

DIFFERENTIAL OUTCOMES FACILITATE RELATIONAL ASSOCIATIONS

Except where reference is made to the work of others, the work described in this thesis is my own or was done in collaboration with my advisory committee. This thesis does not include proprietary or classified information.

Kelly Ann Schmidtke

Certificate of Approval:

Martha C. Escobar
Associate Professor
Psychology

Jeffrey S. Katz
Associate Professor
Psychology

Lewis Barker
Professor
Psychology

Joe F. Pittman
Interim Dean
Graduate School

DIFFERENTIAL OUTCOMES FACILITATES RELATIONAL ASSOCIATIONS

Kelly Ann Schmidtke

A Thesis

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Master of Arts

Auburn, Alabama
May 10, 2008

DIFFERENTIAL OUTCOMES FACILITATES RELATIONAL ASSOCIATIONS

Kelly Ann Schmidtke

Permission is granted to Auburn University to make copies of this thesis at its discretion, upon request of individuals or institutions and at their expense. The author reserves all publication rights.

Signature of Author

Date of Graduation

THESIS ABSTRACT

DIFFERENTIAL OUTCOMES FACILITATES RELATIONAL ASSOCIATIONS

Kelly Ann Schmidtke

Master of Arts, May 10, 2008
(B.A., University of Minnesota-Twin Cities, 2005)

72 Typed Pages

Directed by Jeffrey S. Katz

The Differential Outcomes Procedure facilitates the acquisition and performance of tasks that encourage item specific strategies, but its affects on relational strategies is unknown. Here, 8 pigeons (*columba livia*) were trained to use a relational strategy with the abstract concept same/different. Of these pigeons, 4 received differential outcomes and 4 received same outcomes. Both groups acquired the tasks at equal rates but the pigeons that received differential outcomes evidenced partial concept use sooner than those pigeons that received the same outcome.

Style manual used: Publication Manual of the American Psychological Association
(5th edition)

Computer software used: Microsoft Word 2000®
Microsoft Excel 2000®
SPSS 11.5®
Visual Basic

TABLE OF CONTENTS

LIST OF FIGURES	viii
CHAPTER 1. INTRODUCTION	
Classic Experiments	2
Extensions of DOE	8
Comparing Item Specific and Relational Strategies	9
Gaining Conditional Control over Relational Strategies	12
The DOE in Cook's Two-Hopper Procedure	19
A Better Suited Same/Different Procedure	21
CHAPTER 2: EXPERIMENT	
Introduction	25
Methods	
Participants	26
Apparatus	26
Procedure	27
Results	31
Discussion	44
REFERENCES	49
APPENDIX A- SPSS OUTPUTS FOR SIGNIFICANT ANALYSES	54

LIST OF FIGURES

Figure 1	Brodigan and Peterson's experimental design	4
Figure 2.	Sample displays from the Cook, Cavoto & Cavoto (1995) same different Experiment	12
Figure 3.	Sample displays from Cook, Katz, Cavoto's (1997) same different experiment	14
Figure 4	Sample displays from Cook's (2002) same/different experiment	15
Figure 5	Sample Displays from Young, Wasserman, and Garner (1997) same/different experiment	17
Figure 6	Sample displays From Katz and Wright's (2006) same/different experiment	22
Figure 7	Pigeons performance between baseline and transfer and across set sizes	23
Figure 8	Flow chart of both a same and different trial in the present experiment	29
Figure 9	Each bird's average reaction times for same and different correct, first choices, within each session, over the 8 item set-size acquisition.	32

Figure 10	Each bird's average reaction times for same and different correct, first choices, within each session, over the 512 item set-size acquisition.	33
Figure 11	The left graph displays the mean difference reaction times for the last three sessions in the 8 item set-size acquisition for each bird, and the right graph shows the same information by outcome group. Error bars represent the standard error of the mean.	34
Figure 12	The left graph displays the mean difference reaction time for the last three sessions in the 512 item set-size acquisition for each bird, and the right graph shows the same information, but by type group. Error bars represent the standard error of the mean	35
Figure 13	The number of trials to criteria across set-sizes for individual birds. The Y-axis was held constant for each bird to make visual comparisons easier. For Nickolai's 64, 128 and 512 acquisition the trials to criteria were 7700, 7300, and 10700 respectively.	37
Figure 14	The average number of trials required to acquire each set-size by groups. Bars represent the standard error of the mean	38
Figure 15	Individual performance on baseline and transfer trials by set size. Closed circles represent the baseline performance and open circles represent transfer performance. The bars represented the standard error of the mean.	40

Figure 16	The average performance at each set size for baseline trials. The bars represent the standard error of the mean.	41
Figure 17	The average performance at each set size for transfer trials. The bars represent the standard error of the mean.	42
Figure 18	The average percent difference (baseline trial percent correct – transfer trial percent correct) across set sizes. The bars represent the standard error of the mean.	43

CHAPTER 1

LITERATURE REVIEW

In the learning of conditional discriminations, the differential outcome effect (DOE) describes the increased acquisition rate, retention and/or terminal accuracy that occurs when each sample and its comparison are correlated with a unique outcome (Edwards, Jagielo, Zentall & Hogan, 1982; Kruse, Overmier, Konz & Rokke, 1983; Maki-Kahn, Overmier, Delos & Gutmann, 1995). This report will present an early history of the DOE in the learning of stimulus specific conditional discriminations and then describe an experiment using the differential outcomes procedure to teach a relational discrimination. If the DOE has the same effect for training a relational discrimination as it does for stimulus-specific discriminations, then it would be a powerful training procedure for enhancing abstract concept learning.

Furthermore it would provide evidence to the ability of pigeons to form and use relational strategies to guide choice behavior as well as provide support for expectancy formations, which will be described later (Desmarse & Urcuioli, 2005; Trapold 1970). Finally, the differential outcomes procedure might not be limited to pigeons in operant chambers, but could be extended to humans, and be used to create more effective real-world training procedures.

Classic Experiments

One of the earliest publications describing the DOE was an experiment done in 1970 by Trapold (1970). Trapold's rats learned conditional discriminations that involved hearing a clicker or tone and then making a choice response to a right or left lever. In a between groups design he manipulated whether the rats received one outcome for both choices or a particular outcome for each choice. The different outcomes group received consistently different outcomes for each correct response. For example, upon hearing a clicker the rat would press the right lever and receive pellets; if the rat heard the tone, it would press the left lever and receive sucrose. Two control groups were employed that received consistently similar outcomes, either pellets or sucrose for both correct choices. To compare the groups' performance, Trapold measured acquisition rate and terminal accuracy.

The different outcomes group acquired the discriminations faster and maintained higher terminal accuracy, over 21 sessions, than either control group. Trapold believed that the rats that experienced differential outcomes developed different expectancies linked to particular sample stimuli. The ability to expect a particular outcome upon hearing a sample provided the different outcomes group with an additional cue that the same outcome groups lacked.

Brodigan and Peterson (1976) replicated and extended Trapold's (1970) work. One goal of their study was to determine whether the DOE Trapold found was the result of expectancies or attention. In Trapold's experiment, either factor may have mediated the different outcomes group's performance. The *expectancy hypothesis* states that it was

the predictive relationship between the stimuli and the outcomes that mediated the DOE. That is, each sample was predictive of a particular outcome. For example, pressing the right lever after the clicker always resulted in pellets and pressing the left lever after the tone always resulted in sucrose. In contrast, the same outcomes groups could not use the samples to predict a particular outcome, because both discriminations resulted in the same outcome. The *differential attention hypothesis* states that having more than one outcome might increase the subjects' attention to the stimuli that precede them, regardless of the predictive relationship between the sample and outcome.

To test the differential attention hypothesis, Brodigan and Peterson used a control group that received inconsistently different outcomes. These pigeons could not use expectancies to predict a particular outcome because each sample had the same probability of leading to either outcome. If the consistently different outcomes group acquires the discrimination faster and with higher terminal accuracy than the inconsistently different outcomes group, then the differential attention hypothesis must yield to the expectancy hypothesis.

Brodigan and Peterson (1976) split their pigeons into four groups in a two by two factorial design. The first factor was the pigeons experience in the world, wild or domestic and the second factor was the outcome condition, consistently different outcomes or inconsistently different outcomes. All the pigeons learned the same conditional discriminations with visual stimuli. Figure 1 depicts of the outcome contingencies. For example, for the consistently different groups, when a red stimulus appears, the pigeon should choose a vertical line to receive water. When a green stimulus

appears the pigeon should choose a horizontal line to receive food. For the inconsistently different outcomes groups, if a red stimulus appears the pigeon should choose the vertical line to receive randomly water or food. The same contingency was in place for the green stimulus. If the differential attention hypothesis is correct, then there should be no differences between the groups' performance. However, if the expectancy hypothesis is correct, then the consistently different outcomes group should perform better than the inconsistent group.

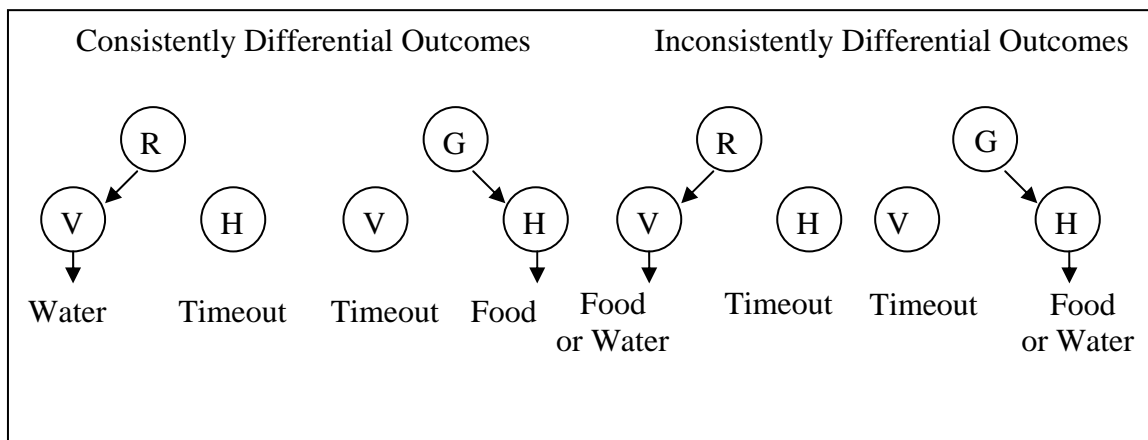


Figure 1. Brodigan and Peterson's (1976) Experimental Design

In addition to measuring acquisition rate and terminal accuracy, they also looked at retention. They hypothesized that if the expectancy hypothesis was correct, then the consistently different outcomes groups should perform more accurately over longer delays between a sample offset and the comparisons onset than the inconsistently different outcomes groups, by facilitating a memory of either the sample or comparison. Three different delay durations were used (0 s, 3 s, 15 s), and choice accuracies were measured at each delay.

Another variable of interest was the response topography. If the expectancy hypothesis is correct, then the consistently different outcomes groups should anticipate the outcome and accordingly prepared to consume it (Jenkins & Moore, 1973). Specifically, the consistently different outcomes groups should peck softly after the discrimination followed by water and peck sharply after the discrimination followed by food. In contrast, the pigeons in the inconsistently different outcomes groups should not be able to anticipate the outcome and therefore would be unable to prepare specific response topographies.

Their results corroborated the expectancy hypothesis. Regarding acquisition rate and terminal accuracy, the consistently different outcomes groups learned the discriminations faster and reached higher terminal accuracies than the inconsistently different outcomes groups. As for retention, the consistently different outcomes groups showed greater retention over 0 s and 3 s, delays than the inconsistently different outcomes groups. The wild pigeons that experienced consistently different outcomes showed greater retention at the longest delays. That the wild but not domestic pigeons showed enhanced performance at the longest delay suggests the differential outcome procedure's effect on retention might depends on life experience. (What "life experience" exactly entails is not known, but it is an interesting factor that should be examined.) Lastly, the topography of responses, as observed descriptively by the experimenters, differed only for the consistently different outcomes groups; pigeons scooped softly for water and pecked sharply for food. Like Trapold (1970), Brodigan and Peterson (1976) concluded that expectancies are the mechanisms that underlie DOE.

Extensions of the DOE

Learning mechanisms provide a law-like structure through which a class of events is explained. The larger the class of events is, the more powerful that mechanism is thought to be. The DOE has been successfully demonstrated with different species, outcome types, and methodologies. This suggests that expectancies may be a general learning mechanism.

The DOE has been observed in a wide range of nonhuman species, such as, rats (Carlson & Wielkiewicz, 1976; Trapold, 1970), pigeons (Brodigan & Peterson 1976; Urcuioli, DeMarse & Lionello-DeNolf, 2001), dogs (Overmier, Bull & Trapold, 1971) and horses (Miyashita, Nakajima & Imada, 2000). The DOE has also been demonstrated in various human populations; such as, language-deficit children (Janssen & Guess, 1978; Hewitt, 1965), typical children (Estevaz, Fuentes, Mari-Beffa, Gonzalez & Alvarez, 2001), typical adults with alcohol induced amnesia (Hochhalter, Sweeney, Bakke, Holub & Overmier, 2000), and recently with typical college students (Miller, Waugh & Chambers, 2002).

The DOE also emerges with a variety of outcome types. The outcomes used are not restricted to reinforcers with different qualitative natures, like food and water. Different quantities of one reinforcer also yield a DOE, such as one versus five food pellets (Carlson & Wielkiewicz, 1976). Other experiments have used one biological and one conditional reinforcer, such as food and a hopper light (Urcuioli, DeMarse &

Lionello-DeNolf, 2001). Some experiments, particularly those with humans, use only conditional reinforcers, like different colored tokens (Estevez, Fuentes, Overmier & Gonzalez, 2003).

The DOE also emerges with a variety of methodologies. Most experiments have followed the between subjects design example given by Trapold (1970) and Brodigan and Peterson (1976). However, Alling, Nickel and Poling (1991a) found the DOE using an ABAB methodology. When pigeons were in the differential condition their performance exceeded that of even their highest performance in the nondifferential condition. This experiment offers compelling evidence that the DOE is a product of environmental contingencies that can be manipulated at the experimenter's discretion.

In summary, the differential outcomes procedure enhances learning across species, reinforcers, and methodologies, suggesting that expectancies are a general learning mechanism. The benefits of using the differential outcomes procedure encourages its use where and when it can be employed. To employ the differential outcomes procedure the task must include more than one thing to be learned simultaneously. It might be argued that there is rarely a case in the applied world where only one behavior is targeted for improvement. So, with a little ingenuity, training procedures that have traditionally used same outcomes could be modified to include different outcomes.

Comparing Item Specific and Relational Strategies

When subjects use an *item specific strategy*, choice responses are bound by the absolute properties of the training stimuli. For example, subjects may learn a matching-to-sample discrimination between red and green: if the sample is red then they should choose the red comparison, but if the sample is green choose the green comparison. To respond correctly, subjects can memorize the two types of item specific associations (or four if the configurations control behavior). However when a novel stimulus is presented, a yellow sample, subjects that memorized two concrete rules will be unable to choose the correct yellow comparison. By contrast, when subjects use *relational strategies*, also called abstract concepts in which the subject response to the relationship between two present stimuli, their choice responses can transfer beyond the training exemplars. Given the aforementioned matching-to-sample task, if subjects learned to use a relational strategy, identity, then they will transfer that relation to the novel yellow stimulus. The ability to transfer relational rules makes procedures that encourage relational strategies more powerful than those that encourage item specific strategies.

Although sparse, there have been attempts to use the differential outcomes procedure to train populations with cognitive disabilities on conditional discriminations. For example, Estévez and Fuentes (2003) and Estévez, Fuentes, Overmier and González (2003) used a differential outcomes procedure to successfully improve the performance of children and adults with Down syndrome on a conditional discrimination task. These subjects learned conditional discriminations like, if the sample is a plus sign then choose

the star. These experiments and those with nonhuman animals used artificial stimuli and subjects appeared to have used item specific strategies. Thus, what the subjects learned was bound to the stimuli with which they were trained. Furthermore, what they learned may not even be useful or make sense in the real-world.

There are many practical reasons to use artificial stimuli. For example, artificial stimuli are easy to equate for perceptual complexity and their experimental pairings have little historical relevance for the participants. These factors allow the experimenter to be more certain the obtained results are caused by the arranged experimental contingencies.

Unfortunately, reliance on artificial stimuli may create an illusion that the differential outcomes procedure is only useful for studying underlying mechanisms of laboratory learning, and that it has little relevance for the real-world. In the real-world, responses are often based on relational strategies with natural stimuli. For instance, when teaching number concepts, a teacher may give a child one piece of candy and ask them “How many pieces of candy will you have if I give you another?” If the child responds, “two” then the child is correct. However, unless the child spontaneously transfers this response, the teacher cannot conclude the child learned a relational association. The absence or presence of a relational strategy is a noteworthy distinction because the child’s ability to use the item specific strategy to identify one piece of candy is not nearly as useful as the child’s ability to use the relational association to identify any single object.

Because the differential outcomes procedure has proven an efficacious procedure to train item specific strategies, it may also be advantageous to use it to train relational strategies. One such relation is the same/different concept.

William James wrote that the “sense of sameness is the very keel and backbone of our thinking.” Just sentences later, James emphasizes that “our” means mankind’s. He continues, “Not all psychic life need be assumed to have the sense of sameness developed in this way. In the consciousness of worms and polyps, though the same realities may frequently impress it, the feeling of sameness may seldom emerge,” (1950/1890, p. 459). However, René Descartes had earlier proposed that relational associations, like same/different, are dependent on language. Therefore, Descartes believed, language-deficient animals could not have such thoughts (Descartes, trans. 1646, p. 207). Unfortunately, Descartes and James spoke more philosophically than experimentally on these matters.

Premack (1983a) offered empirical support for Descartes’ claim, by showing that a language-trained primate could use relational strategies that were beyond the reach of her language-deficient conspecifics. These results suggest that subjects’ linguistic ability is a critical factor for their learning to use a relational strategy. However, Premack noted (1983b) that the inability of animals to demonstrate a same/different discrimination may be an artifact of the experiment and not the animals’ intelligence.

Premack (1983b) observed that when novel items were presented with familiar items in a matching to sample task, his chimps choose the familiar item regardless of the sample. He reasoned that,

“While learning to choose the alternative that matched the sample, the animal also learned to choose toys, i.e. a class of items with certain properties. Learning does not take place only on the relational (or absolute) level but on both levels. If apes and pigeons differ in this regard, it may be in the relative weights the two species assign the factors; the relational factor being weighted more heavily in the ape... Although apes learn on both levels, I have not been able to gain conditional control of the two levels and thus, in effect instruct the ape to “Pay attention to the relations” or “Pay attention to the details”” (Premack, 1983b p. 356).

This passage suggests that language-deficient pigeons might have the ability to learn a relational strategy, but that relational strategies are seldom observed because pigeons have a predisposition to use item specific strategies. Thus, it might be more difficult for experimenters to gain conditional control over relational than item-specific strategies.

Gaining Conditional Control over Relational Strategies

Cook (2002) lists five criteria to determine if a participant is using a relational strategy:

1. Accurate discrimination with training items.
2. Accurate discriminations with novel items.
3. Accurate discrimination occurs across different training items.
4. Accurate discrimination of novel items and training items.
5. Alternative strategies are ruled out.

Although all of these criteria are important, the second, accurate discrimination of novel items is most often cited. The following experiments are presented to portray how difficult it is, even for Cook himself, to satisfy his criteria. Furthermore, as will be discussed at the end of this section, these experiments provide a good way to start thinking about how the differential outcomes procedure might be applied to relational strategies.

Cook, Cavoto and Cavoto (1995) attempted to teach pigeons a same/different discrimination with visual textures. On same trials, a display contained 81 instances of the same small colored shape, while on different trials a display contained a target and a distracter region with a contrasting color or shape. Figure 2 gives examples of these displays. In total there were 81 same and 1,296 different displays. When a trial began, a display appeared for a fixed time of 4 s, followed by the illumination of two side hoppers. Each hopper was tied to a particular relation, same or different. The pigeons indicated their choice by placing their head into one of the hoppers. Head entries were detected by a LED sensor. A correct choice raised the hopper making mixed grain accessible for 2 s. An incorrect entry resulted in a timeout.

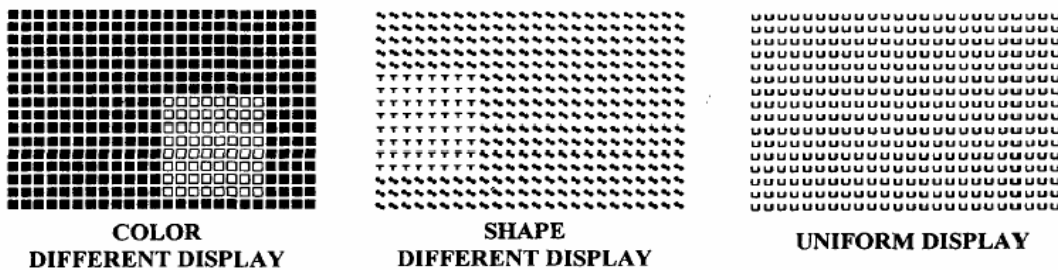


Figure 2. Sample displays from the Cook, Cavoto & Cavoto (1995) Same/Different experiment.

Adjustments were made, because the “birds appeared to not learn with this initial procedure” (Cook et al., 1995, p.255). The first modification provided different types of food in the left and right hoppers, peanut hearts and safflower seeds (i.e. differential outcomes). Accuracy did not increase. Later a change replaced the fixed time display interval with a FR 10 requirement to the screen. Accuracy remained low. Lastly, an adjustment was made to the nature of the FR requirement. Now, instead of just 10 pecks anywhere on the different display, the pigeons were required to peck the target region 5 times. With this adjustment, the pigeons were able to successfully discriminate the training exemplars (Cook’s first criterion).

Two transfer tests were done to rule out use of an item specific strategy. The first set of tests included 16 displays that contained a combination of one of two novel colors and shapes (Cook’s fourth criterion). The second set of tests contained displays that combined the features used in the former set of test displays with those used in training. Although performance accuracy with the tests displays was below baseline, it was above chance (Cook’s second criterion). This accurate performance with novel stimuli helps to rule out an item specific strategy and provides support for the relational strategy.

One might stop there, and assume that the pigeons were using a relational strategy to discriminate same and different. However, Cook’s fifth criteria must be examined. An alternative explanation is that the pigeons used a lower level perceptual mechanism. The *perceptual mechanism hypothesis* posits that the pigeons made their choices by recognizing perceptual differences within the displays. Cognitively speaking, using a perceptual mechanism is a lower level process than using an abstract concept. A

perceptual difference within the different feature displays is at the edges between the target and distracter regions, which appear in both the training and test displays. This means that the test trials were not seen as novel by the pigeons (Cook's second criteria).

To address this alternative explanation another experiment was performed. This experiment used the same two-hopper procedure, but included additional display types so no single perceptual feature could maintain accuracy (Cook, Katz & Cavoto, 1997). Four different types of displays were concurrently examined; texture, feature, geometric, and object (see Figure 3). If the pigeons were using a perceptual feature to guide their choices, they should have found it more difficult to discriminate the geometric and object displays than the feature and geometric displays, because the geometric and object displays have less distinct edges. Their results showed that despite the differences between the display types, acquisition across the displays was similar. This suggests the pigeons were using a single strategy, such as same/different, that could be applied consistently to all display types.

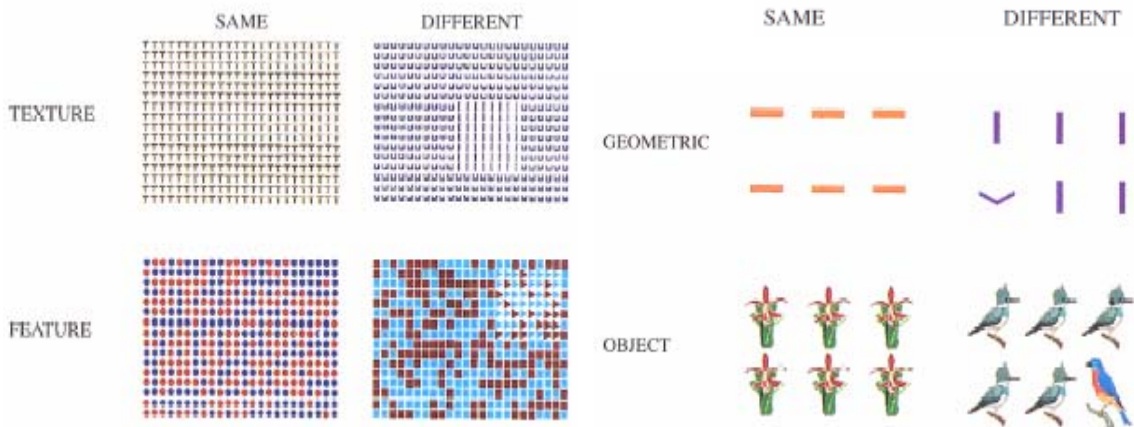


Figure 3. Sample Displays from Cook, Katz, Cavoto's (1997) Same/Different Experiment

Further negating the perceptual mechanism hypothesis, in 2002b, Cook used realistic photographs, in addition to texture, geometric, and objects displays, to train the same/different discrimination. Examples of photograph displays are shown in Figure 4. The same two-hopper procedure was used, except this experiment included not only transfer tests to differentiate whether the pigeons employed a relational or item specific strategy, but an inconsistent correct response location control group over acquisition.

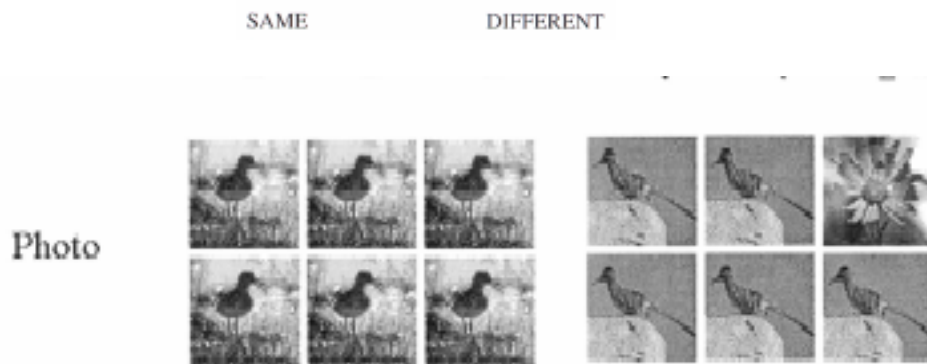


Figure 4. Sample Displays from Cook's (2002) Same/Different Experiment

For the consistent correct response location group, the relational aspects of the displays were tied to a particular hopper (e.g., if same then left; if different then right). To respond correctly, the consistent group could use the relational association. For the inconsistent correct response location group, each display itself was tied to a particular hopper. Half of the same displays required the pigeon to enter the left hopper and the remaining same displays required pigeons to enter the right hopper (e.g., if six red birds then left hopper; if six blue birds then right hopper). The different displays' correct hopper locations were split in a similar manner. To respond correctly the inconsistent group had to use item specific strategy.

If the consistent and inconsistent groups learned the task at the same rate and both failed to transfer to novel displays, then one could reason that both groups used an item specific strategy. However, if the learning rates differed and only the consistent group transferred to novel displays, this would provide evidence that different strategies were employed by each group. Specifically, it would suggest, the consistent group used a relational strategy and the inconsistent group used an item specific strategy.

Not only did the consistent group learn the task faster than the inconsistent group, but the consistent group transferred to novel displays, whereas the inconsistent group did not. These results support the notion that the consistent group used a relational strategy.

Although these experiments provide evidence that the pigeons need not rely on a perceptual mechanism to learn a same/different task, other alternative explanations still exist. For example, Young, Wasserman, and Garner (1997) noted that if the pigeons' choices resulted from a relational strategy of same/different, then the pigeons' accuracy should not depend on the number of items in the display. That is, an alternative explanation for the pigeons' performance is that their responses in a same/different task are based on the amount of variability present in a particular display, called entropy. The *entropy hypothesis* posits that as the variability within a display decreases, the tendency to treat a different display as if it were a same display increases.

Young et. al.'s (1997) tested the entropy hypothesis. In the initial phase of their experiment, they trained pigeons to successfully discriminate between displays with 16 same and 16 different items where, according to their equation, entropy equals 0 and 4 respectively. In the next phase, they presented displays with fewer items, 12, 8, 4 and 2,

where the entropy for different displays decreases as they contain fewer items, 3.6, 3, 2, and 1. The entropy for same displays remains 0 regardless of the number of items.

Examples of these displays are presented in Figure 5.

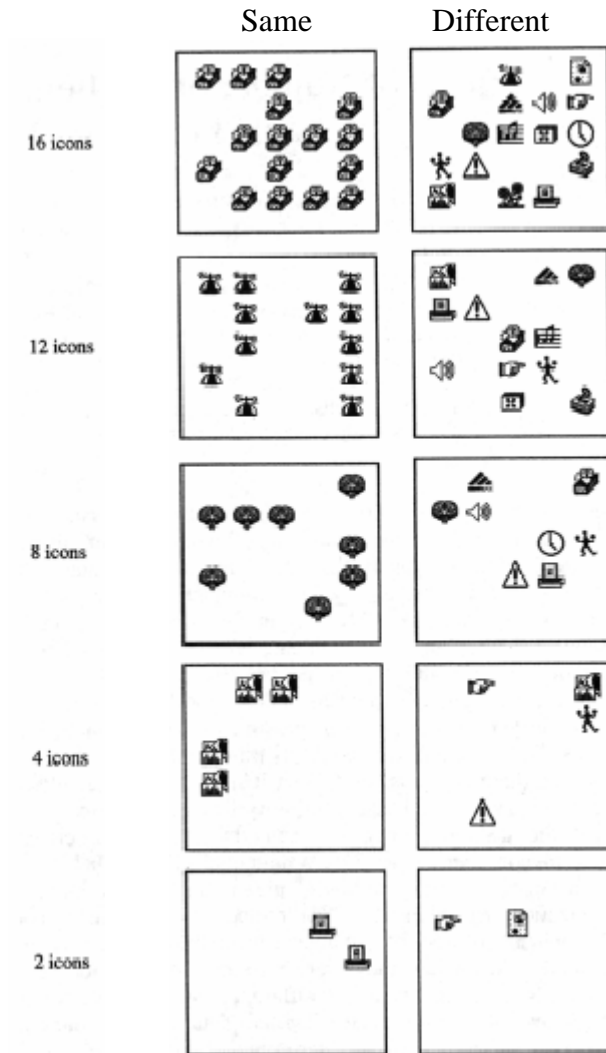


Figure 5. Sample Displays from Young, Wasserman, and Garner (1997) Same/Different experiment.

Using either the entropy within the display or same/different relation, the pigeons should treat same displays as "same." However, if the entropy hypothesis is correct, as the number of items in the display decrease, the likelihood of classifying a different

display as same should increase. This was in fact the case. In Young, Wasserman, and Garner's (1997) experiment, pigeons treated displays with two items exclusively as if they were same displays. These results suggest that when pigeons are trained with multiple item displays their responses may not be based on the relational association of same/different, but instead on the entropy of total display.

In 1997, Cook et al. addressed the entropy hypothesis. They point out that the displays used in their task differ from those used by Young et al. (1997). That is, all of the Cook et al. displays contain only 1 distracter area or item; whereas, Young et al.'s largest displays contain 16 distracter items. This may make entropy a more salient feature for the Young et al. displays than Cook et al.'s. Furthermore, the predictions made by the entropy hypothesis do materialize in the Cook et al. experiment. Due to local variation among elements within feature displays, Cook et al.'s same feature displays have entropy of one. The other three same display types have entropies ranging from 0.59 - 0.65. The entropy differences between display types should have hindered the pigeons' ability to discriminate the feature displays relative to the other types. However, all the display types were acquired at the same rate, which supports the view that a single strategy could be applied equally across the types.

Cook (2002b) suggested that the pigeons might respond with regard to the presence or absence of an odd item (i.e., a concept of oddity). In fact, requiring pecks to the target region may encourage such a concept. Responding based on oddity or consistency would be a relational strategy that applies equally to all the displays. However, that relational strategy does not exemplify a same/different discrimination.

Many of the above criticisms pertain not to the pigeons but the procedure used to train or test the pigeons. Like Premack, these critics pick out something about Cook's procedure (stimuli, hoppers, FR) that may hinder his ability to bring out a same/different relational strategy. Instead, the above experiments might produce a different type of relational strategy. Strategies that are guided by perceptual mechanisms or the entropy of the entire display are relational, but they are not the same as strategies guided by the same/different concept. Perhaps the procedure itself must be revamped if one hopes to observe the use of a relational strategy and conclusively state the presence of the same/different concept.

The DOE in Cook's Two-Hopper Procedure

The Cook et al. (1995) experiment the only same/different experiment that makes explicit use of the differential outcomes procedure. A series of tests were run to assess the contributions of the DOE on the pigeons' same/different discriminations. This testing took place in two parts. First, the location of the peanut hearts and safflower seeds were reversed. Second, both types were replaced with mixed grain. Neither the reversal nor removal of the different outcomes decreased the pigeons' accuracy. This led Cook et al. (1995) to assert that the differential outcomes treatment did not play a role in the pigeons' same/different discriminations. However, at least four points can be made that emphasize why one cannot conclusively say that the differential outcomes procedure did or did not affect the pigeons' performance.

First the experiment lacks an acquisition control group, which would experience either the same or inconsistent outcomes over acquisition. That is, there is no way to

assess if the differential outcomes may have enhanced, or hindered acquisition of the discrimination. Second, the tests of the DOE lasted two sessions. The differential expectancies that may promote acquisition over many training sessions may have disappeared in two sessions. Such a within subject analysis may need to include a longer removal of the differential outcomes to see such effects.

Third, although the differential food types were removed, the locations of the correct choices remained consistent with their respective relation, same or different. The consistently different locations may have acted as differential outcomes procedure. If this were true, then the Cook et al. (1997) and Cook (2002b) experiments also included differential outcomes. Justifiably, one might note that when only hopper location was available as a differential outcome the pigeons failed to discriminate same and different (Cook et al., 1995). Accuracy only increased after the addition the FR 5 requirement to the target region. However, that addition is confounded with the increasing number of training sessions. It is also possible that the additional adjustments increased the salience of the consistently different locations. However, lack of a control group leaves one unable to make definite statements.

Fourth, and most troubling for assessing any affects of the differential outcomes procedure, the differential responses are fused to the different locations rendering a good control group over acquisition impossible. A same or inconsistent outcome group would not provide an appropriate control condition, because to demonstrate the same/different discrimination there must be two ways to respond. An inconsistently different correct response group was used as a control in Cook (2002b), and it was useful to determine

whether the pigeons are using an item specific or relational strategy. But it did not control for the effects of differential outcomes on a relational associations alone. To assess the impact differential outcomes may have for promoting relational strategies it is necessary to separate how the responses are made from the outcomes that follow.

A Better Suited Same/Different Procedure

Wright and Katz (2006) questioned not only whether nonhuman animals can use a relational strategy, but also how fully such a strategy is employed. They note that once a concept is fully formed, there should be no decrement in performance when novel stimuli are introduced. While transfer above chance but below baseline (like that observed in Cook et al., 1995; 1997; 2002) provides evidence for a same/different concept, that concept is not the only thing that controls choice behavior. At best, one can only say a partial concept exists. A full concept is evidenced by statistically similar performance between the training and novel test stimuli.

Wright and Katz (2006) suggest that the decrement typically observed in transfer tests may be due to procedural restrictions in the number of discriminable training exemplars. Such is most often the case with artificial stimuli, as there are typically only so many different combinations of shapes or colors. A solution would be to create a procedure that includes naturalistic stimuli, which are limited only by the amount the experimenter is willing to collect.

Katz and Wright (2006) attempted to generate a full same/different concept by using increasingly more training exemplars. They alleviated past restrictions on the number of discriminable training exemplars by using naturalistic stimuli, colored

photographs. On a given trial, the pigeons first saw a picture at the top of a computer monitor. The pigeons pecked this picture and a second picture appeared below with a white rectangle to the right. Figure 6 offers examples of these displays. To respond correctly, when the second picture was the same as the top picture the pigeons pecked the bottom picture and when the second picture was different the pigeons pecked the white rectangle. After acquiring the discrimination, the pigeons were tested with novel pictures.



Figure 6. Sample Displays From Katz and Wright's (2006) Same/Different Experiment

They hypothesized that the number of exemplars with which the pigeons were trained would predict the pigeons' ability to apply the relational strategy, the same/different concept, to novel items. While a concept may only be partial at a small set size, as more photographs are added to the training set a full concept should emerge. When tested, if the pigeons had learned the relational association, they should perform on baseline and novel displays with equal accuracy. In contrast, if the pigeons had learned to use an item specific strategy, accuracy on novel displays should drop to chance levels. Performance between chance and baseline would indicate a partial concept.

The pigeons were initially trained with a small set size of 8 different photographs. After reaching criterion, they were tested over six sessions each containing 90 baseline and 10 novel displays. In these test sessions, accuracy with novel displays fell to chance.

while baseline accuracy remained high, indicating reliance on the item specific strategy. After the transfer tests, the number of photographs used in the next series of training sessions was expanded to 16 and then 32 items. When the pigeon met criterion on the 32 item set, the next six test sessions commenced. The pigeons preformed better than chance on these transfer tests, 59.6%, but the accuracy was below baseline.

The set size was expanded in a similar manner from 32 to 64, 128, 256, 512 and 1024 photographs, after reaching criteria on each set size the pigeons were tested. The greater the set size used the more accurate transfer became (see Figure 7). At the next intermediate set sizes, 64 and 128, transfer performance remained between baseline and chance. This indicates that at the intermittent set sizes the pigeons had developed a partial same/different concept. However, at the largest set sizes, 256, 512, and 1024, there was no statistical difference between baseline and transfer performance, indicating that a full concept of same/different had emerged. The set size expansion procedure is better suited to study relational associations than past procedures because it allows one to observe how the use of a concept develops from its partial to full form.

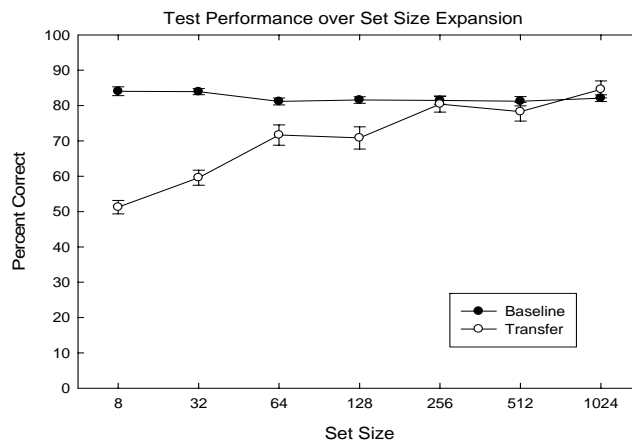


Figure 7. Performance between Baseline and Transfer and across Set Sizes

Since the differential outcomes procedure increases the speed of acquisition for conditional discriminations, one might wonder, if Katz and Wright (2006) had used differential outcomes procedure, would the pigeons have transferred to novel stimuli at a smaller training set size. The Katz and Wright procedure is better suited than Cook's two-hopper procedure to study the affects of the differential outcomes procedures on a same/different discrimination, because it is possible to separate the choice responses from the outcomes. The choice responses are made to the screen, to either the bottom picture or a white box, and the outcomes are delivered at a single hopper. The outcomes could be different if the duration of food was manipulated. If the differential outcomes procedure facilitates concept formation, it would expand the interpretation of outcome expectancies from item specific strategies to include relational strategies. However if the differential outcomes procedure only enhances the memory of particular items, the emergence of the same/different concept may require a larger set size than observed with same outcomes (i.e., block relational learning).

CHAPTER 2

EXPERIMENT

Introduction

In the present experiment, different outcomes were associated with correct responses to a relational association, same/different. There were two primary groups, a same outcome group and different outcomes group. In the same outcome group, the same duration of food was accompanied by the same tone provided after each correct response. For the different outcomes group, grain was made available for 5 s while sound one played for one type of correct choice and 1 s along with sound two for the other type of correct choice. The different outcomes group was divided into subgroups, where one received long access for same choices and one received long access for different choices. The response times, acquisition rate and accuracy in baseline and transfer across set-sizes were compared.

If the differential outcome procedure enhances only item-specific strategy use, then the pigeons that received different outcomes should learn the task faster, but the transfer to novel pictures should take longer. In contrast, if the differential outcome procedure can enhance relational strategy use, then the pigeons that received different outcomes should learn the task at the same rate, but the transfer to novel pictures should occur faster.

The following experiment has three hypotheses. The first hypothesis states that if pigeons use expectancies to guide choice behavior, then those expectancies should be reflected in choice response latencies. The second hypothesis states that if the differential outcome procedure facilitates use of an item specific strategy then the pigeons in the different outcomes group will acquire the discriminations with fewer sessions than the same outcome group. The third hypothesis states that if the differential outcomes procedure facilitates the use of a relational strategy then transfer to novel stimuli should occur at a smaller set-size than same outcomes groups.

Methods

Subjects

Eight pigeons (which were experimentally naive) were maintained at 85% of their free-feeding weight. Grit and water were available at all times when pigeons were in their individual home cage. The colony room was maintained on a 14/10 hr light-dark cycle.

Apparatus

Chambers: The chambers used were the same as those used by Katz and Wright (2006). The wooden chamber is 35.9 cm wide X 45.7 cm deep X 51.4 cm high. A fan (Dayton 5C115A, Niles, IL) located at the back wall of each chamber provides ventilation and white noise. A 28 V (No. 1829, Chicago Miniature, Hackensack, NJ) houselight is located in the center of the ceiling the chambers. Mixed grain is delivered from a custom built hopper, through a 5.1 X 5.7 cm opening in the center of a 3.8 cm panel above the floor. An infrared touch screen (17" Unitouch, Carrol Touch, Rond

Rock, TX), detects the pigeons' pecks. The touch screen fits into a cutout in the front, center of the chamber (40.6 X 32.1 cm) 7.7 cm from the chamber's ceiling.

Stimuli: The colored photographs used were the same as those used by Katz and Wright (2006). These are travel-slide pictures digitized with a Howtek photomaster (no.87RU, Hudson, NH) camera and a Truevision TARGA-16 processing card (Indianapolis, IN) in a 256 X 256 resolution. The stimuli were displayed on a black background. They were presented as 5.7 X 3.8 cm large, digitized, color images. The sample pictures appeared at the top center of the screen and the comparison pictures appeared 1.28 cm directly below it. A white rectangle 2.5 X 2.4 cm appeared on the right side of the comparison picture. This experiment also included audio stimuli; a tone, the Windows opening theme and a short ditty that were all emitted from the same location.

Experimental control: Events were controlled and recorded with custom software written in Visual Basic on a Pentium personal computer. A video card (ATI 3D Rage Pro AGP 2X, Ontario, Canada) controlled the monitor and a computer controlled relay interface (Model no. PI0-12, Metrabyte, Taunton, MA), and operated the food hopper and houselight.

Procedure

The procedure used here is the same as that used by Katz and Wright (2006), with the exception of the food access durations and sounds provided for correct responses for the different outcomes group.

Preliminary training: Pigeons were trained to eat from a hopper and then their responses were auto-shaped to peck two white rectangles. The left rectangle was replaced with the comparison pictures (the same choice) appeared and the right rectangle was replaced with the white rectangle (the different choice) appeared. These rectangles appeared randomly but equally often over 100 trials. If a rectangle was pecked, it disappeared and food was presented. If the rectangle was not pecked, it remained on for 10 s after which food was presented. Food presentations lasted 3 s and were accompanied by a .5 s 660.6 Hz tone and a lit hopper light. A 50 s inter-trial-interval (ITI) followed each food presentation. When a pigeon was reliably pecking the rectangles, same/different training with or without different outcomes commenced.

Same/Different training: Figure 8 diagrams the temporal flow of a correct choice for a same trial on the left and a different trial on the right. A trial started with the sample photograph at the top of the screen. After a FR10 response requirement to the sample photograph, the bottom photograph and white rectangle appeared simultaneously. After a FR1 choice response, the entire display extinguished. Correct choices resulted in grain access and a sound. Incorrect responses resulted in a 15 s timeout. After the fifth session a correction procedure was introduced, the trial repeated until a correct response was emitted, but only first choice responses were tallied in the percent correct. Both correct and incorrect responses were followed by a 15 s ITI.

Correct choices resulted in grain access and a sound that varied according to each pigeon's group. The same outcome group (Boris, Nikolai, Lev and Dimitri) always experienced the same outcome within a session (2 – 5 s access to food) and tone

regardless of the correct choice response. For two of the pigeons in the different outcomes group (Emil and Max), correct same responses were followed by 5-s accesses to grain and sound one (the Windows theme) and correct different responses were followed by 1 s access to grain and sound two (the ditty). These outcomes were reversed for the remaining two pigeons in the different outcomes group (Sinclair and Hesse).

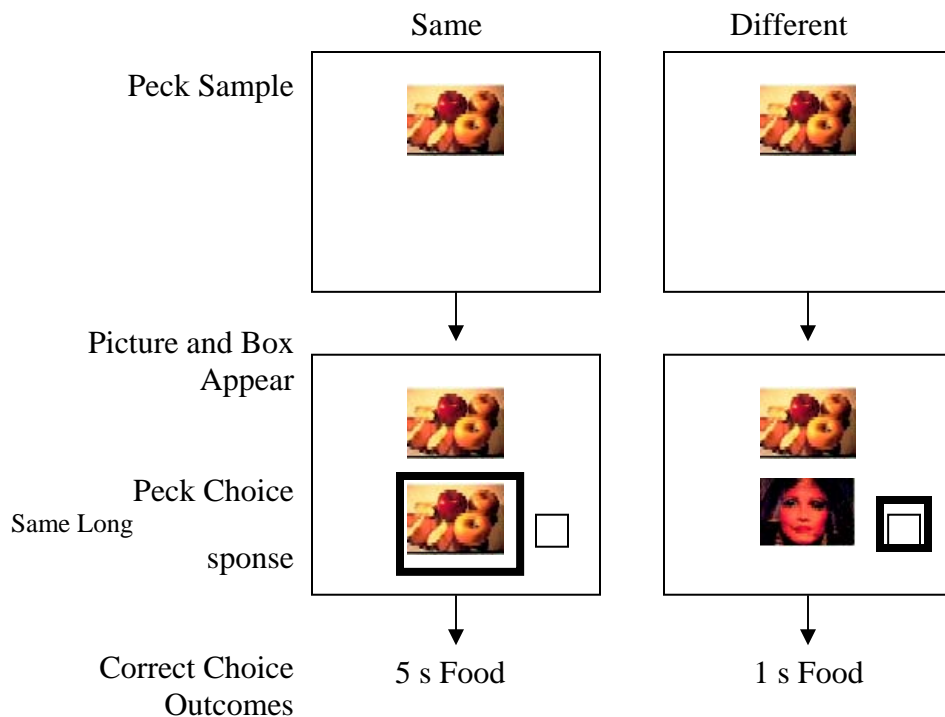


Figure 8. Flow chart of same/different trial in the Present Experiment

The initial training sessions included 8 different photographs selected at random with replacement over 100 trials (50 same and 50 different). Sessions took place 6 days a week and those sessions that were not completed in a single day were continued the following day. When a pigeon met acquisition criterion (80% over 3 consecutive days with the correction procedure and then 80% for three consecutive days without the correction procedure) same/different testing commenced.

Same/Different Testing: Six test sessions were conducted after each set-size was acquired. Test sessions were like training sessions, except that they included transfer trials with novel photographs. These photographs were novel in the sense that they had not been used in training or to test concept learning at a different set size (i.e., transfer photos were novel with no repetition). Each test session contained 100 trials, 90 baseline (45 same and 45 different) and 10 transfer (5 same and 5 different). Baseline trial photographs were selected at random and with replacement from the previously acquired set-size. Transfer trials appeared randomly between trials 8 and 93. Baseline and transfer trials were reinforced in the same manner as they were in training.

Set-Size Expansion: The last day of same/different testing was followed by a set-size expansion, where the number of photographs in the training set doubled to 16. Upon meeting criterion, 85% over 2 consecutive days, the correction procedure was removed. The pigeons then had to meet this criterion again without the correction procedure.

After the 16 item set-size was acquired, the set-size was expanded to 32. When the criterion was reached, at least 3 days one of which was 85% or higher, testing commenced for the 32 item set-size. After testing, the set-size continued to expand over 64, 128, 256, 512 and 1024. Testing occurred after each set-size was acquired.

Results

The first hypothesis stated that if pigeons use expectancies to guide choice behavior, then those expectancies should be reflected in their choice response latencies. For this experiment, the reaction time analysis was complicated because the spatial distance from pecking the sample to a different choice was greater than the spatial distance to a same choice. The spatial disparity alone should cause longer reaction times for different choices than same choices. To account for this confound, difference scores (different choice response time – same choice response time) were computed. If spatial disparity was the only factor that affected reaction time, then the difference scores should be similar for all the groups. To obtain the purest analysis of reaction time only first, correct responses were included, which omits choice reaction times made on correction trials and rash incorrect choices. Outliers, defined as response times beyond two standard deviations from the mean for each trial type, within each session, were removed. The analyses were run on the 8 and 512 item set-sizes, allowing one to see how reaction times change over acquisition.

Figure 9 shows each individual's reaction times for each response type, each session, over the 8 item set-size acquisition. Typically, different choice response times were longer than same choice response times. Figure 10 shows each individual's reaction times for each response type, each session, over the 512 item set-size acquisition. Compared to the 8 item set-size, the 512 item set-size reaction times generally appeared faster and less variable for both response types.

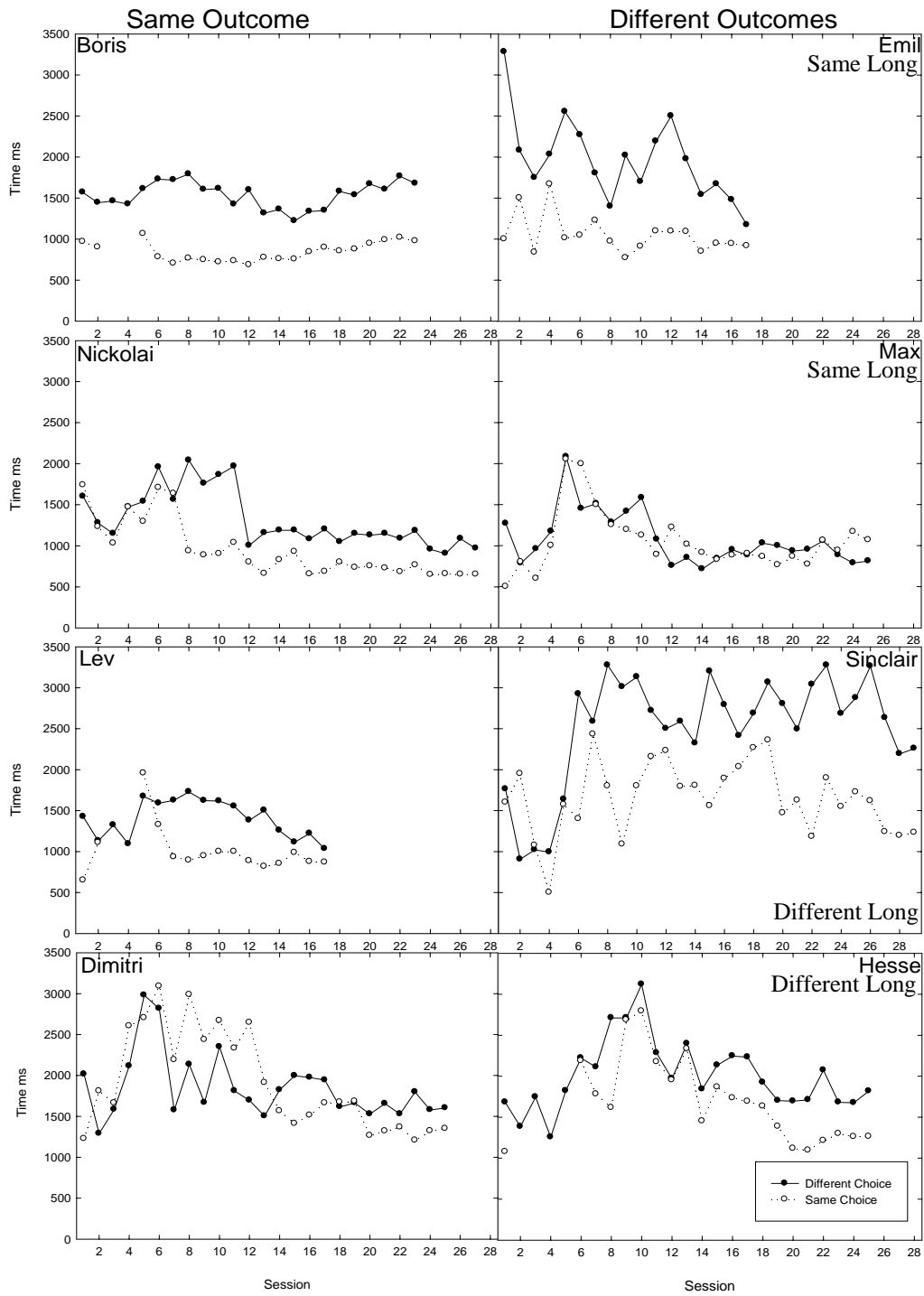


Figure 9. Each bird's average reaction times for same and different correct, first choices, within each session, over the 8 item set-size acquisition.

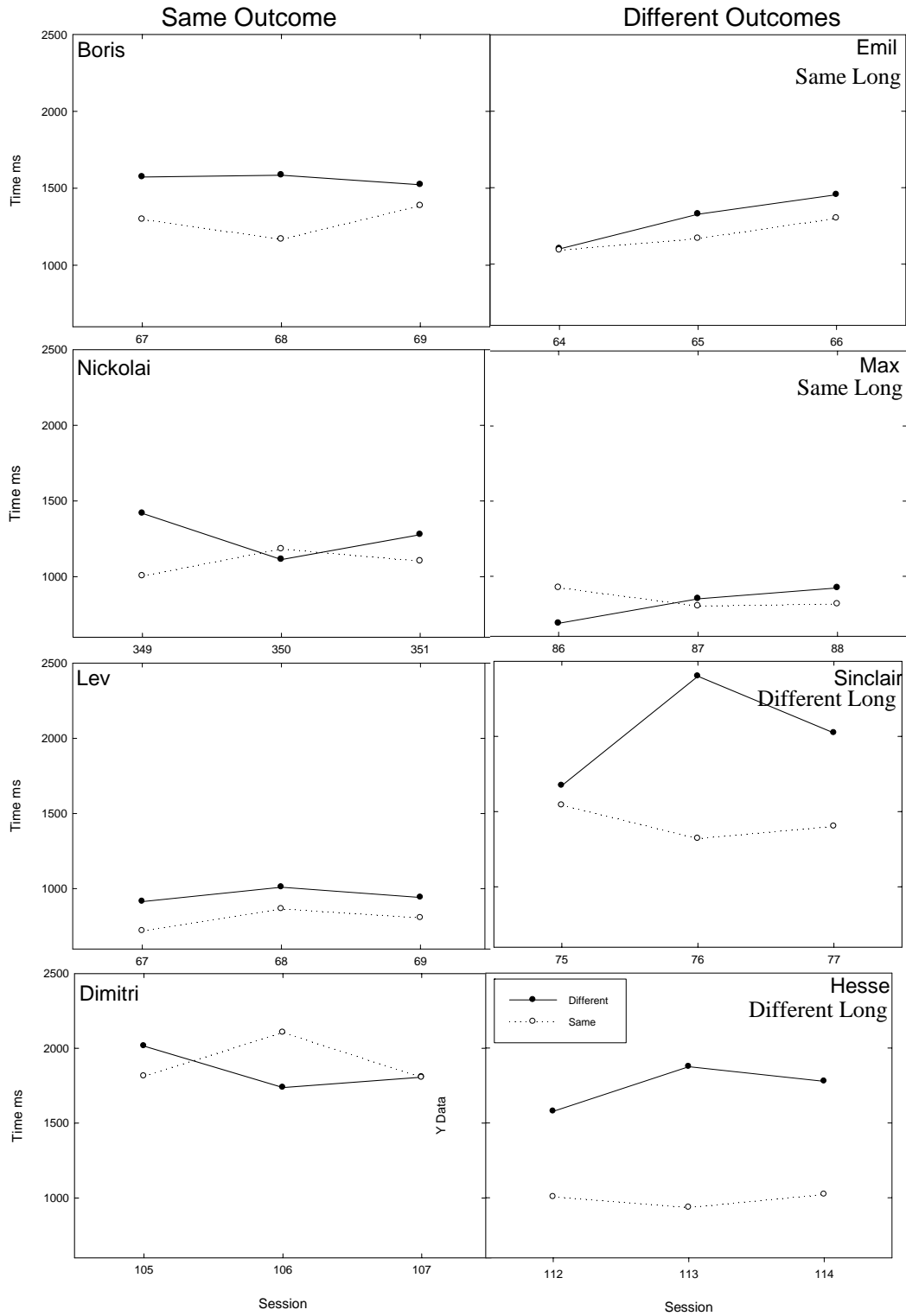


Figure 10. Each bird's average reaction times for same and different correct, first choices, within each session, over the 512 item set-size acquisition.

Reaction time was analyzed as difference scores (different choice response time minus same choice response time) for the last three sessions of the 8 item and then the 512 item set-sizes acquisitions using a one-way ANOVA with group as a between subjects factor. Figure 11 displays the average difference scores for the 8 item set-size acquisition, for each bird on the left and then for each group on the right. These groups were statistically similar $F(2, 5) = 1.71, p > .05$. However, a pattern emerged where the pigeons that experienced different outcomes with long access for same choices have the smallest difference scores, followed by the same outcome group and then the different outcomes group that experienced long access for different choices.

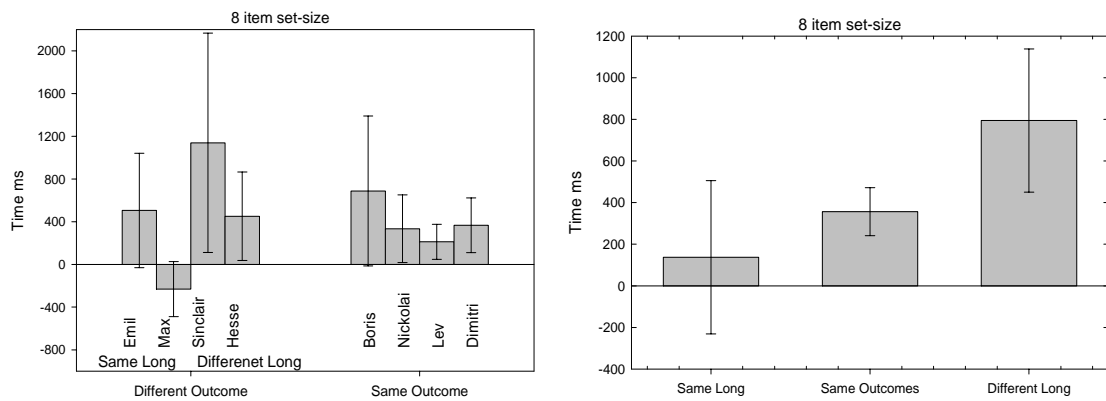


Figure 11. The left graph displays the mean difference reaction times for the last three sessions in the 8 item set-size acquisition for each bird, and the right graph shows the same information by outcome group. Error bars represent the standard error of the mean.

Figure 12 displays the average difference scores for the 512 item set-size acquisition, for each bird on the left and then for each group on the right. Statistically, the groups were different, $F(2, 5) = 27.53, p < .05$. The same pattern of reaction times seen

in the 8 item set-size repeated itself in the 512 item set-size. The SPSS outputs for the reaction time analyses are included in Appendix A, and the mean differences in reaction time are in Appendix B.

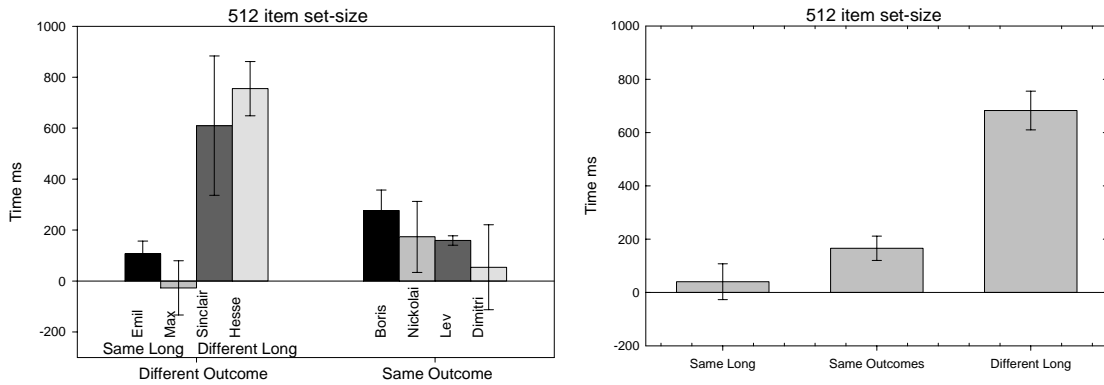


Figure 12. The left graph displays the mean difference reaction time for the last three sessions in the 512 item set-size acquisition for each bird, and the right graph shows the same information, but by type group. Error bars represent the standard error of the mean.

In the reaction time analyses the different outcomes group was split into two subgroups, one with long access to grain for same choices and one with long access to grain for different choices. For acquisition and accuracy, the same outcome group and different outcomes groups were analyzed to assess if they could be combined. The interaction between the subgroups and set-size was not significant for acquisition rate ($F(7, 7) = .143, p > .05$), baseline accuracy ($F(6, 6) = .029, p > .05$) or transfer accuracy ($F(6, 6) = .268, p > .05$). In the subsequent analyses, the two different outcomes groups will be combined. The SPSS outputs for the analyses for reaction time can be found in Appendix A, the mean differences in reaction time can be found in Appendix B.

The second hypothesis stated that if using the differential outcomes procedure enhances an item specific strategy, then the pigeons in the different outcomes group will acquire the discrimination with fewer sessions than the same outcome group. Figure 13 shows the number of trials to criteria across the set-sizes for each bird. In general, the birds tended to acquire larger set-sizes with fewer trials than smaller set-sizes. Yet, some pigeons' graphs, most notably Nickolai's, display large spikes. A spike indicates that acquiring that larger set-size took more sessions than a smaller one had. Spikes might be accounted for by the increased time needed to memorize more items or difficulty consistently applying a conceptual rule. The specific number of trials taken to acquire each set-size can be found in Appendix B.

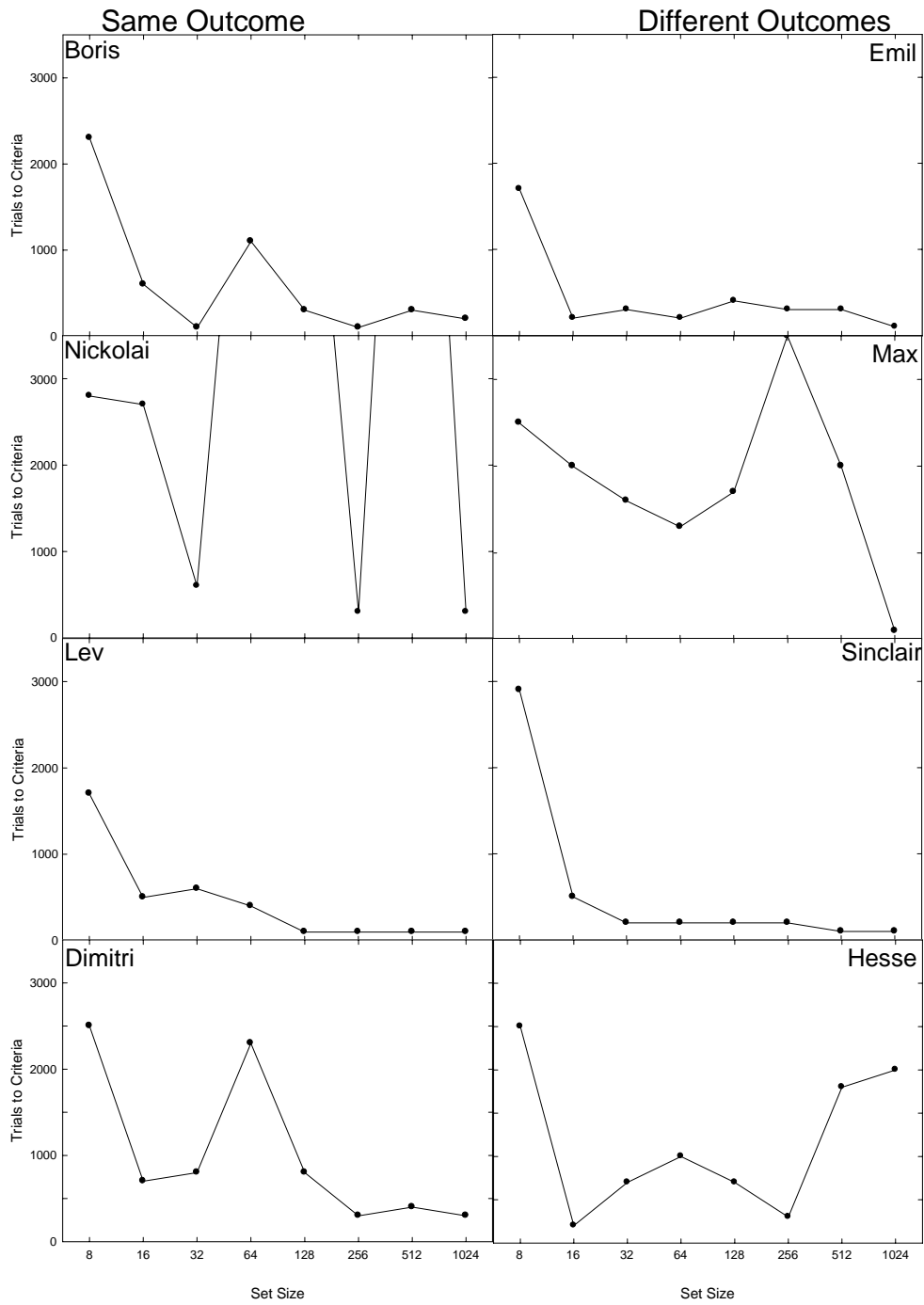


Figure 13. The number of trials to criteria across set-sizes for individual birds. The Y-axis was held constant for each bird to make visual comparisons easier. For Nickolai’s 64, 128 and 512 acquisition the trials to criteria were 7700, 7300, and 10700 respectively.

Figure 14 shows the average number of trials to criteria at each set-size, for each group. The acquisition results were analyzed using a two-way repeated measures ANOVA with set-size as a within subjects factor and group as a between subjects factor. For the within analysis, neither the set-size nor the interaction between set-size and group were significant (all ($F_s(7, 42) < 1.8, p > .05$). For the between groups analysis, the groups were not significantly different, $F(1, 6) = .35, p > .05$). The SPSS outputs for this analysis are in Appendix A.

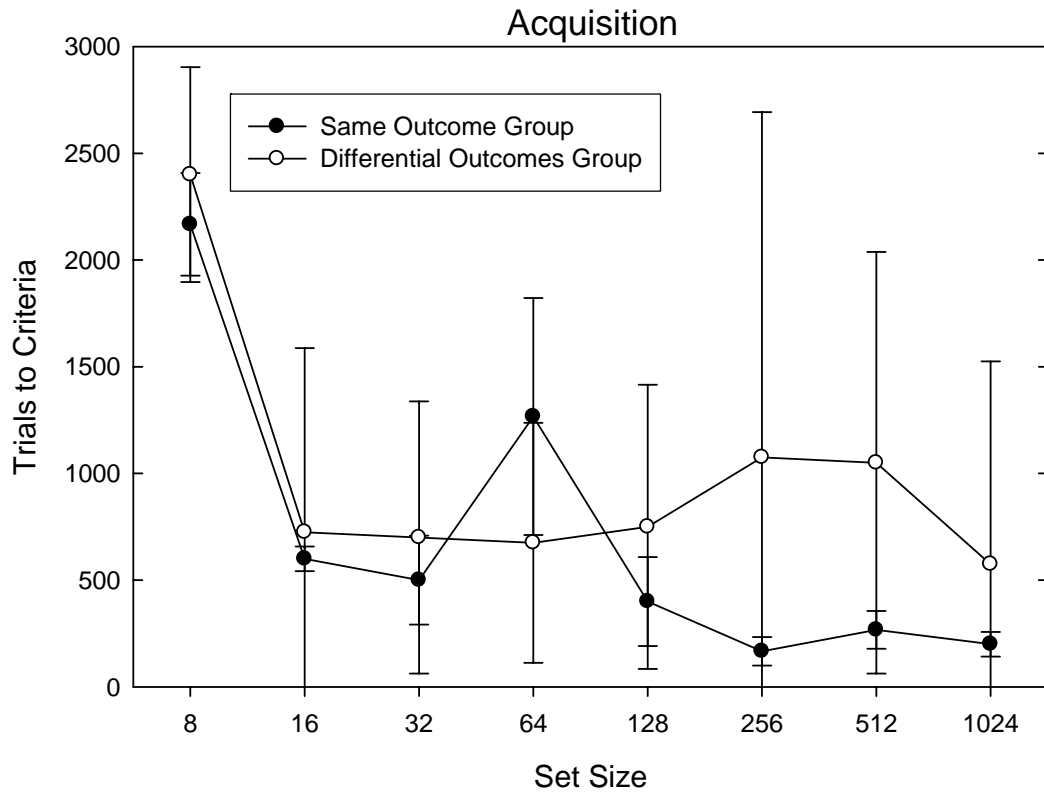


Figure 14. The average number of trials required to acquire each set-size by groups. Bars represent the standard error of the mean.

The third hypothesis stated that if using different outcomes facilitates a relational strategy, then transfer should occur at a smaller set-size. Figure 15 displays each individual's performance on baseline and transfer trials at each set-size. To determine if all the birds were utilizing a same/different concept, a repeated measures ANOVA was run with set-size and trial type (baseline or transfer) as within subjects factors. An interaction emerged between set-size and trial type; where at smaller set-sizes, baseline and transfer performance were different, but at larger set-sizes, baseline and transfer performance were similar, $F(6, 36) = 52.90, p < .05$. This effect reflects differences between baseline and transfer performance at the smaller set-sizes. If only the first set-sizes 8, 32 and 64 are included in the analysis, the interaction between set-size and trial type remains significant, $F(2, 12) = 40.17, p > .05$. However, if only the latter set-sizes are included the interaction disappears, $F(3, 18) = .73, p > .05$. With an independent sample t test, analyzing the difference between baseline and transfer trials at each session suggests that full concept learning emerges at set-size 128, $t(14) = 1.2, p > .05$.

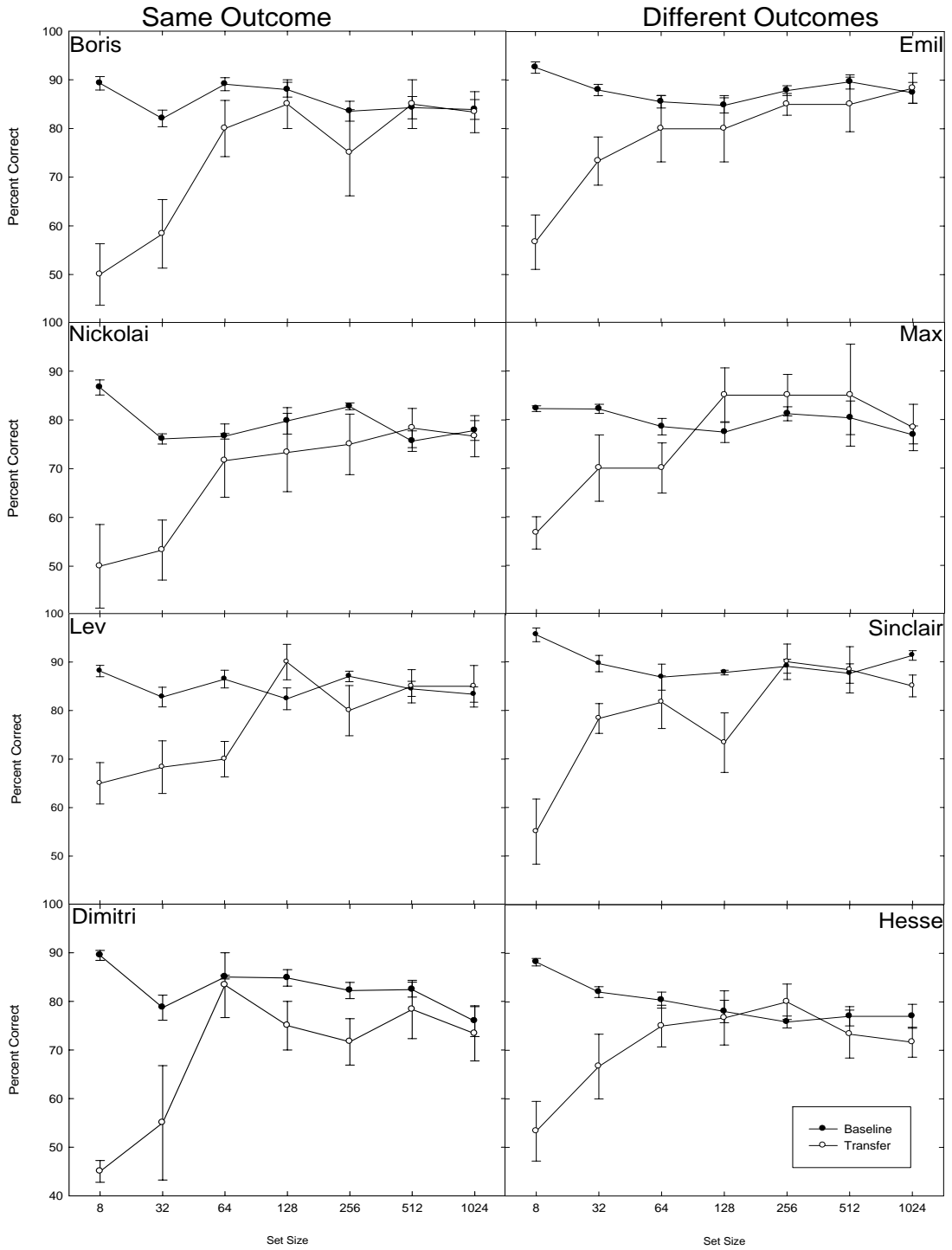


Figure 15. Individual performance on baseline and transfer trials by set size. The bars represent the standard error of the mean.

Figure 16 shows the results averaged for each group's baseline accuracy. To compare groups, the results were analyzed using a two-way repeated measures ANOVA with group as a between subjects factor, and set-size as a within subjects factor. An interaction emerged between set-size and group ($F(6, 30) = 3.09, p < .05$) meaning that the groups were different, but this difference was mediated by differences at specific set-sizes. This effect was strongest at the 32 item set-size ($F(1, 5) = 4.89, p = .06$). The SPSS outputs of these analyses are found in Appendix A, the specific percentages are in Appendix B.

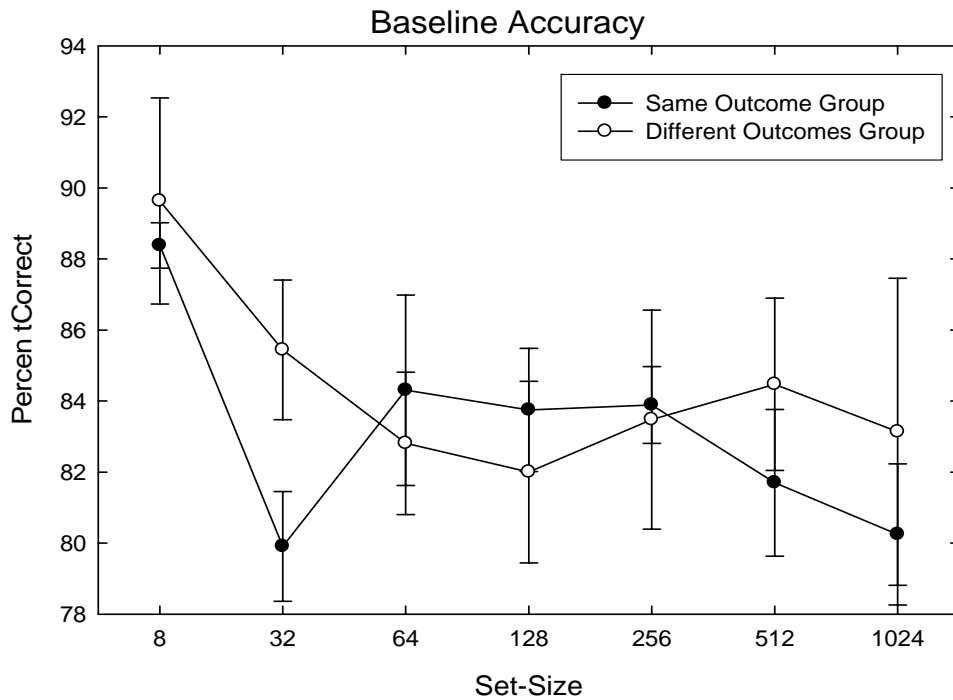


Figure 16. The average performance at each set size for baseline trials. The bars represent the standard error of the mean.

Figure 17 shows the results averaged for each group's transfer accuracy. To compare groups, the transfer accuracy results were analyzed using a two-way repeated

measures ANOVA with group as a between subjects factor, and set-size as a within subjects factor. An interaction emerged between set-size and group ($F(6, 30) = 42.31, p < .05$). This effect was strongest at the 32 and 256 item set-size (all F 's (1, 5) $> 10.17, p$'s $< .05$). The SPSS outputs of these analyses are found in Appendix A, the specific percentages are in Appendix B.

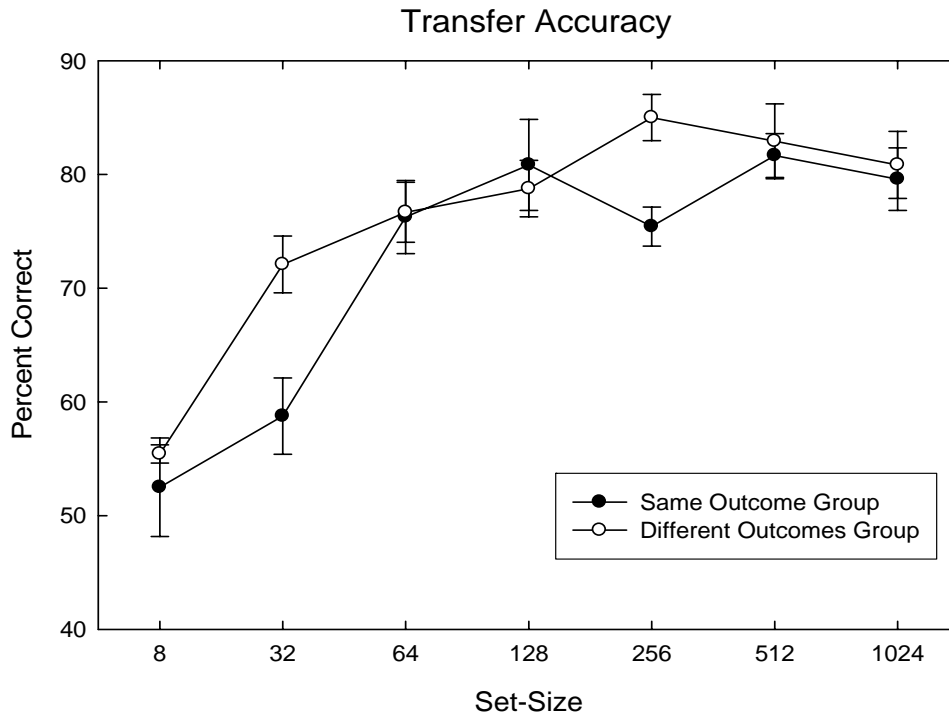


Figure 17. The average performance at each set size for transfer trials. The bars represent the standard error of the mean.

In the introduction, it was stated that full concept learning is exemplified when performance with baseline and transfer items is statistically similar. Because individual pigeons had different baseline performances, one might expect there to be differences in their transfer performances as well. These differences could be attributed to an individual's ability to use a concept. To consider this, the individual was held constant in

the following analysis by using difference scores between the training and novel items (baseline trial percent correct minus transfer trial percent correct). When a concept fully emerges, this number should be or be less than zero. Figure 17 shows the averaged percent correct difference scores for each group across set-sizes. Results were analyzed using a two-way repeated measures ANOVA with group as a between subjects factor and set-size as a within subjects factor. In this analysis the previous interaction between group and set-size disappeared, ($F(6, 30) = 1.893, p > .05$). However, there was still a significant group effect when the groups were compared at individual set-sizes 32 and 256 (all $F_s(1, 5) > 10.16, p_s < .05$), the set-sizes where the strongest effects were previously noticed for transfer accuracy. These SPSS outputs analyses are located in Appendix A.

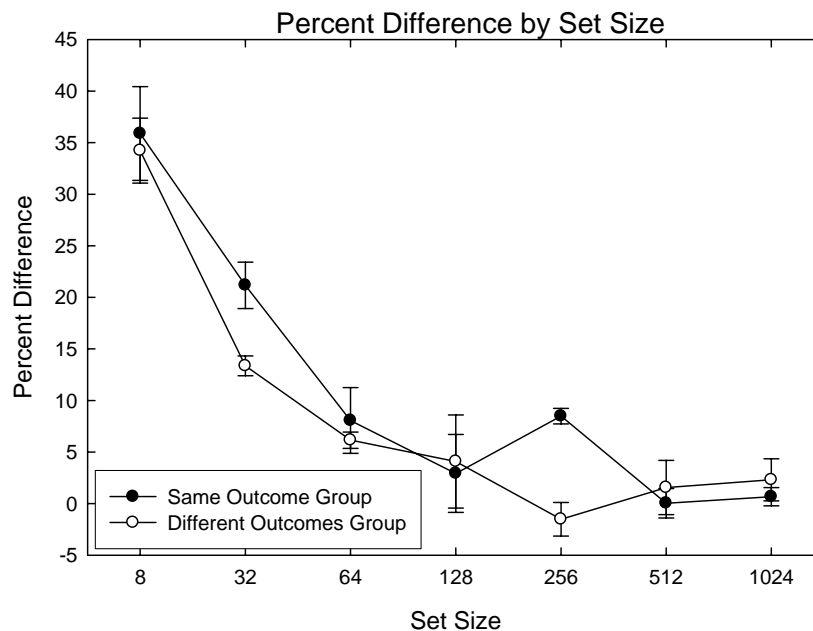


Figure 18. The average percent difference (baseline trial percent correct – transfer trial percent correct) across set sizes. The bars represent the standard error of the mean.

Discussion

In the present experiment, different outcomes were associated with correct responses to a relational association, same/different. The data were analyzed with regard to response time, acquisition and accuracy. The differential outcomes procedure affected response time and accuracy, but not acquisition. For response time, the differential outcomes procedure affected response time at larger set-sizes. With regard to accuracy, the differential outcomes procedure produced partial concept learning at a smaller set-size. For acquisition the differential outcomes procedure did not help subjects learn the task faster.

Regarding response time, the differential outcomes procedure had a greater effect at the 512 item set-size than the 8 item set-size. This could be accounted for by the increased exposure to the outcomes over training or by a sudden shift in the pigeons' response strategy from one guided by memorization to one guided by a relational association. These results favor exposure as an explanation, because the pattern between the groups remained consistent at the 8 and 512 item set-sizes. In other words, the differences in the pigeons' response times became more extreme over training.

Acquisition was not affected by the different outcomes procedure. Birds that experienced the different outcomes did not acquire new set-sizes at a faster rate than birds that experienced the same outcome. Failure to find a the differential outcomes procedure's effect on acquisition might be explained by the increased complexity of the same/different task compared to more frequently used matching-to-sample tasks.

Whereas a two color matching to sample task contains 4 configurations, the

same/different task with 8 items contains 64 configurations. Why the differential outcomes procedure did not speed acquisition, even at the initial set-size, is unclear. This question could be addressed with future experiments by manipulating the number of configurations a pigeon must learn with or without the differential outcomes procedure.

Baseline accuracy was superior for the different outcomes group at set-size 32. Transfer accuracy was superior for the different outcomes group at set-sizes 32 and 256. The percent correct difference scores were significantly lower for the different outcomes group at set-sizes 32 and 256.

The results of this experiment are notable in at least two ways. First, the results add a relational association to the DOE literature. A recent work by Estévez, Vivas, Alonso, Marí'-Beffa, Fuentes, Overmier and González (2007), also explored how the DOE might fare when participants were asked to use a relational association, greater than - less than. In their first experiment human participants were presented with a positive number relation, such as " $4.09 < 4.33$," on a computer screen and asked to press K if the relation was correct and J if the relation was incorrect. They found that the differential outcomes procedure enabled participants to respond faster; however it did not affect participants' accuracy. In the second experiment, the participants were presented with positive and negative numbers alone or in combination. Here they found that the differential outcomes procedure enabled participants to respond more accurately, but only when two negative numbers were presented (e.g., $-1 < -3$). The previous effect on reaction time disappeared. While this experiment also adds a relation strategy to the DOE literature, it does not require them to learn the association. The pigeons in the present

experiment enter the experiment without, or at least not using, the same/different concept. So, unlike the human experiment, the present experiment allows one to look at how concepts are formed over time. The present results suggest one can use the differential outcomes procedure to help relational strategies emerge rather than to help them be performed faster or more accurately. This is indicated by a stable response time pattern and the significant difference in transfer at set-size 32 and 256, but not at 512 or 1024.

Second, these results expand the literature on abstract concept learning. Previous research suggested that the number of items used to train a concept is a critical factor as to whether subjects form a concept, with more items leading to fuller concept formation. These results suggest that the expectancies a subject ties to each relation may be another factor that mediates concept formation.

Some researchers have questioned whether using increasingly more items facilitates concept learning or generalization. Two of these theorists include McLaren and Mackintosh (2000 & 2002) with their Stimulus Sampling Theory (SST). SST posits that when subjects sense a stimulus, they sense not the whole stimulus, but rather parts (a sample) of it. These parts are referred to as the micro-elements that constitute the configuration that is the whole stimulus. Generalization occurs because some stimuli contain the same micro-elements. Therefore, the greater number of stimuli with which one has been presented, the greater the chance he/she will generalize accurately from those stimuli to novel stimuli. SST also states that, the more the same stimulus is viewed the more solidified that stimulus' micro-elements become, leaving fewer micro-elements available to generalize across stimuli.

Using the framework provided by SST, one might suggest that in transfer sessions while the whole picture in the present experiment might be novel, the micro-elements are not. One way to refute the SST's generalization explanation is to look at transfer performance where the set-size took an unusual number of trials to acquire. One might theorize that it took longer to acquire those set-sizes because the bird was attempting to memorize instead of using the relational associations between the items. Because of this, each item is viewed more frequently, meaning each item's micro-elements are more tightly bound. For example Nickolai took 10,700 trials to acquire the 512 item set-size. If Nickolai had been memorizing pictures and the micro-elements were tightly bond, than his transfer performance should have fallen to chance. This was not the case, as Nickolai's transfer remained well above chance ($M = 78.33\%$). Unfortunately, this evidence is only partial, because SST never specifies how large a micro-element is or how long it takes micro-elements to bond together. Furthermore, other theories of generalization claim that generalization can occur across entire configurations (see, Pearce, 1987).

Regardless of how the pigeons are performing the present same/different task, using a concept or generalizing, the different outcomes group transferred more accurately at a smaller set-size than the same outcome group. Since this finding is new, replication is important to assess this finding's reliability. Replications should be done not only with the same/different concept, but also with other concepts, such as relative size and/or numerosity. The outcomes used are the experimenter's choice. The effect was obtained here with different durations of food access and sounds. However, better effects may be

found with more explicit differences. For example, one might use a 1:10 second difference instead of 1:5, or they may use a different modality such as visual cues (e.g., a light blinking at different rates or a different colored background). Lastly, if this effect proves reliable, it should not stay confined in cold laboratory walls. Rather, creative minds should pounce on the opportunity to use the differential outcomes procedure to help people (cognitively disabled and normal students) learn difficult tasks.

REFERENCES

- Alling, K., Nickel, M., & Poling, A. (1991a). The effects of differential and nondifferential outcomes on response rates and accuracy under a delayed-matching-to-sample procedure. *The Psychological Record, 41*, 537-549.
- Carlson, J. G., & Wielkiewicz, R. M. (1976). Mediators of the effects of magnitude of reinforcement. *Learning and Motivation, 7*, 184-196.
- Colwill, R. M., & Rescorla, R. A. (1985). Postconditioning devaluation of a reinforcer affects instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 120-132.
- Cook, R. G., Cavoto, K. K. & Cavoto, B. R. (1995). Same/different texture discrimination and concept learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 21*, 253-260.
- Cook, R. G., Katz, J. S., & Cavoto, B. R. (1997). Pigeon same/different concept learning with multiple stimulus classes. *Journal of Experimental Psychology: Animal Behavior Processes, 23*, 417-433.
- Cook, R. G. "Same/Different concept formation in pigeons." Chapter in *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. edited by Marc Bekoff, Colin Allen and Gordon M. Burghardt, with an afterward by Donald Griffin (2002a) MIT Press.

- Cook, R. G. (2002b). The structure of pigeon multiple-class same/different learning. *Journal of the Experimental Analysis of Behavior*, 78, 345-364.
- Descartes, R. (1970). *Descartes's philosophical letters* (A. Kenny, Ed. and Trans.). Oxford, United Kingdom: Clarendon. (Originally written, 1646).
- Edwards, C. A., Jagielo, J. A., Zentall, T. R., Hogan, D. E. (1982). Acquired equivalence and distinctiveness in matching to sample by pigeons: Mediation by reinforcer-specific expectancies. *Journal of Experimental Psychology: Animal Behavior Processes*, 8, 244-25.
- Estevaz, A. F., Fuentes, L., Mari-Beffa, P., Gonzalez, C., and Alvarez, D. (2001). The differential outcome effect as a useful tool to improve conditional discrimination learning in children. *Learning and Motivation*, 32, 48-64.
- Estevez, A. F., Fuentes, L. (2003). Differential outcomes in four-year-old children. *Psicológica*, 24, 159-167.
- Estévez, A. F., Fuentes, L. J., Overmier, B. J., & González, C. (2003). Differential outcomes effect in children and adults with Down Syndrome. *American Journal on Mental Retardation*, 108, 108–116.
- Estévez, A. F., Vivas, A. B., Alonso, D., Marí'-Beffa, P., Fuentes, L. J., Overmier, B. J., & González, C. (2007). Enhancing challenged students' recognition of mathematical relations through differential outcomes training. *The Quarterly Journal of Experimental Psychology*, 60(4), 571–580.
- Hewitt, F.M. (1965). Teaching speech to an autistic child through operant conditioning. *American Journal of Orthopsychiatry*, 35, 927-936.

- Hochhalter, A.K., Sweeney, W.A., Bakke, B.L., Holub, R.J., & Overmier, J.B. (2000). Improving face recognition in Alcohol Dementia. *Clinical Gerontologist*, 22, 3-18.
- James, W. (1950). *The principles of psychology* (Vol. 1). New York: Dover Publications. (Original work published 1890).
- Jansen, C., & Guess D. (1978). Use of function as a consequence in training receptive labeling to severely and profoundly retarded individuals. *Association for the Education of the Severely/Profoundly Handicapped Review*, 3, 246-258.
- Jenkins, H. M., & Moore, B. R. (1973). The form of the auto-shaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior*, 20, 163–181.
- Katz, J. S., Wright, A. A. & Bachevalier, J. (2002). Mechanisms of same/different abstract- concept learning by rhesus monkeys (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, 28, 358-368.
- Katz, J. S., & Wright, A.A. (2006). Same/different abstract-concept learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 32, 80-86.
- Kruse, J. M., Overmier, J. B., Konz, W. A., & Rokke, E. (1983). Pavlovian CSs effects upon instrumental choice behavior are reinforcer specific. *Learning and Motivation*, 14, 165-181.
- Maki, P., Overmier, B. J., Delos, S., & Gutmann, A. J. (1995). Expectancies as factors influencing conditional discrimination performance of children. *The Psychological Record*, 45, 45–72.

- McLaren, I. P. L., and Mackintosh, N. J. (2000). An elemental model of associative learning: Latent inhibition and perceptual learning. *Animal Learning and Behavior*, 28, 211-246.
- McLaren, I. P. L., Mackintosh, N. J., (2002). Associative learning and elemental representation: II. Generalization and discrimination *Animal Learning and Behavior*, 30(3), 177-200.
- Miller, O., Waugh, K.M. & Chambers, K. (2002). Differential outcomes effect: Increased accuracy in adults learning kanji with stimulus-specific rewards. *The Psychological Record*, 51, 315-323.
- Miyashita, Y., Nakajima, S., & Imada, H. (2000). Differential outcomes effect in horses. *Journal of the Experimental Analysis of Behavior*, 74, 245–253.
- Overmier, J. B., Bull, J. A., III, & Trapold, M. A. (1971). Discriminative cue properties of different fears and their role in response selection in dogs. *Journal of Comparative and Physiological Psychology*, 76, 478-482.
- Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, 94(1), 61–73.
- Peterson, G. B., & Trapold, M. A. (1982). Expectancy mediation of concurrent conditional discriminations. *American Journal of Psychology*, 95, 571-580.
- Premack, D. (1983a). The codes of man and beasts. *The Behavioral and Brain Sciences*, 6, 125-167.

- Premack, D. (1983b). Animal cognition. *Annual Review of Psychology*, 34, 351-362.
- Rescorla, R. A. & Colwill, R. M. (1989). Associations with anticipated and obtained outcomes in instrumental learning. *Animal Learning & Behavior*, 17, 291-303.
- Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different? *Learning and Motivation*, 1, 129-140.
- Trapold, M. A. & Overmier, J. B. (1972). The secondary learning process in instrumental learning. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning: Vol 2. Current research and theory* (pp. 427-452). New York: Appleton.
- Young, M. E., Wasserman, E. A., & Garner, K. L. (1997). Effects of number of items on the pigeon's discrimination of same from different visual displays. *Journal of Experimental Psychology: Animal Behavior Processes*, 23, 491-501.

APPENDIX A

SPSS OUTPUTS FOR SELECT ANALYSES

Reaction Time Analyses:

8 item set-size acquisition

Tests of Between-Subjects Effects
 Dependent Variable: RT_8 item set-size

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	455337	2	227668	1.71	.27	.41
Error	666374	5	133274			
Corrected Total	1121711	7				

512 item set-size acquisition

Tests of Between-Subjects Effects
 Dependent Variable: Reaction time_512 item set-size

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	488883	2	244441.744	27.53	.00	.92
Error	44403	5	8880.705			
Corrected Total	533287	7				

Equivalence of the Different Outcomes Groups across Set-Sizes:

ACQUISITION

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	661250	1	661250.000	.143	.742	.067
Error	9280000	2	4640000.000			

BASELINE

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	8.016	1	8.016	.029	.880	.014
Error	545.300	2	272.650			

TRANSFER

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	32	1	32.142	.268	.656	.118
Error	240	2	120.038			

Acquisition Analyses:

Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
SET_SIZE	29313593	7	4187656.250	1.831	.106	.234
SET_SIZE * GROUP	17614843	7	2516406.250	1.100	.381	.155
Error(SET_SIZE)	96052812	42	2286971.726			

Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	4676406.250	1	4676406.250	.356	.572	.056
Error	78769687.500	6	13128281.250			

Accuracy Analyses - Baseline:

Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Set Size	274.179	6	45.696	9.96	.000	.624
Set Size * GROUP	85.289	6	14.215	3.09	.015	.341
Error	165.158	36	4.588			

Tests of Between-Subjects Effects

Dependent Variable: BASE_32

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	61.215	1	61.215	4.891	.069	.449
Error Corrected Total	75.102 136.317	6 7	12.517			

Accuracy Analyses - Transfer:

Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Set Size	5286.010	6	881.002	42.31	.000	.876
Set Size * GROUP	368.353	6	61.392	2.948	.019	.329
Error	749.603	36	20.822			

Tests of Between-Subjects Effects

Dependent Variable: TRAN_32

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	355.556	1	355.556	10.172	.019	.629
Error	209.721	6	34.954			
Corrected Total	565.278	7				

Tests of Between-Subjects Effects

Dependent Variable: TRAN_256

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	183.680	1	183.680	12.90	.011	.683
Error	85.416	6	14.236			
Corrected Total	269.097	7				

Accuracy Analyses- Difference Scores:

Tests of Within-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
SET_SIZE	7443.242	6	1240.540	53.360	.000	.899
SET_SIZE * GROUP	264.103	6	44.017	1.893	.109	.240
Error(SET_SIZE)	836.951	36	23.249			

Tests of Between-Subjects Effects

Dependent Variable: set size 32

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	121.709	1	121.709	10.164	.019	.629
Error	71.846	6	11.974			
Corrected Total	193.555	7				

Tests of Between-Subjects Effects

Dependent Variable: set size 256

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
GROUP	199.815	1	199.815	31.103	.001	.838
Error	38.545	6	6.424			
Corrected Total	238.360	7				

APPENDIX B

DATA

Reaction Time Analyses (Mean Difference Scores in milliseconds):

8 item set-size

Different Outcomes		Same Outcome	
Same Long	Different Long		
Emil 505.3333	Sinclair 1138.0000	Boris 688.0000	
Max -231.0000	Hesse 450.0000	Nickolai 317.0000	
		Lev 164.0000	
		Dimitri 256.0000	

512 item set-size

Different Outcomes		Same Outcome	
Same Long	Different Long		
Emil 107.6670	Sinclair 610.0000	Boris 276.6670	
Max -26.6667	Hesse 755.3333	Nickolai 173.6670	
		Lev 159.3330	
		Dimitri 54.3333	

Acquisition (Trials to Criteria)

Set-Size	Different Outcomes				Same Outcomes			
	Emil	Max	Sinclair	Hesse	Boris	Nickolai	Lev	Dimitri
8	1700	2500	2900	2500	2300	2800	1700	2500
16	200	2000	500	200	600	2700	500	700
32	300	1600	200	700	100	600	600	800
64	200	1300	200	1000	1100	7700	400	2300
128	400	1700	200	700	300	7300	100	800
256	300	3500	200	300	100	300	100	300
512	300	2000	100	1800	300	10700	100	400
1024	100	100	100	2000	200	100	100	300

Baseline Accuracy

Set-Size	Different Outcomes				Same Outcomes			
	Emil	Max	Sinclair	Hesse	Boris	Nickolai	Lev	Dimitri
8	92.59	82.22	95.56	88.17	89.26	86.67	88.15	89.44
32	87.96	82.17	89.63	82.00	82.04	76.11	82.78	78.70
64	85.56	78.52	86.83	80.33	89.07	76.67	86.48	85.00
128	84.81	77.41	87.78	78.00	87.96	79.81	82.41	84.81
256	87.83	81.17	89.07	75.83	83.52	82.78	87.04	82.22
512	89.63	80.33	87.59	80.33	84.26	75.67	84.46	82.41
1024	87.41	74.83	91.30	77.00	83.89	77.83	83.33	75.93

Transfer Accuracy

Set-Size	Different Outcomes				Same Outcomes			
	Emil	Max	Sinclair	Hesse	Boris	Nickolai	Lev	Dimitri
8	56.67	56.67	55.00	53.33	50.00	50.00	65.00	45.00
32	73.33	70.00	78.33	66.67	58.33	53.33	68.33	55.00
64	80.00	70.00	81.67	75.00	80.00	71.67	70.00	83.33
128	80.00	85.00	73.33	76.67	85.00	73.33	90.00	75.00
256	85.00	85.00	90.01	80.00	75.00	75.00	80.00	71.67
512	85.00	85.00	88.33	73.33	85.00	78.33	85.00	78.33
1024	88.33	78.33	85.00	71.67	83.33	76.67	85.00	73.33