EFFECTS OF SET-SIZE ON ABSTRACT CONCEPT LEARNING IN RATS USING MATCH/NON-MATCH TO SAMPLE PROCEDURES

Lucia Lazarowski

A Thesis Submitted to the University of North Carolina Wilmington in Partial Fulfillment of the Requirements for the Degree of Master of Arts

Department of Psychology
University North Carolina Wilmington
2010

Approved by

Advisory Committee

Mark Galizio Raymond C. Pitts

Kate Bruce
Chair

Accepted by

Dean, Graduate School
TABLE OF CONTENTS

ABSTRACT.................................................................................................................... iv
ACKNOWLEDGEMENTS................................................................................................. v
LIST OF TABLES ........................................................................................................... vi
LIST OF FIGURES .......................................................................................................... vii
INTRODUCTION ............................................................................................................. 1
METHOD ......................................................................................................................... 19
   Subjects ....................................................................................................................... 19
   Apparatus ............................................................................................................... 19
   Stimuli ...................................................................................................................... 22
   Procedure. .............................................................................................................. 24
   Data Analysis ....................................................................................................... 28
RESULTS ....................................................................................................................... 29
   Acquisition ........................................................................................................... 29
   Transfer Performance .......................................................................................... 35
DISCUSSION .................................................................................................................. 41
LITERATURE CITED .................................................................................................... 54
APPENDIX ..................................................................................................................... 61
ABSTRACT

Match (MTS) and Non-Match-to-sample (NMTS) procedures are used to assess concepts of identity and oddity across species and are measured by transfer performance to novel stimuli. Number of exemplars used in training (set-size) has been shown to affect learning. Larger set-sizes have been shown to promote concept learning in several species. Type of procedure (MTS vs. NMTS) may affect acquisition, with mixed findings on which procedure is learned faster. The present study explored the effects of set-size and procedure on concept learning in rats using olfactory stimuli. Rats were trained to either MTS (n=15) or NMTS (n=10) with 2 (n=17) or 10 (n=8) stimuli, and then tested for concept learning by presenting 10 novel stimuli. No difference was found in acquisition or transfer between MTS and NMTS, but rats trained with 10 stimuli performed better on novel transfer tests than rats trained with 2. When set-size was expanded from 2 to 10 and rats were re-tested with 10 novel stimuli, performance increased demonstrating that training with multiple exemplars facilitates learning.
ACKNOWLEDGEMENTS

I am extremely grateful to Dr. Kate Bruce, who has been an ideal and exceptional mentor, for her continual guidance, support, and kindness throughout this process. I am also greatly appreciative of committee member Dr. Mark Galizio, who has served as a supplementary mentor offering constant advice and direction. I am also very indebted to committee member Dr. Ray Pitts for his feedback on this work, and for his teaching and advice which led me to pursue a Masters degree in experimental psychology.

I would also like to thank the many members of the animal behavior lab for their integral help in the data collection process. Specifically I would like to thank Rachel Eure, Alex McLean, Aly Mack, Preston Stakias, Aaron Ward, Jackie Beyth, and especially Mallory Gleason and Adam Goodman for their ideas and contributions to this project.

Finally, this work would not be possible without the enduring support of my parents and role models, Eduardo and Alicia Lazarowski, who have always provided me with unconditional faith and encouragement.
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Olfactory stimuli used</td>
<td>23</td>
</tr>
<tr>
<td>2.</td>
<td>Number of sessions to each criterion</td>
<td>34</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Modified operant chamber.</td>
<td>21</td>
</tr>
<tr>
<td>2.</td>
<td>Individual data for subject P2 in the 10-stimuli, NMTS condition.</td>
<td>32</td>
</tr>
<tr>
<td>3.</td>
<td>Number of sessions before receiving the first transfer test for individual subjects as well as means.</td>
<td>33</td>
</tr>
<tr>
<td>4.</td>
<td>Percent correct on the first novel transfer test for each subject as a function of number of stimuli used in training.</td>
<td>37</td>
</tr>
<tr>
<td>5.</td>
<td>Percent correct on the second novel transfer test as a function of original number of stimuli used in training.</td>
<td>38</td>
</tr>
<tr>
<td>6.</td>
<td>Individual and group performance for all conditions on each transfer test.</td>
<td>40</td>
</tr>
</tbody>
</table>
INTRODUCTION

Concepts are formed based on past experiences and are fundamental to the ability to generalize what was learned to novel situations. The incapacity to do so would mean that every object and situation encountered would have to be individually processed (Roitbalt & von Fersen, 1992). The ability of an individual to behave conceptually is clearly adaptive; being capable of identifying patterns in everyday situations and applying these experiences in order to adapt to new conditions is surely critical to the survival of any organism (Cook, Kelly, & Katz, 2003; Kastak & Schusterman, 1994). Animals must make judgments about patterns in order to survive and reproduce, and doing so without some general model that may be applied to future instances would be inefficient. Instead, unique relations would have to be individually processed, which would be time-consuming and uneconomical (Lombardi, Fachinelli, & Delius, 1984).

Abstract concepts are based on relationships between stimuli (i.e., sameness or identity) instead of absolute features of stimuli; behavior is able to transcend the fixed features of a stimulus and instead rely on the relation among them (Katz, Wright, & Bodily, 2007; Mauck & Dehnhardt, 2005; Wright, Rivera, Katz, & Bachevalier, 2003). When an abstract concept is formed, behavior is said to be released from control by the particular stimulus and the previous reinforcement associated with it, becoming flexible and adapting to novel situations (Cook & Wasserman, 2006).

Abstract concepts are contrasted with natural (also known as perceptual) concepts which involve grouping items into categories that share specific physical features such as shapes and colors (Katz et al., 2007). Items in natural concepts are constrained by fixed stimulus properties and perceptual similarities (Katz, Wright, & Bachevalier, 2002). The basis of natural concepts is believed to rely on basic stimulus generalization, in which grouping is based on high perceptual
similarities between stimuli (Zentall, Wasserman, Lazareva, Thompson, & Ratterman, 2008). With such categorization, absolute features of the stimuli are what control responding and learning the relationship between stimuli is not necessary (Zentall, Galizio, & Critchfield, 2002).

Keller and Schoenfeld (1950) proposed a more operational definition which allows for the experimental analysis of conceptual behavior. They defined conceptual behavior as generalization within classes of stimuli and discrimination between classes of stimuli. Lazareva and Wasserman (2008) proposed that, in addition to responding similarly to members of one stimulus class and differently to members of another, a good definition of conceptual behavior, or relational stimulus control, should also include transfer of accurate responding to novel members of the class.

Historically, researchers believed that non-human animals are not capable of more sophisticated cognitive tasks and that abstract concepts in particular are unique to humans. This assumption was largely due to the close link between the ability to form concepts and language acquisition (Roitblat & von Fersen, 1992). However, recent evidence supporting abstract concept learning has been demonstrated in a variety of species such as dolphins (Herman, Hovancik, Gory, & Bradshaw, 1989), sea lions (Kastak & Schusterman, 1994), monkeys (D'Amato, Salmon & Colombo, 1985; Wright et al., 2003), pigeons (Bodily, Katz & Wright, 2008; Wright, Cook, Rivera, Sands, & Delius, 1988), rats (Peña, Pitts, & Galizio, 2006) and honeybees (Giurfa, Zhang, & Jenett, 2001).

Lazareva and Wasserman (2008) discussed different types of abstract concept learning that have been demonstrated in animals such as the concepts of smaller/larger, numerosity, and sameness/difference, with perhaps the most widely studied concept in non-humans being that of sameness/difference or an “identity concept.” The capability to discriminate whether two things
are same or different plays an important part in the development of abstract cognitive thinking (Wright & Katz, 2006). Moreover, the ability to form an identity concept is thought to be valuable because the capacity to judge relations of sameness and difference may be at the core of such abilities as language acquisition, mathematics, analogical reasoning, social relationships, and art (Cook et al., 2003).

The experimental paradigms most commonly used to test the concept of identity are the matching-to-sample (MTS) and same/different discrimination tasks (Lazareva & Wasserman, 2008). MTS procedures are conditional-discrimination tasks that typically involve the presentation of a sample stimulus followed by the presentation of two or more comparison stimuli in which responses to the correct comparison are reinforced. In MTS procedures, a correct response is a response to the comparison stimulus that is physically identical to the sample (Lazareva & Wasserman, 2008). Variations of the procedure include non-matching (also known as oddity from sample or oddity matching) in which responses to the comparison that is different than the sample are reinforced and arbitrary matching-to-sample in which responses to specific comparison stimuli arbitrarily designated correct by the experimenter are reinforced.

MTS procedures have been widely adapted for use with a variety of species by manipulating the procedure and stimulus modality to better suit the species in question. For example, Herman and Gordon (1974) adapted a MTS procedure for dolphins in which the dolphins were trained to MTS using auditory stimuli projected by underwater speakers. Later, Herman et al. (1989) incorporated visual stimuli along with auditory commands to form a MTS paradigm. A dolphin was trained to match auditory commands with corresponding objects (arbitrary matching). A sample object was shown for 3 s followed by an underwater sound indicating a previously trained action to be performed on the corresponding object. Two objects
were then displayed, one matching the sample object. The dolphin was required to swim to the comparison object matching the sample and perform the previously indicated action. Each correct action performed on the correct object was reinforced with fish and social interaction. High accuracy was sustained, leading the authors to conclude that dolphins of capable of MTS with auditory as well as visual stimuli. Similarly, sea lions were able to match with high accuracy when presented with visual stimuli (Kastak & Schusterman, 1994). Two female California sea lions were trained to MTS with visual stimuli consisting of black shapes painted on white backgrounds on wooden panels. A sample image in the center panel was shown for 4 seconds after which the two side doors opened revealing the comparison stimuli. After a correct response of a nose poke to the stimulus matching the sample, the subject received a reinforcer (fish). Sea lions, too, were able to maintain high performance throughout the study. Pigeons and monkeys have also excelled in visual MTS procedures (Bodily et al., 2008; Katz et al., 2002). In these procedures stimuli such as different colored lights or images are presented and the animal is required to peck or touch the sample (called an observing response). This response produces the two comparison stimuli, and a peck or touch to the comparison matching the sample is reinforced. These studies illustrate the adaptability of the MTS paradigm and the diversity of species capable of excelling at the basic task.

However, matching with high accuracy is not sufficient to conclude that behavior is under relational stimulus control, or that an abstract concept such as identity has been learned, because other sources of stimulus control are possible (Peña et al., 2006). Specifically, Carter and Werner (1978) proposed that MTS behavior in pigeons may come under the control of a number of different stimulus-response relations. They described three forms of stimulus control for conditional-discrimination learning that pigeons may be using to MTS that result in either
stimulus-bound or relational learning. The configuration model states that responding may come under the control of specific stimulus arrangements; pigeons learn specific responses to different configurations of stimulus presentations. That is, the visual display of all the stimuli together forms one whole stimulus, and specific responses to each arrangement are learned (Katz, Bodily, & Wright, 2008). For example, a pigeon trained to MTS with different colored keys may learn each combination of the key arrangements separately, and learn the appropriate response for each combination. For instance, a pigeon may learn to peck the left key when presented with an array of red-red-green; or, peck the right key when presented with red-green-green. Such stimulus control requires no relational identity learning of the sample and comparisons.

Alternatively, or perhaps additionally, Carter and Werner (1978) proposed that pigeons may learn multiple rule models or “if-then” rules. In this case, specific stimulus-response chains are learned between a sample and the corresponding stimulus (Katz et al., 2008). For example, a pigeon may learn to respond to a blue light when presented with a blue sample, and respond to a red light when presented with a red sample. However this type of associative learning is connected the specific stimuli and accuracy may not withstand when novel colors are presented.

The configuration and response-chain alternatives do not indicate abstract concept learning because they do not go beyond the training stimuli and do not involve relational stimulus control. Thus, to assess whether the identity relation between sample and comparison is what controls responding, novel stimuli must be used (Zentall et al., 2008). Testing for transfer of performance with novel stimuli is necessary to gauge whether behavior is independent of prior reinforcement history with those specific stimuli (Katz et al., 2007). So the third model of conditional discrimination proposed by Carter and Werner (1978), the single rule model/concept learning/relational stimulus control, states that subjects will continue to respond correctly in a
new situation as long as the new problem fits the rule that has been learned; that is, behavior is under the control of the relationship between the stimuli and not bound to the specific items.

In light of varied definitions for abstract concept learning used in the past, Katz et al. (2007) proposed specific criteria to rule out inconclusive results. Along with the requirement that transfer test stimuli be novel, the novel stimuli should be different enough from the training stimuli to rule out simple stimulus generalization based on similarity. Moreover, an adequate test for generalized identity matching should employ novel stimuli as both samples and comparisons. If a novel stimulus is presented as a comparison in combination with a familiar stimulus on a test trial, previous history of stimulus presentations may control responding by exclusion of the familiar or novel stimulus instead of the identity relation (Dube, McIlvane, & Green, 1992; Katz et al., 2007). Further, Katz et al. advised that novel stimuli should not be repeated due to possible re-learning within a session. Lastly, Katz at al. proposed that behavior should only be considered full abstract concept learning when transfer performance is equal to or better than baseline (training before transfer test) performance. If that which is learned in training transfers to a new circumstance, that is, high accuracy is maintained with novel stimuli and these criteria are satisfied, it can be concluded that the stimuli have come under control of their relation to one another and what was learned in training has generalized to the novel stimuli.

A variety of species have demonstrated the ability to learn identity relations. In the Kastak and Schusterman (1994) example, transfer tests were conducted to assess stimulus control of the identity relation. In these transfer tests, trials consisting of all novel stimuli were introduced into baseline sessions. First-trial performance, when stimuli were presented for the first time, was analyzed for purposes of generalized identity matching assessment. One of the sea lions performed significantly better than chance when novel stimuli were presented (80% or
better), while the other sea lion averaged 70% correct with novel stimuli, which was just slightly below significance. Thus, Kastak and Schusterman concluded that what was learned in MTS training had transferred to the novel problems.

Similarly, Katz et al. (2002) demonstrated that rhesus monkeys were capable of demonstrating identity abstract concept learning. Nine rhesus monkeys were trained on an 8-item same/different (S/D) procedure, a procedure comparable to MTS that allows for testing of the identity relation. The monkeys were shown digitized pictures on a color monitor and trained to correctly identify pairs of stimuli as same or different. A trial began with the presentation of two pictures displayed vertically with a white rectangle adjacent to the pictures. If the two pictures were identical, touching the lower picture produced reinforcement. If the two pictures were different, touching the white rectangle resulted in reinforcement. Upon acquisition of the S/D task, the monkeys were tested for transfer with 10 novel stimuli in which the novel stimuli were introduced into a baseline session. Thereafter, 8 novel stimuli were added each time a criterion of 85% correct or better for three sessions was met. The rhesus monkeys eventually transferred performance to novel stimuli with accuracy equal to or greater than baseline, although extensive training (including increasing the number of stimuli used in training, which will be further discussed below) was required. These results were duplicated under comparable experimental procedures by Wright et al. (2003) with capuchin monkeys, showing that both old and new world monkeys are capable of abstract concept learning.

For many years there was much debate over whether or not pigeons were capable of abstract concept learning with several studies reporting conflicting evidence of relational stimulus control in pigeons. Katz et al. (2007) suggested that the Wright et al. (1988) study was the first to demonstrate relational concept learning in pigeons that satisfied the Katz et al.
suggested criteria. Wright et al. (1988) trained two groups of pigeons to MTS with visual stimuli. One group was trained with a baseline of 152 different stimuli, while the other group was trained with a baseline of just two stimuli. For the two-stimulus group, the stimuli were repeated on every trial, while the 152-stimuli group received sessions in which stimuli were trial unique (stimuli were only used once and not repeated on any trials thereafter). After meeting a criterion of 75% correct or better, both groups received transfer tests with novel stimuli. The two-stimulus group performed at chance levels on the transfer test, while the 152-stimuli group performed as well as they had with the baseline stimuli, thus satisfying the Katz et al. (2007) criterion for abstract concept learning.

As previously mentioned, the modality of the stimulus is crucial to the ability of an animal to acquire MTS and transfer performance to novel stimuli. Oden, Thompson, and Premack (1988) noted that when auditory instead of visual stimuli were used with monkeys, their ability to demonstrate generalized matching was “fragile at best.” Van Hest and Steckler (1995) suggested that for different species, stimuli may vary in salience according to modality. This may be why studies of visual MTS in rats have failed to show generalized identity matching (Iversen, 1993; Iversen, 1997). In one experiment, Iversen (1993) trained three rats to MTS with visual stimuli. Rats were tested in an operant chamber and trained to make a nose poke to illuminated keys. Stimuli used were keys that projected either a steady white light or a blinking light. The sample could be either a blinking or steady light and the comparison keys were always one steady light and one blinking light. A nose poke to a key corresponding to the sample stimulus in the center produced two comparison stimuli on either side of the sample. A nose poke to the comparison key that matched the sample key then produced a food pellet. After at least 45 sessions, all three rats acquired MTS with a stable performance of 90% correct or better. A zero-
delay procedure was then introduced in which the sample was turned off when the comparisons appeared. Performance deteriorated when the rats were shifted to the zero-delay procedure. Iversen (1993) proposed that the rats were merely acquiring a discrimination of specific stimuli and their position (configural learning), leading him to conclude that the results from this study added to the growing evidence that pigeons, monkeys, and rats do not learn a “matching concept,” but instead learn to discriminate separate stimulus configurations, as proposed by Carter and Werner (1978).

In a follow-up study, Iversen (1997) aimed to determine whether responding was simply under the control of the spatial location of the stimuli and not the relation between the sample and the comparison stimuli. The procedure was the same as the 1993 study except that the sample could appear in the center or on either of the comparison keys. Performance deteriorated and the rats did not accurately match when the same stimuli appeared in new locations. Iversen (1997) concluded that the performance could not be described as MTS based only on the identity relation between sample and comparisons, but that the physical properties of the stimuli and their spatial locations were controlling behavior.

However when olfactory rather than visual stimuli are used, rats perform with much higher accuracy (Dudchenko, Wood, & Eichenbaum, 2000; Lu, Slotnick, & Silberberg, 1993; Slotnick, 2001). Slotnick (2001) discusses the importance of selecting a stimulus modality that is appropriate for the species being studied, and reviewed a variety of studies demonstrating that rats in particular can excel at a number of cognitive tasks when olfactory stimuli are used. For example, Lu et al. (1993) trained rats to MTS with odors selected from pure chemicals, foods, and other commercial products using an olfactometer. Rats were trained on a go/no go procedure in which a sample was presented followed by the presentation of either the correct or incorrect
comparison. A correct response (“go”), defined as a lick to the odor identical to the sample, was reinforced with access to water. Not responding (“no go”) when an odor different from the sample was presented was scored as a correct rejection. High accuracy in the rats’ performance led Lu et al. (1993) to conclude that rats can readily learn an olfactory MTS task. However, first trial performance was not presented and so it is not clear whether performance transferred to the novel stimuli or instead rapid learning of the new configurations occurred.

Dudchenko et al. (2000) studied the contributions of the hippocampus to memory capacity in rats using odor and location span tasks. Rats were trained to respond by digging into cups of scented sand. Plain sand was mixed with .5 gm of different household spices to create olfactory stimuli in which the rats dug to receive a reinforcer. In the first phase of the experiment, the odor span task, rats were introduced to novel odors and reinforced for responding to the novel stimulus. On the first trial, rats were presented with one cup containing scented sand and digging in the cup produced reinforcement. On the second trial, a new cup with a novel odor was added, along with the cup from the first trial, and digging only in the novel cup was reinforced. On the third trial, the first two odors remained and a third odor was added, and only responses to the novel cup were rewarded. This continued until there were 24 odors in the set. Dudchenko et al. (2000) found that the rats were able to perform without error up to a span of 24 odors, demonstrating that rats can excel at a variety of olfactory tasks.

Peña et al. (2006) adapted the Dudchenko et al. (2000) procedure to assess identity concept learning rats using MTS procedures and olfactory stimuli (household spices). In a modified operant chamber, trays were inserted to reveal cups of scented sand. Once the rat dug in the sample to retrieve a sugar pellet (while being exposed to the scent of the stimulus), the tray was further inserted to reveal the two comparison cups, one that contained the scent identical to
the sample and one that contained a different scent. Reinforcement was provided for digging in the cup matching the sample. Training began with only three stimuli and occurred until a criterion of 90% correct or better on two consecutive sessions was met. At this point, two novel stimuli were added to the training set and abstract concept learning was assessed by determining performance on trials in which the novel stimuli appeared for the first time. Two novel stimuli were subsequently added to the baseline every time criterion was reached. Criterion for generalized matching was defined as correctly matching on at least five out of six consecutive novel test trials. Rats rapidly acquired high levels of accurate matching, and three of the four subjects transferred performance and met criteria for generalized matching, supporting the interpretation that responding was under control of the identity relation between sample and comparison. Additionally, several controls were used to ensure that the stimulus relation was controlling responding rather than other factors. To control for pellet scent detection, for example, pellets were placed into both the correct and incorrect cups in control trials.

The number of different stimuli that the subject encounters during training is also related to acquisition of relational stimulus control. When trained with a small number of stimuli, generalized MTS generally does not emerge because responding tends to come under the control of specific features of the training stimuli, as described by Carter and Werner (1978). In the Peña et al. (2006) study, as the experiment progressed and more novel stimuli were added, fewer trials were required to meet generalized matching criteria. The authors concluded that these results provided evidence that the control by identity relation was facilitated by training with large numbers of stimuli. Oden, Thompson, and Premack (1990) found that after training on a same/different procedure using only six objects, chimpanzees failed to respond correctly when presented with novel exemplars. The authors suggested that training with trial-unique stimuli
would prevent stimulus-bound learning such as response-chain or configurational learning, which is what presumably occurred with the chimpanzees. Kastak and Schusterman (1994) also discussed the idea that the large number of exemplars used in their study probably influenced the strong performance by their subjects.

Multiple exemplars, or multiple examples of the rule, have recently been shown to have an effect on abstract concept learning in several species. Studies with set-size (number of exemplars or different baseline stimuli used in training) have shown that a larger set-size facilitates higher accuracy in performance on tests with novel stimuli in several species. Both Katz et al. (2002) and Wright et al. (2003) found that a larger set-size promoted abstract concept learning in rhesus and capuchin monkeys. As previously described, S/D testing began with an 8-stimuli set of digitized pictures. Following transfer testing, the 8-item set was doubled to 16 stimuli, then 32, 64, and 128. With a training set-size of 8 stimuli, there was little to no transfer performance to novel stimuli. As the set-size increased, transfer performance increased and after the training set-size had reached a sufficiently large size, full abstract-concept learning was seen. Mean transfer performance rose from 52% correct at a set-size of 8 to 87% correct at a set-size of 128.

These results were duplicated by Katz and Wright (2006) using similar stimuli and procedures with pigeons. Pigeons were trained on the same/different task beginning with an 8-item set-size that was subsequently expanded. For pigeons, 256 different pictures were required before transfer to novel stimuli was seen. Similarly, Bodily et al. (2008) tested the effects of systematically increasing the number of training stimuli in a MTS task with pigeons. Five male pigeons were trained to MTS with color, computer-drawn cartoon images. A peck to the sample on the screen produced two comparison images, and a peck to the one matching the sample
produced reinforcement. Training began with a stimulus set-size of three images repeated over 96 trials in each session. After a performance criterion of 85% correct on one session was met, transfer tests were introduced. Twelve transfer trials in which novel stimuli were introduced were quasi-randomly inserted into a regular training session replacing 12 training trials. After each transfer test, the set-size was systematically increased (3 to 6, 12, 24, 48, 96, 192, 384, and 768). As the training set-size increased, performance on novel transfer tests also increased. Transfer performance was 55% correct at a set-size of three, and increased to 82.5% correct at a set-size of 768.

However, Nakamura, Wright, Katz, and Bodily (2009) found that if testing begins with a sufficiently large set-size, pigeons may not require as large of a set-size before showing transfer to novel stimuli as previously shown. Using the same procedures at Katz and Wright (2006), Nakamura et al. (2009) trained pigeons with an initial 64-item set in order to compare transfer performance to the Katz and Wright pigeons that began training with only 8 items. Nakamura et al. found that a 64-item set was sufficient for novel transfer when training began with 64 items, as opposed to pigeons that began with an 8-item set and did not show transfer when the set-size was increased to 64. Thus, rhesus and capuchin monkeys may not require as many as 128 stimuli as suggested by Katz et al. (2002) and Wright et al. (2003) if training begins with a sufficiently large set.

According to Oden et al. (1988), much of the previous literature suggests that most organisms are predisposed to attend to absolute stimulus features, but that when the environment is structured so that these features are not as salient (i.e., when a large training set is employed), some species will learn to match using relational features. van Hest and Steckler (1995) proposed that when a large number of stimuli are used, recency effects are reduced due to minimal
exposure to each stimulus. Alternatively, when a limited number of stimuli are used, the subject must discriminate the current stimulus presentation on the given trial from previously repeated presentations of the same stimulus. When more stimuli are used, each stimulus presentation occurs less often, thus lessening such proactive interference within the session (van Hest & Steckler, 1995). While it may be simple to learn about individual properties of a small number of stimuli, this becomes impractical as the number of stimuli become too great and makes item-specific learning, as opposed to learning a more general concept, an uneconomic strategy (Lazareva & Wasserman, 2008; Lombardi et al., 1984; Wright & Katz, 2007). The number of stimuli used in training (set-size) is an important consideration influencing concept learning and to my knowledge its effects on abstract concept learning have not yet been explored systematically in rats, which was the main purpose of this study.

Other experimental parameters in addition to stimulus modality and set-size may influence learning. Of theoretical interest is whether rats show a differential performance between MTS procedures and non-match-to-sample (NMTS) procedures, two procedures that are commonly used interchangeably. Wright and Delius (2005) defined the oddity preference effect (OPE) as a preference for the comparison stimulus that does not match the sample. Davenport and Menzel (1960) found that when simultaneously presented with a set of three stimuli, two identical and one unique, chimpanzees chose the unique stimulus significantly more than the other two in the absence of any reinforcement contingencies. Studies with various species of birds have found a superior performance on NMTS tasks compared to MTS tasks. Wilson, Mackintosh, and Boakes (1985b) compared NMTS to MTS performance between pigeons and jays. Birds were trained to either NMTS or MTS with pairs of colored disks. The performance of the jays was significantly better than pigeons, but within each species, oddity performance was
higher than matching. This effect was also found in a separate experiment in which pigeons alone were tested, and the authors attributed this finding to an inherent bias towards the non-matching stimulus (Wilson, Mackintosh, & Boakes, 1985a). This effect had previously been seen when Ginsburg (1957) trained pigeons on MTS and NMTS procedures. A peck to a sample stimulus (a colored disk) was followed by the presentation of two simultaneous disks, one that matched the color of the sample and one that was a different color. In the matching condition, pecking the comparison that matched the sample was rewarded. In the non-match condition pecks to the comparison that did not match the sample were rewarded. Training continued until a criterion of 80% correct in a session was reached. Ginsburg (1957) found that the pigeons in the non-matching condition learned the procedure and reached criterion significantly faster than those in the matching condition. The OPE has also been shown in rhesus monkeys under certain procedural conditions. Mishkin & Delacour (1975) trained monkeys on MTS or NTMS procedures, and found that monkeys trained on non-matching acquired criterion performance faster than those trained on matching. However this effect was only seen when stimuli were trial-unique.

Carter and Werner (1978) discussed the occurrence of an oddity preference early in (N)MTS training, which dissipated over time. They found that pigeons showed a tendency to peck the non-matching stimulus in early sessions. Berryman, Cumming, Cohen, and Johnson (1965) found a similar effect. Berryman et al. (1965) suggested that the effect occurred in pigeons due to extinction. They discussed procedures in which pigeons were required to begin a trial by pecking the sample key several times, while consuming reinforcers (typically seeds) that were placed on top of the sample stimulus. After pecking the sample stimulus and consuming the seed placed on top of the sample, subsequent pecks were no longer reinforced. Thus,
responding to the sample stimulus is eventually extinguished and may decrease the probability that the subject responds to that same stimulus again. This may explain the significantly faster acquisition of NMTS than MTS by pigeons in Lombardi (2008), in which a grain of wheat reinforcer was placed on top of the sample stimulus. However this theory would only be applicable to procedures that reinforce a response to the sample stimulus and allow the subject to continue to respond after reinforcement has been consumed.

Similarly, Zentall and Hogan (1974) suggested that initial preferences for the non-matching stimulus that have been seen in the literature may be a result of pre-training, specifically the observing response requirement. That is, in procedures that require an observing response to the sample before the comparisons are presented, responses to the sample key are not directly reinforced. Not reinforcing the observing response may function to extinguish responding to that particular stimulus, thus minimizing the likelihood of responding to that same stimulus when it appears again as a comparison. In essence, this would encourage oddity learning (D’Amato et al, 1985; Zentall & Hogan, 1974). If so, the reverse should be true for matching to sample: procedures in which a response to the sample is reinforced should favor identity learning. This was the logic behind an attempt to explore the cause of the OPE by Wright and Delius (2005). They proposed that the effect is not a predisposed preference but rather may be due to aspects of the experimental procedure. Specifically, they suggested that differences in sample reinforcement may have to do with the OPE in pigeons. Their reasoning was that often in typical MTS or NMTS procedures with pigeons, pecking the sample produces the two comparison stimuli, but no primary reinforcement. As Zentall and Hogan (1974) and D’Amato et al. (1985) suggested, not reinforcing responses to the sample presumably begins an extinction process related to responding to that stimulus, thus weakening the tendency to choose
that particular stimulus again. When the sample response is rewarded, responding to that particular stimulus is presumably strengthened.

To test this hypothesis, Wright and Delius (2005) trained pigeons to either MTS or NMTS while also manipulating observing response reinforcement. Four experimental conditions were employed: MTS with sample-response reinforcement, MTS with no sample-response reinforcement, NTMS with sample-response reinforcement, and NTMS with no sample-response reinforcement. In the first experiment, eight pigeons were trained to dig in ceramic pots with different kinds of gravel under one of the four conditions and performance was compared across groups. Wright and Delius (2005) found that the sample reinforcement did in fact affect acquisition of the procedure. Pigeons in the NTMS/no sample reinforcement acquired the procedure the fastest, followed by MTS/sample reinforcement. The researchers concluded that rewarding the sample enhanced the acquisition of MTS while retarding the acquisition of NMTS, while not rewarding the sample strengthened the acquisition of NMTS and retarded the acquisition of MTS. A second experiment was conducted to determine whether or not the pigeons from the first experiment learned the relational concept of identity and could transfer performance, and if sample reinforcement had an effect. The same pigeons were tested for transfer with novel stimuli which were mixed in with baseline trials. Transfer performance varied directly with the rate of learning. Wright and Delius (2005) concluded that sample reinforcement did have an effect on learning, and interestingly performance on the oddity task paired with no sample reinforcement was superior to the other groups. They further concluded that a task that is rapidly learned will be more readily generalized to novel stimuli, an effect that had not been shown before.
Alternatively, when Zentall and Hogan (1974) compared matching and oddity performances, no major difference were found in rates of learning over time. Pigeons were trained on either MTS or NMTS tasks with colored keys. Responding to the center key turned on the comparison keys that were positioned on either side of the center key. They found that the non-matching was learned at a slightly faster rate than matching in early sessions, but by the end of training there was no apparent difference between the two groups. In a second experiment, stimuli were four different brightness values. In this experiment, matching performance was superior to oddity across training sessions. However, on the transfer tests in which novel stimuli were introduced, performance on the non-matching task was superior to matching. Zentall and Hogan (1974) concluded that non-matching performance may be superior to matching initially, but matching may be learned faster when a more stringent learning criterion is used. It was suggested that the slight initially-superior oddity performance over identity performance was attributable to extinction of responses to a particular stimulus after a non-reinforced response to the sample, as previously mentioned. Berryman et al. (1965) also found that despite an initial superior performance under NMTS conditions, MTS was acquired at a faster rate and eventually performances under the two procedures were comparable. Similarly, Smirnova, Lazareva, and Zorina (2000) trained crows to either MTS or NMTS with colors, shapes, and numbers or elements. When matching and oddity performance were compared, no significant difference in acquisition between the groups was found. Clearly, there is mixed evidence of performance on these procedures and across species, and data are inconclusive.

The purpose of the present study was to assess abstract concept learning, specifically identity and oddity learning, in rats. Training set-size was manipulated to compare abstract learning performance between large (10 stimuli) and small (2 stimuli) set-sizes. In addition to
set-size, procedures were manipulated to compare performance on MTS and NMTS conditions. Observing response reinforcement was not manipulated between groups; instead there was a 50% chance of sample reinforcement throughout sessions for all conditions. Based on previous studies investigating the effect of set-size on concept learning with pigeons and monkeys, I hypothesized that limited, if any, transfer would occur after training with two stimuli. Further, rats trained with ten stimuli should show full abstract concept learning, performing as high on novel transfer tests as they did during baseline. Based on previous studies, I predicted that rats, too, will show a superior performance when trained on oddity versus matching, although the evidence for this effect is not strong.

METHOD

Subjects

Subjects were 25 experimentally naive male Sprague-Dawley rats from Harlan Laboratories in Indianapolis, IN. Access to water was unrestricted, and rats were fed approximately 1 g of Purina Lab Chow every day such that they were maintained at approximately 85% of their free feeding weight. Rats were approximately 30-90 days old at the start of the experiment and individually housed in temperature and humidity controlled environments on a reversed 12:12 hr light-dark cycle. Experimental sessions lasted approximately 1 hr and were typically conducted 5 days a week during the dark cycle. Subjects were fed approximately 20 minutes following testing sessions.

Apparatus

The apparatus used (shown in Figure 1) was a modified operant chamber similar to that used in Peña et al. (2006). The chamber measured 28 cm long x 26 cm wide x 30 cm high. The front and rear walls of the chamber were made of transparent Plexiglas while the side walls were
stainless steel, with stainless steel grids spaced 1.3 cm apart as the floor of the chamber. A 4-cm section was removed from the bottom of the front wall so that a plastic tray could be inserted into the chamber to present the stimuli. Two trays were used, one to present the sample stimulus and one to simultaneously present two comparison stimuli. The sample tray had a 5-cm hole drilled into the top of the tray, approximately 3 cm from the front of the tray and 10 cm from the sides. The comparison tray contained two holes, identical to that of the sample tray, drilled adjacent to each other and approximately 8 cm apart. Four screws were drilled around the edges of the hole, forming a square around the hole, to hold the stimulus lids in place. The screws allowed for the lids to slide back and forth when pushed.
Figure 1. Modified operant chamber, shown with the comparison tray inserted into the chamber.
Stimuli

Stimuli were square Plexiglas lids that were scented with household spices (The Great American Spice Co.). See Table 1 for a list of the scents used. A total of 30 spices were used. Spices were arbitrarily divided to create three sets consisting of 10 spices. Lids were stored overnight in stacks inside plastic containers with about 1 tablespoon of the corresponding powdered spice in the bottom of the container for saturation of the odor. The lids sat on a 1-inch plastic separator that was affixed to the bottom of the container to ensure that the lids did not come into direct contact with the actual spice, to prevent oversaturation. Four identical lids were used for each spice and were rotated on every session and within the session. The lids were separated from each other by about 1 cm to ensure proper dispersion of the odor during storage. During the experimental session, the lids were held in place by the 4 screws on the presentation tray and covered a 2 oz translucent condiment cup that was filled to about 1 cm below the rim with plain sand (The Home Depot). The cups sat inside the holes drilled into the sample and comparison trays. A sucrose pellet for reinforcement was placed, by an experimenter wearing latex gloves and using tweezers, into the cups and buried just beneath the surface of the sand.
Table 1

*Olfactory stimuli used*

<table>
<thead>
<tr>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sage</td>
<td>Rosemary</td>
<td>Fennel</td>
</tr>
<tr>
<td>Bay</td>
<td>Caraway</td>
<td>Paprika</td>
</tr>
<tr>
<td>Nutmeg</td>
<td>Spinach</td>
<td>Beet</td>
</tr>
<tr>
<td>Tomato</td>
<td>Cinnamon</td>
<td>Savory</td>
</tr>
<tr>
<td>Lime</td>
<td>Garlic</td>
<td>Celery</td>
</tr>
<tr>
<td>Thyme</td>
<td>Turmeric</td>
<td>Anise</td>
</tr>
<tr>
<td>Sumac</td>
<td>Clove</td>
<td>Coriander</td>
</tr>
<tr>
<td>Cumin</td>
<td>Dill</td>
<td>Mustard</td>
</tr>
<tr>
<td>Onion</td>
<td>Allspice</td>
<td>Marjoram</td>
</tr>
<tr>
<td>Raspberry</td>
<td>Oregano</td>
<td>Carob</td>
</tr>
</tbody>
</table>
Procedure.

Training occurred five days a week (M-F) with one session per day lasting approximately 1 hr. Sessions were conducted in the presence of soft (<70 dB) white noise. There were 30 trials in a session with an inter-trial interval of 30 s. If no response was made to either the sample or a comparison within 2 m, the trial was ended and the next trial began after the ITI. Rats were trained with the same stimuli each session until they were given a novel transfer test.

Preliminary training. Rats were first handled and adapted to the chamber. They were habituated to the chamber for 2-3 days and allowed access to sugar pellets until they consumed them reliably. Rats were trained to dig in the sand to retrieve the buried sugar pellet. The target response, lid pushing, was then shaped by reinforcing successive approximations of lid movement until rats were sliding the lid back when it was completely covering the cup beneath. Shaping began with a pellet buried in the sand-filled cup with an unscented lid placed so that it was only covering 25% of the cup, then 50%, and then 75%, until the cup was completely covered by the lid and the rat was sliding the lid far enough to obtain the buried pellet. Once unscented lid removal behavior was reliable, scented lids were introduced and the experimental procedure began.

Experimental conditions. After preliminary shaping, subjects were randomly assigned to either Match-to-sample (MTS) or Non-match-to-sample (NMTS) conditions, and to either a large (10-stimuli) set or small (2-stimuli) set-size. Rats were randomly assigned to one of the three olfactory stimuli sets for baseline training (Table 1). Rats in the 10-stimuli group trained with the full set of 10 while rats in the 2-stimuli group trained with two stimuli randomly selected from the set of 10. One of the two remaining stimulus sets was used for the first transfer test. After the first transfer test, subjects were trained to criterion on the new set (set-size expansion) and then
tested again for transfer (using the last stimulus set), allowing for within-subject comparisons. Thus there were three independent variables in a 2x2x2 mixed design: set-size, procedure, and transfer test.

Match to sample training (MTS). Typical matching-to-sample procedures were trained in which reinforcement was provided for responding to the comparison that matched the scent of the sample. Thus, the correct comparison was conditional upon the sample stimulus on the current trial. A trial began by inserting the sample tray into the chamber to allow access to the sample stimulus. Forward movement of the lid exposed plain sand inside the cup with a sugar pellet buried just below the surface such that the pellet could not be seen. After digging and consuming the pellet, the sample tray was removed, and the comparison tray was immediately placed at the tray entrance slot simultaneously presenting the two comparison odors (zero-delay procedure). The comparison tray was held just out of reach of the rat for two 2 s, after which the tray was inserted revealing the two comparison cups. This pause was established as an attempt to standardize tray presentation times across experimenters. Although there was a “delay” in which the rat was not able to make a response, the comparisons were nevertheless available during the pause allowing the rat to sample the odors before making a response. Thus, this was not a delay-procedure.

One of the comparison cups was topped with a lid scented with the same scent as the sample (S+), and the other with a lid with a different scent. A response to the sample stimulus was operationally defined as the lid being pushed with either the paws or snout past the first two screws, ensuring adequate exposure to the sample scent. A response to a comparison was defined as any displacement of the lid such that the cup beneath was exposed. Comparison cups with the lid that matched the sample scent contained a sugar pellet for reinforcement, and trials in which
the first response was to this cup were scored as a correct trial. The other cup that did not match the sample (S-) contained no pellet and trials in which a response was first made to this cup were scored as incorrect.

Two types of correction procedures were used throughout the experiment. In the non-correction procedure, used twice a week, the tray was removed immediately upon an incorrect response and the trial was terminated. Because of likely loss of responding, especially early in training, a full correction procedure was also used every other day (3 times a week) in which the trial continued after an incorrect response until the S+ pellet was consumed. To control for scent detection of the pellet, double-bait sessions were conducted twice a week in which testing occurred as usual except that there was also a pellet buried in the incorrect comparison. Because of the risk of incorrect responses being reinforced, the tray was removed immediately upon an incorrect response before the pellet could be obtained. Thus, double-bait sessions occurred on the same day as non-correction days. After termination of the trial, there was a 30 s inter-trial interval after which the next trial began.

Non match to sample training (NMTS). Sessions were identical to MTS sessions, except that a correct response was scored as a response to the cup that was scented differently from that of the sample.

Set-size. Within MTS or NTMS conditions, subjects were randomly assigned to the 10-stimuli group or the 2-stimuli group. Upon beginning the experimental procedure, each rat was randomly assigned to one of the stimulus sets for baseline training. In the 10-stimuli condition, rats were presented with stimuli drawn from a pool of 10 different stimuli throughout the session. On each trial, two stimuli from the set of 10 were used. Each individual scent from the set of 10 was presented 3 times as the sample, correct, and incorrect comparison throughout the session.
the 2-stimuli condition, only two scents were used in the entire session. Each scent served as the sample, correct, and incorrect comparison 15 times each. In both 10-stimuli and 2-stimuli conditions, the number of times in which the correct and incorrect comparisons appeared on the left and right sides of the comparison tray were balanced. The number of times each scent appeared on the left and right was also balanced across the session.

Sample Reinforcement Reduction. Early sessions contained trials in which responses to the sample were reinforced on every trial. To ensure that responding was not under the control of reinforcement history of the sample, but instead on the relation between sample and comparisons, the probability that the sample cup contained a pellet was reduced, a procedure used similar to that in Peña et al (2006). Once accuracy of 90% correct or better was reached for one session, the sample reinforcement probability was reduced to 75%. Thus, 25% of the samples in the session were not baited and the rest were. The un-baited sample trials were pseudo-randomly selected balancing the number of times an un-baited sample trial contained the S+ on the left and right sides. No more than three trials in a row contained an un-baited sample. Subsequently, when accuracy reached 90% correct or better in the 75% sample reinforcement phase for two sessions in a row, the sample reinforcement probability was then reduced to 50%. At this point, half of the trials contained a baited sample and half contained an un-baited sample. Based on results of Peña et al. (2006), sample reinforcement probability was not lowered any further due to possible disruption in performance and loss of responding. This also allowed for testing of the oddity preference effect in the absence of sample reinforcement differences.

If by the 50th session of the 100% sample phase accuracy was not within 10% of criterion or showing trends towards improvement, the subject was dropped from the study.
Transfer testing. After reaching a performance criterion of 90% correct or better for two consecutive sessions in the 50% sample reinforcement phase, rats were tested for generalized MTS or NMTS. Novel transfer tests consisted of a session with the original (baseline) training scents as well as 10 novel stimuli. Novel transfer sessions were always double-baited and non-correction. The first 10 trials served as warm-up trials and contained only the baseline scents. If the rat failed to respond correctly on at least 7 of the warm-up 10 trials, in which errors were not made in trials 8, 9, and 10, consecutively, the transfer test did not occur and a regular baseline session resumed. Otherwise, the transfer session continued and trials containing novel scents began to appear after trial 10. The 10 novel trials appeared pseudo-randomly such that no more than 2 novel trials appeared in a row. A transfer trial consisted of presenting a scent as the sample that had never appeared as a sample before, with another scent from the set of 10 novel scents serving as the correct/incorrect comparison (depending on MTS or NMTS condition). Novel scents were never paired with a scent from the baseline training set. Novel trial samples never contained reinforcement. Only performance on novel trials only was analyzed for transfer performance. After the transfer test, testing resumed using the stimuli from the previous transfer test and the baseline training stimuli were dropped. Testing continued with the new set of stimuli until criterion was reached again. Thus, the 2-stimuli group underwent a set-size expansion. The 10-stimuli group continued training with the new set of 10 to control for experience. A second novel transfer test, identical to the first except with a new set of 10 stimuli, was then conducted.

Data Analysis

Abstract concept learning was assessed with a novel transfer test after performance criterion was reached. Performance on the 10 novel stimulus trials introduced during a transfer test session was analyzed to compute a transfer test score, which was the dependent variable
(percent correct). Novel transfer performance was compared between 10-stimuli and 2-stimuli groups within the MTS or NMTS procedures. A second novel transfer test allowed for comparison of performance between the first and second tests after set-size expansion in the 2-stimuli group. Acquisition, which was analyzed both in terms of sessions to criterion as well as errors to criterion, was also a dependent variable. A 3-way Mixed ANOVA was used to analyze differences between groups as well as on first and second transfer tests. Double bait data were analyzed for each rat using a related samples t-test by comparing the 5 pairs of double bait and non double bait sessions preceding a the first novel transfer test.

RESULTS

Acquisition

Figure 2 illustrates the acquisition of the MTS task across reinforcement reduction phases for subject P2 in the 10-stimuli NMTS condition. Responding for this and all subjects began around chance levels (50% correct). Accuracy gradually increased above chance and P2 met the first performance criterion of 90% correct for the first time after only 7 sessions, which was the fewest number of sessions required among all subjects. At this point, reinforcement probability for digging in the sample was reduced to 75%. Performance dipped slightly to 86% correct on the first day of this phase but immediately increased on the next two sessions to 90% correct, at which point the next performance criterion was met (90% correct or better for two consecutive days). The sample probability was then reduced to 50% and performance did not suffer, remaining above the criterion level. The final criterion for a novel transfer test (two consecutive days at or above 90%) was met at session 14, at which point the next session was a transfer test. See Appendix A for individual data for all subjects. Figure 3 shows acquisition of the matching/non-matching to sample task as sessions to first transfer test for subjects in all
conditions. Number of sessions before reaching each criterion of sample reinforcement and first transfer test is shown in Table 2. As seen in Table 2, there was much variability in acquisition between subjects.

On average, the group with the fewest sessions required before receiving the first transfer test was the 10-stimuli NMTS (NMTS-10) group ($M=33.75$, $SEM=10.05$), followed by MTS-10 ($M=36.73$, $SEM=6.14$), MTS-2 ($M=38.25$, $SEM=26.06$), and NMTS-2 ($M=44$, $SEM=10.79$); however, there were no statistically significant differences across groups. The main effect of set-size on sessions to acquisition was not significant, $F(1,12)=.232$, $p>.05$, and neither was the main effect of procedure, $F(1,12)=.013$, $p>.05$. The interaction of these two factors on sessions to acquisition was also not significant, $F(1,12)=.129$, $p>.05$.

Three out of four subjects in the 2-stim MTS group met criterion in less than 30 sessions; however, one subject (V4) was tested for 93 sessions before receiving a transfer test. Acquisition for this subject was slow; however, performance was within 10% of criterion by session 17. Similarly, three out of four subjects in the 2-stimuli NMTS group met criterion for a transfer test within 40 sessions, while one subject required 76. Two subjects in the 10-stim NMTS condition met criterion within 18 sessions, while the other two required over 40.

Dropped subjects. Because nine subjects were dropped from the study for failing to meet criterion, they were not included in the previous analyses of sessions to criterion. However, a separate analysis of number of errors made until dropped that included the nine dropped subjects revealed that rats trained with 2 stimuli made significantly more errors ($M=559.521$, $SEM=57.79$) than those trained with 10 stimuli ($M=321.87$, $SEM=80.515$), $F(1,21)=5.75$, $p<.05$. All nine dropped subjects were in the 2-stimulus condition. No significant differences were found in
errors made between procedures, $F(1,21) = .369, p > .05$; however, seven of the nine dropped subjects were in the MTS condition.
Figure 2. Individual data for subject P2 in the 10-stimuli, NMTS condition. Percent correct is shown across sessions. Sample reinforcement phases indicated as a proportion, with each phase separated by vertical lines. Triangle symbols show overall percent correct on a novel transfer test session, with a fraction in parenthesis below indicating number correct out of the 10 novel trials.
Figure 3. Number of sessions during acquisition before receiving the first transfer test for individual subjects as well as means (horizontal bars) for MTS (grey) and NMTS (black). Sessions consisted of 30 trials. Only data for subjects that met criterion for a transfer test is shown (dropped subjects not included).
Table 2.
*Number of sessions to each criterion for subjects that met criterion for a transfer test.*

<table>
<thead>
<tr>
<th>Subject &amp; Condition</th>
<th>75% Criterion</th>
<th>50% Criterion</th>
<th>Transfer Test Criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>S13 (MTS-2)</td>
<td>20</td>
<td>23</td>
<td>27</td>
</tr>
<tr>
<td>T15 (MTS-2)</td>
<td>7</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>W27 (MTS-2)</td>
<td>8</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>V4 (MTS-2)</td>
<td>68</td>
<td>90</td>
<td>93</td>
</tr>
<tr>
<td>P1 (NMTS-2)</td>
<td>21</td>
<td>23</td>
<td>29</td>
</tr>
<tr>
<td>P8 (NMTS-2)</td>
<td>26</td>
<td>31</td>
<td>34</td>
</tr>
<tr>
<td>T11 (NMTS-2)</td>
<td>51</td>
<td>73</td>
<td>76</td>
</tr>
<tr>
<td>T3 (NMTS-2)</td>
<td>24</td>
<td>28</td>
<td>37</td>
</tr>
<tr>
<td>S4 (MTS-10)</td>
<td>24</td>
<td>28</td>
<td>33</td>
</tr>
<tr>
<td>S2 (MTS-10)</td>
<td>26</td>
<td>39</td>
<td>46</td>
</tr>
<tr>
<td>U12 (MTS-10)</td>
<td>15</td>
<td>40</td>
<td>47</td>
</tr>
<tr>
<td>T1 (MTS-10)</td>
<td>9</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>P2 (NMTS-10)</td>
<td>7</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>S3 (NMTS-10)</td>
<td>8</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td>U9 (NMTS-10)</td>
<td>39</td>
<td>43</td>
<td>54</td>
</tr>
<tr>
<td>T20 (NMTS-10)</td>
<td>39</td>
<td>43</td>
<td>48</td>
</tr>
</tbody>
</table>
Transfer Performance

Figures 4 and 5 show mean and individual performance separately for the first and second novel transfer tests for all conditions. In general, it appears that, regardless of type of procedure (MTS and NMTS), rats’ performance was better if trained with 10 exemplars and at the second transfer test.

Percent correct scores on novel trials for both transfer tests were subjected to a 3-way mixed ANOVA which showed a significant interaction between set-size and transfer test, $F(1, 12)=7.119, p<.02$, as well as a significant main effect of set-size, $F(1, 12) = 10.711, p<.007$; no other effects were significant (see Appendix B). Therefore, data were collapsed across procedure (MTS and NMTS) and analyzed with a two-way mixed ANOVA with set-size and transfer test (1 or 2) as factors.

A two-way ANOVA confirmed significant main effects of set-size [$F(1, 14) = 10.602, p<.006$], transfer test [$F(1, 14) = 5.15, p<.04$], and a significant interaction between transfer test and set-size, $F(1, 14) = 9.356, p<.009$. Post-hoc comparisons of performance at Transfer Test 1 and Transfer Test 2 were conducted.

Transfer Test 1. As seen in Figure 4, on average subjects in the 2-stimuli condition ($M=48.75$ percent correct, $SEM=11.41$) performed significantly worse than subjects in the 10-stimuli condition ($M=90$, $SEM=2.67$; $t(14) = 3.35, p<.004$) on the first transfer test.

2-stimuli condition. Of the eight subjects in the 2-stimuli condition, only one subject (P8, NMTS) responded significantly above chance during the first transfer test (80% correct or above, $p<.05$, binomial test). Of the remaining subjects originally trained with two stimuli, two responded correctly on 70% of the novel trials (V4, MTS; and T11, NMTS). Three of the remaining subjects (S13, T15, and W27, all MTS) responded with an accuracy of 30%, 40%, and
60%, respectively. Two subjects (P1 and T3, both NMTS) failed to respond during the allotted time on some of the novel trials. For data analysis purposes, a non-response trial was scored as incorrect. P1 responded on only three novel trials (2/3 correct) and T3 did not respond to any of the 10 novel trials. Thus, performance for these subjects was far below chance (20% and 0% correct, respectively).

10-stimuli condition. All eight subjects in the 10-stimulus condition responded significantly above chance during the first transfer test (8/8 performed at 80% correct or above, p< .05, binomial test). Two of the eight (S4, MTS and P2, NMTS) performed without error on all 10 novel trials. Four subjects (S2 and T1, MTS and U9 and T20, NMTS) performed correctly on 90% of the novel trials, while two subjects (U12, MTS and S3, NMTS) performed with an accuracy of 80%.

Set-size Expansion. Following Transfer Test 1, subjects continued testing with only the 10 stimuli used in the transfer test. Subjects were required to again meet a criterion of 90% or better for two consecutive days before receiving a second novel transfer test. All subjects met criterion for a second transfer test in 3-7 sessions.

Transfer Test 2. All subjects received a second novel transfer test after reaching criterion again with the new stimulus set. On average, performance between subjects in the 2-stimuli condition (M= 82.5, SEM= 3.66) and subjects in the 10-stimuli condition (M= 85, SEM= 1.89) was comparable on Transfer Test 2 after set-size expansion, t (14) =.32, 0.76). Figure 5 shows individual and mean performance on the second transfer tests for all groups. After set-size expansion, six of the eight subjects originally trained with two stimuli performed significantly above chance on Transfer Test 2 (6/8 performed at 80% correct or above, p< .05, binomial test), while the other two performed just below significance (70% correct).
Figure 4. Percent correct on the first novel transfer test for each subject as a function of number of stimuli used in training for both MTS (gray circles) and NMTS conditions (black squares). Group means indicated by horizontal bar.
Figure 5. Percent correct on the second novel transfer test as a function of original number of stimuli used in training for both MTS (gray circles) and NMTS groups (black squares). Means shown as horizontal bar.
Comparison of Transfer Test Performance. On average, group performance significantly increased from Transfer Test 1 ($M = 48.75$, $SEM = 11.41$) to Transfer Test 2 ($M = 82.5$, $SEM = 3.66$) for subjects in the 2-stimuli group, $t(7) = 2.76$, $p < 0.02$. In contrast, performance for all subjects in the 10-stimuli condition was similar at both transfer tests, $t(7) = 1.53$, $p > 0.05$. Performance remained significantly above chance on Transfer Test 2 (8/8 performed at 80% correct or above, $p < .05$, binomial test, $M = 85$, $SEM = 1.89$), just as it did on Transfer Test 1 ($M = 90$, $SEM = 2.67$). Figure 6 compares mean and individual performance for all subjects on both transfer tests.

Controls. Double bait sessions in which both of the comparison cups contained a sugar pellet were analyzed using a related samples t-test. Pairs of double bait and non double bait days for the 5 days prior to the first novel transfer test were matched for each subject. Comparisons revealed that accuracy on days in which only the correct cup contained a pellet ($M = 83.61$, $SEM = 1.08$) did not differ significantly from double bait sessions ($M = 85.97$, $SEM = 1.06$), $t(79) = -1.971$, $p > 0.05$, ruling out the possibility that performance relied on pellet detection.

Reliability. Video for each novel transfer session for subjects in which video was available ($n = 10$) as well as 3 additional sessions randomly selected for each subject was reviewed to ensure correct scoring. Of the videos reviewed, there was a 98.7% agreement between original scoring and video review. When a scoring error was found while reexamining the videos, the data were corrected.
Figure 6. Individual and group performance for all conditions on each transfer test. MTS shown in grey, upper panels, and NMTS shown in black, lower panels. Left column shows transfer performance for the 2-stimuli group, 10-stimuli transfer performance shown in the right column. Each symbol corresponds to an individual subject’s performance on Transfer Test 1 and Transfer Test 2. Group means represented by horizontal bars. Asterisks (*) represent sessions containing trials in which a subject did not make a response and were scored as an incorrect trial.
DISCUSSION

No differences were found between matching and non-matching procedures on either acquisition or novel transfer performance in this study. These findings are inconsistent with the OPE effect reported by Wright & Delius (2005) in which pigeons trained to MTS acquired the task faster than pigeons trained to NMTS when responses to the sample were reinforced. Although sample reinforcement was not manipulated between groups in the current study, all subjects began training with a 100% sample reinforcement probability, meaning responses to the sample were always reinforced until the first criterion was met. According to the OPE which is based on basic principles of reinforcement and extinction, reinforcing a response to a particular sample would thereby increase the likelihood of responding to that same stimulus when it appeared next as a comparison. This effect would not benefit subjects trained to NMTS, who would have to work against these reinforcement contingencies. Thus, it would be expected that rats trained to MTS would acquire the task faster than rats trained to NMTS when the sample was always reinforced, an effect that was not seen in this study. Nevertheless, there was a considerable amount of variability in acquisition, and more subjects should be tested to further examine differences in procedures.

Results from the present study support previous findings that training with a large number of exemplars facilitates abstract concept learning, and extends these findings to rats. Rats originally trained with 10 stimuli performed with high accuracy on the first transfer test, while rats originally trained with 2 stimuli generally did not perform significantly better than chance. Further, transfer performance for rats originally trained with 2 stimuli improved after set-size expansion on Transfer Test 2.
On Transfer Test 1, all rats trained with 10 stimuli performed above 80% correct, which was significantly better than chance. Six of the eight rats performed above baseline levels (90% correct), satisfying the Katz et al. (2007) criteria for abstract concept learning that transfer performance should equal baseline, while the other two subjects performed at 80% correct. As expected, most subjects originally trained with 2 stimuli did not perform significantly above chance on Transfer Test 1. These results corroborate previous findings that training with multiple exemplars facilitates novel transfer performance (Katz et al., 2002; Wright et al., 2003).

However, there was much variability among transfer performance within the group of subjects originally trained with 2 stimuli. One subject performed significantly better than chance and above baseline levels, satisfying the Katz et al. (2007) criteria for abstract concept learning. Of the seven remaining subjects in this group, two scored 70% correct and one scored 60% correct on the first novel transfer test. According to Katz et al. (2007), these subjects satisfied the criteria for “partial concept learning”, meaning that performance was above chance levels (50% correct) but below baseline (90% correct). This suggests that, although a full abstract concept was not learned and behavior was not solely under relational control, there may have been competing sources of stimulus control. If behavior was controlled purely by absolute stimulus cues, transfer performance should have been close to chance. This is further evidenced by the rapid reacquisition of these three subjects; only 3-4 sessions were required to meet the criterion for a second novel transfer test after set-size expansion. If no relational learning had occurred and behavior was stimulus bound, the rats should have required much more exposure to the new stimuli before seeing high matching accuracy with the new set. Thus, it is possible that the behavior of these rats was under some combination of relational and absolute stimulus control. In contrast though, the other four rats in this group scored below 50% correct on Transfer Test 1,
and presumably did not learn any relational concept. Type of procedure was not relevant, with two of the four subjects that performed above chance levels in the MTS condition and two in the NMTS condition.

On Transfer Test 2, all eight subjects originally trained with 10 stimuli remained significantly above chance levels, as expected. Subjects originally trained with 2 stimuli showed an increase from Transfer 1 to Transfer 2 following set-size expansion. Four of these eight subjects performed significantly above chance on Transfer Test 2, with the other two scoring 70% correct (partial transfer). Thus, it can be concluded that responding shifted from absolute stimulus control or partial concept learning to responding based on stimulus relations once the set-size was expanded. However, as Nakamura et al. (2009) found, the number of stimuli used during initial training may have a significant impact on subsequent transfer performance. As mentioned, pigeons were trained on a same/different task with a 32-item stimulus set, using digital picture stimuli, and then tested for transfer. The transfer performance was compared to the performance of pigeons in the Katz et al. (2006) study, using the same stimuli and procedure, in which training began with an 8-item set-size which was expanded to 16 and then 32 items after testing for transfer at each set-size. Pigeons trained with an initial 32-item set showed 73.8% correct transfer to novel stimuli, while pigeons originally trained with 8 stimuli showed only 59.6% correct after reaching the 32-item set, indicating that initial training with few exemplars may have hindered concept formation after set-size expansion. These findings may account for the fact that some rats originally trained with 2 stimuli in the current study still only met partial concept learning criteria on Transfer 2, and on average were not as accurate as rats originally trained with 10. Interestingly, the subjects in the 2-stimuli condition that were still only at partial concept learning levels after set-size expansion were also the subjects that had scored below
chance on Transfer 1, and the subjects that were at partial levels on Transfer 1 increased to above baseline levels on Transfer 2 (except for one subject who remained at 70% correct on both tests). These data agree with the Nakamura et al. (2009) conclusion that initial learning may interfere with subsequent performance.

As mentioned, one subject (P8) in the 2 stimuli condition performed above baseline levels on Transfer 1, responding correctly on all 10 novel stimuli trials. A few studies have reported obtaining transfer after training with a small number of stimuli with pigeons and chimpanzees (Lombardi 2008; Oden et al., 1988; Wright 1997). The successful transfer after training with such few exemplars was attributed to experimental conditions which may have promoted extensive processing of the stimuli, which will be discussed later. However, it is not clear why this would only occur for one subject out of eight trained with two stimuli when procedural conditions were consistent across the subjects.

Two subjects, P1 and T3 (both in the 2-stim NMTS condition), completely avoided responding to most of the novel stimuli when they were presented during the first transfer test and thus have inconclusive data for Transfer Test 1. P1 only made a response on four out of the ten the novel trials (the first three and the very last of the session), failing to respond before the timeout period elapsed on the other six novel trials. However, he readily responded to the baseline trials in between the novel trials, indicating that it was the novelty of the stimulus that disrupted responding. T3 incorrectly responded to the first novel trial, subsequently failing to respond on the remaining nine novel trials. Notably, both of these subjects were in the 2-stimuli condition. This behavior can likely be explained by the common neophobia observed in rats (D’Amato et al., 1985; Jackson & Pegram, 1970). When presented with new stimuli after extensive experience with a fixed set of items, the novelty of the test stimuli may cause
behavioral disruption and thus cause an animal to fail to transfer. D’Amato et al. (1985) further suggested that before an animal can optimally respond to a novel stimulus, some perceptual pre-learning may be necessary. Oden et al. (1988) allowed chimpanzees to become familiarized with objects in a free-play situation prior to transfer tests in order to minimize potential disruption of performance caused by stimulus novelty. In the present study, pre-exposing subjects to novel scents prior to transfer tests may have controlled for neophobia; however, care would need to be taken to ensure all stimuli were equally presented and that none of the novel scents were associated with any type of reinforcement. Aside from the notion that using a large number of stimuli promotes relational learning by providing a less costly alternative to stimulus-specific learning, it is also likely that the use of many exemplars helps diminish any innate neophobia by exposing subjects to a large number of frequently varying stimuli. This is supported by the fact that both P1 and T3 readily responded to novel trials on their second novel transfer test after their set-sizes had been expanded.

There have been a few instances in which transfer to novel stimuli has occurred after training with a few exemplars, and these findings were likely due to features of the procedure. Wright (1997) trained pigeons to MTS using three color cartoons. The stimuli were presented from the floor of the chamber so that pigeons pecked downward at the stimuli. The observing response was varied between groups; pigeons either pecked the sample 0, 1, 10, or 20 times before the comparisons were presented. After acquisition of the MTS task, tests with novel stimuli were introduced. Transfer performance increased as a function of number of pecks required to the sample. Pigeons that were required to peck the sample 20 times performed as accurately on novel stimuli tests as they did in training, whereas pigeons required to peck 0, 1, or 10 times did not, demonstrating that the observing response had a great effect on abstract concept
learning. However, the effect of the observing response may have been enhanced by the use of horizontally positioned stimuli, which has been shown to affect concept learning (Katz et al., 2008). Similar studies which were comparable in other aspects except using vertically positioned stimuli have failed to demonstrate abstract concept learning (Bodily et al., 2008).

Lombardi (2008) also found that pigeons showed transfer after training with a small set-size using horizontally positioned stimuli. Six pigeons were trained to either MTS or NMTS with square cardboard lids of different colors that could slide to uncover grain reinforcer underneath. Pigeons were divided into two different stimulus arrangement groups. In one condition, stimuli were presented adjacently in a horizontal line, with the two comparisons on either side of the sample. In the second condition, all conditions were equal except that comparison lids were 1 cm apart from the sample lid. After training with two stimuli (red and yellow lids), pigeons were tested with novel colors and performed as accurately as or better than baseline performance. However, this was only seen for pigeons trained in the separate-lid condition. Pigeons in the adjacent-lid condition performed around chance on novel transfer tests. The authors proposed that the stimulus arrangements had an effect on learning. When the stimuli were adjacent to each other, it is likely that the three lids were viewed as one entire arrangement and behavior probably was under control of the different color configurations. On the other hand, separating the stimuli by 1 cm apparently prompted individual processing of the sample and comparisons, thus facilitating relational learning. Further, the authors suggested that the fact that the sample lid was a different size than the comparison lids in the adjacent-lid condition may have facilitated more distinctive processing of the sample, and previous failures to find transfer with a small set-size may have been due to minimal processing of the sample by the subjects (Lombardi, 2008). This latter explanation may have been why Oden et al. (1988) found that chimpanzees were able to
correctly match novel objects after MTS training with only two stimuli. Infant chimpanzees were handed a sample object by an experimenter and allowed to handle the object and then place it into a tin. They were then presented with two alternatives, one which matched the sample and one which did not, and were rewarded for placing the matching object into the tin with the sample. After meeting the criterion level of making 10/12 correct responses, subjects were tested for transfer by presenting them with two novel objects. Although only two stimuli were used in training, all subjects transferred matching performance to the 12 novel stimuli. This may have been due to the fact that subjects were allowed to extensively handle the objects and thus able to thoroughly handle the stimuli, as well as the aforementioned pre-exposure that may have helped reduce novelty-induced disruption.

The results of this study also contribute to findings that ecologically relevant conditions may facilitate learning. Wright and Delius (1994) discussed the importance of developing procedures that employ more natural responses in the advancement of comparative cognition research. They suggested that a natural response to which an animal may be predisposed could be adapted to a commonly used laboratory procedure. On this premise, pigeons were trained to MTS or NMTS by digging in pots of gravel for buried reinforcers. Stimuli were pots filled with white or grey gravel in which grain reinforcer was buried. After digging in the sample pot and obtaining the buried seed reinforcer, two comparison pots containing matching and non-matching gravel were exposed and pigeons responded by digging in the comparison pot. Learning of this task occurred much faster than typical pigeon MTS procedures involving pecking two-dimensional stimuli. Criterion of 90% accuracy was achieved in about 27 trials, whereas classic key-peck settings have required hundreds. Pigeons failed to transfer to novel colors of gravel, but this finding was attributed to the small number of training stimuli. The
authors suggested that the gravel-digging setting resembled pigeons’ natural environments in which they must rapidly learn to forage for food and was therefore a more suitable setting to obtain optimal performance. Likewise, several studies have required pigeons to peck downward at stimuli on the ground rather than in front of them because of the suggestion that this more natural situation was analogous to pigeons pecking on the ground for seeds, thus encouraging performance (Lombardi, 2008; Wright, 1997; Wright et al., 1988). In the present study, stimuli were presented horizontally at the base of the chamber and rats attended to stimuli much in the way they would encounter objects in a natural environment. Moreover, the requirement that the rats used their paws and/or snout to push the stimulus lids and dig in the sand for buried food pellets may have reflected natural foraging behavior. Additionally, the use of scented stimuli undoubtedly may have capitalized on the rat’s superior olfactory system, a likely assumption given previous failed attempts to obtain abstract concept learning in rats with visual stimuli (Iversen, 1993; Iversen, 1997). These factors may account for why Thomas and Noble (1988) did not find transfer in rats trained with olfactory stimuli. Rats were trained to NMTS with odor-saturated ping-pong balls, using a pool of 16 scents. Three ping-pong balls were presented together, and the rat was required to displace the non-matching scented ping-pong ball to uncover a food well beneath. Although rats were able to meet a criterion of 90% correct on the training problems, performance on novel stimuli trials was around chance. Given that rats in the current study were able to transfer to novel stimuli after training with only 10 stimuli, it is surprising that with an olfactory procedure and employing a larger number of stimuli, rats in the Thomas and Noble (1988) study did not. Thus, it is likely that some aspect of the current procedure, perhaps the digging requirement, facilitated relational stimulus control.
Wright and Delius (1994) compared the rapid learning by pigeons in the gravel-digging task to that of the natural environment of a pigeon, in which learning to find food and avoid predators must occur quickly. The authors concluded that the contact with and manipulation of three-dimensional stimuli as well as discovering the buried reinforcer were crucial contributors to the rapid acquisition, and that this task may have been better suited for testing cognitive processes in pigeons. Thus, the vast difference between number of exemplars required for transfer between rats (10 odors or fewer, present study and Peña et al., 2006), monkeys (128 items, Wright et al., 2003), and pigeons (64 items, Nakamura et al., 2009 to over 700 items, Bodily et al., 2008) may essentially be due to using procedures that take advantage of natural behavior therefore permitting optimal performance, and not any innate species learning differences.

It is noteworthy that under similar procedures and using a similar apparatus, subjects in Peña et al. (2006) reached the same criterion level used in the current study with two stimuli much faster than most subjects trained with two stimuli in the current study. All four subjects in the Peña et al. (2006) study met criterion in 35 sessions or less. Most subjects in the 2-stimuli condition in current study required a similar number of sessions, but some required over 70. One explanation for this more rapid acquisition is the sample exposure in the Peña et al. (2006) study. Rats were required to dig in scented sand as a response requirement, ensuring thorough exposure to the sample. This procedure contrasts with that of the present study in which the sand was not scented but was covered by a scented lid. In this case, responses (removing the lid from the cup) were typically made very rapidly, minimizing exposure to the actual scent. Requiring longer exposure to the sample has in fact been shown to speed learning. Sacks, Alan, & Kamil (1972) trained pigeons on a zero-delay MTS procedure in which either 1, 10, 20, or 40 responses to the
sample were required in order to produce the comparison stimuli. Speed of acquisition increased as a direct function of the sample response requirement. Increasing the sample response requirement also facilitated performance on tests with novel stimuli. In spite of these findings, scented sand was eliminated in the present study due to complications noted by Peña et al. (2006). These included such as unclear response definitions, which led to difficulty scoring the dependent variable.

Another advantage of the procedure used by Peña et al. (2006) was that simultaneous matching was employed. In the current study, a zero delay procedure was used because of the aforementioned scoring difficulties. The apparatus in the current study was modified from that used in Peña et al. (2006) to allow room for the stimulus lids, which were developed for ease of scoring responses. However this procedure required the use of two separate trays- one for the sample and one for the comparisons. This necessarily created a zero-delay procedure. Although use of the new apparatus enabled a more reliably defined response, the zero-delay as opposed to the simultaneous procedure may have impaired learning. Wilson et al. (1985a) examined effects of procedure by training pigeons to MTS on either a simultaneous delay or zero delay procedure and found that learning occurred more rapidly when sample and comparisons were presented simultaneously than with a zero-delay. Another advantage of the Peña et al. (2006) procedure was the proximity of the sample to the comparisons. Wilson et al. (1985a) also manipulated closeness of the sample and comparison keys in MTS and NMTS learning in pigeons. Sample and comparison keys were .3 cm apart in the close-key condition and 3.5 cm apart in the spaced-key condition. During acquisition, learning occurred more rapidly when sample and comparisons were closer together rather than spaced apart. Despite the advantage of a simultaneous
procedure, learning and transfer was obtained in the present study despite the use of a zero-delay
procedure.

Worthy of discussion is the fact that eight subjects in the 2-stimuli condition were
dropped from the study for failing to perform above chance after 50 sessions. It is puzzling that
these subjects failed to even learn stimulus-specific associations as is commonly reported in the
literature for subjects trained with few stimuli. A likely explanation is that most of these subjects
developed exclusive side biased responding, rarely straying from chance performance each
session. Even subjects U8 and T5 who met criterion early in training and moved to the 75%
sample reinforcement phase developed side-bias responding. Thus, for these rats, side-bias
responding may have been more easily learned, interfering with relational stimulus learning or
even stimulus-specific associations. Parameters of the procedure in the present study made side-
biased responding a low-cost strategy which still allowed for intermittent reinforcement. Because
sessions were arranged such that trials were counterbalanced, ensuring equal reinforcement on
the left and right side throughout the session, exclusive side-bias responding on non-correction
days would still result in reinforcement on a minimum of 50% of the trials. Since one position
never contained the correct comparison for more than two consecutive trials, side-biased
responding would result in no more than two non-reinforced trials in a row. On correction
procedure days, trials continued until a correct response was made and so, although delayed on
some trials, reinforcement would eventually be obtained on 100% of the trials. Thus, the cost for
exclusive side-biased responding on either correction or non-correction days was low, allowing
the possibility that side biased responding could be learned very easily. Notably, none of the
dropped subjects were from the 10-stimuli condition, suggesting that the use of multiple
exemplars created conditions that favored relational learning over side-biased responding.
It should also be noted that according to the definition for abstract concept learning proposed by Katz et al. (2007), not all rats in the current study showed abstract concept learning. The Katz et al. criterion required that transfer performance be as accurate as or higher than baseline performance. In the present study, baseline criterion was defined as 90% correct or better, and so in order to satisfy the Katz et al. criterion, transfer performance would need to be 90% correct or higher. In the 10-stimuli group, all subjects had an accuracy of 80-100% on both novel transfer tests. According to the Katz et al. (2007) criteria, subjects who had an accuracy of 80% were only at partial transfer. However, because there were only 10 transfer trials, this is a difference of just one trial. Moreover, the 90% criterion used in the current study was higher than the criteria employed in many other studies. If a slightly less stringent criterion of 75-80% had been used instead, all of these rats would have met the abstract concept learning criterion proposed by Katz et al. (2007). Further, the response definition used in the present study was very strict. Any movement of the lid past the rim of the cup was considered a response. Subjects typically came into contact with both comparisons before making a choice, frequently rejecting the incorrect comparison and then switching to the correct one. Often this brief sampling caused movement of the lid, inadvertently causing a response to the incorrect comparison even though the rats did not search for the pellet in the incorrect comparison. Yet, these instances had to be scored as incorrect for the sake of the response definition. Thus several responses that were scored as incorrect involved a difference of only millimeters. If a less stringent response definition had been used such as a movement of the lid past the center of the cup, much higher accuracies would have been reached in much fewer sessions.

A considerable limitation of the current procedure was the use of a manual, non-automated apparatus. As with any procedure, automation provides better experimental control.
and reduces the likelihood of experimenter error or cuing. However, the inter-rater reliability data collected support the validity of the present results. Although attempts were made to standardize several aspects of the procedure, there are features that may have been inconsistent. For instance, although intertrial intervals (ITI) were standardized, it is possible that there was variation within trials. Because the sample and comparison trays were manually inserted into the chamber, the speed at which they were presented may have varied across subjects or within sessions. Also, although this procedure was not explicitly a delayed MTS procedure, the lack of automation inevitably creates a delay in between the sample and comparison. Efforts were made such that this delay was kept as close to 0-sec as possible.

The substantial amount of variability within sessions to criterion across subjects in the 2-stimuli condition may be attributable to the use of randomly selected stimuli. Pairs of scents to be used in training for rats in the 2-stimuli condition were randomly selected from the set; each rat was trained with different baseline scents. It is possible that some scents are more similar to each other or more differentiable than others, and therefore some pairs may be more easily discriminated. Perhaps training each rat with the same stimuli would allow for better comparisons. Further, the use of olfactory stimuli presents the issue of difficulty in controlling the quantity of the stimuli.

The ability to behave conceptually provides the means to learn efficiently. That is, past learning can be applied to a new situation and by doing so, re-learning and adapting to each new situation is not necessary (Zentall et al., 2008). Concepts are valuable for solving problems and adjusting to new circumstances by learning universal relationships, rather than the precise features of a single situation. Comparative research has clearly shown that the ability to behave conceptually is not unique to humans, with evidence of conceptual behavior demonstrated in a
variety of species. Further, the present results add to the evidence that tasks such as matching-to-sample may be solved either relationally or by stimulus-specific learning, and that experimental parameters may influence which type of learning takes place.

A theoretical implication of this field of research is the possibility that the cognitive abilities of humans and animals have a common origin— that we expect to see a continuum of learning across species because ancestral species showed relational learning that was critical to survival. The mechanisms by which animals learn these tasks may be shared by the way humans learn to solve problems, and investigating how these processes differ across species may advance our understanding of animal and human cognitive processing (Katz et al., 2007). Another possibility is that similar performance may be an example of convergent evolution (Lazareva & Wasserman, 2008). Studying differences in concept-learning ability across species may provide evidence for the possibility of evolutionary differences in the basic mechanisms necessary to perform these higher-order tasks (Cook & Wasserman, 2007; Wright & Katz, 2007).
LITERATURE CITED


significant impairments on spatial span, recognition, and alternation. *The Journal of Neuroscience, 20*(8), 2964-2977.


APPENDIX

APPENDIX A. Session-by-session acquisition of MTS or NMTS task for individual subjects.
T15 (MTS-2)

V4 (MTS-2)

T11 (NMTS-2)
S3 (NMTS-10)

100
90
80
70
60
50
40
30
20
10
0

Percent Correct

Consecutive Sessions

p = 1

p = .75

p = .5

(8/10)

(10/10)
APPENDIX B. Analysis of Variance tables.

### Analysis of Variance of Transfer Test

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfer</td>
<td>1</td>
<td>3.919</td>
<td>0.071</td>
</tr>
<tr>
<td>Transfer*Set-size</td>
<td>1</td>
<td>7.119</td>
<td>0.02*</td>
</tr>
<tr>
<td>Transfer*Procedure</td>
<td>1</td>
<td>0.67</td>
<td>0.801</td>
</tr>
<tr>
<td>Transfer<em>Procedure</em>Set-size</td>
<td>1</td>
<td>0.007</td>
<td>0.933</td>
</tr>
</tbody>
</table>

### Analysis of Variance of Between-subjects Effects

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>593.968</td>
<td>0</td>
</tr>
<tr>
<td>Set-size</td>
<td>1</td>
<td>10.602</td>
<td>0.006*</td>
</tr>
</tbody>
</table>

### Analysis of Variance of Within-subjects Contrasts

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfer test</td>
<td>1</td>
<td>5.15</td>
<td>0.040*</td>
</tr>
<tr>
<td>Transfer test*Set-size</td>
<td>1</td>
<td>9.356</td>
<td>0.009*</td>
</tr>
</tbody>
</table>

### Analysis of Variance of Acquisition

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set-size</td>
<td>1</td>
<td>0.232</td>
<td>0.639</td>
</tr>
<tr>
<td>Procedure</td>
<td>1</td>
<td>0.013</td>
<td>0.912</td>
</tr>
<tr>
<td>Set-size*Procedure</td>
<td>1</td>
<td>0.129</td>
<td>0.726</td>
</tr>
</tbody>
</table>