performance on participants' final-block strategy preferences[8], and conducted a no-intercept mixed-effects logistic regression analysis (Table 2.6). This model was composed of 19 fixed effects. The first 18 fixed effects each represented a group (e.g., sequential presentation rule-final learner categorizing ambiguous objects) that was dummy coded (using 0s and 1s). The final fixed effect was working memory, which acted as a covariate in the model. Overall, the model accounted for 36.98% of the total variance (conditional R^2 = .37).

[8] The interested reader can find similar analyses with participants classified according to their average strategy preferences in Appendix A.

Table 2.6 Probabilities of Correct Classification of Transfer Objects as a Function of Final-block Strategy Preferences in Experiment 1

Object Type	Presentation Mode	Final-block Strategy Preference		
		Rule	Memory	Equal
Ambiguous	Sequential	.74 [.70, .78]	.27 [.21, .33]	.39 [.34, .45]
	Simultaneous	.61 [.57, .66]	.33 [.27, .39]	.16 [.11, .21]
Rule-favored	Sequential	.83 [.80, .87]	.60 [.53, .67]	.71 [.66, .76]
	Simultaneous	.79 [.75, .83]	.59 [.54, .65]	.60 [.54, .66]
Memory-favored	Sequential	.77 [.73, 81]	.78 [.72, .84]	.77 [.72, .82]
	Simultaneous	.86 [.82, .89]	.82 [.78, .87]	.87 [.82, .91]

Note: Performance on ambiguous and rule-favored objects reflect probabilities of categorization according to the correct rule. For ambiguous objects, performance of 1.00 would indicate perfect categorization according to the correct rule, and performance of 0.00 would indicate perfect categorization according to memory for perceptually similar training objects. Performance on Memory-favored objects reflects probabilities of correct categorization based on perceptually similar trained items. Margins of error for 95% confidence intervals are displayed in brackets.

We then computed a series of theoretically motivated linear combination hypothesis tests to determine significance of any effects (Figure 2.7). When going from the memory-final to the rule-final preference group, the odds of categorizing ambiguous objects according to the rule increased by a factor of 43.97 (z = 8.27, p < .001) and the odds of an accurate response on the rule-favored transfer task increased by a factor of 14.65 (z = 5.83, p < .001). There were no differences between the preference groups, however, in their odds of an accurate response on the memory-favored transfer task (OR = 1.69, z = 1.09, p = .28).

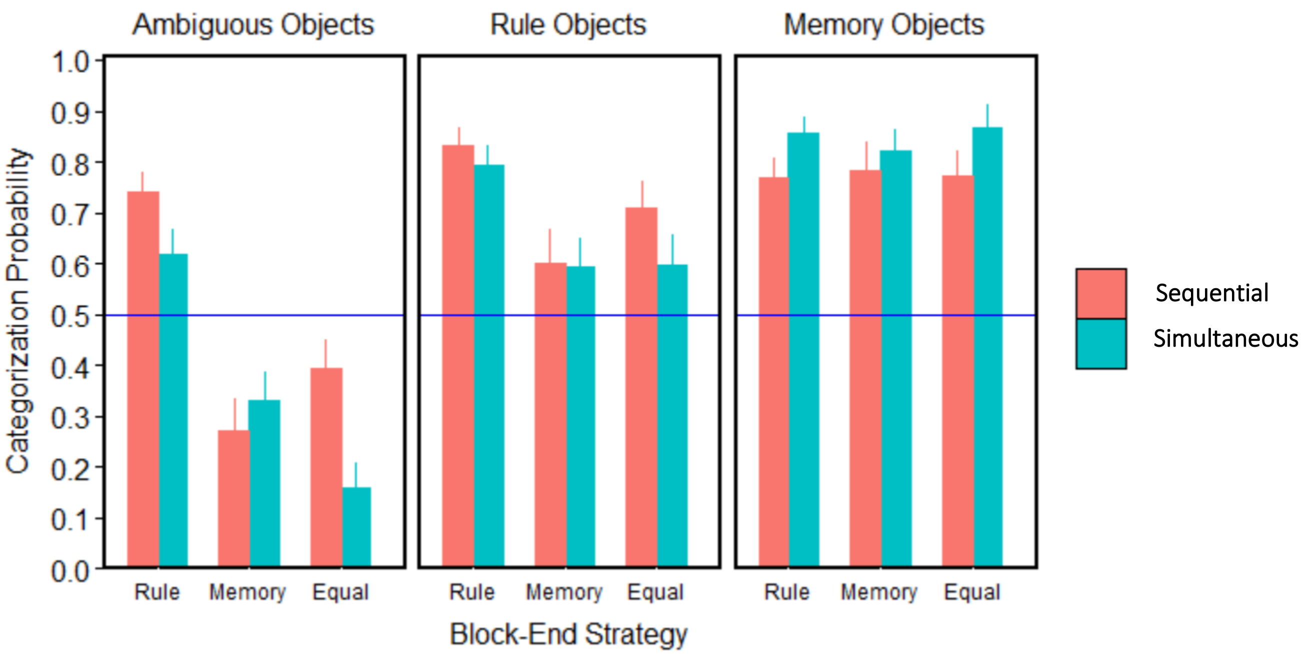

Figure 2.7. Probability of correct categorization of ambiguous, rule-favored, and memory-favored transfer objects in Experiment 1, conditionalized on participants' final-block strategies (rule, memory, and equal preference) and presentation mode condition (sequential in red, simultaneous in teal).

Relative to those who endorsed both strategies equally after the final block of training, participants in the rule-final group had significantly greater odds of categorizing ambiguous objects according to the rule (*OR* = 63.69, z = 9.21, p < .001). Participants in the rule-final group also had greater odds of an accurate response on the rule-favored task (*OR* = 8.76, z = 4.86, p < .001), but the two groups did not differ in their odds of an accurate response on the memory-favored objects (*OR* = 1.34, z = .62, p = 54). There were no differences in the categorization of ambiguous, rule-favored, or memory-favored objects between those who ended with a memory preference versus those who ended with an equal preference, all *p*s > .05.

The presentation mode x rule-final versus memory-final strategy preference group interaction was significant only for ambiguous objects (*OR* = .39, z = -2.04, p = .04), but not for rule-favored or memory-favored objects (both *p*s > .05). Specifically, participants in the sequential presentation condition had greater odds of categorizing ambiguous objects according to the rule especially when they were in the rule-final group. The presentation mode x rule-final versus equal preference-final group interactions were nonsignificant across all transfer objects type, all *p*s > .05. Lastly, the presentation mode x memory-final versus equal preference-final strategy group interaction was significant only for the ambiguous objects ($OR = .75$, $z = 3.18$, $p = .001$), such that the odds of categorizing ambiguous objects according to the rule were especially lower for equal preference-final participants in the simultaneous condition.

2.3 Discussion

To our knowledge, Experiment 1 was the first investigation into the effects of simultaneous (relative to sequential) presentation on both categorization of novel transfer items and categorization strategy preferences. With respect to the effects of simultaneous presentation on rule acquisition and categorization performance, the two presentation modes have sometimes resulted in similar transfer performance, but when differences have emerged, simultaneous presentation has improved generalization. Researchers have theorized that these benefits were due to simultaneous presentation (especially when the display contained stimuli from all to-be-learned categories) lowering working memory demands, and thus directly benefiting hypothesis testing and rule abstraction (Kang & Pashler, 2012; Meagher et al., 2017). If this were the case, participants in the simultaneous condition would have benefitted from their training in their categorization of rule-favored objects, especially if they had low working memory capacities.

However, the findings of Experiment 1 directly contradicted this prediction. We found similar frequencies of learners in both conditions across rule acquisition groups (full, partial, and incorrect rule), suggesting that simultaneous presentation did not improve learners' ability to acquire the correct bi-dimensional disjunctive rule that defined "Blickets" and "Daxes." Further, when conditionalizing transfer performance on degree of rule acquisition, the patterns strongly indicated that participants' simultaneous presentation training had in fact improved memory for studied stimuli (relative to those in the sequential condition). Those in the sequential condition were better able to categorize objects according to the rule (especially if they were able to identify the full rule), whereas participants in the simultaneous condition fared better on memory-favored objects. And finally, working memory capacity was not related to participants' performance on rule-favored transfer tasks in either condition, suggesting that previous theories about the benefits of simultaneous presentation might require revision.

With respect to strategy preferences, we outlined competing hypotheses in the introduction regarding whether individual differences in strategy preferences would emerge in the present study, and what downstream effects these strategy preference patterns would have on strategy dynamics and categorization performance. Individual differences in strategy preferences emerged at various points of training (following the first block, following the final block, and averaged across all 12 blocks), suggesting that presentation mode in our study was not a task-level factor that could drive all participants towards a single strategy. Furthermore, even in the presence of individual differences, there were some differences in strategy preferences across those in the simultaneous versus sequential condition. Following the initial training block under their respective presentation conditions, participants differed systematically in their preferred strategies. Specifically, those receiving simultaneous presentation preferred to engage in exemplar memorization, while those learning in the more typical sequential paradigm showed an initial preference for rule abstraction. These findings suggest that, when presenting novel to-be-learned stimuli only once, educators should be cognizant of presentation mode, as it could impact the categorization strategy learners choose to adopt.

However, by the end of training, learners across both presentation modes displayed similar patterns of individual differences in strategy preferences (Table 2). Thus, it might be the case that extensive repetition of a relatively limited set of to-be-learned stimuli encouraged some learners to vary in their strategy use, as both strategies were viable and effective. Indeed, there was clear evidence of strategy shifting, with almost a third of individuals switching between strategies at least four times during training. As predicted, however, the observational nature of the experiment (i.e., absence of direct corrective feedback to help learners monitor their error rates) contributed to decreased switching rates relative to Gouravajhala et al. (2019).

Taken together, the findings in Experiment 1 suggest that the two presentation modes confer different benefits. Thus, educators should consider their specific goals for students (i.e., whether they want students to abstract a rule for to-be-learned categories or memorize trained stimuli and categorize novel exemplars based on perceptual similarity) prior to choosing a particular mode of display.

While there were clear differences in transfer performance across the two presentation modes, we were intrigued that that these differences were not always reflected in participants' self-reports. One possible explanation for this discrepancy was that the stimuli used in Experiment 1 were too simple and allowed participants to readily choose either a rule- or memory-based strategy, even though their respective presentation modes offered different benefits. Thus, in Experiment 2, we manipulated the to-be-learned stimuli in an effort to direct all participants (especially those in the simultaneous presentation condition) towards a memory-based strategy, and we examined whether there would be more alignment between participants' strategy preferences and categorization performance.

Chapter 3: Experiment 2

3.1 Introduction

Experiment 1 focused on the effects of simultaneous presentation on the category learning of previously used stimuli ("Blickets" and "Daxes," as developed in Little & McDaniel, 2015). In Experiment 2, we aimed to extend these findings by investigating what patterns would emerge when using more complex stimuli. Specifically, categories from Experiment 1 were modified to be more idiosyncratic in nature, with multi-dimensional stimuli that now varied along three salient but irrelevant dimensions: internal shape pattern, external shape pattern, and antennae (see Figure 3.1 for examples of training stimuli).

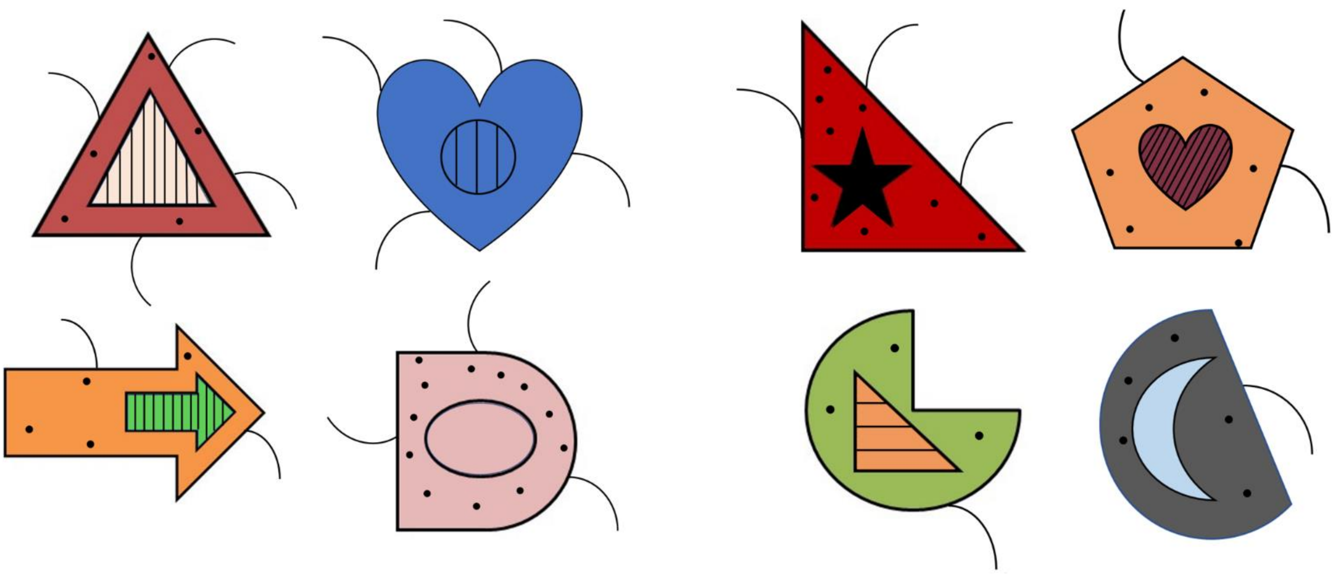

Figure 3.1. Eight of the 12 training stimuli used in Experiment 2. The four objects on the left belong to the "Blicket" category, and they include inner and outer shapes that share the same color or form. The four objects on the right belong to the "Dax" category, and they include inner and outer shapes that share neither the same color nor form. The additional dimensions of internal stripe pattern, external dot pattern, and antennae are irrelevant for correct classification.

Critically, the underlying category-defining rule – the inside and outside shapes of "Blickets" match in either form or color, whereas "Daxes" match in neither – remained the same in this experiment. Because the three additional dimensions were not diagnostic, objects within and across both categories shared certain dimension values. For example, objects from both “Blicket” and “Dax” categories might share a striped pattern in the internal shape. To ensure that participants do not immediately discount each of these dimensions as being irrelevant to the categorization task, the shared features were not exactly the same across all members of both categories. To refer back to the striped inner shape example, the number and orientation of stripes might differ from object to object.

The stimulus modifications in Experiment 2 were motivated by several factors. For example, one primary objective was to create stimuli that would more accurately reflect the complexity existent in many real-world categories, thus improving upon the ecological validity of the present study (Murphy, 2003).

Another objective was to extend Experiment 1 findings regarding the benefits of simultaneous presentation for memory of trained stimuli. We had reason to believe that these patterns would be replicated. Specifically, by constructing a highly dimensional stimulus set, we increased the idiosyncrasy and distinctiveness of the stimuli, thus rendering them easier to store in long-term memory (Konkle et al., 2010; Regehr & Brooks, 1993). However, prior research on stimulus complexity has only been conducted under feedback learning conditions (with all to-be-learned stimuli presented sequentially), and thus the present experiment directly tested whether these patterns would emerge under observational and simultaneous presentation conditions.

Finally, we wished to extend previous findings regarding participants’ strategy preferences. The same training procedures from Experiment 1 were used in this experiment, such that the relatively limited set of to-be-learned stimuli were repeated across 12 blocks. Thus, both rule-based and exemplar-based strategies were equally viable in this experiment. However, we expected that the additional irrelevant dimensions would cause confusion during hypothesis testing and increase the prevalence of incorrect rules, especially when participants attempted to incorporate all five dimensions into their rules (Vong et al., 2019). Thus, we believed that participants across both conditions would preferentially endorse exemplar memorization during the task. If simultaneous presentation was still beneficial for memory-based learning, then we would have greater alignment between transfer performance and participants' strategy preferences than in Experiment 1.

Below, we describe our predictions regarding how sequential versus simultaneous presentation of these complex stimuli might impact rule acquisition, categorization of novel transfer objects, and categorization strategy preferences.

3.1.1 Effects of Presentation Mode on Rule Acquisition and Transfer using Complex Stimuli

Because the addition of irrelevant dimensions was expected to impair hypothesis testing and rule abstraction, we expected fewer participants across both conditions would obtain the correct rule and that rule-based transfer performance would be poorer in this experiment (relative to Experiment 1). With respect to differences across the two conditions, there are several potential outcomes.

Previously, we found that simultaneous presentation, relative to sequential presentation, promoted memory for the trained items (as evidenced by participants' performance on the transfer tasks), rather than benefiting the acquisition and use of the bi-dimensional rule, as had been previously theorized by other researchers (Andrews et al., 2011; Meagher et al., 2017). Since no procedural details of the presentation modes were altered in Experiment 2, we believed it likely that the simultaneous presentation condition would again benefit participants' memory for trained stimuli.

Alternatively, it is possible that the organized simultaneous display would in fact promote rule-based learning (in contrast to Experiment 1 findings) by alerting participants to the irrelevant dimensions (especially since certain feature values were shared by members of both categories) more readily than sequential presentation would be expected to. If this were the case, then we would expect a greater number of participants in the simultaneous condition to acquire the correct rule and exhibit higher accuracy on the rule-based transfer tasks. Additionally, learners find the multi-dimensional stimuli too overwhelming to memorize in the simultaneous condition, and instead attempt to abstract a rule to increase efficiency.

Moving to a secondary question about the relationship between working memory capacity and rule-based transfer performance, one possibility is that Experiment 1 findings would be replicated, such that working memory capacity would have no effect on categorization performance. Another possibility, however, is that low working memory capacity learners in the sequential presentation condition would have especially poor rule-based transfer performance, as they would be at the greatest disadvantage when hypothesis testing and abstracting a rule (Andrews et al., 2011; Bourne, 1974).

3.1.2 Presentation Mode and Categorization Strategy Preferences

In Experiment 1, there were some differences across the two presentation modes in terms of strategy preferences. Specifically, learners in the simultaneous presentation condition were more likely to exhibit a memorization preference following the first training block, while those in the sequential presentation condition had greater odds of endorsing a rule abstraction strategy. However, individual differences in strategy preferences abounded, and condition differences disappeared by the end of training, and on average across all training blocks. As previously mentioned, a primary objective of Experiment 2 was to, by utilizing complex stimuli that comprised numerous irrelevant dimensions, determine if participants would preferentially endorse an exemplar-based strategy throughout training. Thus, one possibility is that participants in Experiment 2 (especially in the simultaneous condition) would all exhibit a persistent exemplar memorization strategy preference. If this were the case, then we would expect there to be very few strategy switches during training, and for transfer performance to be largely memory-favored.

Another possibility is that the same individual difference patterns as in Experiment 1 again emerge: although participants in the two conditions might exhibit a strong preference for one of the strategies following the first block, other task-level factors (such as stimulus set size and repetition during training) might lead to the emergence of individual differences in strategy preferences in subsequent blocks. Transfer performance would then be expected to follow Experiment 1 in this case.

With respect to strategy dynamics, if individual differences in strategy preferences develop, one possibility is that participants in Experiment 2 will shift strategies with a similar frequency as participants in the first experiment. However, it is also possible that the complexity of the stimuli would encourage more strategy shifting than in Experiment 1. Specifically, we would expect that Experiment 2 stimuli would be more difficult to learn (either through memorization or rule abstraction), and thus participants might struggle more during their training than did participants in Experiment 1. Thus, learners in the present experiment (if they were attuned to their category learning during the task) might be more motivated to shift between strategies. In support of this point, there is evidence suggesting that younger adults have enough metacognitive awareness to evaluate their learning and moderate their use of general study strategies (Brigham & Pressley, 1988).

3.2 Method

As with Experiment 1, all stimulus materials, programming scripts used for analysis, and raw anonymized data are available to the interested reader on the OSF (https://osf.io/he48n/). This experiment was approved by the Institutional Review Board of Washington University in St. Louis, and administered using Collector, a PHP-based software.

3.2.1 Participants and Design

The participants were 160 younger adults at Washington University in St. Louis who received partial course credit for their participation. Participants were again divided into two between-subjects conditions: sequential presentation (N = 80) and simultaneous presentation (N = 80). Participants were recruited through the University's cloud-based participant management software system, SONA, and completed the experiment online. As before, participants were instructed to complete the tasks in one sitting, and without the aid of other people or electronic devices. Lastly, Experiment 2 was conducted in temporal conjunction with Experiment 1, and all participants from the first experiment were excluded from participating in the present experiment.

3.2.2 Procedure

As mentioned above, the critical difference between Experiment 1 and Experiment 2 was that the category stimuli in the present experiment were modifications of Experiment 1 stimuli, with the addition of internal and external shape, as well as antennae. New "Blickets" and "Daxes" were also created for each of the three transfer tasks. As shown in Figure 3.2 (left panel), ambiguous transfer objects were those whose external shape (including any external patterns and antennae) was the same as that of a trained object, but whose internal shape differed in such a way that classification according to the rule yielded the opposite category label than classification according to perceptual similarity. Novel rule objects (i.e., items that were perceptually dissimilar to any objects previously presented to participants in the study; see Figure 3.2, middle panel) and memory objects (i.e., items whose external shapes, including any pattern and antennae, were the same as in trained stimuli; see Figure 3.2, right panel) were also presented to the participant.

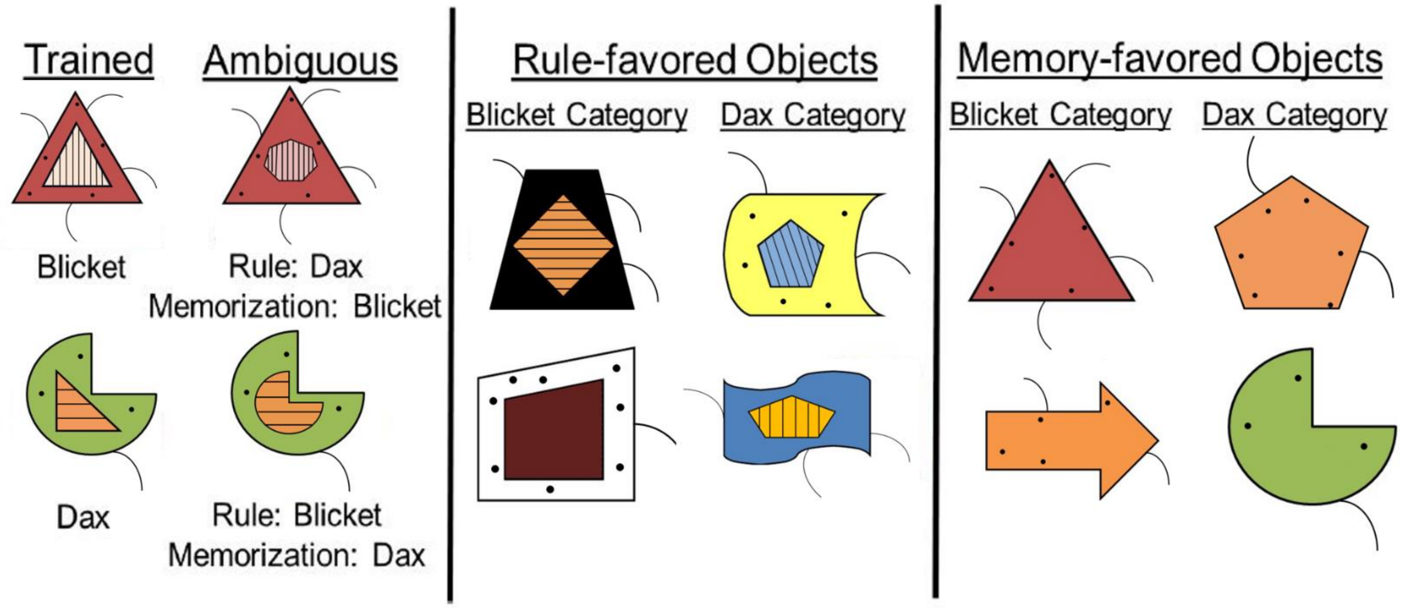

Figure 3.2. Left panel: Examples of training stimuli (left column) and their respective ambiguous transfer objects (right column). The ambiguous objects place rule-abstraction and memorization strategies in opposition to one another. Therefore, categorization of an ambiguous object according to the rule results in classification opposite that to which the perceptually similar training items belonged. Middle panel: Examples of novel objects used to assess rule-abstraction for categories of training objects. Right panel: Examples of novel objects used to assess memory for categories of training objects. In comparison to Experiment 1, these stimuli have been modified to include the nondiagnostic dimensions of internal stripe pattern, external dot pattern, and antennae.

All other procedural details, from task instructions and working memory tasks to training blocks and the post-transfer strategy probes, followed Experiment 1 exactly.

3.3 Results

Experiment 1 scoring procedures were used in the present experiment when calculating participants' performance on the backward digit span and the operation span tasks. Furthermore, the same heuristics as before were used to classify participants according to their rule reports on the global strategy probe questionnaire, as well as according to their block-by-block strategy reports (first-block, final-block, and average).

3.3.1 Working Memory Capacity

3.3.1.1 Backward Digit Span

There were no differences in performance on this task between those in the sequential presentation condition (M = 6.09, SD = 2.30) and those in the simultaneous condition (M = 6.40, SD = 2.15), $t(157.3) = -.88$, $p = .38$.

3.3.1.2 Operation Span

As with the backward digit span, there were no differences in performance between those in the sequential presentation condition (M = 48.23, SD = 16.88) and those in the simultaneous condition (M = 46.71, SD = 16.21), $t(157.74) = .57$, $p = .57$.

3.3.1.3 Composite Working Memory Measure

Having established that there were no differences in working memory task performance between conditions, we again calculated the correlation between the two working memory measures collapsed across presentation modes, and found it to be significant: $r = .47$, $t(158) = 6.44$, $p <$ .001. Thus, as before, we created a working memory composite score for each participant by averaging their standardized performance on both tasks (Figure 3.3). These composite scores were used in all future analyses involving working memory capacity.

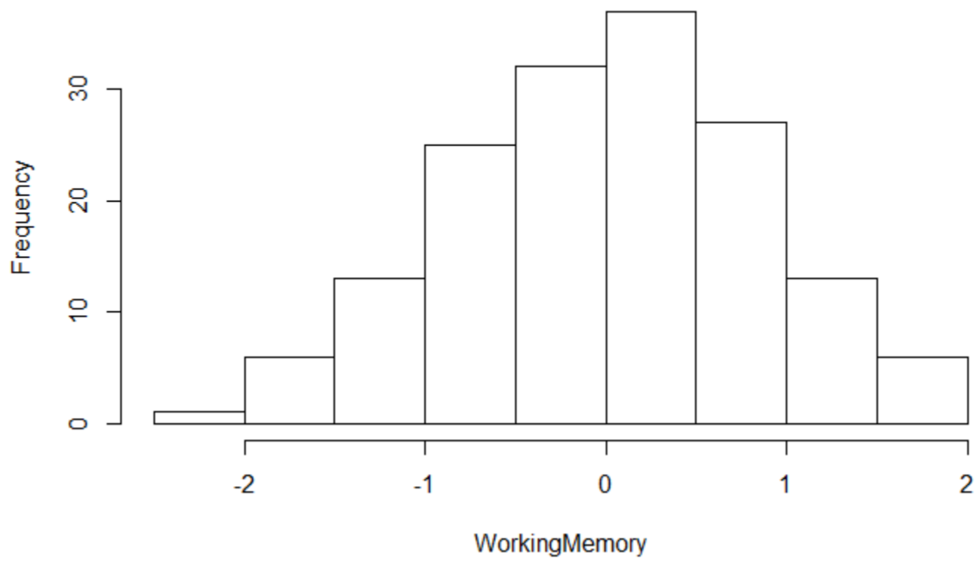

Figure 3.3 Histogram of working memory composite scores in Experiment 2.

3.3.2 Effects of Presentation Mode on Rule Acquisition and Transfer

We first coded participants' responses on the rule question of the global strategy questionnaire. Again, the author and a research assistant independently coded rule responses into the aforementioned categories of "full rule", "partial rule", and "incorrect rule". The two raters exhibited excellent initial agreement (97.5% of responses), and they resolved disagreements through discussion. The frequencies of participants classified according to their degree of rule acquisition are displayed in Table 3.1.

Table 3.1 Frequencies of Rule Acquisition in Experiment 2

	Degree of Rule Acquisition		
Condition	Full	Partial	Incorrect/None
Sequential	10	17	53
Simultaneous	7	10	63

We then conducted a 2 (presentation mode: sequential vs. simultaneous) x 3 (rule acquisition: full vs. partial vs. incorrect) chi-square test of independence, which again revealed a nonsignificant presentation mode x rule acquisition interaction, $\chi^2(2) = 3.21$, $p = .20$. As predicted, far fewer participants in the present experiment were able to correctly identify the full rule ($n = 17$), relative to in Experiment 1 ($n = 54$).

We then conditionalized participants' transfer performance on their degree of rule acquisition, and we conducted a no-intercept logistic mixed effect regression model to look for differences across condition and rule acquisition groups (Table 3.2). The model was composed of 19 fixed effects, 18 of which each corresponded a dummy-coded group (e.g., simultaneous presentation full rule learner categorizing memory-favored objects), and the last of which was our working memory composite. Overall, the model accounted for 35.94% of the variance (conditional $R^2 = .36$).

Table 3.2 Probabilities of Correct Categorization of Transfer Objects as a Function of Rule Acquisition in Experiment 2

		Degree of Rule Acquisition		
Object Type	Presentation Mode	Full	Partial	Incorrect/None
Ambiguous	Sequential	.88 [.83, .93]	.36 [.30, .43]	.25 [.21, .28]
	Simultaneous	.80 [.70, .88]	.39 [.31, .48]	.21 [.18, .24]
Rule-favored	Sequential	.99 [.98, 1.00]	.68 [.62, .75]	.51 [.47, .55]
	Simultaneous	.98 [.94, 1.00]	.70 [.62, .78]	.50 [.45, .52]
Memory-favored	Sequential	.84 [.78, .90]	.81 [.76, .87]	.80 [.77, .83]
	Simultaneous	.73 [.63, .82]	.85 [.78, .91]	.79 [.76, .82]

Note: Performance on ambiguous and rule-favored objects reflect probabilities of categorization according to the correct rule. For ambiguous objects, performance of 1.00 would indicate perfect categorization according to the correct rule, and performance of 0.00 would indicate perfect categorization according to memory for perceptually similar training objects. Performance on Memory-favored objects reflects probabilities of correct categorization based on perceptually similar trained items. Margins of error for 95% confidence intervals are displayed in brackets.

We then performed a series of linear combination tests of significance on the transfer data (see Figure 3.4). Collapsing across rule acquisition levels, participants in the simultaneous and sequential conditions did not differ in their odds of correctly classifying ambiguous, rule-favored, or memory-favored objects, all *p*s > .05.

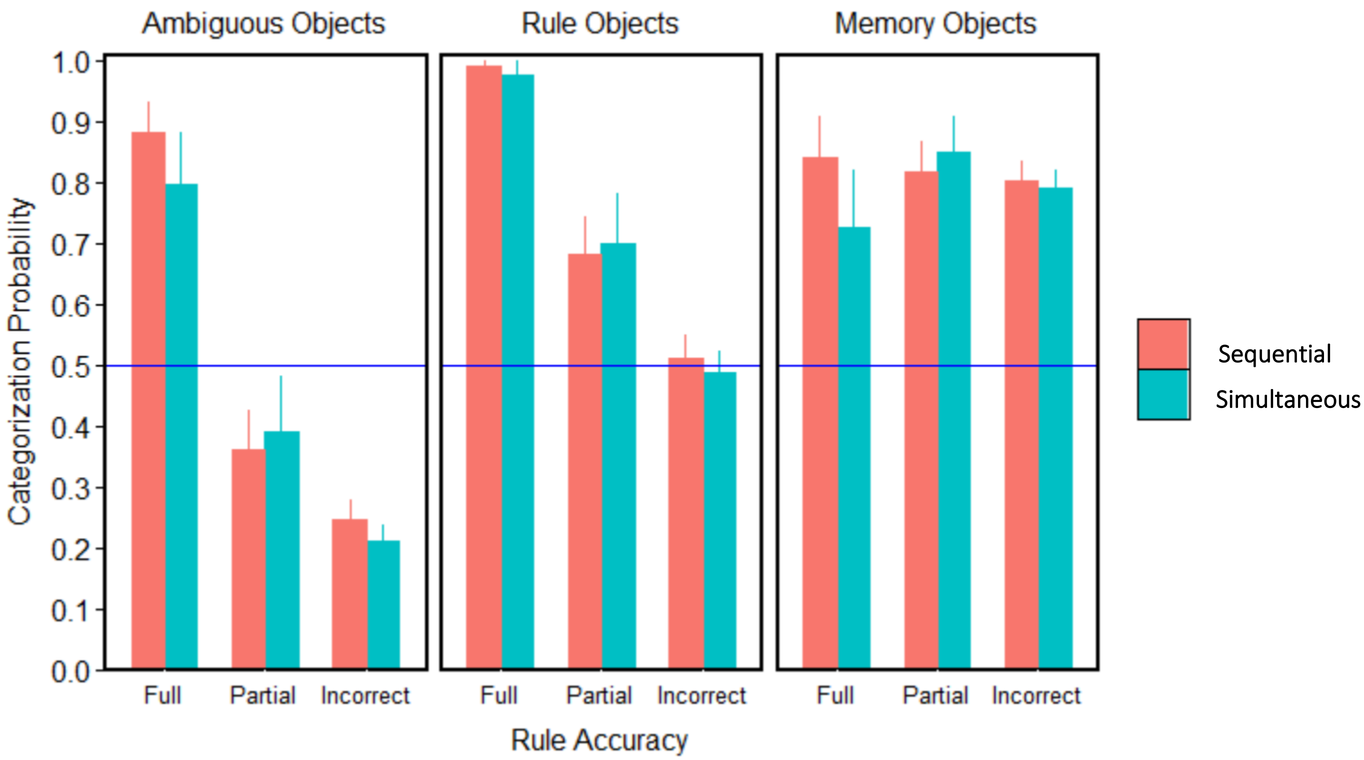

Figure 3.4. Probability of correct classification of ambiguous, rule-favored, and memory-favored transfer objects in Experiment 2, conditionalized on the degree of participants' rule acquisition (full, partial, and incorrect) and presentation mode condition (sequential in red, simultaneous in teal).

When going from the partial rule to the full rule group, the odds of categorizing ambiguous objects according to the rule increased by a factor of 87.20 (z = 8.98, p < .001), and the odds of an accurate response on the rule-favored transfer task increased by a factor of 1024.20 (z = 5.46, p < .001). The two groups did not differ in their odds of accurately categorizing the memory-favored objects (OR = .53, z = -1.24, p = .22).

Next, as a validity check for our scoring rubric, we determined whether participants in each of the rule acquisition groups differed in their categorization of the transfer objects. The odds of categorizing ambiguous objects according to the rule increased by a factor of 376.71 going from the incorrect rule to the full rule group (z = 13.35, p < .001). Similarly, the odds of an accurate response when categorizing rule-favored objects increased by a factor of 5383.98

between the two groups (z = 6.89 p < .001). There were no differences in their categorization of memory-favored objects (z = -.22, p = .83). Relative to the incorrect rule group, participants in the partial rule group had greater odds of categorizing more ambiguous objects according to the rule (OR = 4.32, z = 4.91, p < .001), and of having an accurate response when categorizing rule-favored (OR = 5.25, z = 5.52, p < .001) but not memory-favored objects (OR = 1.71, z = 1.48, p = .14).

The presentation mode x full rule versus partial rule group interaction was only marginally significant for memory-favored objects (OR = .39, z = -1.85, p = .06), but not significant for either ambiguous or rule-favored objects (both ps > .05). While participants in both conditions who acquired the partial rule performed similarly on the memory transfer task, full rule learners in the simultaneous condition had lower odds of an accurate response on the task. No other interactions were significant, all ps > .05.

3.3.2.1 Effects of Working Memory Capacity on Rule-based Transfer

We next investigated the relationship between working memory and participants' mean performance on the rule-based transfer tasks (where they categorized either ambiguous or rule-favored objects). We conducted a linear regression on mean ambiguous transfer performance, with presentation mode, working memory, and presentation mode x working memory included as factors. The model accounted for -.3% of the variance (adjusted R^2 = -.003). There were no significant effects, all ps > .05. We then conducted a linear regression on mean accuracy on the rule-favored objects, including the same factors as above in the model. This model accounted for 1.51% of the total variance (adjusted R^2 = .015). Again, there were no significant effects, all ps > .05.

3.3.2.2 Cross-Experimental Comparison of Transfer Performance

One of the goals of Experiment 2 was to create more complex stimuli (relative to Experiment 1) that would discourage participants from endorsing rule-based strategies. To verify whether Experiment 2 stimuli did indeed yield worse rule-based transfer performance, we conducted an independent samples t-test comparing mean performance on categorization of rule-favored objects across experiments. As predicted, participants exhibited significantly lower accuracy when categorizing these novel stimuli in the present experiment (M = .58, SD = .21), relative to Experiment 1 (M = .72, SD = .23), $t(317.07) = 5.50$, $p < .001$. Further in line with our predictions, a t-test comparing participants' accuracy on memory-favored objects across experiments (M_{E1} = .81, SD_{E1} = .17; M_{E2} = .80, SD_{E2} = .21) was not significant, $t(307.56) = .34$, p =.74.

3.3.3 Categorization Strategy Preferences during Training

Next, we addressed whether individual differences in strategy preferences would emerge in the present experiment (with its complex stimuli), or if participants in both presentation modes would be driven towards a single strategy. As in Experiment 1, we assessed learners' strategy preferences at three points of training.

3.3.3.1 First-block Strategy Preference

We classified participants' strategy preferences following the first training block. As shown in Table 3.3, the frequencies of participants classified according to this first-block method clearly reveal individual differences in strategy preferences. To determine if the two presentation mode conditions differed in their preferences, we computed a 2 (presentation mode: sequential vs. simultaneous) x 3 (first-block strategy: rule-first vs. memory-first vs. equal preference-first) chi-square test of independence, which revealed a marginally significant presentation mode x first-block strategy interaction, $\chi^2(2) = 4.74$, $p = .09$.

Table 3.3 Frequencies of First-block Strategy Preferences in Experiment 2

	First-block Strategy Preference		
Condition	Rule	Memory	Equal
Sequential	38	28	14
Simultaneous	26	41	13

We then conducted a generalized linear model (using a Poisson error distribution) to predict frequency by presentation mode (dummy coded using 0s and 1s) and first-block strategy (with group membership dummy coded using 0s and 1s). As before, we also included three presentation mode x first-block strategy interaction terms to create a fully saturated model that accounted for all variability in frequencies. Follow-up linear combination tests of significance revealed that the two presentation modes did not differ in their odds of endorsing a rule-first training strategy (OR = .94, z = -.25, p = .80). The odds of endorsing a memory-first strategy, however, increased by a factor of 1.67 (z = -2.03, p = .04) and the odds of endorsing an equal preference-start strategy decreased by a factor of 2 (z = -1.70, p = .09) going from the sequential presentation to the simultaneous presentation condition.

Lastly, we conducted a linear regression on working memory with rule-first, memory-first, and equal preference-first included as factors (dummy coded using 0s and 1s for group membership), collapsed across condition. The results of our subsequent linear combination tests of significance revealed no significant effects, all ps > .05.

3.3.3.2 Final-block Strategy Preference

In line with Experiment 1, we next classified individuals according to their strategy preferences following the final training block. The frequencies of participants according to this final-block method are displayed in Table 3.4, and they revealed a mix of strategy preferences amongst participants in both conditions.

Table 3.4 Frequencies of Final-block Strategy Preferences in Experiment 2

	Final-block Strategy Preference		
Condition	Rule	Memory	Equal
Sequential	28	25	27
Simultaneous	20	36	24

We then conducted a 2 (presentation mode: sequential vs. simultaneous) x 3 (final-block strategy: rule-final vs. memory-final vs. equal preference-final) chi-square test of independence. The presentation mode x final-block strategy preference interaction was not significant, $\chi^2(2) = 3.49$, $p = .17$. We also conducted a linear regression to determine if participants' final-block strategy preferences (with each group dummy coded using 0s and 1s) predicted their working memory ability. Again, linear combination tests revealed no significant effects, all ps > .05.

3.3.3.3 Average Strategy Preference

The frequencies of participants classified according to their average strategy are displayed in Table 3.5, again revealing individual differences in preference. We conducted a 2 (presentation mode: sequential vs. simultaneous) x 3 (average strategy: rule vs. memory vs. equal) chi-square test of independence. There was no significant presentation mode x average strategy interaction,

$\chi^2(2) = 1.05$, $p = .59$. Finally, we conducted a linear regression to help illuminate the relationship between participants' average strategy preferences and working memory ability. Follow-up linear combination tests revealed no significant effects, all *p*s > .05.

Table 3.5 Frequencies of Average Strategy Preferences in Experiment 2

	Average Strategy Preference		
Condition	Rule	Memory	Equal
Sequential	20	35	25
Simultaneous	19	30	31

3.3.4 Strategy Preference Dynamics during Training

Next, we calculated the total number of times each participant switched from one of the strategies (rule, memory, equal preference) to another on a consecutive block (see Figure 3.5). While participants in both conditions switched between strategies throughout training, there was no difference in the total number of shifts by participants in the sequential condition ($M = 3$, $SD = 2.43$) and simultaneous condition ($M = 3.51$, $SD = 2.42$), $t(157.98) = -1.33$, $p = .19$.

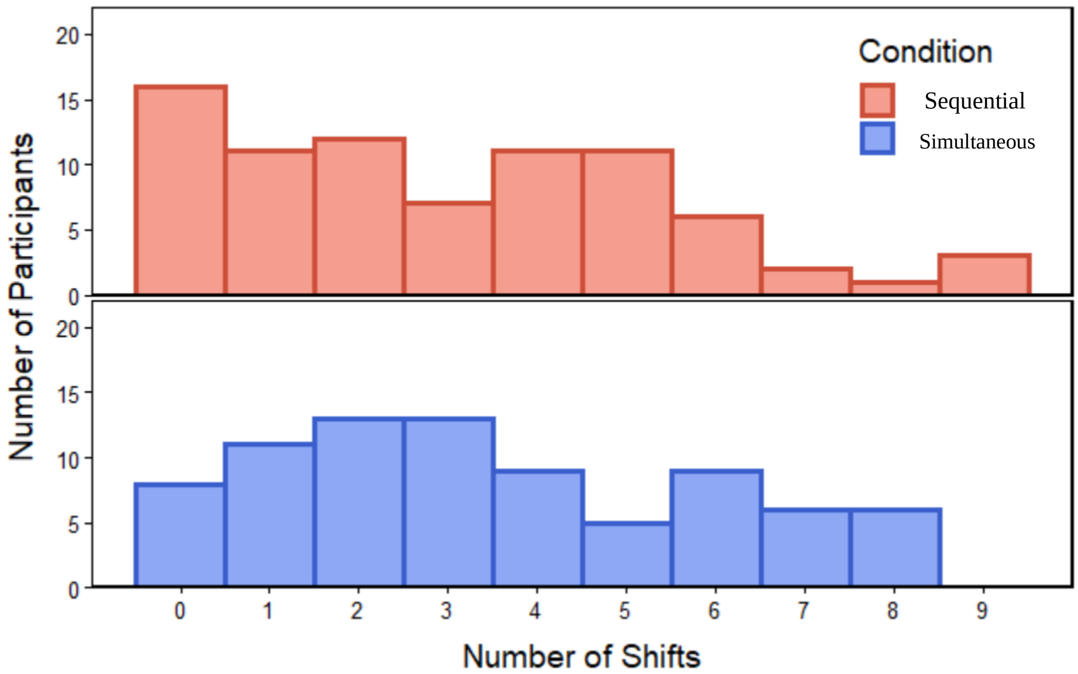

Figure 3.5 Histogram depicting the number of times, in Experiment 2, participants in the sequential presentation condition (top panel, in red) and the simultaneous condition (bottom panel, in blue) shifted from one strategy to another on consecutive blocks.

Collapsing across presentation modes, there were, as predicted, significantly more shifts in the present experiment (M = 3.26, SD = 2.42) relative to the first experiment (M = 2.54, SD = 2.45), $t(311.04) = -2.80$, $p = .005$. Lastly, the correlation between working memory and total number of shifts, collapsed across condition, was nonsignificant, $r = -.01$, $t(158) = -.19$, $p = .85$.

3.3.5 Effects of Participants' Strategy Preferences on Transfer

As was the case in Experiment 1, we were interested in whether the individual differences that emerged in participants' strategy preferences impacted their transfer performance. Given the prevalence of strategy shifting in this experiment – 85% of participants shifted at least once during training, and 43% shifted at least four times – we did not conditionalize transfer performance on participants' first-block strategy preferences. To decide between the other two classification methods, we conditionalized participants' transfer performance on both methods and conducted separate linear discriminant analyses. Conditionalizing transfer performance according to participants' average strategy preference was slightly more effective (canonical R^2 = .231) than when participants were classified into their final-block strategy groups (canonical R^2 = .198). Thus, we chose to conditionalize categorization transfer performance on participants' average strategy preferences[9] (Table 3.6).

Table 3.6 Probabilities of Correct Categorization of Transfer Objects as a Function of Average Strategy Preferences in Experiment 2

		Average Strategy Preference		
Object Type	Presentation Mode	Rule	Memory	Equal
Ambiguous	Sequential	.53 [.47, .59]	.24 [.20, .28]	.36 [.31, .42]
	Simultaneous	.43 [.37, .50]	.16 [.13, .20]	.31 [.26, .36]
Rule-favored	Sequential	.73 [.68, .79]	.54 [.50, .59]	.60 [.54, .65]
	Simultaneous	.68 [.61, .74]	.48 [.43, .53]	.56 [.52, .61]
Memory-favored	Sequential	.81 [.76, .86]	.87 [.83, .90]	.73 [.68, .78]
	Simultaneous	.75 [.69, .80]	.88 [.84, .91]	.74 [.70, .78]

Note: Performance on ambiguous and rule-favored objects reflect probabilities of categorization according to the correct rule. For ambiguous objects, performance of 1.00 would indicate perfect categorization according to the correct rule, and performance of 0.00 would indicate perfect categorization according to memory for perceptually similar training objects. Performance on Memory-favored objects reflects probabilities of correct categorization based on perceptually similar trained items. Margins of error for 95% confidence intervals are displayed in brackets.

[9] The interested reader can find analyses of participants' transfer performance conditionalized on their final-block strategies in Appendix B.

To this end, we conducted a no-intercept mixed-effects logistic regression model, which comprised 19 fixed effects. The first 18 fixed effects each represented a dummy-coded group (e.g., simultaneous presentation average rule learner categorizing memory-favored objects), and the final fixed effect was the covariate working memory. Overall, the model accounted for 32.99% of the total variance (conditional $R^2 = .33$).

Follow-up linear combination tests of significance were then conducted (Figure 3.6). Going from the memory average to the rule average group, the odds of categorizing ambiguous objects according the rule increased by a factor of 16.62 ($z = 7.41$, $p < .001$), the odds of an accurate response on the rule-favored task increased by a factor of 7.19 ($z = 5.30$, $p < .001$), and the odds of an accurate response when categorizing memory-favored objects decreased by a factor of 2.86 ($z = -2.53$, $p = .01$). Meanwhile, when comparing the equal average to the rule average group, participants in the rule average group had greater odds of categorizing ambiguous objects according to the rule ($OR = 3.84$, $z = 3.58$, $p < .001$), greater odds of an accurate response on the rule-favored task ($OR = 3.87$, $z = 3.54$, $p < .001$), and marginally lower odds of an accurate response on the memory-favored task ($OR = 2.09$, $z = 1.82$, $p = .07$). When going from the equal average to the memory average group, the odds of categorizing ambiguous objects to the rule decreased by a factor of 4.35 ($z = -4.29$, $p < .001$) and the odds of an accurate response on the rule-favored task decreased by a factor of 1.85 ($z = -4.29$, $p < .001$), but the odds of an accurate response on the memory-favored task increased by a factor of 5.97 ($z = 4.85$, $p < .001$).

Figure 3.6 Probability of correct categorization of ambiguous, rule-favored, and memory-favored transfer objects in Experiment 2, conditionalized on participants' average strategies (rule, memory, and equal preference) and presentation mode condition (sequential in red, simultaneous in teal).

The presentation mode x rule versus memory average strategy preference group interactions were not significant for ambiguous, rule-favored, or memory-favored objects, all *p*s > .05. The presentation mode x rule versus equal average preference interaction, however, was marginally significant for memory-favored objects ($OR = .51$, $z = -1.65$, $p = .10$), but not for the other two transfer object types, both *p*s > .05. Although participants in both presentation conditions showed similar accuracy when they had an equal average strategy preference, participants in the simultaneous rule average group had especially lower odds of an accurate response on the memory-favored task. The presentation mode x memory versus equal average strategy preference interactions were not significant for ambiguous, rule-favored, or memory-favored objects, all *p*s > .05.

3.4 Discussion

The primary objective of Experiment 2 was to extend the investigation of the relationship between presentation mode and category learning with the use of complex (multi-dimensional) stimuli. Prior research has suggested that increased stimulus complexity should lead to more benefit from and broader adoption of memory-based categorization strategies (Nosofsky et al., 2018; Regehr & Brooks, 1993). The complexity of the stimuli would render rule-learning more difficult, and the distinctiveness of each individual to-be-learned object would be expected to more memorable. When using simple stimuli (in Experiment 1), we found that simultaneous presentation improved memory-based transfer performance for participants in that condition, despite prevalent individual differences in their strategy preferences. Thus, we hypothesized that simultaneous presentation would confer the same benefits as before in terms of categorization performance, and that all participants in the experiment would preferentially endorse exemplar memorization as their strategy throughout training. Our findings, however, did not wholly support these predictions.

With respect to categorization performance when categorizing novel transfer objects, participants in this experiment fared worse than those in the previous one, as predicted. Underscoring the difficulty in hypothesis testing with complex stimuli (that contained three salient but irrelevant dimensions), only 10.63% of participants in the present study acquired the full rule, compared to 48.75% in Experiment 1. Moreover, when participants were grouped according to their degree of rule acquisition, there emerged no consistent condition differences in the categorization of novel transfer objects, suggesting that neither presentation mode was particularly beneficial for the learning of complex stimuli. These findings were in line with previous work that found no differences in novel categorization performance between sequential and simultaneous presentation modes (Andrews et al., 2011, Wahlheim et al., 2011).

As in Experiment 1, participants' patterns of self-reported block-level strategy preferences revealed mixed use of strategies throughout training, as opposed to endorsement of a single strategy. For example, following the initial training block, participants in both conditions exhibited similar frequencies of different strategy preferences to Experiment 1 (as shown by a visual comparison across Tables 3 and 9). Comparing across presentation modes, however, the trends were more aligned with our prediction. Unlike in Experiment 1, participants in the sequential condition did not show a bias for a rule-based strategy preference (relative to simultaneous presentation learners), but instead trended towards equal endorsement of rule- and memory-based strategy preferences. Furthermore, participants in the simultaneous presentation condition were significantly more likely to prefer exemplar memorization in the first training block. However, in further replication of Experiment 1, these first-block strategy preference differences did not persist through the final block of training. In other words, participants across the two presentation modes did not differ in their endorsement of the strategies, nor was there an overarching preference for exemplar memorization over rule abstraction. As previously mentioned, these findings directly contradict prior research, and strongly suggest that more research must be conducted to illuminate the circumstances in which participants are or are not directed towards an exemplar memorization preference when categorizing complex stimuli.

Finally, moving to strategy dynamics during training, participants in Experiment 2 switched strategies more often than their counterparts in Experiment 1. This was presumably because the complexity of the stimuli rendered them more difficult to learn through either rule- or exemplar-based strategy use, and so the frequency of switching increased even in the absence of direct feedback. Relatedly, participants’ increased propensity for switching strategies could help explain why their transfer performance was better captured by their average strategy preferences, rather than their final block strategy reports. After all, the former classification method took strategy switches throughout training into consideration, while the latter focused only on strategy preferences from one block. Together, these findings strongly suggest that future laboratory studies should consider the complexity of stimulus materials when developing measures used to capture learners’ strategy preferences.

Chapter 4: General Discussion

In the present study, we examined how sequential versus simultaneous presentation during training impacted participants' categorization of novel transfer objects (including degree of rule acquisition) and strategy preferences as they learned to categorize both simple (Experiment 1) and complex (Experiment 2) stimuli. The "Blickets" and "Daxes" used in both experiments followed a bi-dimensional disjunctive rule, where "Blickets" were objects whose inside and outside shapes shared either form or color, and "Daxes" shared neither (Little & McDaniel, 2015). We utilized Gouravajhala et al.'s (2019) block-by-block strategy probes to identify participants' strategy preferences and dynamics during training, and also incorporated a global strategy questionnaire following training to obtain participants' rules. Participants' categorization of novel ambiguous, rule-favored, and memory-favored transfer objects was first conditionalized on their degree of rule acquisition, and then on their block-level strategy preferences. Below, we address the key findings and implications of the present study.

4.1 Categorization of Novel Transfer Objects

The first major objective of the present study was to extend prior research on how presentation mode impacts the categorization of transfer objects. As outlined in the introduction, findings have been inconclusive thus far. For example, some researchers have found no difference between the two conditions in transfer performance (Andrews et al., 2011; Wahlheim et al., 2011). Others, however, have found a benefit of simultaneous presentation in the categorization of novel objects (Higgins & Ross, 2011; Kang & Pashler, 2012; Meagher et al., 2017). Rather than settling the debate, the present study also offered mixed findings.

When simple stimuli were used (Experiment 1), our data clearly suggested that the fully organized simultaneous display benefited exemplar-based category learning (i.e., categorization of novel exemplars in accordance with similarity to objects stored in memory), collapsed across all strategy preferences. Such a pattern vastly diverged from any of the previous findings, and so we developed several potential explanations for this discrepancy. As theorized in the introduction, it is possible that our specific design gave participants in the simultaneous condition more of an opportunity to self-regulate their allocation of time and attention for any given training object (relative to those in the sequential condition).

Another possibility is that prior research that has been assumed to show rule-favored benefits of simultaneous presentation was more aligned with exemplar-based learning than previously considered. After all, many researchers have not required participants to self-report their strategy use or utilized transfer items that objectively differentiated memorizers versus rule abstractors (as was the case with our ambiguous transfer task). In a related vein, sometimes the category learning stimuli used in studies can also complicate the issue. For example, Kang and Pashler (2012) used landscape paintings as their stimuli; because these paintings have no underlying category defining rule, it is difficult to confirm exactly how participants categorized novel objects during transfer.

A final possibility is that the design of our simultaneous display directly resulted in the discrepancy between our findings and previous research. Excepting Meagher et al. (2017), all previous studies comparing sequential and simultaneous presentation utilized only partial simultaneous presentations (i.e., presenting participants with only some of the to-be-trained stimuli at once). Thus, it is possible that the full simultaneous display of the present study was too large to allow for easy comparisons within and across categories. Instead, simultaneous presentation allowed participants to more efficiently and effectively (relative to sequential presentation) memorize training stimuli. Future research should manipulate the degree of simultaneous presentation and examine the effects on learners' categorization of novel ambiguous, rule-favored, and memory-favored objects.

When we utilized more complex stimuli (Experiment 2), there were no differences in categorization performance between those in the sequential and simultaneous presentation conditions. One possibility is that simultaneous presentation could not stimulate exemplar-based learning in this experiment because the stimuli contained too many dimensions for effective storage in long-term memory. In support of this view, previous work with similar findings (i.e., no condition differences in transfer performance) has utilized artificial organism-like aliens (Andrews et al., 2011) and naturalistic bird species (Wahlheim et al., 2011), both of which were highly complex. However, Higgins and Ross (2011) also used fictitious aliens as their stimuli and found that participants in their simultaneous presentation condition outperformed their sequential presentation counterparts on the transfer task, and so we caution against drawing strong conclusions.

Across all previous studies, the theorized benefits of simultaneous presentation focused on allowing participants to compare and contrast between exemplars from different categories (in an effort to identify diagnostic characteristics) without needing to hold details in working memory. To directly test this theory, we measured working memory ability using the backward digit span and the operation span in the present study. Interestingly, participants' working memory capacity was not related to their categorization performance in either presentation mode, whether simple or complex stimuli were used. Given that we also did not demonstrate a benefit of simultaneous presentation for rule-based learning, we do not find the lack of relationship between working memory capacity and transfer performance surprising.

At a broader level, our findings also contradict previous relevant research showing a positive relationship between working memory capacity and categorization performance on rule-based transfer tasks (DeCaro et al., 2008; Lewandowsky, 2011; Lewandowsky et al., 2012). Importantly, these previous studies not only presented objects sequentially, but they also incorporated feedback learning paradigms, in which participants were tested on their ability to categorize each to-be-learned stimulus before being presented with its correct category label. For participants abstracting a rule, they would be expected to update existing hypotheses and reorganize their category structures based on the feedback they received, both of which rely on working memory. In the present study, however, participants engaged in observational learning and to-be-learned stimuli (whether presented sequentially or simultaneously) were always accompanied by their category labels throughout training. Perhaps then the training format – feedback versus observational – impacted the relationship between working memory and rule-based transfer performance.

To address this question, we subdivided data in both experiments to only include participants with a rule-final preference in the sequential presentation condition, as they were direct counterparts of participants in the aforementioned studies. We then conducted separate linear regressions examining whether participants' working memory capacities predicted mean performance on the ambiguous transfer and the rule-favored transfer tasks in each experiment. We found no significant relationships between working memory capacity and rule-based transfer performance, all *p*s > .05, in either task or in either experiment. Thus, it appears as though, even when limiting our scope to include only rule participants who received sequential presentation, working memory did not positively predict rule-based transfer performance when using observational training. One potential explanation is that, in the absence of testing in the form of feedback learning, participants in the present study were less motivated to use their working memory capacities as often during training.

Another possible explanation for our findings – one that includes all participants, regardless of presentation mode condition or strategy preference – relates to the stimuli used in the present study. In support of this, Little and McDaniel (2015) also used “Blickets” and “Daxes” to test the relationship between working memory (as measured by performance on the operation span) and categorization performance. While they found that performance on a fluid intelligence task predicted categorization of ambiguous objects according to the rule (for participants who had a rule-based preference), there was no correlation between working memory capacity and transfer task performance. The present study extended these findings by uncovering the same lack of relationship even when using simultaneous presentation in an observational learning paradigm. More research using a variety of category learning stimuli should be conducted in the future to help illuminate these and other possibilities.

4.1.1 Presentation Mode and Rule Acquisition

A related major finding of the present study was that participants in the two presentation modes did not differ in the degree to which they acquired the correct bi-dimensional disjunctive rule. However, as previously mentioned, participants in the sequential condition exhibited better rule-favored transfer performance, at least when categorizing simple stimuli. Furthermore, a comparison between sequential and simultaneous presentation learners who acquired the full rule in Experiments 1 and 2 (see Tables 2.1 and 3.1) revealed a numerical trend favoring participants in the sequential presentation.

To determine whether our two experiments were simply underpowered to detect a significant difference in the frequency of the three different rule acquisition groups (full, partial, and incorrect) between the two presentation conditions, we collapsed across rule acquisition data from both experiments (see Table 4.1 for combined frequencies). A chi-square test of independence and subsequent log-linear model[10] revealed a marginally significant effect: following sequential presentation, participants were more likely to acquire the full rule. Participants who received sequential presentation were also more likely to acquire the partial rule. These findings revealed that presentation mode might in fact impact the degree to which participants are able to acquire a bi-dimensional disjunctive rule during 12 blocks of training, and thus should be taken into consideration in future studies. However, because the overall interaction was only marginally significant, we caution against drawing stronger conclusions regarding this relationship.

Table 4.1 Combined Frequencies of Rule Acquisition across Experiments 1 and 2

	Degree of Rule Acquisition		
Condition	Full	Partial	Incorrect/None
Sequential	44	31	88
Simultaneous	27	27	103

[10] We conducted a 2 (presentation mode: sequential vs. simultaneous) x 3 (rule acquisition: full vs. partial vs. incorrect) chi-square test of independence, which revealed a marginally significant presentation mode x rule acquisition interaction, $\chi^2(2) = 5.41$, $p = .07$. Follow-up linear combination tests of significance showed that there were significantly more sequential learners who acquired the full rule ($z = 3.17$, $p = .002$). Additionally, participants in the sequential condition were more likely to acquire the partial rule ($z = 2.92$, $p = .003$). There was no difference in the frequency of sequential versus simultaneous participants in the incorrect rule group, $p > .05$.

A final secondary analysis of interest in the present study was related to the memory probe, where participants were asked to classify how many training objects they had memorized (with “All,” “Some,” and “None,” serving as possible response options. The majority of participants in the present study (72.5% in Experiment 1 and 81.76% in Experiment 2) reported having memorized either some or all of the stimuli by the end of training. Interestingly, participants who acquired the full rule also reported having memorized some or all of the objects with high frequency (67.92% in Experiment 1 and 64.71% in Experiment 2). These findings align with theoretical frameworks of category learning which attest that rule abstractors simultaneously acquire memory for particular instances, as opposed to storing only the rule (Erikson & Kruschke, 1998, Experiment 2; Jacoby, 1991). Thus, it was unsurprising that even rule learners in the present study self-reported having a good memory for trained stimuli.

4.2 Presentation Mode and Categorization Strategy Preferences

The second major objective of the present study was to investigate whether presentation mode affected participants’ categorization strategy preferences, a question that no published study to date has addressed. While previous studies have identified the emergence of individual differences in strategy preferences in a sequentially presented rule-based task (Little & McDaniel, 2014), we were interested in whether simultaneous presentation of simple and complex stimuli would instead drive all participants toward one particular strategy.

A major contribution of the present study is that individual differences in strategy preferences emerged in both the sequential and simultaneous presentation conditions at various points of training (following the first block, following the final block, and on average across all blocks). Importantly, however, participants in the two presentation modes sometimes differed in their preference patterns.

Specifically, participants in the simultaneous presentation condition had greater odds of endorsing a memorization strategy following the first training block across both experiments, whereas those in the sequential presentation had greater odds of initially preferring a rule-based strategy (for simple stimuli) or a tendency to adopt both strategies equally (for complex stimuli). We believed that these strategy preference patterns largely emerged due to the salience of the relatively limited stimulus set size (12 training objects in total) for participants in the simultaneous condition. Prior research using "Blickets" and "Daxes" has trained younger adult participants on similar set sizes and found that participants were able to effectively use memorization as a categorization strategy (Little & McDaniel, 2015; Wahlheim et al., 2016). This outcome suggests that the 12-item training set in the present study was not too large for memorization to be a viable strategy and importantly, stimulus set size has impacted strategy preferences in previous studies, such that smaller set sizes were more conducive for exemplar memorization (Homa et al., 1981; Little & McDaniel, 2013). One might question why participants in the sequential condition did not exhibit such a memory-first bias across both experiments, considering that participants across both conditions were presented with the same number of items per block. To this, we note that it was unlikely that learners in the sequential presentation condition were counting the objects as they were presented; therefore, we would expect less susceptibility on their part to the effects of set size following the first training block. To check the plausibility of this explanation, we suggest that future experiments should, after the first block, ask participants in both conditions to report how many objects were presented to them.

Moving to the rest of training, there were no differences across presentation mode conditions in final-block or average strategy preferences, suggesting that other task- and individual-level factors influenced participants' strategy preferences. For example, prior to the category learning task, all participants were informed of the viability of both rule-based and memory-based strategies, which might have encouraged them to vary in their preferences throughout training. Furthermore, the repetitive nature of training (i.e., each stimulus was shown to the participant 12 times) might also have driven some participants to adopt different strategies over time. Another possibility is that while presentation mode might have moderated participants' initial task approach, they subsequently reverted to their general preference tendencies which led to the emergence of typical individual difference patterns by the end of training (Little & McDaniel, 2015; McDaniel et al., 2014; Wahlheim et al., 2016).

A secondary finding of the present study was that working memory capacity (as measured by performance on the backward digit span and operation span) was not related to participants' strategy preferences when categorizing simple or complex stimuli. Previous findings regarding on the relationship between working memory and strategy preferences have been mixed. For example, McDaniel et al. (2014) found that high working memory capacity (as measured by performance on the operation span) predicted a rule-based strategy preference on a function learning task, whereas Wahlheim et al. (2016) found that high working memory capacity (as measured by performance on the reading span and operation span) correlated with an exemplar memorization strategy preference in a rule-based category task similar to the one used in the present study. These studies used feedback learning paradigms that, as mentioned previously, might have brought participants' working memory ability to the fore more than observational training would be expected to. However, Little and McDaniel (2015) also found no relationship between performance on the operation span and participants' training strategy preferences using feedback learning on the same rule-based task as Wahlheim et al. (2016). Thus, more research is needed to help determine the conditions in which working memory might play a role in participants' strategy preferences.

Lastly, when conditionalizing transfer performance on their block-level strategy preferences (final-block in Experiment 1 and average block in Experiment 2), we found that participants' preferences generally aligned with how they categorized ambiguous, rule-favored, and memory-favored objects. These findings further validate Gouravajhala et al. (2019)'s block-by-block strategy probes. Interestingly, some presentation mode x strategy preference interactions emerged across both studies, all of which generally showing that simultaneous presentation improved memory for trained items, rather than facilitating rule use in transfer.

4.2.1 Categorization Strategy Preference Dynamics during Training

Prior research has shown that performance error during training precipitates strategy switching during rule-based category learning tasks using feedback learning (Gouravajhala et al., 2019). The present study extended this work by revealing that participants exhibit switching even in the absence of direct feedback, which was previously believed to be a necessary factor in inducing strategy switches (Kalish et al., 2005). Though participants switched less than in previous studies, our findings nevertheless suggest that participants have the capacity to monitor the effectiveness of their current strategy in their learning. Indeed, participants' tendency to switch more often when presented with more difficult and complex stimuli highlighted their metacognitive awareness of the extent to which they were learning the categories.

In further extension of previous work, we were interested in whether working memory capacity was an additional factor that could help explain participants' switching behavior. To this point, we found that working memory as not predictive of total number of strategy switches in either experiment. This lack of relationship was unsurprising in light of a previous finding with respect to working memory and strategy preferences. If there had been a significant correlation between working memory and a particular strategy preference, then perhaps participants with high working memory capacities would be expected to switch less often (as they would instead exhibit a stable strategy preference). Importantly, our study was the first, to our knowledge, to investigate this relationship, and so we caution against drawing strong conclusions that working memory might never be an important factor in participants' switching behavior. Instead, we encourage future researchers to incorporate block-by-block strategy probes (or other measures of strategy dynamics) to continue exploring mechanisms underlying strategy switching.

4.3 Educational Implications and Conclusion

As mentioned in the Introduction, the present study was partially motivated by a consideration of presentation mode in educational settings. Specifically, we aimed to determine whether simultaneous presentation (often utilized in textbooks and other classroom materials) would yield different category learning benefits in comparison to sequential presentation (often used in laboratory settings). Based on our findings of different first-block strategy preferences in the sequential and simultaneous conditions, we advise educators to be particularly mindful of the presentation mode of to-be-learned stimuli in lecture settings.

If professors wish for students to have improved memory for previously learned items, then a fully simultaneous display of to-be-learned stimuli would be expected to be more effective than sequential presentation. On the contrary, if educators' goal is for learners to abstract a rule, then a more traditional sequential presentation might be more beneficial. A mismatch in presentation mode and course aim might disadvantage students. For example, imagine a geology major presented with rock categories in a manual prior to a test where they must identify underlying rules (see Figure 4.1). If the to-be-learned rocks are presented simultaneously (with category labels accompanying each one), then the display might instead facilitate memory for the rocks, rather than supporting rule abstraction. Relatedly, it is also possible that the presentation of to-be-learned stimuli in educational settings might motivate different strategy preferences in students. After all, it is likely that most lectures will present learners with information only once or twice (as opposed to repeating the same information 12 times).

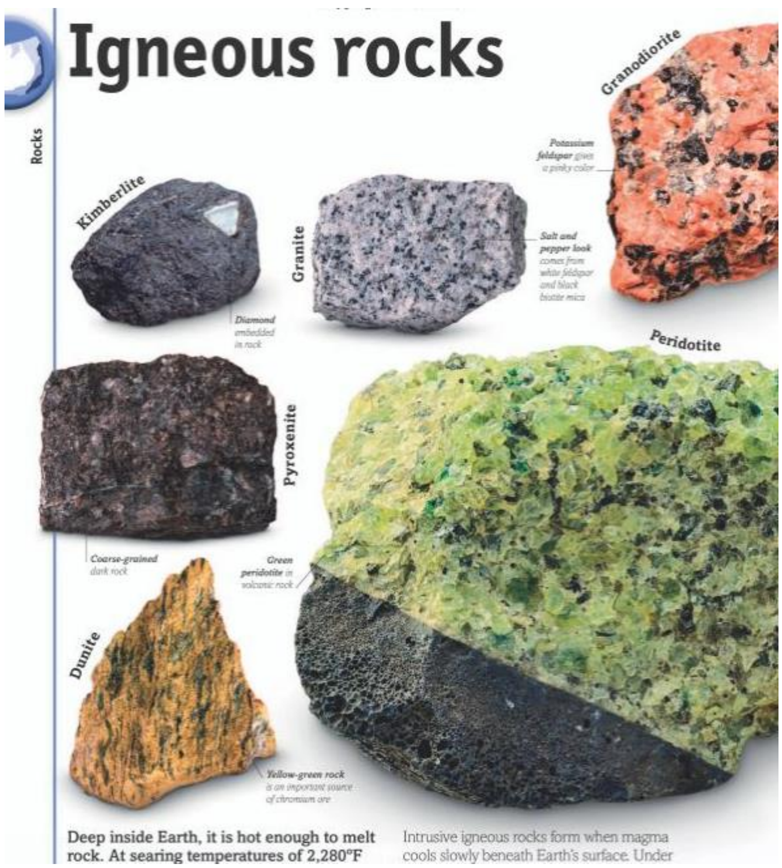

Figure 4.1 Examples of simultaneously presented igneous rocks in a geology manual. Taken from *The Rock and Gem Book: And Other Treasures of the Natural World*.

Much of prior category learning research has only utilized sequential presentation. Because much of category learning in the real world often diverges from this laboratory norm, we believed it worthy of further investigation. The present study was the first, to our knowledge, to examine the effects of simultaneous presentation (relative to sequential) on the learning of both simple and complex rule-based category stimuli. We focused on learners' categorization of three types of transfer objects, their acquisition of a bi-dimensional disjunctive rule, and their self-reported categorization strategy preferences and dynamics during observational training.

We identified several important areas of divergence between participants across the two presentation modes. When presented with simple rule-based stimuli, participants in the simultaneous condition categorized transfer objects according to memory for trained items, rather than according to an abstracted rule. Further, following the first training block, participants in this condition preferentially endorsed an exemplar memorization strategy, while those in the sequential presentation mode had greater odds of a rule abstraction preference. With more complex stimuli, though there were no differences in transfer performance, participants in the simultaneous condition again showed greater odds of a memory-based strategy preference after the first training block. Finally, when collapsing across both experiments, participants in the sequential presentation condition had greater odds of abstracting the correct bi-dimensional disjunctive rule. Thus, we strongly advocate for the incorporation of simultaneous displays into more future studies, so that we are better able to identify the underlying reasons for these differences.

References

1. Alfonso-Reese, L. A. (1996). Dynamics of category learning. *Unpublished doctoral dissertation.* University of California, Santa Barbara.

2. Anderson, J. R., & Betz, J. (2001). A hybrid model of categorization. *Psychonomic Bulletin & Review, 8,* 629-647. doi: 10.3758/BF03196200

3. Andrews, J. K., Livingston, K. R., & Kurtz, K. J. (2011). Category learning in the context of co-presented items. *Cognitive Processing, 12,* 161-175. doi: 10.1007/s10339-010-0377-5

4. Asbhy, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences, 5*, 204-210. doi: 10.1016/S1364-6613(00)01624-7

5. Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56,* 149-178. doi: 10.1146/annurev.psych.56.091103.070217

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

7. Ashby, F. G., Maddox, W. T., & Bohil, C. J. (2002). Observational versus feedback training in rule-based and information-integration category learning. *Memory & Cognition, 30*, 666-677. doi: 10.3758/BF03196423

8. Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear Mixed-Effects Models Using Eigen and S4. *R* Package 1.1-10.

9. Bourne, L. E., Jr. (1974). An inference model of conceptual rule learning. In R. L. Solso (Ed.), *Theories in Cognitive Psychology: The Loyola Symposium* (pp. 231-256). Potomac, MD: Erlbaum.

10. Bourne, L. E., Jr., Healy, A. F., Parker, J. T., & Rickard, T. C. (1999). The strategic basis for performance in binary classification tasks: Strategy choices and strategy transitions. *Journal of Memory and Language, 41*, 223-252. doi: 10.1006/jmla.1999.2647

11. Brigham, M. C., & Pressley, M. (1988). Cognitive monitoring and strategy choice in younger and older adults. *Psychology and Aging, 3,* 249–257. doi: 10.1037/0882-7974.3.3.249

12. Carvalho, P. F., & Goldstone, R. L. (2012). Category structure modulates interleaving and blocking advantage in inductive category acquisition. *Proceedings of the Annual Meeting of the Cognitive Science Society, 34,* 186-191.

13. Carvalho, P. F., & Goldstone, R. L. (2014). Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory & Cognition, 42,* 481-485. doi: 10.3758/s13421-013-0371-0

14. Craig, S., & Lewandowsky, S. (2012). Whichever way you choose to categorize, working memory helps you learn. *The Quarterly Journal of Experimental Psychology, 65,* 439-464. doi: 10.1080/17470218.2011.608854

15. DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition, 107*, 284-294. doi: 10.1016/j.cognition.2007.07.001

16. Erikson, M. A. (2008). Executive attention and task switching in category learning: Evidence for stimulus-dependent representation. *Memory & Cognition, 36*, 749-761. doi: 10.3758/mc.36.4.749

17. Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127*, 107-140. doi: 10.1037/0096-3445.127.2.107

18. Floyd, T. (2008). *Smithsonian Field Guide to the Birds of North America*. New York, NY: Collins.

19. Gao, Z., Xu, X., Chen, Z., Yin, J., Shen, M., & Shui, R. (2011). Contralateral delay activity tracks object identity information in visual short-term memory. *Brain Research, 1406,* 30–42. doi: 10.1016/j.brainres.2011.06.049

20. Green, D. (2016). *The Rock and Gem Book: And Other Treasures of the Natural World*. New York, New York: DK Children.

21. Gouravajhala, R., Wahlheim, C. N., & McDaniel, M. A. (2019). Individual and age differences in block-by-block dynamics of category learning strategies. *The Quarterly Journal of Experimental Psychology*, 73, 578-593. doi: 10.1177/1747021819892584

22. Higgins, E. J., & Ross, B. H. (2011). Comparisons in category learning: How best to compare for what. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1388–1393). Austin, TX: Cognitive Science Society.

23. Hoffmann, J. A., von Helverson, B., & Rieskamp, J. (2016). Similar task features shape judgment and categorization processes. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *42*, 1193-1127. doi: 10.1037/xlm0000241

24. Homa, D., Sterling, S., and Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory,* 7, 418-439. doi: 10.1037/0278-7393.7.6.418

25. Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513-541.

26. Johansen, M. K., & Palmeri, T. J. (2002). Are there representational shifts during category learning?, *Cognitive Psychology, 45*, 347 – 354. doi: 10.1016/S0010-0285(02)00505-4

27. Jones, M., Love, B. C., & Maddox, W. T. (2006). Recency effects as a window to generalization: Separating decisional and perceptual sequential effects in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 316-332. doi: 10.1037/0278-7393.32.3.316

28. Kalish, M. L., Lewandowsky, S., & Davies, M. (2005). Error-driven knowledge restructuring in categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 846-861. doi: 10.1037/0278-7393.31.5.846

29. Kang, S. H., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology, 26,* 97-103. doi: 10.1002/acp.1801

30.

31. Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science, 21*, 1551-1556. doi: 10.1177/0956797610385359

32. Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22-44. doi: 10.1037/0033-295X.99.1.22

33. Levine, M. (1975). *A cognitive theory of learning: Research on hypothesis testing*. Hillsdale, NJ: Erlbaum.

34. Lewandowsky, S. (2011). Working memory capacity and categorization: individual differences and modeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 720-738. doi: 10.1037/a0022639

35. Lewandowsky, S., Yang, L., Newell., B. R., & Kalish, M.L. (2012). Working memory does not dissociate between different perceptual categorization tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 881-904. doi: 10.1037/a0027298

36. Little, J. L., & McDaniel, M. A. (2015). Individual differences in category learning: Memorization versus rule abstraction. *Memory & Cognition*, *43*, 283-297. doi: 10.3758/s13421-014-0475-1

37. Luria, R., & Vogel, E. K. (2014). Come together, right now: dynamic overwriting of an object's history through common fate. *Journal of Cognitive Neuroscience, 26,* 1819–1826. doi: 10.1162/jocn_a_00584

38. Magen, H., & Emmanouil, T. A. (2019). Spatial organization in self-initiated visual working memory. *Frontiers in Psychology, 10,* 1 – 17. doi: 10.3389/fpsyg.2019.02734

39. 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*, 668 – 693. doi: doi.org/10.1037/a0032963

40. Meagher, B. J., Carvalho, P. F., Goldstone, R. L., & Nosofsky, R. M. (2017). Organized simultaneous displays facilitate learning of complex natural science categories. *Psychonomic Bulletin & Review, 24*, 1987-1994. doi: 10.3758/s13423-017-1251-6

41. Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238. doi: 10.1037/0033-295X.85.3.207

42. Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: the effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 775-799. doi: 10.1037/0278-7393.27.3.775

43. Murphy, K. R. (2003). *Validity Generalization: A Critical Review*. Mahwah, New Jersey: Erlbaum.

44. Nosofsky, R.M., Palmeri, T.J., & McKinley, S. (1994). Rule-plus-exception model of classification learning. *Psychological Review, 104*, 266-300. doi: 10.1037/0033-295X.101.1.53

45. Nosofsky, R. M., Sanders, C. A., Zhu, X., & McDaniel, M. A. (2018). Model-guided search for optimal natural-science-category training exemplars: A work in progress. *Psychonomic Bulletin & Review, 26,* 1-29. doi: 10.3758/s13423-018-1508-8

46. Regehr, G., & Brooks, L. R. (1993). Perceptual manifestations of an analytic structure: The priority of holistic individuation. *Journal of Experimental Psychology: General, 122*, 92-114. doi: 10.1037/0096-3445.122.1.92

47. Salatas, H., & Bourne, L. E. (1974). Learning conceptual rules: III. Processes contributing to rule difficulty. *Memory & Cognition, 2,* 549-553. doi: 10.3758/BF03196919

48. Sana, F., Yan, V. X., & Kim, J. A. (2017). Study sequence matters for the inductive learning of cognitive concepts. *Journal of Educational Psychology, 109*, 84-98. doi: 10.1037/edu0000119

49. Stewart, N., & Brown, G. D. A. (2004). Sequence effects in the categorization of tones varying in frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 416-430. doi: 10.1037/0278-7393.30.2.416

50. Trabasso, T., & Bower, G. H. (1968). *Attention in learning: Theory and research*. New York, NY: Wiley.

51. Vong, W. K., Hendrickson, A. T., Navarro, D. J., & Perfors, A. (2019). Do additional features help or hurt category learning? The curse of dimensionality in human learners. *Cognitive Science, 43,* e12724-e127749. doi: 10.1111/cogs.12724

52. Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011). Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory & Cognition, 39,* 750-763. doi: 10.3758/s13421-010-0063-y

53. Wahlheim, C.N., McDaniel, M. A., & Little, J. L. (2016). Category learning strategies in younger and older adults: Rule abstraction and memorization. *Psychology and Aging, 31*, 346-357. doi: 10.1037/pag0000083

54. Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences, 104,* 18766-18771. doi: 10.1073/pnas.0705618104

Appendix A

We conducted a no-intercept mixed-effects logistic regression model which comprised 19 fixed effects. The first 18 fixed effects each represented a dummy-coded group (e.g., simultaneous presentation average rule learner categorizing memory-favored objects), and the last fixed effect was working memory. Overall, the model accounted for 37.06% of the total variance (conditional $R^2 = .371$; see Table A.1).

Table A.1 Probabilities of Correct Categorization of Transfer Objects as a Function of Average Strategy Preferences in Experiment 1

		Average Strategy Preference		
Object Type	Presentation Mode	Rule	Memory	Equal
Ambiguous	Sequential	.72 [.68, .76]	.25 [.20, .31]	.40 [.33, .46]
	Simultaneous	.60 [.55, .65]	.29 [.23, .35]	.28 [.23, .33]
Rule-favored	Sequential	.85 [.82, .88]	.59 [.52, .65]	.67 [.61, .73]
	Simultaneous	.80 [.76, .84]	.56 [.50, .62]	.63 [.58, .69]
Memory-favored	Sequential	.78 [.75, .81]	.80 [.75, .86]	.73 [.67, .78]
	Simultaneous	.84 [.81, .88]	.83 [.50, .63]	.88 [.58, .69]

Follow-up linear combination tests of significance were conducted. Going from the average rule to the average memory group, the odds of categorizing ambiguous objects according to the rule increased by a factor of 45.65 ($z = 8.11$, $p < .001$), and the odds of an accurate response on the rule-favored task increased by a factor of 19.59 ($z = 6.30$, $p < .001$). There were no differences, however, between these groups in their categorization of memory-favored objects ($OR = 1.34$, $z = .58$, $p = .56$). The odds of categorizing ambiguous objects according to the rule was 23.71 times greater for the average rule group than for those who endorsed rule and memory strategies equally ($z = 7.12$, $p < .001$). Further, the odds of an accurate response when categorizing rule-favored objects were increased by a factor of 9.66 going from the average rule to the average equal preference group ($z = 4.99$, $p < .001$). The odds of accurate categorization on the memory-favored task, however, were the same across both groups ($OR = 1.37$, $z = .66$, $p = .51$). There were no performance differences between those with an average memory preference and those who, on average, endorsed rule and memory strategies equally, all *p*s > .05.

The presentation mode x rule versus memory average strategy group interaction was marginally significant only for ambiguous objects ($OR = .45$, $z = -1.66$, $p = .10$). While participants in both conditions categorized ambiguous objects similarly in the memory average group, sequential average rule learners had greater odds of categorizing ambiguous objects according to the rule than participants in the simultaneous average rule group. None of the presentation mode x rule versus equal average strategy preference group interactions were significant, all *p*s > .05. Lastly, the presentation mode x memory versus equal average strategy group interaction was marginally significant only for memory-favored objects ($OR = .39$, $z = -1.70$, $p = .09$). In this instance, average memory learners across sequential and simultaneous presentation modes exhibited similar performance when categorizing memory-favored objects, but equal average strategy participants in the simultaneous condition had greater odds of an accurate response compared to their sequential counterparts.

Appendix B

We conducted a no-intercept mixed-effects logistic regression model which was composed of 19 fixed effects. The first 18 fixed effects each represented a dummy-coded group (e.g., sequential presentation rule-final learner categorizing ambiguous objects), and the final fixed effect was working memory. Overall, the model accounted for 33.16% of the total variance (conditional R^2 = .33; see Table B.1).

Table B.1 Probabilities of Correct Categorization of Transfer Objects as a Function of Final-block Strategy Preferences in Experiment 2

		Final-block Strategy Preference		
Transfer Objects	Mode	Rule	Memory	Equal
Ambiguous	Sequential	.48 [.42, .53]	.24 [.19, .29]	.32 [.28, .38]
	Simultaneous	.50 [.43, .56]	.18 [.14, .21]	.27 [.22, .32]
Rule-favored	Sequential	.70 [.65, .75]	.53 [.48, .59]	.58 [.52, .63]
	Simultaneous	.65 [.59, .71]	.50 [.46, .55]	.56 [.50, .61]
Memory-favored	Sequential	.82 [.78, .86]	.87 [.83, .90]	.75 [.70, .79]
	Simultaneous	.69 [.63, .75]	.85 [.82, .89]	.79 [.74, .83]

We then conducted linear combination hypothesis tests of significance. Relative to memory-final participants, those in the rule-final group had 16.65 times greater odds of categorizing ambiguous objects according to the rule (z = -2.98, p = .003). Their odds of accurate categorization were 4.84 times greater on the rule-favored task (z = 4.29, p < .001), and 3.23

times lower on the memory-favored task ($z = -1.78$, $p = .07$). Going from the equal-final to the rule-final group, the odds of categorizing ambiguous objects according to the rule increased by a factor of 16.65 ($z = 7.61$, $p < .001$). Further, the odds of an accurate response on the rule-favored task increased by a factor of 4.84 ($z = 4.41$, $p < .001$), whereas the odds of an accurate response when categorizing memory-favored objects decreased by a factor of 3.23 ($z = -2.94$, $p = .003$). The odds of categorizing ambiguous objects according to the rule increased by a factor of 2.94, going from memory-final to equal-final groups ($z = -2.96$, $p = .003$). Both groups exhibited equal odds of correctly categorizing both rule-favored objects ($OR = .67$, $z = -1.16$, $p = .25$), but the memory-final group had 3.85 times greater odds of an accurate response when categorizing memory-favored objects ($z = 3.46$, $p < .001$).

The presentation mode x rule-final versus memory-final strategy preference group interaction was marginally significant for memory-favored objects ($OR = .48$, $z = -1.8$, $p = .07$), but not for ambiguous or rule-favored objects. In this case, participants in the sequential rule-final group had greater odds of an accurate response on the memory-favored task than did simultaneous rule-final participants, but participants in both conditions exhibited similar levels of accuracy in the memory-final group.

The presentation mode x rule-final versus equal preference-final group interaction was significant only for memory-favored objects ($OR = .33$, $z = -2.81$, $p = .005$). Though participants in both conditions had similar odds of correct categorization if they were in the equal performance-end group, sequential rule-final participants had greater odds relative to simultaneous rule-final learners. No other interactions were significant, all *p*s > .05.