THE DIFFERENTIAL OUTCOME EFFECT WITH TYPICAL ADULT HUMANS

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

Kelly Ann Schmidtke

Certificate of Approval:

Chris J. Correia
Associate Professor
Psychology

Jeffrey S. Katz, Chair
Alumni Associate Professor
Psychology

Martha C. Escobar
Associate Professor
Psychology

Ana Franco-Watkins
Assistant Professor
Psychology

George T. Flowers
Dean
Graduate School

THE DIFFERENTIAL OUTCOME EFFECT WITH TYPICAL ADULT HUMANS

Kelly Ann Schmidtke

A Dissertation

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Doctorate of Philosophy

Auburn, Alabama

August 10, 2009

THE DIFFERENTIAL OUTCOME EFFECT WITH TYPICAL ADULT HUMANS

Kelly Ann Schmidtke

____
Signature of Author

____
Date of Graduation

DISSERTATION ABSTRACT

THE DIFFERENTIAL OUTCOME EFFECT WITH TYPICAL ADULT HUMANS

Kelly Ann Schmidtke

Doctor of Philosophy, August 10, 2009
(M.S., Auburn University, 2008)
(B.A., University of Minnesota-Twin Cities, 2005)

100 Typed Pages

Directed by Jeffrey S. Katz

A mechanism that affects nonhuman animal learning is expectation, as explored with experiments that produce the differential outcome effect (DOE). To obtain the DOE, animals are trained using a differential outcomes procedure where each correct choice response is correlated with a particular outcome. Animals trained with the differential outcomes procedure learn to "expect" a particular reinforcer, and this "expectation" helps them learn tasks faster and perform them more accurately than animals that cannot "expect" a particular reinforcer. While the DOE appears when nonhuman animals are trained using the differential outcomes procedure, it has been more elusive when typical adult humans are trained using the same procedure. Contrary to past experiments, the present experiments demonstrate that expectancies do affect typical adult human learning. Previous failures to obtain the DOE with typical adult humans result because previous experiments have employed tasks that were too simple and participants acquired the associations quickly regardless of the expectancies they could or could not form. In the present Experiment 1 participants saw a Kanji character (a Japanese word), and then 9 English words. If the participant selected the correct English word, then they saw either a correlated or uncorrelated outcome followed by a second correlated or uncorrelated outcome. If the participants choose the wrong word, then they saw either corrective or noncorrective feedback. Because Kanji are less discriminable and 15 associations were required, this task proved difficult for participants and the DOE was obtained. In Experiment 2, after a participant made an incorrect choice response they saw noncorrective feedback. Participants in Experiment 1 acquired the information faster and more completely than participants in Experiment 2 but the DOE was obtained regardless of the type of feedback given. In Experiment 3, after a participant responded correctly they saw an entry into a prize lottery and then saw a picture. Those participants who saw a correlated prize acquired the kanji's meanings faster than those who saw an uncorrelated prize, but the picture condition had no effect. These results suggest that the temporal placement of the outcomes matter. In Experiment 4, the participants learned who painted 15 abstract paintings. The DOE was not obtained in this experiment because all participants learned the task quickly. However, an additional post test indicates that these participants also learned about proximal outcomes, suggesting that associations are learned not only between stimuli and responses but also between responses and outcomes. These experiments add to the DOE literature and comparative psychology by asserting that expectancies are a general mechanism that affects species learning.

APA Style

Visual Basic 6.0

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

One mechanism that has been reliably shown to affect nonhuman animal learning is expectation as explored with experiments that produce the differential outcome effect (DOE). For these experiments subjects are trained to perform conditional discriminations using different procedures. One common type of conditional discrimination is matching-to-sample (MTS). In a MTS task with color stimuli, the subject might see a red sample with a red and green comparison. The correct response in this case would be the red comparisons. However, if the sample was green, then the correct response would be the green comparisons. The discrimination is conditional because the correct comparison response is always conditional upon the sample presented.

In typical conditional discrimination training procedures, all correct choices are followed by the same outcome, a nondifferential outcomes procedure. For example, when a pigeon learns a matching to sample task with red and green samples, all correct choices might be followed with 3 s of grain access. Alternatively, to obtain the DOE, subjects are trained using a differential outcomes procedure where each correct choice response is correlated with a particular outcome. For example, the pigeons correct choice responses after the red sample might be followed by 1 s of grain access while correct choice responses after the green sample might be followed by 5 s of grain access. Theoretically, subjects trained with the differential outcomes procedure can learn to "expect" a particular reinforcer, and this "expectation" helps them learn tasks faster and perform conditional discriminations more accurately over increasingly longer delays than subjects that cannot "expect" a particular reinforcer (Edwards, Jagielo, Zentall, & Hogan, 1982; Kruse, Overmier, Konz, & Rokke 1983; Maki-Kahn, Overmier, Delos, & Gutmann, 1995). The DOE is an interesting phenomenon because it enhances performance without requiring much physical effort from subjects and because of its beneficial effects researchers and educators cannot help but concern themselves with why it occurs.

Experiments have employed the differential outcomes procedure with various reinforcers, methodologies and species. The DOE has been obtained with reinforcers having different qualitative natures (food vs. water, Trapold, 1970), different quantities (one vs. five food pellets, Carlson, & Wielkiewicz, 1976) and conditional reinforcers (hopper light vs. food, Urcuioli, DeMarse, & Lionello-DeNolf, 2001; token colors, Estévez, Fuentes, Overmier, & González, 2003). The DOE has been demonstrated with both between-subjects (Trapold, 1970; Brodigan, & Peterson, 1976) and within-subject designs (Alling, Nickel, & Poling, 1991a). The different nonhuman animals tested include 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). Also, while examples of the DOE with typical adult human populations are almost nonexistent, the DOE has been demonstrated in various other human populations, such as language-deficit children (Janssen, & Guess, 1978; Hewitt, 1965), typical children (Estevaz, Fuentes, Marý-Beffa, González, & Alvarez, 2001), and adults with alcohol induced dementia (Hochhalter, Sweeney, Bakke, Holub, & Overmier, 2000). These findings suggest that the DOE is a general effect.

DOE Theories

Urcuioli (1990) discusses three to theoretical reasons why the differential outcome procedure may enhance subjects' performance. First, the differential outcomes procedure may enhance the stimuli's discriminability, particularly the samples. This was first suggested by Peterson and Trapold (1980) when they stated that the "differential outcomes procedure facilitates conditional discrimination learning by enhancing the difference between the functional conditional cues" (p. 572). Second, the differential outcome training procedure may create additional cues, also known as expectancies, that a nondifferential outcomes training procedure does not. If this is true then subjects trained with the differential outcomes procedure may use both the enhanced stimuli's discriminability and expectancies to respond more accurately than subjects trained with a nondifferential outcomes procedure. Third, the expectancies formed during training with the differential outcomes procedures may not be just additional cues, but in fact the expectancies may overshadow any effects of enhanced stimuli's discriminability. Any of these three possibilities, 1) stimulus discriminability, 2) a combination of stimulus discriminability and expectancies or 3) expectancies alone, could explain the DOE's occurrence.

With pigeon subjects, Urcuioli (1990) attempted to determine whether the DOE resulted from solely enhanced sample discriminability or the combination of discriminability and expectancies, in a many-to-one conditional discrimination task. A many-to-one conditional discrimination is similar to the MTS task. In the MTS task one sample is correlated with one comparison. (e.g., when the sample is a vertical line then the correct comparison is red), in the many-to-one task multiple samples are correlated with a one comparison (e.g., when the sample is a vertical line or blue then the correct comparison is red).

While holding the samples' correlation with the outcome constant, Urchioli (1990) manipulated whether the outcomes were correlated or uncorrelated with the comparisons. In the correlated group, a correct red choice response was followed by food 100% of the time, whereas a correct green choice was followed by food 20% of the time. In the uncorrelated group, a correct red choice was followed by food 100% of the time if the sample was a vertical line, but 20% of the time if the sample was blue. The green choice was followed by reinforcement 20% of the time if the sample was a horizontal line, but 100% of the time if the sample was yellow. Put another way, in the correlated group the probability of an outcome given the sample is constant, but in the uncorrelated group the probability of an outcome given the sample is variable (Figure 1).

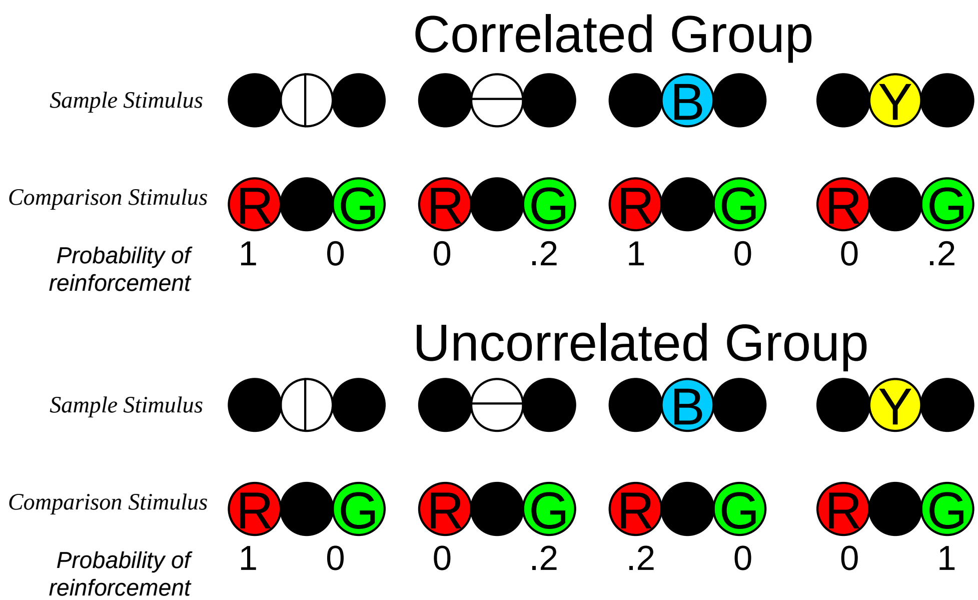

Figure 1. DOE: Example of the stimulus displays used by Urcuioli (1990)

In this experiment, any difference between the correlated and uncorrelated groups would be due to expectancies, not enhanced sample discriminability. This is because, for the correlated group, the expectancy produced by the sample was predictive of the correct choice; and in contrast, for the uncorrelated group, the expectancy produced by the sample was not predictive of the correct choice. For both groups, the sample discriminability should have been equally affected by the differential outcomes training procedure, meaning that any difference between groups cannot be attributed to increased sample discriminability.

Urcuioli's results (1990) support the view that the expectancies, not the enhanced sample's discriminability, cause the superior performance of subjects trained using the differential outcomes procedure. The pigeons in the correlated group met acquisition criterion more than twice as fast as those pigeons in the uncorrelated group. In tests where retention intervals were inserted (1 s, 2 s, and 4 s) the correlated group's accuracy declined less and remained higher across all intervals then the uncorrelated group's accuracy. For example, for the 4 s delay, the correlated group's accuracy was about 85% compared to the uncorrelated group's accuracy of about 65%. Since all the groups experienced the same degree of enhanced sample discriminability (if any) any effect of the increased sample discriminability could not explain differences between the groups. Again, the only difference between the groups was the reliability of the expectancies, and therefore the expectancies are the proposed reason for the correlated group's more accurate performance (the DOE).

While the previous experiment suggests that expectancies affect pigeon's performance, the experiment did not directly examine how the discriminability of the samples affects the differential outcomes training procedure. This meant that the DOE might be the product of either a combination of the expectancies and the enhanced sample discriminability or the expectancies alone. To test if the enhanced samples' discriminability contributed to the DOE, in Urcuioli's (1990) next experiment the differential outcomes correlated with each sample were removed. This was accomplished by having the pigeons experience all correct choice comparisons after the samples leading to the same probability of food, 60%. First the pigeons experienced 10 sessions where pecks to just the sample were reinforced with the same probability and then the former conditional discrimination task was reintroduced and correct comparison choices were reinforced with the same probability.

If the expectancies overshadowed the effects of enhanced samples' discriminability, then the sessions of experience with similar outcomes for the samples should have resulted in the correlated group's accuracy decreasing more than the uncorrelated group's performance across the retention intervals. This is because, for the correlated group during the original training each sample and its comparison predicted a particular outcome, but now all the samples and comparisons would predict the same outcome. In contrast, for the uncorrelated group the expectancies were never predictive of the correct comparison choice. Put another way, if expectancies overshadow the possibly effects of enhanced sample discriminability making the reinforcement probabilities the same changed more factors for the correlated group then for the uncorrelated group. In contrast, if the effects of expectancies do not overshadow the possibly enhanced sample discriminability, then both groups should be equally affected by the decreased sample discriminability.

The results revealed that the pigeons in the correlated group performed with a sizable decreased accuracy from their previous performance in delay tests while the uncorrelated group's performance for the delay tests were similar. As the correlated group was more greatly affected by the similar outcome experience than the uncorrelated group, these results suggest that expectancies may overshadow enhanced sample discriminability. However, Urcuioli's finding is complicated because despite the correlated group's greater decrease in performance accuracy then the uncorrelated group, the correlated group's performance was still more accurate overall than the uncorrelated group's performance. The enhanced overall accuracy of the correlated group suggests that the expectancies may not overshadow the enhanced sample discriminability. Further analyses suggested that the correlated group's superior performance was due to overtraining. Since the correlated group preformed more accurately in previous sessions, the pigeons in this group were reinforced more for the correct responses than the uncorrelated group. In summary, the results of both these experiments suggest that the DOE occurs because of both the presence of expectancies and enhanced sample discriminability.

To further explore why the DOE occurs, in a later experiment, Urcuioli (1991) again trained pigeons to perform a zero-delay many-to-one conditional discrimination. This time he included three groups: one where both the samples and the comparisons were correlated with the outcomes (CORR), one where the samples were correlated with the outcomes but the comparisons were uncorrelated with the outcomes (UNC) and a completely nondifferential outcome group where both the samples and the comparisons were uncorrelated with the outcomes (NDF). If the DOE occurred because the differential outcomes procedure enhanced the samples' discriminability, then the group trained with the nondifferential outcomes should learn the task slower than the groups where the samples were correlated with the outcomes, regardless of whether or not the comparisons were correlated with the outcomes.

He found that the NDF group actually acquired the task faster than the UNC group, Figure 2. If the differential outcome procedure did enhance the distinctiveness of the samples, then the UNC group should have learned the task faster than the NDF. From these results Urcuioli concluded that the differential outcomes procedure does not "affect conditional discrimination learning merely by enhancing the discriminability or distinctiveness of the samples" (p. 29), but instead the expectations generated by the differential outcomes procedure matter.

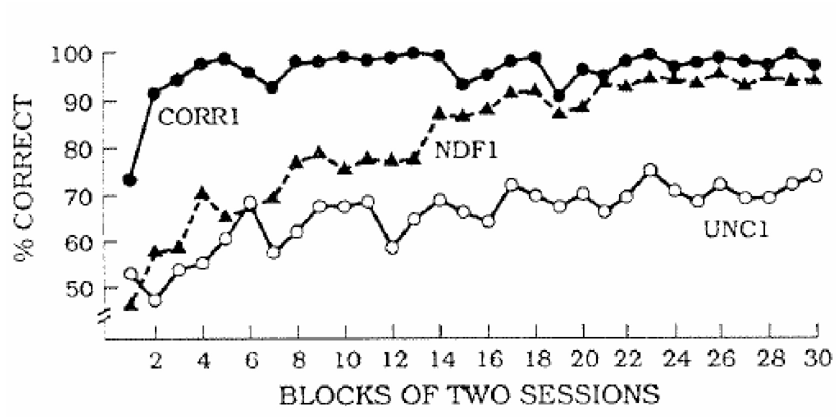

Figure 2. DOE: Figure obtained from Urcuioli (1991, p. 32). CORR indicates the samples and comparison items were correlated with the outcomes, NDF indicates that the samples and the comparisons were uncorrelated with the outcomes, and UNC indicates that samples were correlated but the comparisons were uncorrelated with the outcomes.

An explanation not offered by Urcuioli, but put forth by others, is that the differential outcome procedure enhances subjects' performance because it creates unique response topographies (Peterson, Wheeler, & Trapold, 1980). For example, subjects may respond slower to those samples and comparisons correlated with less food and faster to those correlated with more food. Supporting topographical differences, Brodigan and Peterson (1976) noted a pecking topography difference when they reinforced one discrimination with food and another discrimination with water. Specifically, the pigeons pecked sharply on the key followed by food and scooped gently on the key followed by water.

Also addressing response topography, DeLong and Wasserman (1981) taught pigeons a conditional discrimination where the first peck after 5 s to the red or green sample extinguished the sample and presented slanted and vertical line comparisons. For the differential outcomes group, one correct choice was reinforced with food 20% of the time and the other comparison was reinforced 100% of the time. For the nondifferential outcomes group, both correct comparison choices resulted in food reinforcement 60% of the time. The pigeons in the differential outcome group not only acquired the task faster, but they also responded at a faster rate to the sample correlated with 100% food than the sample correlated with 20% food.

While response topography may contribute to the DOE it should be noted that the only reason the responses differ is because of the outcome expectations the differential outcomes procedure creates. If the pigeons had the same expectations for both samples, then their rate of responding to each sample would not differ. Moreover, the DOE has been obtained when conditional reinforcers are used and response topographies do not differ. For example, Estévez and Fuentes (2003) obtained the DOE by providing four-year-old children with red or green tokens following correct choices. The effects of directly manipulated response topographies on the differential outcomes procedure would be an interesting area of research. For example, one could manipulate the number or rate of responding required for the conditional discriminations being trained.

In summary, there has been much debate as to whether the DOE occurs because the differential outcomes procedure enhances the discriminability of the samples, provides additional expectancies or provides expectancies alone. Experiments suggest that the DOE emerges due to the expectancies the differential outcomes procedure produces alone and that the differential outcomes procedure may not increase samples' discriminability. This is significant because it indicates that behavior can be affected by mental representations of future events. Other experimenters have pointed out that the differential outcomes procedure often creates different response topographies. The DOE is an fascinating phenomenon because it enhances performance without requiring much physical effort from subjects, and again because it is advantageous researchers and educators cannot help but concern ourselves with why the DOE occurs.

Species Comparisons

While the DOE has been widely examined with nonhuman animals, obtaining the DOE has been difficult with typical adult humans. This is of interest to comparative psychologists, because many experiments in comparative psychology investigate the similarities and differences between nonhuman and human animal’s cognitive abilities. Any differences these experiments find may be attributed to qualitative or quantitative factors, i.e., in kind or in degree. However, while there may be qualitative differences between nonhuman and human animals’ cognitive abilities, it is presumptuous to believe that when a nonhuman animal does not readily display a human cognitive ability that that animal lacks the cognitive ability. Instead, the difference may be quantitative and emerge because the methods used to obtain effects with humans may be inadequate to do so with other animals.

Regarding animal intelligence, Premack (1983) notes how the methods used to obtain cognitive effects with different animals has lead experimenters to misinterpret animals' cognitive abilities. In his work, he observed that when novel items were presented with familiar items in a matching-to-sample task, chimpanzees choose the familiar item regardless of the sample. From those results, an experimenter could reason that those results were indicative of a qualitative difference between humans and chimpanzees, being that chimpanzees do not possess the cognitive ability to match-to-sample that humans readily demonstrate. But instead Premack discusses possible quantitative differences between the animals; explaining that “while learning to choose the alternative that matched the sample, the animals also learned to choose toys, i.e., a class of items with certain properties…” and that "Although apes learn on both levels [item specific and relational], 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, 1983, p. 356). This passage suggests that Premack believed the traditional methods used to obtain matching-to-sample with apes had led experimenters to mistakenly underrate apes’ cognitive abilities.

Along similar lines, it is also important that experimenters do not overrate human’s cognitive abilities when the effects reliably obtained with nonhuman animals do not emerge with humans. Specifically relevant to the present report, although the DOE is a reliable finding with nonhuman animal subjects, it has been difficult to obtain with typical adult human participants. From those results, one may proposed that humans are qualitatively different from other animals. However, in the present report I will argue that the difference is quantitative and that expectations do affect humans learning. Furthermore, the DOE has been difficult to obtain with typical adult humans because the tasks used have been inadequate to do so. The inadequacy being that those tasks used are too easy, resulting in spectacular performance regardless of the method by which humans are trained (i.e., a ceiling effect). By increasing the task’s difficulty (Experiments 1, 2, and 3) and introducing a new method to test for the effect (Experiment 4) I aim to obtain the DOE with typical adult human subjects.

Before further discussing the present experiments, this report will explore three related issues. First, I will discuss the effects of stimulus discriminability on nonhuman and humans' performance. Second, I will examine the affects of temporal delays within learning tasks on nonhuman animals’ and humans’ ability to form associations between events. These discussions should show how discriminability and temporal manipulations affect species in qualitatively similar fashions. Third, I will review the available literature on the differential outcome procedure with children and cognitively disabled humans. After discussing these issues, the reader should feel confident about the characteristics of the present task and the manipulations used it to obtain and explore the DOE with typical adult humans.

Stimulus Discriminability

Both nonhuman and human animals' performance are affected by stimulus discriminability. For stimuli that differ on a one-dimensional continuum, Shepard's law of generalization states that the probability of responding decreases exponentially with psychological distance between novel and trained stimuli (Shepard, 1987). For example, after subjects are trained to respond to a 180º line, they can be tested to see how they respond to novel line orientations. The experimenter can then compare the way subjects' responds to novel line orientations to the way the subject responds to the trained 180º line. Compared to the subject's responding to the 180º line, more similar responding to novel line orientation is considered stimulus generalization, and more different responding is considered stimulus discrimination. The more perceptually different the novel line orientations and the 180º line are, the faster the subjects will respond differently to the novel line orientation (e.g., after training the subject to touch a 180º line, the subjects may never touch a 90º line). Consequently, in a conditional discrimination task, if the stimuli are not discriminable (e.g., a 180º and a 179.75º line), then subject will never learn to respond differently to them.

While it is essential that the stimuli in conditional discrimination tasks are discriminable, how discriminable those stimuli are is at the experimenter's discretion. Assuming that the two stimuli differ in a perceivable way, the amount of similarity between the stimuli will be positively correlated with the time it takes the subjects to respond differently to those stimuli. As the stimuli become more similar the task becomes more difficult. Accordingly, any procedure that can enhance the discriminiability of the stimuli (as the differential outcome procedure may do to the sample stimuli) should increase the subjects' acquisition rate.

Nonhuman Animals. One experiment that exemplifies the effects of stimulus discriminability on acquisition of conditional discriminations was completed by Urcuioli and Zentall (1986). In their experiment, pigeons learned a zero-delayed conditional discrimination with two very discriminable color stimuli, red and green, and two less discriminable line orientation stimuli, 90º and 180º (i.e., because color information is contained in individual pixels, whereas line orientation information is contained in the pixels' spatial relations to each other, colors are more discriminable than line orientations). The subjects were split into groups where the task included either more discriminable comparisons or more discriminable samples (see Figure 3 for an example of the stimuli displayed).

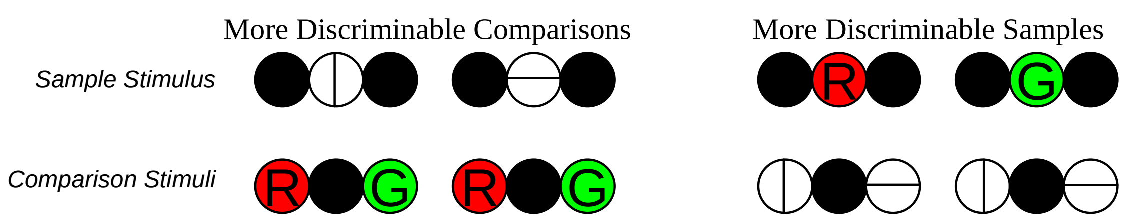

Figure 3. Stimulus Discriminability: Example of the stimulus displays used by Urcuioli and Zentall (1986)

As a brief aside, it should be noted that while this experiment demonstrates the affects of stimulus discriminablity on the acquisition of a conditional discriminations, Urcuioli and Zentall (1986) were actually interested in the memory strategies pigeons use to learn conditional discriminations. They reasoned that if the pigeons used retrospective memories of the samples, then the pigeons that received more discriminable samples would perform more accurately over increasingly long retention intervals (a delay between the sample offset and comparisons onset) than those pigeons that had less discriminable samples. In contrast, if pigeons used prospective memories of the comparisons, then the discriminability of the comparisons should matter and the pigeons that received more discriminable comparisons would perform more accurately over increasingly long retention intervals than those pigeons that had less discriminable comparisons.

When the sample groups were compared, they found that the pigeons that had more discriminable samples performed more accurately across increasingly long retention intervals than those that had less discriminable samples. When the comparison groups were analyzed, they observed no difference between them. However, if the comparison groups' accuracies were collapsed across the retention intervals, the group with the more discriminable comparisons preformed more accurately than the group with the less discriminable comparisons. Urcuioli and Zentall interpreted these results as meaning that pigeons rely more heavily on retrospective memories in conditional discrimination tasks (i.e., memory of the sample). However, since more discriminable stimuli enhanced performance regardless of if they were the sample or the comparison when the data were collapsed across the delays, the discriminability of the comparisons might also matter. More conservatively stated, the results of this experiment suggest that the more discriminable stimuli are the more accurately subjects perform.

Humans. Like pigeons, humans are also affected by stimuli's discriminability. For example, imagine a mother picking up her child from school. To complete this visual search task, the mother must locate her child (i.e., the target) amidst many other children (i.e., the distracters) exiting the school at the same time. Finding a red headed child is easier than finding a brunette (a feature search). Visual searches are similar to conditional discriminations in that how an individual responds is conditional upon which stimulus is present. For example, the mother will respond in a different manner to her child than to a child that is not hers.

To better understand the processes that make visual search tasks more or less difficult, in the laboratory experimenters manipulate the number and discriminability of distracter items on computer displays. In a typical visual search task participants are presented with visual displays of various items and asked to respond whether a target is present or absent. Figure 4 below provides an example of displays used by Sobel and Cave (2002). In their experiment participants were asked to locate the vertical red line target amongst distracters that were red or green with distinct or similar orientations (a conjunctive search).

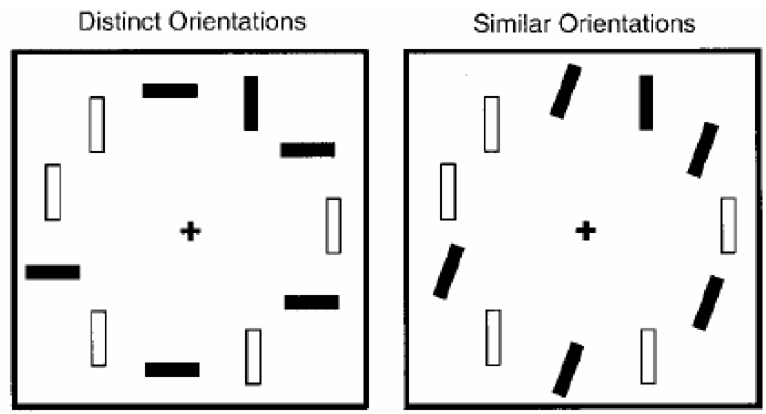

Figure 4. Stimulus Discriminability: Example of the stimulus displays used by Sobel and Cave (2002, p. 1058) Filled rectangles represent red objects, and unfilled rectangles represent green objects.

They found that when the orientations were distinct, the subjects tended to perform faster and more accurately than when they were similar. Figure 5 shows the participants' reaction times and error rates as a function of the number of red distracters. When the distracters had similar orientations, the participants took longer to react than when the distracters had distinct orientations. Also, when the distracters had similar orientations, the participants made more errors, at least when the target was present and the number of distracters was sufficient.

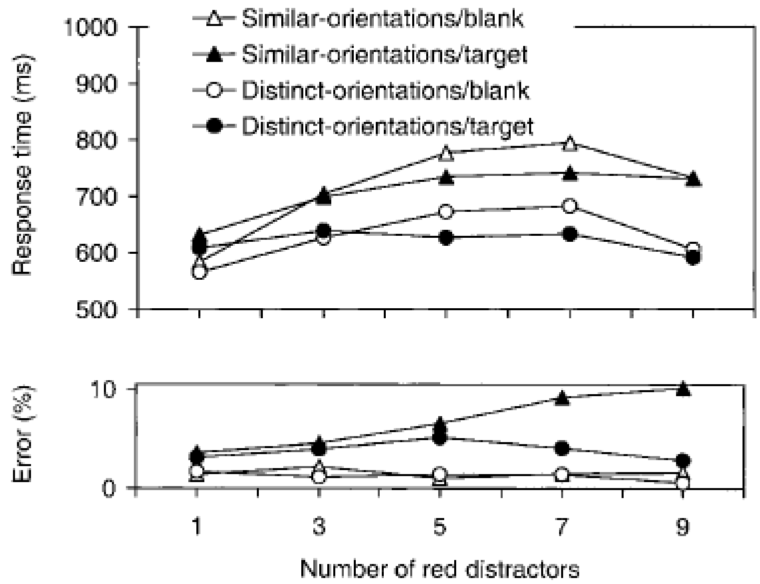

Figure 5. Stimulus Discriminability: Figure obtained from Sobel and Cave (2002, p. 1059)

In summary, both nonhuman animals' and humans' performance are affected by stimulus discriminability. Specifically, pigeons performance was better when highly discriminable stimuli were used (colors) than when less discriminable stimuli were used (line orientations). Human performance was faster and more accurate in a visual search task when the distracters were more discriminable (distinct orientations) than when the distracters were less discriminable (similar orientations). Further, these results stress that if the differential outcomes procedure enhances the discriminability of the samples, then that would, at least partially, explain its ability to create the DOE.

Within Session Delays

The delays within a task also affect nonhuman animals' and humans' performance. To imagine the effects of within session delays, first consider the time course of a typical conditional discrimination task. First, subjects are presented with the sample item which may be extinguished after a specified response requirement or duration. Then there may or may not be a delay where no stimuli are present (a retention interval) before the comparisons are presented. Then the subjects make a choice response to one of the comparisons. After some duration (feedback interval), the subjects are rewarded if the chosen comparison is associated with the sample and punished if it is not associated with the sample. After the contingency is delivered, the subjects must wait a programmed amount of time before the next sample is presented (the intertrial interval). These three types of delays, retention, feedback and intertrial interval, each affect how subjects perform conditional discriminations. Typically shorter intervals between events within one trial facilitate the formation of associations between them and longer intervals between each trial help subjects to dissociate each trial.

Nonhuman animals. An experiment by Sargisson and White (2007) offers insights into both the effects of retention and feedback interval on pigeons' performance in delayed matching-to-sample tasks with different feedback intervals (in their paper they refer to feedback interval as reinforcement delay, and the feedback interval after an incorrect response was always 0 s). In their experiment six pigeons learned a delayed matching-to-sample task with red and green items. On a single trial, a red or green sample was displayed on the center key of a three key array. After pecking the sample five times it extinguished, followed by a programmed retention interval (0, 2, 4, 6, or 8 s) where all the keys were blackened. Then the red and green comparisons appeared on the left and right keys. Correct responses resulted in 3 s of food access after a feedback delay (0, 2, or 4 s) and according to the programmed reinforcement probabilities (.5/.5, .2/.8, and .8/.2). Incorrect responses resulted in a 3-sec blackout immediately. A 15-s intertrial interval followed all reinforcements and timeouts.

Figure 6 below presents three graphs, one for each reinforcement probability. The x-axis illustrates the retention intervals and the lines represent different feedback intervals (i.e., reinforcement delays), and each graph represents a different reinforcement probability condition. Interestingly, unequal reinforcement probabilities (the center and right graphs) created a flatter memory function then equal reinforcement probabilities (the left graph). But importantly, the pigeons' performance decreased as the retention interval increased in all feedback intervals and reinforcement probability conditions. This indicates that both retention interval and feedback interval influence pigeons' performance.

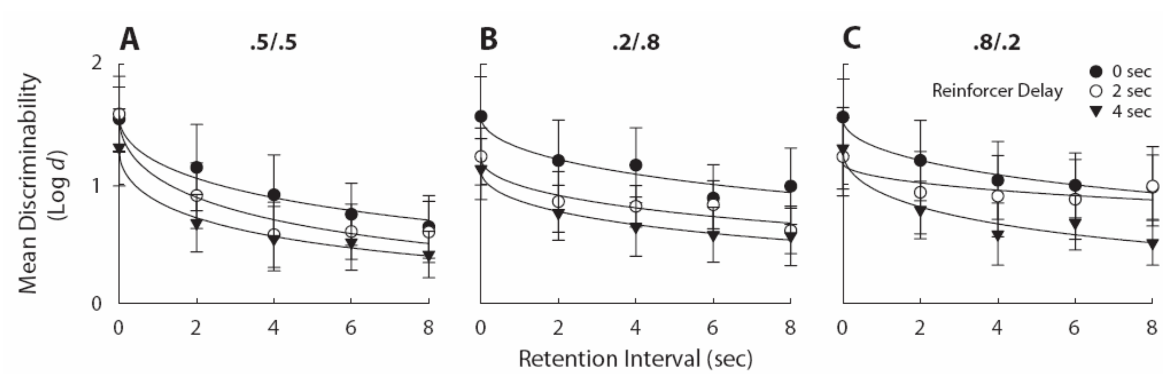

Figure 6. Delays: Figure obtained from Sargisson and White (2007, p. 180)

Pigeons are also affected by the duration of the intertrial interval, as demonstrated in Holt and Shafer (1973) experiment. In their experiment, 12 pigeons acquired a matching-to-sample task with a 0, 5, 15, 25, or 60 s intertrial interval. While both pigeons in the 0 s group and one pigeon in the 5 s group failed to acquire the task, all of the other pigeons with longer intertrial intervals acquired the task. Also interesting, those pigeons with longer intertrial intervals preformed more accurately than those with shorter intertrial intervals (Figure 7).

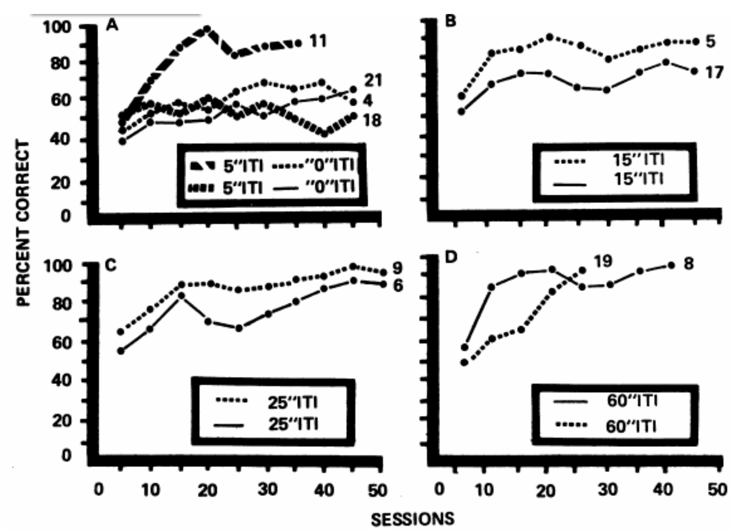

Figure 7. Delays: Figure obtained from Holt and Shafer (1973, p. 183)

Humans. Like pigeons, human performance is also affected by the three discussed delays: retention, feedback and intertrial interval. King, Jones, Pearlman, Tishman, and Felix (2002) found evidence that human performance is negatively affected by longer retention interval's duration. In this experiment, participants were presented first with a warning stimulus and then a line that was either 17, 23, 31, or 43 pixels long followed by a black screen retention interval of .4 or 3.3 ms. A second line was then presented that had either the same or a different length as the first and the participant responded by pressing S on a keyboard if they were the same or D if they were different. They found that longer retention intervals decreased participants' accuracy.

Regarding the effect of feedback delay, Maddox, Ashby, and Bohil (2003) found evidence to support that human performance is negatively affected longer feedback interval. In their experiment, participants separated lines into two categories. On a typical trial participants were presented with a line which they categorized and then saw a mask over the line for 0, 2.5, 5, or 10 s (the feedback interval) before the receiving feedback, the word "Correct" or "Error". They found that longer the feedback interval was the less accurately the participants performed.

Lastly, regarding the effect of intertrial interval, Williams, Johnston, and Saunders (Experiment 2, 2006) compared the accuracies with which adults with mental retardation performed a delayed matching-to-sample task with either a 2 or 8 s intertrial interval. They found that the participants' accuracies were higher across increasingly long retention intervals when the intertrial interval was longer (Figure 8).

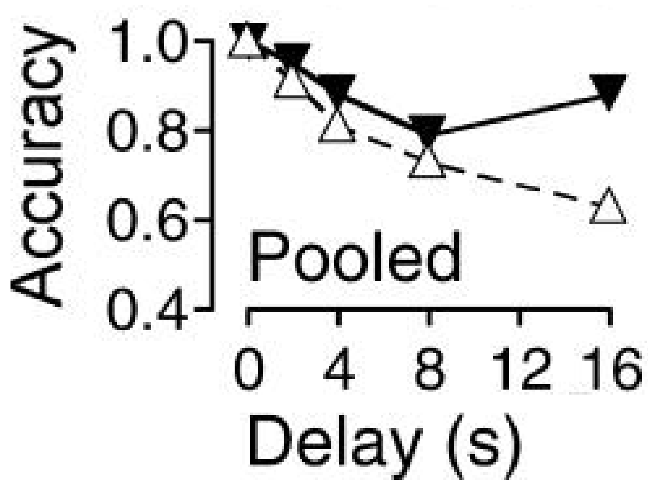

Figure 8. Delays: Figure obtained from Williams, Johnston, and Saunders (2006, p. 263). The black line represents the 8 s intertrial interval group and the dashed line represents the 2 s intertrial interval group.

In conclusion, both nonhuman animals and humans are affected by delays within a task. Although the procedures used to test nonhuman and human animals presented here are different, the effects created when critical factors similar to all the tasks are the same. While longer retention and feedback intervals degrade performance, longer intertrial intervals enhance performance. These results suggest that for the differential outcomes procedure, short intervals between experimental events in each trial and long intervals between each trial may be essential for participants to form effective associations and produce a DOE.

Humans and the DOE

Although most support for the DOE comes from nonhuman animal research, the DOE has also been obtained with human subjects. Unfortunately, this support is mostly limited to children and cognitively disabled populations. If the differential outcomes procedure is a general procedure, it is important that it also affects typical adult humans. Further, if the DOE facilitates typical adult human learning, then the differential outcomes procedure would be an excellent procedure to use in classrooms and other training situations. Before discussing the limited research with typical adult humans, I will review research with mentally challenged and young humans.

Mentally Challenged Humans. Malanga and Poling (1992) obtained the DOE with four mentally challenged participants. In their experiment, the participants learned the association between two sign language letters (e.g., A and E, G and H). The participants experienced both conditions; one where correct responses were followed by either correlated and the other where uncorrelated outcomes (food or verbal praise). The participants' terminal accuracies were higher when correct responses were followed by a correlated reinforcer type than an uncorrelated reinforcer type. These findings suggest that the differential outcomes procedure can be employed to effectively teach discriminations to adults with mental disabilities.

Estévez, Fuentes, Overmier and González (2003) obtained the DOE with children diagnosed with Down's syndrome. In their experiment, 5 to 5.5 -year-old children learned several picture conditional discriminations. During the task, the participants saw three pictures on a page contained in a binder. They were told to guess which bottom picture (the comparisons) "goes with" the top picture (sample). Delays between the sample and two comparisons were gradually introduced so that eventually the sample stimulus appeared on one page, which was then flipped by the experimenter to a blank sheet for a predetermined time and then to a third page to reveal the comparisons. The participants in the differential outcomes group received a different correlated colored token for each type of correct response, while participants in the nondifferential outcomes group received uncorrelated colored tokens for each type of correct response. Red tokens could later be exchanged for food and the green tokens for toys. Examples of these stimuli and how differential outcomes were employed is given in the Figure 9.

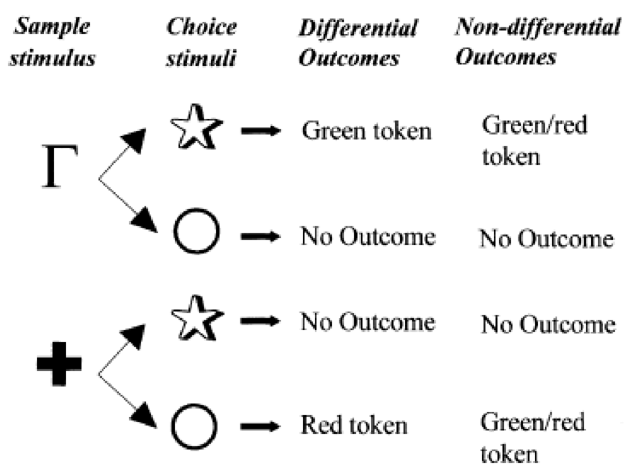

Figure 9. DOE Humans: Figure obtained from Estévez, Fuentes, Overmier and González (2003, p. 153).

Their results, as depicted in Figure 10, revealed that the children in the differential outcomes group learned the discriminations more quickly than those in the nondifferential outcomes group. These findings suggest that the differential outcomes procedure can be employed to effectively teach discriminations to children with mental disabilities.

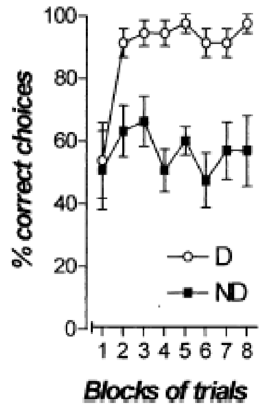

Figure 10. DOE Humans: Figure obtained from Estévez, Fuentes, Overmier and González (2003, p. 156).

Hochhalter, Sweeney, Bakke, and Overmier (2000) obtained the DOE in an experiment using four adult humans with alcohol-induced dementia (AlcDem) and four without AlcDem (control). AlcDem is a progressive and permanent condition resulting from chronic alcohol abuse and its symptoms include cognitive impairment and memory loss. In the experiment, the participants learned conditional discriminations with black and white high school year book photos presented in a binder. The experimenter first presented the sample photo, on one page, then flipped to second blank page for a predetermined retention interval (2 seconds during the training phase and after reaching a criterion 2, 5, 10, or 25 s during the testing phase) and then flipped to a third page with the two comparison photo choices. The participants experienced both conditions, differential and nondifferential outcomes procedures. During the differential outcomes condition, after correct responses the participants received a correlated outcome for each correct choice (a nickel for one type of discrimination and a point displayed on a computer screen for another type of discrimination). And during the nondifferential outcomes condition, participants received an uncorrelated outcome for a correct response.

All the participants quickly acquired the task before advancing to the test phase. However, differences were obtained between the participant populations (AlcDem and Control) and the training procedure (Differential and Nondifferential) used. As depicted in the Figure 11, the control participants remained stable and highly accurate across increasingly long retention intervals. In contrast, the participants with AlcDem decrease in accuracy as the length of the retention intervals increased. The training procedure, did not affect the Control participants’ performance, but did enhance the AlcDem participants' performance across the retention intervals. These findings suggest that the differential outcomes procedure can be employed to effectively teach discriminations to adults with an induced mental disability.

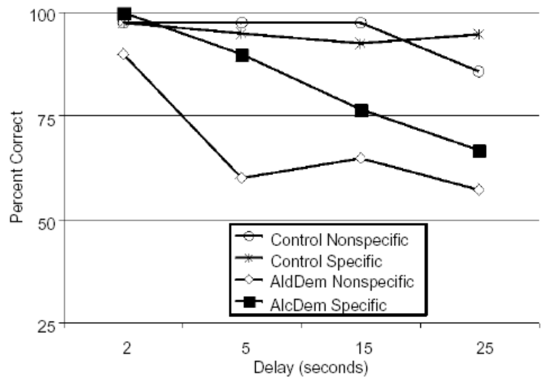

Figure 11. DOE Humans: Figure obtained from Estévez, Fuentes, Overmier and González (2003, p. 11).

Typically Developing Child Humans. In 1995, Maki, Overmier, Delos and Gutmann obtained the DOE with normal 4 to 5 -year-old children. In their task children learned conditional discriminations with picture stimuli. Then to produce evidence that the DOE could be obtained with a broader age range, Estévez, Fuentes, Marý-Beffa, González and Alvarez (2001) replicated their result using a similar task with 4 to 7-year-old children. Their task include two different groups, a differential outcome group that received correlated outcomes and a nondifferential outcome group that received uncorrelated outcomes for correct response (the reinforces included toys or food). Examples of the stimuli used in phase 1 and phase 2 are given in Figure 12.

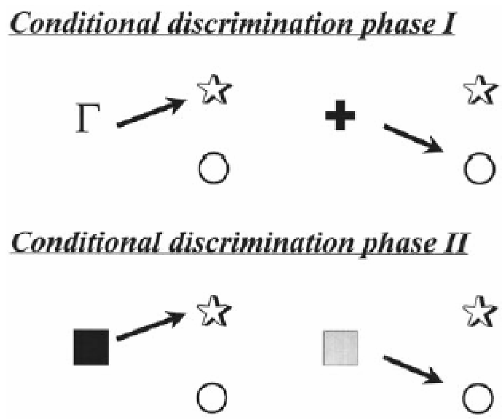

Figure 12. DOE Humans: Figure obtained from Estévez, Fuentes, Marý-Beffa, González and Alvarez (2001, p. 52).

Their results, as shown in Figure 13, were that the group which received differential outcomes acquired discriminations faster and preformed them more accurately than those that received nondifferential outcomes. These results indicate that the differential outcomes procedure also affects older typically developing children.

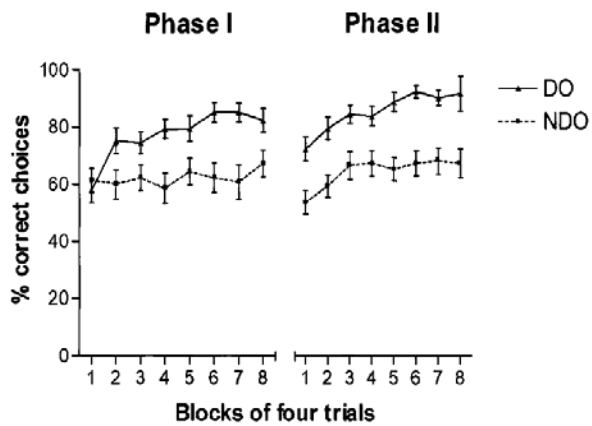

Figure 13. DOE Humans: Figure obtained from Estévez, Fuentes, Marý-Beffa, González and Alvarez (2001, p. 56). DO represent the differential outcome group, and NDO represents for the nondifferential outcome group.

To obtain the DOE with older children, Estévez, et. al. (2001) increased their task's difficulty. Specifically, in their experiment (where the DOE was obtained with children 4 to 7-year-old), the children learned a delayed matching-to-sample task with highly discriminable stimuli and where *two* comparison choices followed a sample's offset. In Experiment 2 (where the DOE was obtained with children 7.5 to 8.5 –year-old), the children learned a delayed matching-to-sample task with less discriminable stimuli and where *four* comparison choices followed a sample's offset. Examples are given in the Figure 14.

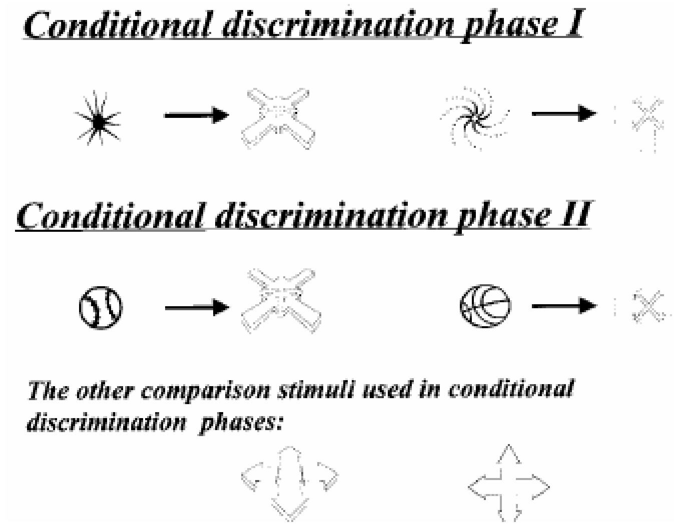

Figure 14. DOE Humans: Figure obtained from Estévez, Fuentes, Marý-Beffa, González and Alvarez (2001, p. 59).

By decreasing the discriminability of the stimuli within the task and increasing the number of distracters the participants had to choose between they obtained the DOE with 7.5 to 8.5 -year-old children (Figure 15). . These findings then suggest that not only does the differential outcomes procedure affect developmentally disabled and young children, but also older children if the task is sufficiently difficult.

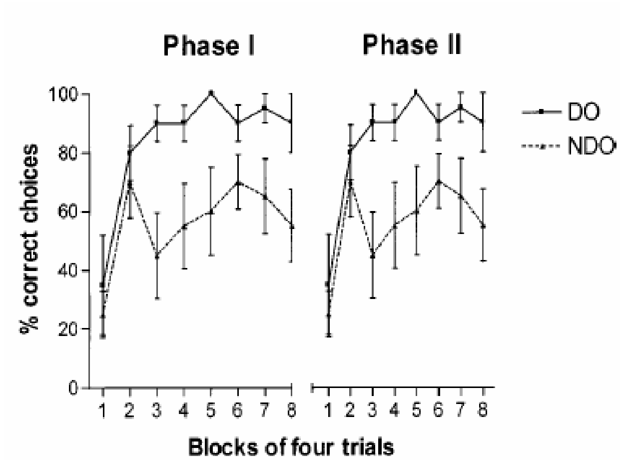

Figure 15. DOE Humans: Figure obtained from Estévez, Fuentes, Marý-Beffa, González and Alvarez (2001, p. 60). DO represents the differential outcome group, and NDO represents the nondifferential outcome group.

Typical Adult Humans. Often, the DOE is not obtained with typical adult humans because the tasks participants are asked to learn are too simple. Task ease creates a ceiling effect, which is a situation where participants trained with the nondifferential outcomes procedure perform so well there is no benefit to using the differential outcomes procedure (Estévez, 2005). The importance of task difficulty for obtaining the DOE should not be surprising at this point. After all, Estévez, et. al. (2001) increased the difficulty of his task to obtain the DOE with older children. Therefore, to obtain the DOE with typical adult humans may require an even more complicated task that that used with older children, possibly with more to-be-learned discriminations and less discriminable stimuli. At least two experiments have been performed which obtained the DOE with typical adult humans by increasing the task’s difficulty.

Estévez, Vivas, Alonso, Marý-Beffa, Fuentes, Overmier and González (2007) attempted to see how the DOE might affect typical adults’ use of a relational association (greater than / less than) to respond to numerical pairs. In their experiment 1 typical adult humans (college students) 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 Experiment 2, 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). Unfortunately, the results of the first and second experiment are inconsistent. That is, the results cause one to question if the differential outcomes procedure affects reaction time (as in the first experiment) or accuracy (as in the second experiment). Furthermore, the participants entered the task already aware of number relations. So this experimental procedure only allows one to explore look at how the differential outcomes procedure affects human performance, not human learning.

Miller, Waugh, and Chambers (2002) more clearly obtained the DOE with typical adult humans (college students) by making their task more difficult. To increase the difficulty they increased the number of to-be-learned associations and decreased the discriminability of the stimuli in the task, by asking the participants to learn the English meaning of 15 different kanji (Japanese characters).

To complete their task, participants were seating in front of a computer and were first presented, with 1 of 15 kanji which were then replaced with 9 of the 15 possible English words (see the left panel of Figure 16 for kanji and their English translations). The participants then used a mouse to select the English word they thought the kanji meant. If the participants selected the correct English word, then they saw 1 of 15 picture (see the right panel of Figure 16) followed by a message informing them they were correct and had received a lottery entry in a draw for 1 of 15 possible prizes. If the participants selected the wrong English word, then they saw a message informing them that they were wrong and what would have been the correct answer (i.e., corrective feedback) (Figure 17).

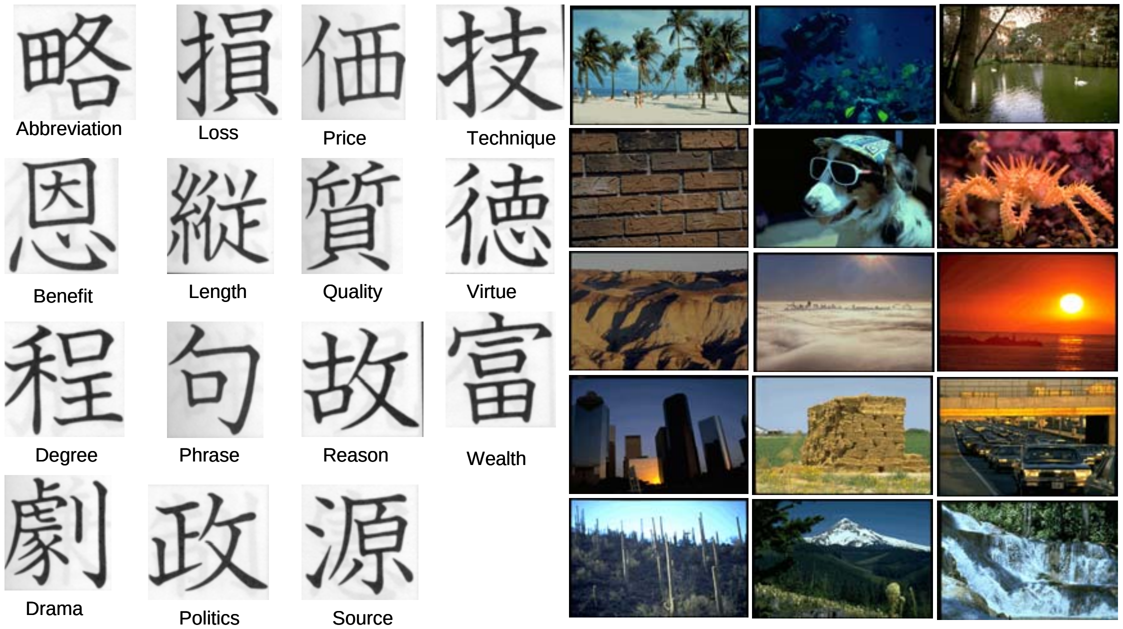

Figure 16. DOE Humans: Examples of the stimuli used in Miller, Waugh, and Chambers (2002) experiment, on the left the kanji (discriminative stimuli) and on the right the picture outcomes (stimulus reinforcers).

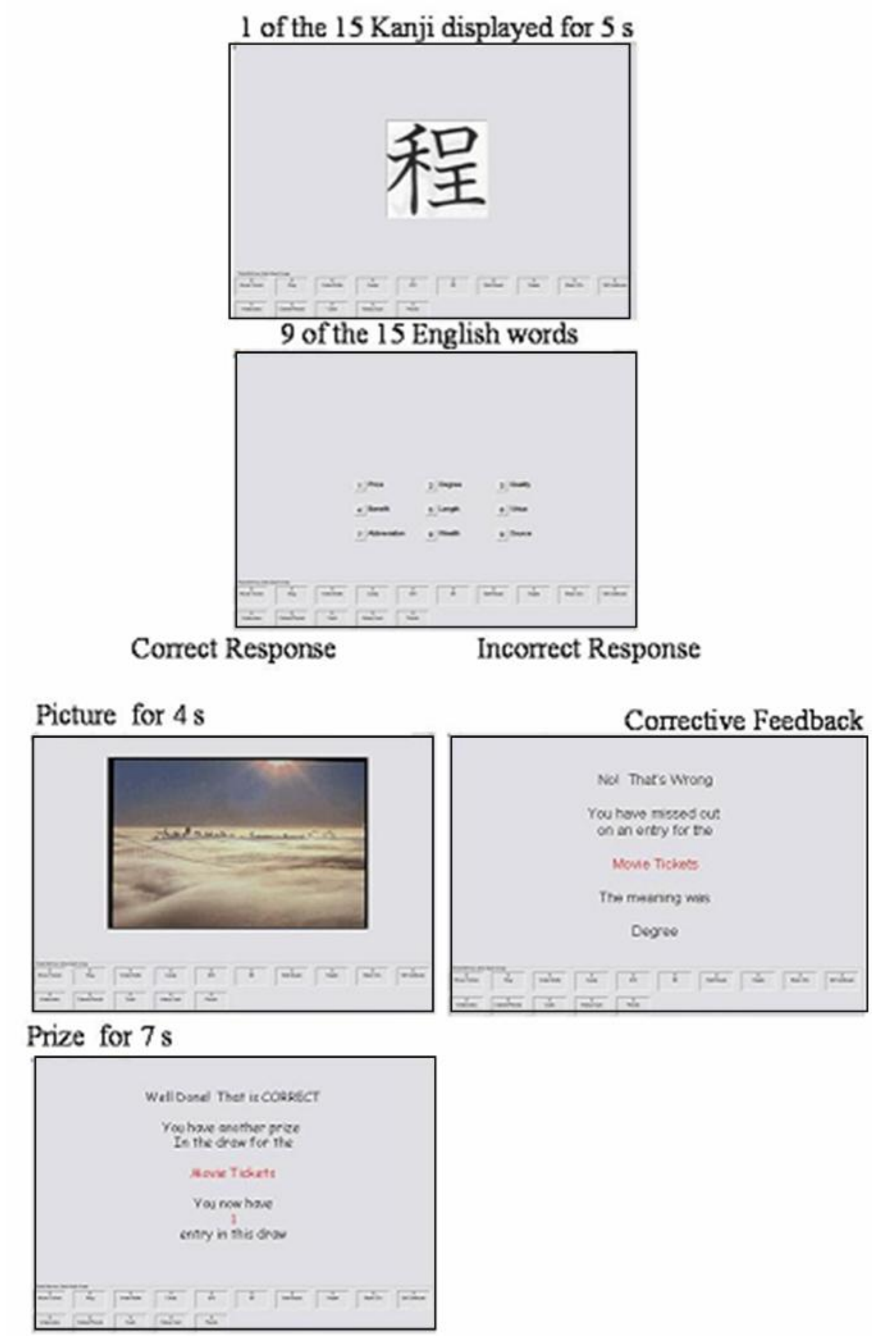

Figure 17. DOE Humans: Examples of a typical flow of a trial in Miller, Waugh, and Chambers (2002) experiment, temporally from the top to bottom when the participant responds correctly. All ISI and ITI 0 s.

Depending on the participant's group, the picture that appeared was either correlated or uncorrelated with the kanji and similarly, the prize entry was either correlated or uncorrelated with the kanji. They ran participants in a randomly determined order in three groups: 1) a differential outcomes group where both the pictures and prizes were correlated, 2) a partially differential outcomes group where the pictures were correlated but the prize entries were uncorrelated, and 3) a nondifferential outcomes group where both the pictures and the prize entries were uncorrelated.

Their 63 college participants ranged from 18 to 38 -year-old and were recruited from the University of Canberra to receive course credit or voluntarily. The participants were run individually by the same experimenter in a small quiet room on a PC computer with a 15" monitor and a standard keyboard and mouse. The participant completed three sessions, each lasting approximately 15 minutes. Sessions were divided into three blocks, and each kanji appeared once without replacement in a randomly determined order. After completing three sessions, the participants rank ordered the prizes and were debriefed on the purpose of the study.

Miller, et. al. (2002) found a session effect which indicates that the participants were learning the kanji's English meanings. Across the sessions, the accuracies increased from about 25% in the first session to 70% by the last session. Note that the mean accuracy remained well below 100%, indicating that their procedure did not suffer from the ominous ceiling effect commonly found in previous human experiments. Comparing the groups' acquisitions, they found that the differential outcomes group acquired the task faster than the partial and nondifferential outcomes groups across the three sessions, there (Figure 18).

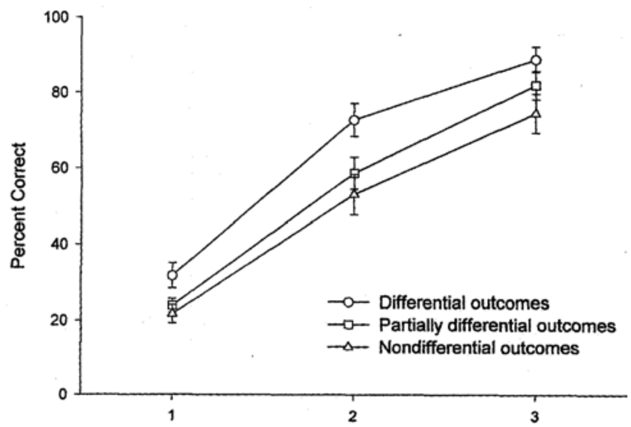

Figure 18. DOE Humans: Figures obtained from Miller, Waugh, and Chambers (2002, p. 320)

Miller, et. al.'s (2002) results add a typical adult human population to the list of species whose learning is affected by the differential outcomes procedure. The ever growing list of affected species suggests that the differential outcome procedure is a general mechanism that can be used to enhance learning. Furthermore, it adds supporting evidence to the idea that there is not a qualitative difference between the way expectations affect nonhuman animal and human learning.

To generate the DOE with typical adult humans, Miller, et. al. manipulated two factors of difficulty. First, their task required participants to learn more information than is typically required (i.e., 15 associations). Second, their task included stimuli that were less discriminable (i.e., kanji). Both or either one of these factors may be sufficient to generate the DOE with typical adult humans, and the specific factor that was necessary to obtain the DOE remains a question for a further research. However, instead of addressing that question, the present experiments aim to replicate their results that demonstrate human learning being affected by expectations and then explore why the effect occurs in some situations and not in others.

In Miller, et. al.'s (2002) discussion they acknowledge two ways to improve their experiment. First, they say that a fourth group could be added that experiences partially differential outcomes with uncorrelated pictures and correlated prizes. Not only would that complete their 2 X 2 factorial design, but it would offer insights into how the temporal placement and the incentive value of the outcomes affect the DOE. In their experiment, the pictures were temporally near the choice response, but have little incentive value. In contrast, the prize entries were temporally delayed, but have more incentive value. Second, they suggest that the effect of corrective feedback should be examined. This might reveal if obtaining the DOE with typical adult humans depends on the inclusion of corrective feedback.

The present experiments replicate Miller, et. al.'s (2002) findings and expand upon them. Experiment 1 replicates the Miller et. al.’s results, with the fourth group included (i.e., pictures uncorrelated and prizes correlated). Since any significant effects obtained in Experiment 1 may depend on either the presence of corrective feedback and/or on the order of the outcomes Experiments 2 and 3 are preformed. To examine the affects of corrective feedback, Experiment 2 omits corrective feedback from the procedure. To examine the affects of the temporally placement of the outcomes, Experiment 3 reverses the order of the outcomes used in the procedure (i.e., prizes then pictures instead of pictures then prizes). Finally, Experiment 4 uses different stimuli and includes a novel method to examine effects that the differential outcomes procedure generates when it fails to generate differences in acquisition.

EXPERIMENT 1

Methods

Participants

Sixty Auburn University undergraduate students, between 18 and 25-year-old were recruited though an internet database (Sona-System). For their participation they received extra credit for a psychology course.

Apparatus

Participants were run in a small room in Auburn University's psychology building. The task was completed on a PC computer and behind the 15" CRT monitor was a cream-colored wall. The computer monitor was positioned 33 cm from the front of the table. The participants sat in a rolling chair at a distance they considered comfortable from the table top (approximately 38 cm from the front of the table). The program was written in Visual Basic 6.0.

Stimuli

Kanji Stimuli. The 15 kanji characters were presented as black characters encased in rectangles ranging from 7.2 - 7.8 cm in height and 6.4 - 7.9 cm in width atop a grey background (Figure 16, left).

English Word Stimuli. Nine of the fifteen English words were selected and displayed in a 3 by 3 matrix. All text was black and the background was grey. The words were: abbreviation, benefit, degree, drama, length, loss, phrase, politics, price, quality, reason, source, technique, virtue, and wealth.

Picture Outcome Stimuli. The 15 pictures were obtained from Eureka Software's Graphics Explosion Pack. To the experimenter, no picture had an explicit affective value. The pictures were of a: beach, brick wall, canyons, cityscape, desert, divers, dog, golden gate bridge, haystack, mountain, pond, sea creature, sunset, traffic, and waterfall (Figure 16, right).

Prize Outcome Stimuli. A bar at the bottom of the computer screen tracked how many entries the participant had won across all three sessions. The 15 lottery prizes were: $5, $10, bath beads, blank CDs, candy, colored pencils, a deck of cards, a gift certificate, a hacky sack, movie tickets, a mug, pencils, a stapler, a water bottle, watercolors. When the experiment was completed 15 winners were drawn and contacted via e-mail to receive their prize.

Procedure

Participants were tested individually in a pseudorandom order, where each condition was randomly ordered 15 times. Before the computer task began, the participants rank ordered the prizes from their least to most preferred prize. Note that this is not the order in which Miller, et. al. (2002) required participants to rank the prizes. Instead, Miller, et. al. informed the participants that they would be ranking the prizes before they ran their sessions, and then upon completing the computer sessions they ranked the prizes. The placement of the prize ranking was altered in the present experiment to ensure that the participants would be aware of the prizes and would have considered the prizes' incentive values prior to acquiring the kanji's English meanings.

Acquisition. Before beginning a session the participants read the following message on the computer screen:

> Welcome to the experiment. Your task is to learn the meaning of 15 kanji. We will present you with one kanji at a time. It will be on the screen for 5 seconds before it is removed and replaced with nine different words. One of the words will mean the same as the kanji you just saw. You will need to select the word you think correctly matches the meaning of the kanji. If you get this correct you go into the draw for a prize (each kanji has its own prize). Every time you get that kanji correct you will get another entry. The more times you enter the better your chances of winning. To start with you will not know any of the kanji characters; therefore you will only get about 1 in 9 correct. But as you learn, you will get more and more entries in the prize draws. WIN WIN WIN. Use the mouse to select the word. Click OK to continue.

Each kanji appeared in a random order within a 15 trial block. Each session included three blocks (45 trials), and participants completed three sessions in one seating (135 trials). Each session typically lasted less than 15 minutes after which the experimenter set up and started the next session (an intersession interval of less than one minute).

The sequence of a trial is presented in Figure 17. First, a kanji appeared for 5 s, and then was replaced with 9 English words. The array of English words included the correct choice word and 8 incorrect choice words randomly selected from the 14 remaining words. If the participant clicked on the correct English word, they saw a 4 s presentation of the picture outcome and then a 7 s presentation of a message that read "Well, done! That is correct. You have another prize entry into the draw for the _______. You now have _____ entries in this draw." After an incorrect choice they did not see a picture or prize entry but instead saw a message that read, "No, that's wrong! The correct answer is ____. You have missed out on the prize entry in the draw for _______." All text was black, except the prize entry information which was in red text. After each trial's completion (correct or incorrect), the next trial's kanji then immediately appeared (0 s intertrial interval).

Groups. The pictures and prizes that followed correct choices depended on the participant's group. There were four different groups. 1) In the *picture and prize correlated* group each kanji was correlated with a particular picture and prize; 2) In the *picture correlated* group each kanji was correlated with a particular picture but uncorrelated with the prize entry; 3) In the *prize correlated* group each kanji was uncorrelated with the pictures, but correlated with a particular prize entry. 4) And in the *picture and prize uncorrelated* group each kanji was uncorrelated with both the pictures and the prizes.

Recall. Immediately after completing all three sessions the participants were asked to complete the Kanji Recall Worksheet. This worksheet displayed all 15 kanji (8 in the left column and 7 in the right column) with blank lines to their right. Participants were instructed to write the English word aside each kanji that they believed that kanji meant. If they did not know what the word was, then they were told to guess. If they could not recall the English words used in the task then they were told to make up a word to avoid frustration and unreasonable perseverance.

Results

Acquisition

The participants' performance in the first block (15 trials) was compared to chance accuracy. Chance in this task is 11.11% (1 in 9 English words is correct on any trial). Overall, participants performed better than chance in the first block, as indicated by a one-sample *t* test, $t(59) = 6.27$, $p < .05$. Next, each group was examined separately. A significant difference emerged for the *picture and prize correlated* group ($M = 18.67$, $SEM = 2.70$), $t(14) = 2.80$, $p < .05$, the *picture correlated* group ($M = 21.78$, $SEM = 2.30$), $t(14) = 4.64$, $p < .05$, and the *prize correlated* group ($M = 17.78$, $SEM = 1.80$), $t(14) = 3.70$, $p < .05$, and marginally the *picture and prize uncorrelated* group ($M = 15.99$, $SEM = 2.59$), $t(14) = 1.886$, $p = .08$. This indicates that all groups were able to perform better than chance in the first block.

One method participants can use to perform better than chance over the first block is to eliminate choices that are indicated as correct on previous trials. This strategy would work, as each kanji was randomly selected without replacement in blocks of 15 trials. For example, if Wealth appeared as a choice response option on both trials one and two, but was indicated as being the correct answer on trial one in the corrective feedback message, it could not be the correct answer on trial two. If the participant eliminates Wealth as an option on trial two, then their chance of being correct on trial two is 1 in 8, instead of 1 in 9. By consequently eliminating choice response words, the participant can increase their chance of being correct.

All groups learned the task. This resulted was confirmed by a two-way repeated measures ANOVA with Group (picture and prize correlated, picture correlated, prize correlated, picture and prize uncorrelated) X Session (1, 2, 3), which showed only a main effect of Session, $F(2, 118) = 299.48$, $p < .05$. Next, the group acquisitions were compared to see if the correlated condition of the picture and/or prize outcomes enhanced groups' accuracies.

Figure 19 shows mean acquisition across sessions grouped by picture correlated conditions on the left and prize correlated conditions on the right. As a result, each graph line combines the mean of two groups. Specifically, in the picture condition graph on the left, the filled circles represent the *picture and prize correlated* group and the *picture correlated* group and the open circles represent the *prize correlated group* and the *picture and prize uncorrelated* group. Note that the groups with correlated pictures performed more accurately than those with uncorrelated pictures. In the prize condition graph on the right, the filled circles represent the *picture and prize correlated* group and the *prize correlated* group and the open circles represent the *picture correlated group* and the *picture and prize uncorrelated* group. Note that the groups with correlated prizes performed with the same accuracy as those with uncorrelated prizes.

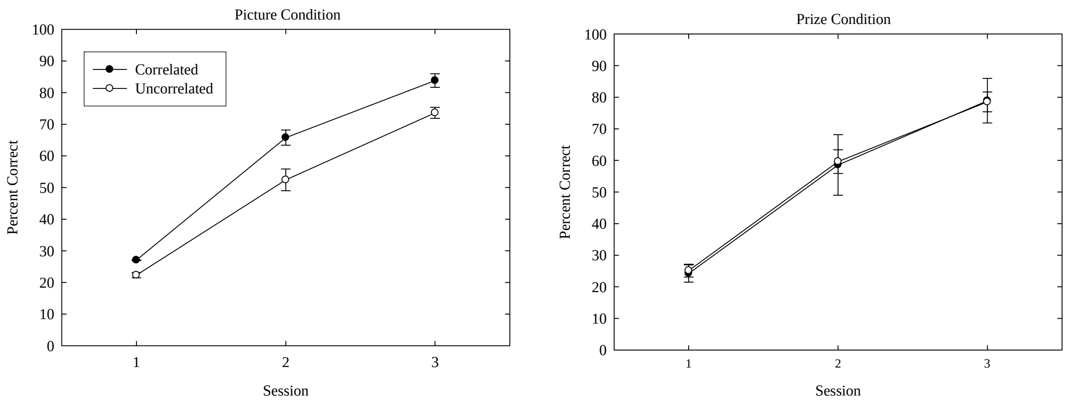

Figure 19. Experiment 1: The left panel shows the acquisition for the picture condition and the right panel shows the acquisition for the prize condition over all three sessions. The bars represent the standard error of the mean.

To compare the participants' acquisitions two separate ANOVAs were conducted to determine whether the different outcomes correlation conditions (picture condition and prize condition) affected group acquisition. A two-way repeated measures ANOVA with Picture Condition (correlated, uncorrelated) X Session (1, 2, 3), found a significant difference for Picture Condition, $F(1, 58) = 4.20$, $p < .05$, and the interaction was not significant $F(2, 116) = 1.34$, $p > .05$. A two-way repeated measures ANOVA with Prize

Condition (correlated and uncorrelated) X Session (1, 2, 3), found that neither the Prize Condition, $F(1, 58) < 1.00$, $p > .05$ nor the interaction $F(2, 116) = .07$, $p > .05$ were significant.

Recall

The groups' performances on the Kanji Recall Worksheet were first descriptively compared. Strong mean differences existed between groups, with the *picture and prize correlated* group performing most accurately ($M = 92.00$, $SEM = 3.63$) followed by the *picture correlated* group ($M = 84.44$, $SEM = 6.26$), then *prize correlated* group ($M = 76.00$, $SEM = 6.76$), and lastly the *picture and prize uncorrelated* group ($M = 76.00$, $SEM = 6.50$). A one-way ANOVA with Picture Condition (correlated and uncorrelated) yielded a significant effect, $F(1, 58) = 4.35$, $p < .05$. In comparison, a one-way ANOVA with Prize Condition (correlated and uncorrelated) yield no significant differences, $F(1, 58) = .39$, $p > .05$.

Stimulus Similarity

One issue concerning the stimuli is that some kanji might be perceptually easier or harder to differentiate then others. In fact, it might be that two kanji are so similar that they create consistent errors for participants labeling them. For example, Drama and Length may look so similar that when the kanji drama is mislabeled on the worksheet the erroneous word may not be randomly selected from the 14 other available words but instead always be Length. Considering this possibility, the accuracy for each kanji's label on the Kanji Recall Worksheet was examined (Table 1). Wealth, Abbreviation, Drama and Length were the easiest kanji to learn with only 5 of the 60 participants labeling them incorrectly. Technique and Reason were the most difficult kanji to recall with 20 people labeling Technique incorrectly and 19 labeling Reason incorrectly. The incorrect labels for Technique included: Abbreviation (2), Benefit (1), Degree (1), Length (1), Lose (1), None (1), Politics (1), Price (1), Reason (4), Technology (1), Tolerance (1), Virtue (2), and Wealth (2). The incorrect labels for Reason included: Benefit (1), Drama (2), Loss (1), Love (1), Order (1), None (1), Phrase (1), Politics (4), Source (1), Technique (3) and Virtue (2). Clearly, neither Technique nor Reason was consistently mislabeled as another word. The same pattern was true for the remaining kanji, suggesting that none of the kanji were perceptually similar enough to create consistent confusion.

Table 2.
Accuracy on the Kanji Recall Worksheet

Kanji	*Picture and Prize Correlated*	*Picture Correlated*	*Prize Correlated*	*Picture and Prize Uncorrelated*	Total
Abbreviation	14	14	13	14	55
Drama	15	15	12	13	55
Wealth	15	13	13	14	55
Length	15	15	13	12	55
Benefit	15	13	13	13	54
Quality	13	12	13	12	50
Phrase	15	11	12	12	50
Loss	13	13	12	11	49
Politics	14	12	11	12	49
Price	15	13	11	10	49
Degree	13	11	11	11	46
Virtue	13	13	10	10	46
Source	11	13	12	10	46
Reason	13	10	11	7	41
Technique	12	13	6	9	40

Note. The number of participants in each group with the correct responses for each kanji on the Kanji Recall Worksheet.

Prize Ranking

Since participants rank ordered the prizes before they completed the acquisition task and recall test, it is possible for the incentive values of the prizes to affect what the participants learn during or recall after the task. Considering this possibility, for the groups with correlated prizes, the accuracy of performance with the kanji correlated with the top three ranked prizes (Reason, Benefit, and Virtue) were compared with the lowest ranked prizes (Length, Source, and Abbreviation) on the across acquisition and on the recall worksheet. Across acquisition the performance was analyzed in a two way repeated measures ANOVA with Incentive Value (high, low) and Block (1 – 9), revealing an effect for Block, $F(8, 464) = 68.60$, $p < .05$, but not for incentive value $F(1, 58) = 3.63$, $p > .05$, or the interaction $F(8, 464) = 68.60$, $p < .05$. The recall worksheet results for the words correlated with the high and low incentive prizes were compared in a dependent samples *t* test, revealing no difference $t(29) = 0$, $p > .05$. This means that not only did the prize correlated condition not enhance performance, but the values of the prizes did not affect performance.

Discussion

Why did the picture condition enhance accuracy while the prize condition did not? First, one possible reason could be that the prize outcomes are temporally delayed after correct choice responses. As previously discussed, the longer the duration between a response and a reinforcer are, the weaker the association between them tends to be. Also related, the prize information is displayed along with the corrective feedback presented after an incorrect choice response. This renders the temporal interval between an incorrect choice response and prize outcome shorter than the interval between a correct choice response and the prize outcome, and may interfere with the correlated prizes' ability to mediate a correct response.

A second possible reason why the correlated prize outcomes did not enhance learning could be that the prize information was less discriminable then the picture information. The prize information was presented as an additional piece of linguistic information along with a lengthier correct response message. Although the duration of the message was long enough for a participant to read entirely, and the prize information was in standout red ink while the rest of the text was in black ink, the verbose message may have been too cumbersome for the participants to digest. This duel presentation (prize information and correct message information) may interfere with the participant's ability to process the prize information, making the prize information less discriminable than the singularly presented visual picture.

A third possible reason that the correlated prize outcomes did not enhance learning could be that pictures and prizes were redundant sources of information. After making a correct choice the participant saw a picture that indicates they were correct and then a prize message which also tells them they were correct. While the participants may care if they are correct and pay attention to the first cue that tells them that they are correct, they may disregard the second cue that indicates the same information. If the participant fails to pay attention to the prize information and the prize is not encoded, then the prize cannot be latter recalled to use as an expectancy to enhance performance.

In summary, the results of Experiment 1 suggest several things. First, they suggest that expectancies, as produced by the differential outcomes procedure, can affect typical, adult human learning. However, Experiment 1 employed corrective feedback, which is not commonly used in the DOE literature. Corrective feedback may or may not be necessary to obtain the DOE with typical adult humans, and Experiment 2 addresses any effects due to corrective feedback by employing noncorrective feedback. If the DOE does not depend on corrective feedback, the DOE should emerge even when corrective feedback is omitted from the task. Second, the results suggest that an immediate outcome (the picture) produces the DOE but a second delayed outcome (the prize) does not. Experiment 3 examines one possible reason that second outcome failed to enhance accuracy by reversing the order in which outcomes occur so that the prizes come first and the pictures are the second outcome. If the temporal order of the outcome matters, than there should be a prize condition effect but not a picture condition effect. However, if the discriminability of the outcomes matter, then the prize condition effect may not emerge.

EXPERIMENT 2

To see if the DOE depends on the inclusion of corrective feedback, in Experiment 2, corrective feedback was omitted from the task. It was expected that the participants would perform less accurately, and such results would not be surprising. The more important question is whether the DOE with typical adult humans depends on the inclusion of corrective feedback. If so, then removing the corrective feedback should cause the DOE to disappear. However, if the DOE does not depend on corrective feedback then the same pattern of results obtained in Experiment 1 should emerge (i.e., the groups with the picture correlated condition should perform better than the groups with the picture uncorrelated condition).

Methods

Participants.

The participants (n = 60) were different students but in all other ways same as Experiment 1.

Apparatus / Stimuli.

The apparatus and stimuli were same as those used in Experiment 1.

Procedure.

The procedure was the same as that used in Experiment 1, with exception of the outcome used for incorrect responses. In Experiment 1, after an incorrect choice participants saw a message that read, “No that’s wrong! The correct answer is ____. You have missed out on the prize entry in the draw for _______.” In Experiment 2, after an incorrect choice the participants saw a message that only read “No that’s wrong!.” This change not only omitted the corrective nature of the feedback, but also prevented prize information from being displayed after an incorrect choice.

Groups.

Same as Experiment 1.

Results and Discussion

Acquisition

Since Experiment 2 did not include any corrective feedback, the participants ability to employ the elimination strategy would be mitigated compared to the participants in Experiment 1. Thus one would expect the participants in Experiment 2 to perform at chance levels in the first block. Indeed, in Experiment 2 the participants performed at chance for the first 15, trials as confirmed by a one sample *t* test, $t(59) = 1.67$, $p > .05$. Next, each group was examined separately. A significant difference did not emerged for the *picture and prize correlated* group ($M = 15.11$, $SEM = 2.39$), $t(14) = 1.68$, $p > .05$, the *picture correlated* group ($M = 12.00$, $SEM = 2.18$), $t(14) < 1.00$,

$p > .05$, the *prize correlated* group ($M = 14.67$, $SEM = 1.86$), $t(14) = 1.91$, $p > .05$, or the *picture and prize uncorrelated* group ($M = 9.78$, $SEM = 1.94$) $t(14) < 1.00$, $p > .05$. This indicates that regardless of the participant's group, they were unable to perform better than chance accuracy in the first block.

Next, to examine the role of feedback over all three sessions, Experiment 1 and Experiment 2 were compared. The results were analyzed in a three-way repeated measures ANOVA with Feedback Condition (corrective and noncorrective), Group (picture and prize correlated, picture correlated, prize correlated, and picture and prize uncorrelated) and Session (1, 2, 3). There was a significant effect of Session, $F(2, 224) = 411.20$, $p < .05$, indicating all groups acquired the task, Feedback Condition $F(1, 112) = 98.59$, $p < .05$ with those groups receiving corrective feedback performing more accurately than those with noncorrective feedback, and Group, $F(3, 112) = 3.75$, $p < .05$ with those groups having correlated pictures performing more accurately than those without correlated pictures, as parsed out in later analyses. There was also a significant interaction between Session and Feedback Condition $F(2, 224) = 53.48$, $p < .05$, due to the difference between the groups' rates of acquisition, where the groups with corrective feedback aquired the task faster than the groups with noncorrective feedback. There was no interaction between Session and Group, $F(6, 224) = 1.57$, $p > .05$, or Feedback Condition and Group, $F(6, 224) = 1.08$, $p > .05$. The three way interaction was also nonsignificant, $F(6, 224) = 1.08$, $p > .05$ (Figure 20).

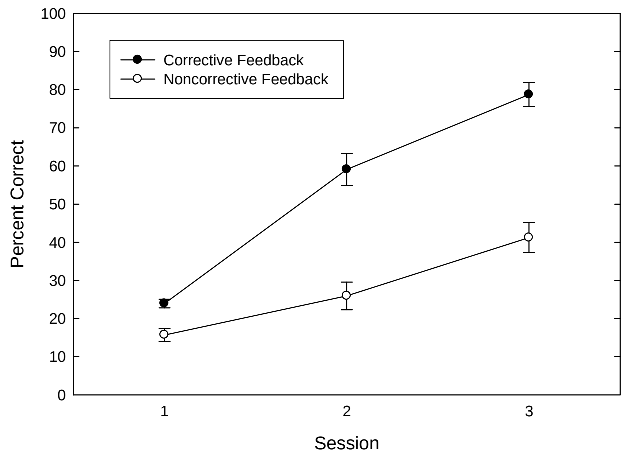

Figure 20. Experiment 2: The acquisitions for participants with and without corrective feedback (Experiment 1 vs. Experiment 2) over all three sessions. The bars represent the standard error of the mean.

To compare how the participants' acquisitions in Experiment 2 were affected by the correlated condition of the outcomes, the groups' acquisitions across all three sessions were again analyzed using two separate, two-way repeated measures ANOVAs. In the first two repeated measures ANOVA, picture condition (correlated and uncorrelated) and Session (1, 2, 3) acted as variables. The results yielded a significant Picture Condition effect, $F(1, 58) = 7.97$, $p < .05$, and Session effect, $F(2, 116) = 118.56$, $p <.05$, but no interaction, $F(2, 116) = 2.40$, $p > .05$ (Figure 21 left). The effect of Prize Condition was analyzed next in a two-way repeated measures ANOVA with Prize Condition (correlated and uncorrelated) and Session (1, 2, 3), yielding a significant Session effect, $F(2, 116) =$

115.07, $p < .05$ but neither a significant Prize Condition effect, $F(1, 58) < 1.00$, $p > .05$ nor a significant interaction, $F(2, 116) < 1.00$, $p > .05$ (Figure 21 right).

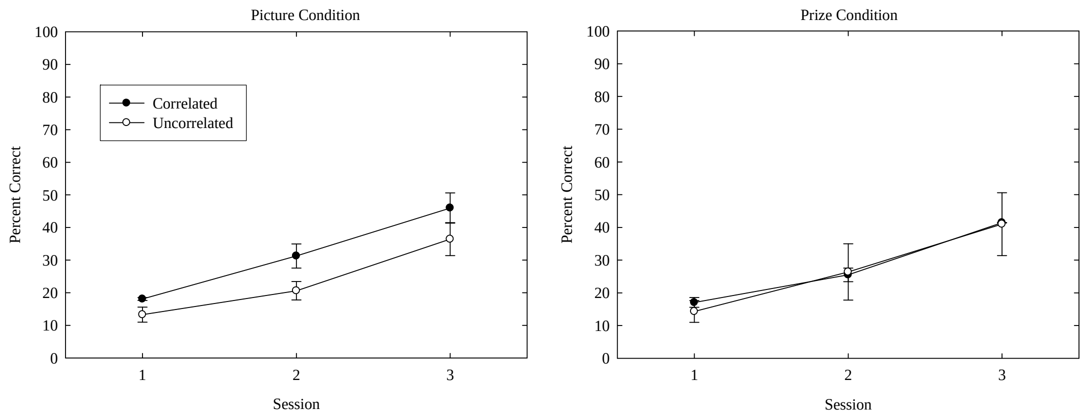

Figure 21. Experiment 2: The left panel shows the acquisition for the picture condition and the right panel shows the acquisition for the prize condition over all three sessions. The bars represent the standard error of the mean.

Recall

The percent correct for each group on the Kanji Recall Worksheet were first descriptively compared. The *picture correlated group* performing most accurately (*M* = 60.88, *SEM* = 7.67) followed by the *prize correlated* group (*M* = 43.70, *SEM* = 4.62), then *picture and prize correlated group* (*M* = 41.48, *SEM* = 5.97), and lastly the *picture and prize uncorrelated group* (*M* = 32.90, *SEM* = 4.58). Run in a one-way ANOVA with Picture Condition a significant effect emerged, $F(1, 58) = 4.46$, $p < .05$. In comparison, a one-way ANOVA for Prize Condition yield no significant difference, $F(1, 58) < 1.00$, $p > .05$. This result replicates that found for Experiment 1.

In summary, the results of Experiment 2 not only emphasize how important corrective feedback is to learning, but importantly also suggest that the DOE does not depend on the presence of corrective feedback.

EXPERIMENT 3

To see if the DOE depends on the proximity of the expected outcome, in Experiment 3 the order of the outcomes were reversed. In so doing the prize became the immediate outcome the picture will became the delayed outcome. However, the prize outcomes were still written inside a longer message making them less discriminable than the picture outcomes. If the immediacy of the outcomes matter, then a prize condition effect should be obtained and the picture condition effect should not. However, if the discriminability of the outcomes affects performance, then the prize condition effect may be less significant and/or the picture condition effect may still occur. But for the present purposes, in Experiment 3 the expected result was a significant prize condition but not picture condition effect.

Methods

Participants.

The participants (n = 60) were different students but in all other ways same as Experiment 1.

Apparatus / Stimuli.

The apparatus and stimuli were same as those used in Experiment 1.

Procedure.

The procedure was the same as that which was used in Experiment 1 with the exception that the picture and prize outcomes occurred in the reverse order.

Groups. Same as Experiment 1.

Results and Discussion

Acquisition

First, the participants' performance in the first block (15 trials) was compared to chance accuracy. As in Experiment 1, one can expect the participants to perform better than chance because they receive corrective feedback and can thus effectively eliminate incorrect choice responses on consecutive trials. All the groups preformed better than chance in the first block, as indicated by a one sample *t* test, $t(59) = 4.56$, $p < .05$. Next each group was examined separately. A significant difference emerged for the following groups individually: the *picture correlated* group ($M = 16.89$, $SEM = 2.20$), $t(14) = 2.58$,

$p < .05$, and the *prize correlated* group ($M = 18.67$, $SEM = 2.92$), $t(14) = 2.58$, $p < .05$, and marginally for the *picture and prize correlated* group ($M = 15.11$, $SEM = 2.00$), $t(14) = 2.00$, $p > .06$ or the *picture and prize uncorrelated* group ($M = 15.11$, $SEM = 2.20$), $t(14) = 1.81$, $p = .09$. This result replicates that found for Experiment 1.

To compare how the participants' acquisitions in Experiment 3 were affected by the correlated condition of the outcomes, the groups' acquisitions across all three sessions were again analyzed using two separate, two-way repeated measures ANOVA. In a two-way repeated measures ANOVA with Picture Condition (correlated, uncorrelated) X Session (1, 2, 3) the results yielded a significant Session effect, $F(2, 116) = 527.72$, $p < .05$. But unlike the previous experiment, the effect of Picture Condition was not significant, $F(1, 58) < 1.00$, $p > .05$. The interaction was also not significant, $F(2, 116) = 1.17$, $p > .05$ (Figure 22 left). The effect of Prize Condition was analyzed next in a two-way repeated measures ANOVA with Prize Condition (correlated, uncorrelated) and Session (1, 2, 3). The results yielded a significant Session effect, $F(2, 116) = 540.37$, $p < .05$. But unlike the previous experiments, there was a significant effect of Prize Condition, $F(1, 58) = 4.09$, $p < .05$. The interaction was also not significant, $F(2, 116) = 2.59$, $p > .05$ (Figure 22 right).

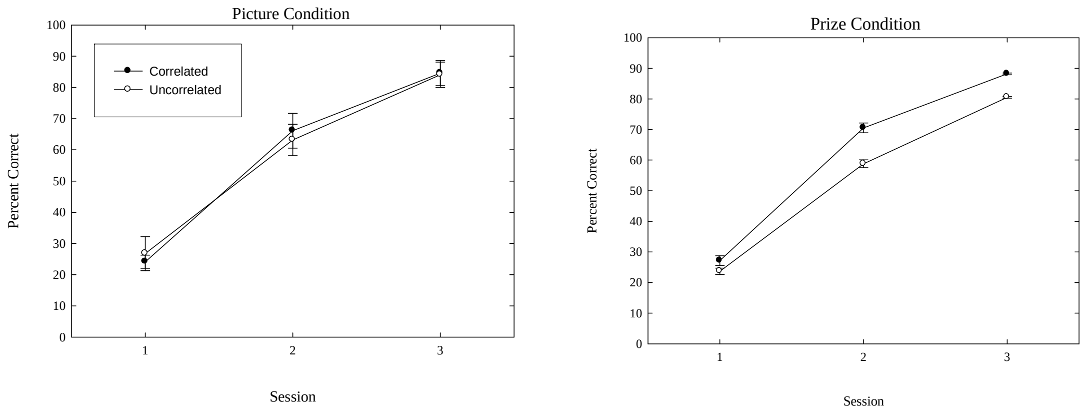

Figure 22. Experiment 3: The left panel shows the acquisition for the picture condition and the right panel shows the acquisition for the prize condition over all three sessions. The bars represent the standard error of the mean.

Recall

The percent correct for each group on the Kanji Recall Worksheet were first descriptively compared. The *prize correlated* group performed most accurately (M = 92.00, SEM = 2.36) followed by the *picture and prize correlated* group (M = 86.67, SEM = 5.40), then *picture correlated* group (M = 84.89, SEM = 5.34), and lastly the *picture and prize uncorrelated* group (M = 82.22, SEM = 5.47). The results run in a one-way ANOVA with picture condition no significant effect emerged, $F(1, 58) < 1.00$, $p > .05$. Likewise, a one-way ANOVA with prize condition yield no significant difference, $F(1, 58) = 1.46$, $p > .05$. The absence of a significant effect here may be due to the fact that the prize outcomes are less discriminable than the picture outcomes.

By reversing the order of the outcomes in Experiment 3 the effect of picture condition found in Experiment 1 disappeared, and the effect of prize condition became significant. These results suggest two things. First they suggest that the discriminability of the outcomes matters, as the more discriminable pictures immediate outcomes in Experiment 1 yielded a larger effect than the less discriminable prize immediate outcomes in Experiment 2. Secondly, they suggest that the immediacy of an outcome matters. The common significant factor in Experiment 1 and 2 is that the most temporally proximal outcome yields the DOE and the temporally distal outcome does not.

Experiment 4

Experiment 4 offers a new way to test for significant effects generated from using the differential outcomes procedure. Specifically after participants acquired the task with either correlated, uncorrelated, or the same outcome(s), then they completed a matching test which asked them to respond to the outcome pictures. An accurate response to the outcome picture was considered the same as the response to the discriminative stimuli that was correlated with that outcome during acquisition. Compared to participants that received uncorrelated or the same outcome(s), if participants who received correlated outcomes were able to more accurately respond to the outcome pictures on the recognition worksheet then this is an implicit learning effect generated by the differential outcomes procedure that has not previously been examined.

If implicit learning does occur it suggests that even when group differences in acquisition do not occur (the result that has commonly occurred in past research with typical adult humans) those participants who experience differential outcomes may still be affected by the differential outcomes procedure. Put another way, the methods used to examine the DOE with typical adult humans in the past may have been insufficient. The post acquisition test used here is proposed as a novel method to look at beneficial effects generated by the differential outcomes procedure.

Experiment 4 employs different stimuli, including 30 abstract paintings by 15 artists (two by each artists, one painting acted as a discriminative stimulus and the other as an outcome for correct responses), and the artists' names. After completing the computer task, participants were asked to complete the Artists Matching Worksheet, which asked participants to identify who painted each of the 15 paintings used as outcomes. If the participants trained with correlated outcomes learned the artists' names faster than those trained with uncorrelated outcomes, then the previous method (Experiment 1, 2 and 3) used to obtain the DOE with typical adult humans is generalizable to new stimuli. However, if the DOE does not emerge, that would be consistent with past experiments and the recognition worksheet will permit one to see another way the participants may have been effected by the differential outcomes procedure.

In Experiment 4, the participants were run in three groups: 1) *correlated paintings* group, 2) *uncorrelated paintings* group and 3) s*miley face* group. In the *correlated paintings* group, after correct responses participants always saw a painting by the same artist as that they had identified in their response (i.e., if they correctly responded Diller, then they saw a second painting by Diller). In the un*correlated paintings* group, after correct responses participants always saw a painting randomly selected from the 15 artists available (i.e., if they correctly responded Diller, then they saw a second painting by a randomly selected artist). In the s*miley face* group, after correct responses participants always saw a smiley face picture, which acted as a control for perceptual similarities between the discriminative stimulus painting and the outcome painting that might assist these participants accuracy on the Artist Matching Worksheet.

The Artists Matching Worksheet was included to explore how much participants learned about the outcomes during the task acquisition. Specifically, in a differential outcomes procedure participants are required to form associations between discriminative stimuli and responses to receive reinforcements (if A then B to receive C), but they are not required to learn associations between experimental events and the reinforcers. Expectation theory hypothesizes that an association between the discriminative stimulus and the reinforcer mediates a correct choice response (when A is presented C is mentally represented and mediates the choice of B). Participants are not required to demonstrate associations between discriminative stimuli and correlated outcomes (A and C) or a responses and a correlated outcomes (B and C) during task acquisition. But just because the performance of those associations is not explicitly required during acquisition, correct responses on the Artist Matching Worksheet would require those associations (implicit or incidental). If participants have learned those associations, then the outcomes may evoke the same responses as the discriminative stimulus had evoked (during acquisition participants learn if A then B to receive C; then the worksheet tests if participants know if C then B).

During task acquisition, the *correlated paintings* group could form specific associations between the artists' names and the painting outcomes. In contrast, the *uncorrelated paintings* group could not form specific associations to use to respond accurately on the worksheet and any associations they did form may act as interference that harms their accuracy. The s*miley face* group would be unable to rely on any specific associations, but they might be able to generalize from perceptual similarities between the discriminative stimulus painting and the outcome painting without interference. It was expected that the participants that received correlated picture outcomes would identify the artists of the outcome pictures more accurately then the other groups.

Methods

Participants.

The participants (n = 34) were different students but in all other ways same as Experiment 1.

Apparatus / Stimuli.

The apparatus and prizes were the same as those used in Experiment 1. The outcome pictures were art works obtained from World Wide Art Resources: Abstract Art (Obtained on January 18, 2007 http://wwar.com/masters/movements/abstract_art.html). Figure 23 displays each of the paintings used for each artist. The top painting was the one the participants were explicitly asked to learn the artist of, and the bottom painting was used as an outcome and on the Artist Matching Worksheet. The smiley face was a yellow circle with dot eyes and a curved mouth.

Figure 23. Experiment 4: The artwork and artists names.

Procedure.

The procedure was the same as that used in Experiment 1, with the exception of the stimuli. In Experiment 1, kanji were used as samples, English words were used as comparisons and pictures were used as outcomes. In Experiment 4, paintings were used as samples, artists names were used as comparisons and paintings by those same artists or a smiley face were used as outcomes.

Groups. Three groups were used in Experiment 4. These groups included: 1) the *correlated paintings* group (N = 10) 2) the *uncorrelated paintings* group (N = 11) and 3) the *smiley face* group (N = 13). The prizes were always uncorrelated. See Figure 24 for a temporal description of a typical trial with a correct response for the three groups.

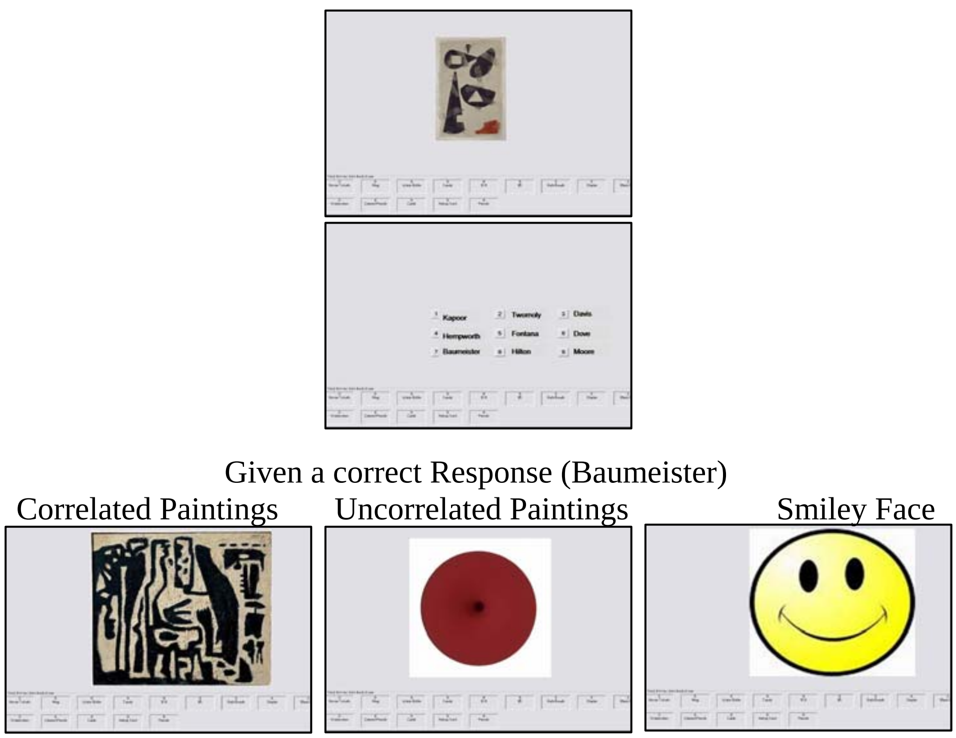

Figure 24. Experiment 4: The sequence of an example trial.

Matching. After completing all three sessions the participants were asked to complete the Artist Matching Worksheet. This worksheet displayed all 15 painting outcomes (three in each five rows) with a blank beneath each, see Figure 25. The 15 artists' names were written on the worksheet and the participants were instructed to write in the name of the artist they believed painted each painting below it.

Figure 25. Experiment 4: The Artist Matching Worksheet.

Results and Discussion

Acquisition

The results were analyzed in an two-way repeated measures ANOVA with Group (*correlated, uncorrelated* and *same*) and Session. Figure 27 shows group acquisition across sessions. All groups acquired the task, $F(2, 62) = 342.42$, $p < .05$. Unlike in Experiment 1, there were no significant differences between the groups, $F(2, 31) = .73$,

$p > .05$, and the interaction was also not significant, $F(4, 62) = 1.17$, $p < .05$. But the failure to obtain the DOE seems to have been caused by the ease of the task, a common finding in the differential outcomes literature with typical adult humans (note the near perfect performance in session 3). As confirmed by a independent samples *t* test, by the second session of acquisition the participants learning about art were performing more accurately ($M = 79.67$, $SEM = 2.99$) than participants learning about kanji in Experiment 1 ($M = 58.35$, $SEM = 3.09$), $t(92) = 4.55$, $p < .05$. Thus, learning who painted each piece of abstract art is easier then learning the English meaning of each kanji.

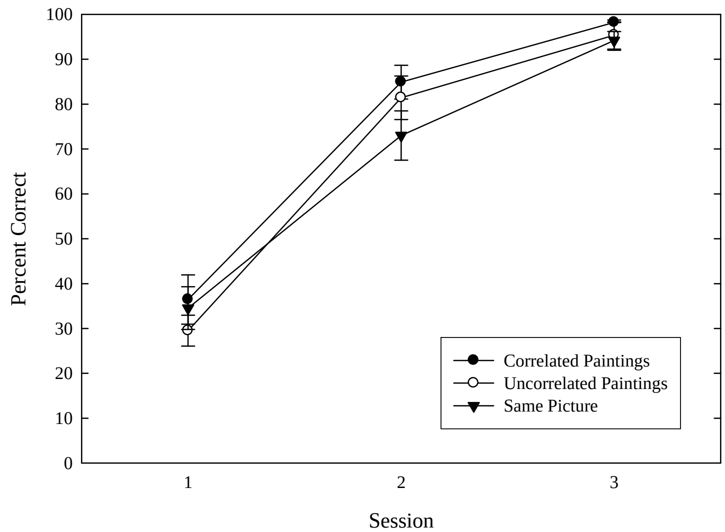

Figure 26. Experiment 4: The percent correct for each group in Experiment 4 across sessions. Bars represent the standard error of the mean.

Matching

Because a participant's accuracy on the Artist Matching Worksheet may depend on his/her exposure to the outcome paintings, an analysis of the worksheet data should exclude those participants who never were exposed to the outcome paintings. During the computer task participants are only exposed to the outcome paintings after a correct response. Thus to measure participants’ exposure to the outcomes paintings one can examine each participants accuracy during the computer task. Fortunately, most participants reached accuracy above 93.33% by the final session of the computer task (42 correct / 45 possible), suggesting that they had been exposed to all the outcome paintings. However, exceptions to this accuracy include one participant in the *paintings uncorrelated group* (62.22%) and three in the s*miley face group* (86.66%, 84.44% and 77.78%). These four participants’ data were removed from the following analyses.

The percent correct for each group on the Artist Matching Worksheet were analyzed using a one-way ANOVA for Group (correlated paintings, uncorrelated paintings, smiley face), revealing a significant difference between the groups, $F(2, 26) = 29.08$, $p < .05$. The *correlated paintings group* preformed the most accurately ($M = 74.00$, $SEM = 8.63$), followed by the *smiley face group* ($M = 23.70$, $SEM = 4.17$), and then the *uncorrelated paintings group* ($M = 16.00$, $SEM = 3.01$). A posthoc tukey test revealed that there was no difference between the *uncorrelated paintings* and *smiley face groups* (Figure 27).

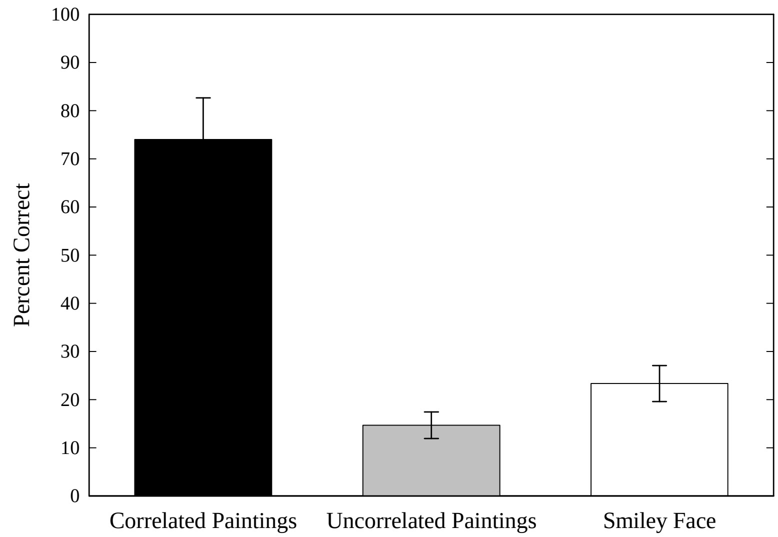

Figure 27. Experiment 4: The percent correct for each group on the Artist Matching Worksheet. The bars represent each group's standard error of the mean.

To further explore the data from the Artist Matching Worksheet, an item analysis was completed. This analysis permits one to see if the aggregated average pattern for groups remains consistent for individual items (Figure 28). Indeed the pattern does remain consistent. For every item the *correlated paintings* group outperformed the other groups. In fact for Nicholson, Diller, and Richter's paintings only participants in the *correlated paintings* group were able to accurately identify the artist. For other items, like Moore, the *smiley face* group's ability to generalize from the original paintings was statistically similar to the *correlated painting* group's performance, but the *uncorrelated paintings* group's performance was notable lower. In summary, the *correlated paintings* group always performed the most accurately followed by the *smiley face* group and the *uncorrelated paintings* group was typically the least accurate.

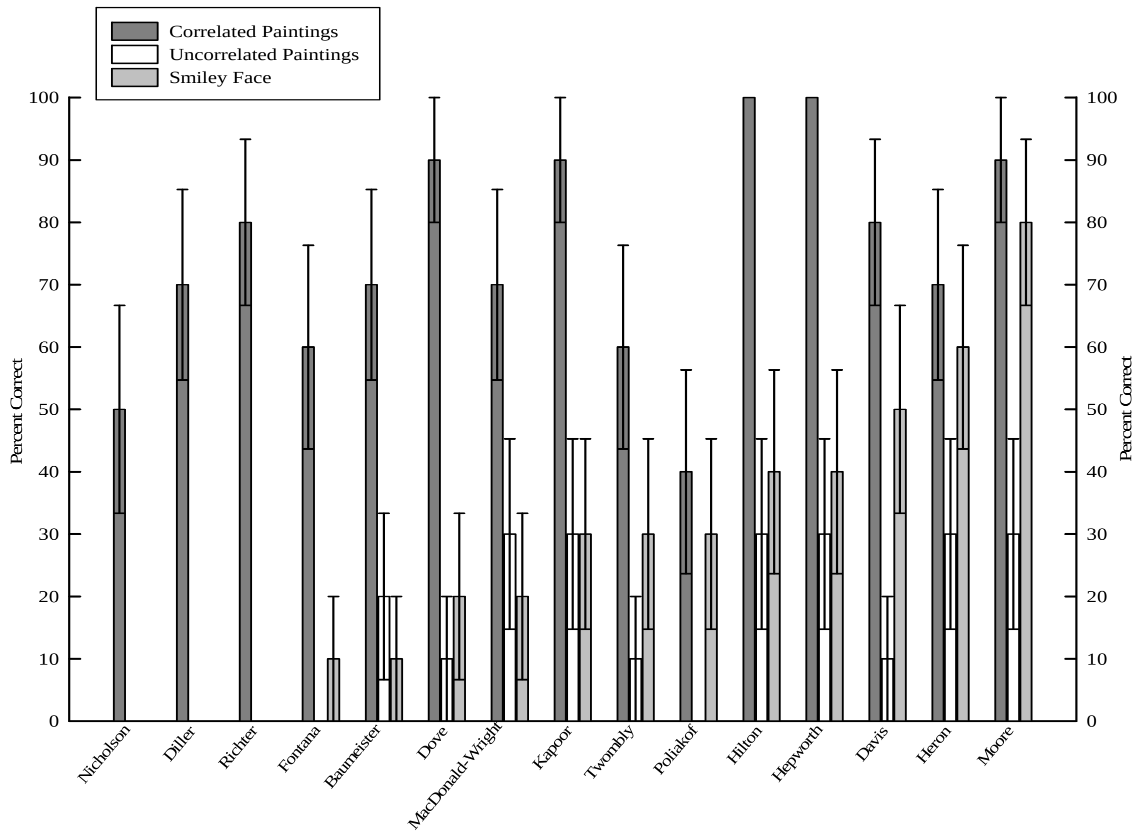

Figure 28. Experiment 4: The percent correct for each item by each group on the Artist Matching Worksheet. The bars represent each group’s standard error of the mean.

While looking at individual items, the concern about perceptual similarity between items raised in Experiment 1 was also examined. For example, in Experiment 4 Davis’ and Diller’s paintings may look so similar then when the Davis’ is mislabeled on the worksheet the erroneous artist selected may not be randomly selected from the 14 other available artists but instead always be mislabeled as Diller. The group most likely to show this affect was the *smiley face* group, because those participants lack experiences with the second pictures that create associations that aid or interfere with labeling and thus their responses on the worksheet are based purely on perceptual similarity. So, only the *smiley face* group's data is considered in the following analyses. Looking at just the accuracies for each item, Moore, Heron and Davis were the easiest items, and thus their two paintings must have been perceived as being very similar, the complete results are shown in Table 2. The most consistently mislabeled artist was Ricter with six of the ten participants' mislabeled responses being Kapoor, but the others saying Nicholson, MacDonald-Wright, Twombly and Davis. Moore, Heron and Ricter were the most often artist used to label a response incorrectly, but never in a consistent manner. So no paintings were so similar that every participant responded to the item in the same way.

Table 2.
Generalization on the Artist Matching Worksheet

Artist	Number of Participants with Correct Responses Smiley Face Group
Moore	8
Heron	6
Davis	5
Hilton	4
Hepworth	4
Poliakof	3
Twombly	3
Kapoor	3
MacDonald-Wright	2
Dove	2
Baumeister	1
Fontana	1
Richter	0
Diller	0
Nicholson	0

Note. The number of participants with the correct responses for each picture item on the Artist Matching Worksheet.

The results of Experiment 4 suggest that in order to obtain the DOE with humans the associations must be sufficiently difficult to learn. However, they also suggest that maybe past experiments have not tested for the DOE in the best manner. The results of the Artist Recognition Worksheet suggest that participants can respond to correlated outcomes in the same manner as they responded to the discriminative stimulus. This suggests that even when group differences do not occur in acquisition, participants may still be able to demonstrate a beneficial implicit learning effect of being trained with the differential outcomes procedure. Because of the implicit learning effect of the differential outcomes procedure, the participants in the *correlated paintings* group learned twice as much information as participants in the *uncorrelated paintings* and *smiley face groups*.

GENERAL DISCUSSION

This series of experiments confidently asserts that the DOE is obtainable with typical adult human participants. Experiment 1 replicated the DOE with typical adult human participants and showed that the correlated status of the pictures was more important than the prizes. Experiment 2 obtained the DOE without corrective feedback, showing that the presence of corrective feedback is not necessary to generate the DOE. Experiment 3 suggested that the placement of the outcomes matter, by reversing the order of the outcomes in Experiment 1 (from pictures than prizes to prizes than pictures), the correlated status of the prizes was more important than the pictures to enhance participants' performance. Experiment 4 failed to obtain the DOE with easily acquired stimuli, but lends support to expectancy theory by showing that associations between responses and outcomes are implicitly acquired.

The task changes necessary to generate the DOE in the present experiments are quantitative in nature, as they altered not the procedure but the difficulty level. These experiments also emphasize the effects of stimulus discriminability (less discriminable stimuli are harder to learn) and temporal placements (the first outcome is effective while the second outcome is ineffective).

While the present experiments show the DOE is obtainable with typical adult humans, the effect generated with kanji was not generalizably to abstract art. The inability to generate the DOE with abstract art would be a problem if the rate at which participants learned about kanji and abstract art were equivalent. However, the participants learned about the abstract art at a much faster rate than the kanji which demonstrated the ceiling effect. So, not finding the DOE with abstract art reemphasizes the difficulty generating the DOE with humans when the task is too simple.

Although not obtaining the DOE with abstract art is not problematic theoretically, one would like to obtain DOE with other stimuli. One suggestion is using the differential outcomes procedure to train doctors to detect cancerous mole. Like kanji for English speaking people, differentiating cancerous and noncancerous moles is difficult. If the differential outcomes procedure facilitates acquisition of cancerous mole detection, that would not only support the generality of the procedure in the DOE literature but also provide an effective method to teach people a beneficial health skill. Another possibility is to train people to identify bird calls. Although some calls are quite distinct, as the number of species increases it becomes increasingly difficult to differentiate them. If the differential outcome procedure facilitates acquisition of call identification then that would support the generality of the procedure in the DOE literature and could also be an advantageous teaching tool for birders or ethnology students.

Aside from expanding the types of stimuli used to generate the DOE, it would be beneficial to expand the populations tested from college students to other typical adult human populations. If the DOE is a general mechanism, then it should not be limited to nonhuman animals and college students. According to The National Center for Higher Education Management Systems (2009), 61.6% of high school graduates go directly to college in the United States, leaving 38.4% unexamined. Also, college age students tend to represent a limited age range, leaving 100% of the more aged populations unexamined. Because of this limitation in the current literature, it would be useful to test those who do not attend college and elderly people and see if they are also affected by the differential outcomes procedure.

Another limitation of the present study is its ability to discuss temporal delays within the task. In Experiment 1, the outcome order was pictures then prizes and in Experiment 2, the order was prizes then pictures. The outcome reversal therefore changes not only the temporal delays from a correct choice response to a particular outcome, but also the redundant nature of the outcome (i.e., in Experiment 1 prizes are redundant and ineffective and in Experiment. As discussed in Experiment 1's results, it could be that the first outcome is the only important outcome and the temporal delay to that first outcome may or may not matter. To better test to see how the temporal delay to the outcome affects the differential outcomes procedure, one could insert different temporal delays (e.g., a blank screen for 1 ms or 5000 ms) between a correct response and a single outcome. If the temporal delay to the outcomes matter, then one could expect that longer delays would hinder the DOE's emergence.

This series of experiments confidently asserts that the DOE is obtainable with typical adult human participants (Experiment 1, 2 & 3). Thus, with regard to the differential outcomes procedure, claims that humans are qualitatively different than other animals are unwarranted. Instead of qualitative differences, the differences between humans and other animals appear quantitative. The task changes necessary to generate the DOE in the present experiments are quantitative in nature, as they altered not the procedure but the difficulty level. These experiments also emphasize the effects of stimulus discriminability (less discriminable stimuli are harder to learn) and temporal placements (the first outcome is effective while the second outcome is ineffective). Furthermore, beyond showing that expectancies do affect typical adult human learning, Experiment 4's results also assert that humans learn associations between responses and outcomes. Furthermore, Experiment 4 offers a new way for experimenters to test for effects the differential outcomes procedure generates.

To close, the present experiments add to the comparative psychology and DOE literature. With regard to species similarities and differences, these experiments indicate that expectancies are a mechanism that affects both nonhuman and human animals learning. Differences previously obtained were due to quantitative not qualitative factors reminding one that while common mechanisms may underlie cognitive abilities the procedures used to generate particular effects may be species specific. More often than not, experimenters may need to adjust procedures (in degree, not kind) to obtain common effects.

REFERENCES

Alling, K., Nickel, M. & Poling, A. (1991). 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.

Brodigan, D. L. & Peterson, G. B. (1976). Two-choice conditional discrimination performance of pigeons as a function of reward expectancy, prechoice delay, and domesticity. *Animal Learning & Behavior, 4,* 121-124.

Carlson, J. G. & Wielkiewicz, R. M. (1976). Mediators of the effects of magnitude of reinforcement. *Learning and Motivation, 7,* 184-196.

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.

Estévez, A. F., Fuentes, L., Marý-Beffa, P., González, C. & Alvarez, D. (2001). The differential outcome effect as a useful tool to improve conditional discrimination leaning in children. *Learning and Motivation, 32,* 48-64.

Estévez, A. F., Fuentes, L. J. (2003). Differential outcomes effect in four-year-old children. *Psicológica, 24,* 159–16.

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. (2005). The differential outcome effect: A useful tool to improve discriminative learning in humans, *The Behavior Analyst Today, 6*, 216 – 220.

Estévez, A. F., Vivas, A. B., Alonso, D., Marý-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.

Holt, G. L., & Shafer, J. M. (1973). Function of the intertrial interval in matching to sample. *Journal of the Experimental Analysis of Behavior, 19*, 181–186.

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.

King, D. L., Jones, F. L., Pearlman, R. C., Tishman, A., & Felix, C. A. (2002). The length of the retention interval, forgetting, and subjective similarity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*, 660-671.

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.

Maddox, W. T., Ashby, F.G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29,* 650-662.

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.

Malanga, P., & Poling. A. (1992). Letter recognition by adults with mental handicaps: Improving performance through differential outcomes. *Developmental Disabilities Bulletin, 20,* 39–48.

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.

Premack, D. (1983). The codes of man and beasts. *Behavioral and Brain Sciences, 6,* 125–167.

Sargisson, R. J., & White, K. G. (2007). Remembering as discrimination in delayed matching to sample: discriminability and bias. *Learning and Behavoir, 35*, 177-183.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*, 1317-1323.

Sobel, K. V., & Cave, K. R. (2002). Roles of Salience and Strategy in Conjunction Search.

Journal of Experimental Psychology: Human Perception and Performance, 28, 1055–1070.

The National Center for Higher Education Management Systems. (2009). College-Going Rates of High School Graduates - Directly from High School**.** Retrieved April 4, 2009, from: http://www.higheredinfo.org/dbrowser/?year=2006&level=nation&mode=data&state=0&submeasure=63

Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different? *Learning and Motivation, 1*, 129-140.

Urcuioli, P. J. & Lionello-DeNolf, K. M. (2001). Some tests of the anticipatory mediated generalization model of acquired sample equivalence in pigeons' many-to- one matching. *Animal Learning & Behavior, 29*, 265–280.

Urcuioli, P. J. (1991). Retardation and facilitation of matching acquisition by differential outcomes. *Animal Learning & Behavior*, *19*, 29-36.

Urcuioli, P. J. (1990). Some relationships between outcome expectancies and sample stimuli in pigeons' delayed matching. *Animal Learning & Behavior*, *18*, 302-314.

Urcuioli, P. J., & Zentall, T. R. (1986). Retrospective coding in pigeons' delayed matching-to-sample. *Journal of Experimental Psychology: Animal Behavior Processes*, *12*, 69-77.

Williams, D. C., Johnston, M. D., & Saunders, K. J. (2006). Intertrial sources of stimulus control and delayed matching-to-sample performance in humans, *Journal of Experimental Analysis of Behavior*, *86*, 253 - 267.