ABSTRACT

GENERALIZATION WITHIN AN IMPLICIT CATEGORIZATION TASK

By Kristin N. Christy

Participants trained on one implicit categorization task performed another implicit categorization task with the same deep underlying structure but different stimulus dimensions. Their performance on the transfer task was compared to two other groups’ performance on the same transfer task. One of these groups was trained on an implicit categorization task with a different underlying structure and the other trained on an explicit categorization task. Evidence of generalization from one implicit categorization task to another with the same deep structure but differing stimulus dimensions was found.
Generalization within an implicit categorization task

A Thesis

Submitted to the

Faculty of Miami University

in partial fulfillment of

the requirements for the degree of

Master of Arts

Department of Psychology

by

Kristin N. Christy

Miami University

Oxford, OH

2003

Advisor

________________________________________
Robin D. Thomas, Ph.D.

Reader

________________________________________
Leonard S. Mark, Ph.D.

Reader

________________________________________
Philip J. Best, Ph.D.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>11</td>
</tr>
<tr>
<td>Participants</td>
<td>11</td>
</tr>
<tr>
<td>Stimuli and Apparatus</td>
<td>11</td>
</tr>
<tr>
<td>Procedure</td>
<td>12</td>
</tr>
<tr>
<td>Results</td>
<td>14</td>
</tr>
<tr>
<td>Conclusion</td>
<td>15</td>
</tr>
<tr>
<td>Discussion</td>
<td>15</td>
</tr>
<tr>
<td>References</td>
<td>19</td>
</tr>
</tbody>
</table>
## List of Illustrations

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Representation of stimuli for the learning task</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Learning task distributions</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>Representation of stimuli for the transfer task</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>Transfer task distribution</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Between- and within-groups comparison</td>
<td>15</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Parameter values</td>
</tr>
</tbody>
</table>
Acknowledgments

First of all, I would like to thank my advisor, Dr. Robin Thomas, for being a supportive mentor and teacher. Next, I would like to thank my committee members, Dr. Phil Best and Dr. Len Mark for their advice and support on this thesis. To my friends back home, thank you for listening, even when my problems seemed so foreign to you. To my friends in Oxford, thank you for drying my tears and listening to me vent, even when we had only just met. To my parents and brothers, thank you for the constant support and love. Finally, it is with much love and adoration, that I dedicate this thesis to my husband, Kevin, for holding an unshakable belief in me and being my best friend through it all.
Categorization is the process that allows us to recognize objects or classify novel objects. This process is essential to our everyday life. Without it we would not be able to drive, to know what to eat, or to recognize anybody or anything. The ability to categorize objects is not only necessary to the functions humans perform in everyday life, but to the survival of all living organisms. In order for a species to survive it must be able to know whether another animal is foe or friend, whether certain food is poisonous or not, and what other organisms it can mate with in order to perpetuate its own species.

Given the importance of this subject, there has been a great deal of research in this area over the years. An ongoing debate in the categorization literature has been whether there exists single or multiple systems of categorization. Early studies on categorization focused primarily on single system models.

**Multiple Versus Single System Models**

*Single System Models*

Three broad theories of categorization have provided most of the groundwork for research in this area over the last 60 years. These are rule-based, prototype, and exemplar theory. The rule-based theory assumes that each category is defined by a certain set of necessary characteristics, or attributes, of the object. These attributes form logical rules that create an algorithm defining that category. For instance, if the object is a mammal plus it has wings plus it has a beak, then it is a bird. While originally this theory, known then as the ‘classical view’, called for very rigid rules, it is now believed that these rules can have exceptions (Nosofsky, Palmeris, and McKinley, 1994).

Both the prototype and exemplar models make up the ‘probabilistic view’. According to this view, categories of concepts are formed by characteristics that are typical of members of that category rather than defining. The category boundaries are fuzzy, allowing for features of objects to overlap with other categories. In the prototype model all new objects are compared to the prototypes of existing categories. The prototype is a perceptual representation of an object made up of all the examples of objects in the category; it is a representation of the ‘average’ or ideal object. New objects are categorized as members of the category whose prototype is most similar. Prototype theory is supported by research findings that prototypes are recognized more easily on transfer tasks than are examples that were viewed in the previous task, even if the prototype was not seen during training (Posner & Kiely, 1968). Because of other shortcomings,
however, this theory has not received much rapport in recent literature (Busemeyer, Dewey, and Medin, 1984; Shin & Nosofsky, 1992). The prototype model doesn’t take into account the variability of examples, the size of the category, correlational information of attributes, or the context with which the example is presented.

In 1978, Medin and Schaffer proposed the context model, in which every example of a category is represented in the mental concept of that category. New objects are compared to all existing examples of each category and placed in the category with the greatest overall similarity. Those models allow for examples of categories that do not possess a significant number of the average features. Exemplar theories are also able to account for studies where novel prototypes were categorized more quickly than previously seen exemplars since the prototypes would be very similar to all of the exemplars in its category. One of the potential problems with exemplar theory is that it calls for an extensive amount of memory for every category.

The generalized context model (GCM) (Nosofsky, 1986) is an extension of Medin and Schaffer’s (1978) context model, the primary difference being the way in which similarity is computed. Here the distance between the values for each dimension is multiplied by a hypothetical attention weight. By taking into account the salience of each dimension on a particular task, GCM is taking into account the affect of the type of task and of the context created by that task. A connectionist version of the model, called ALCOVE (attention-learning covering map) exists (Kruschke, 1992).

After rule-based, prototype, and exemplar theory had been established, more complex models were created that built upon these ideas (Nosofsky, Palmeri, and McKinley, 1994; Anderson et al., 1979; Elio & Anderson, 1981; Gluck & Bower, 1988; and Ashby & Townsend, 1986). One of the most prominent of these is general recognition theory (GRT) or decision bound theory (Maddox & Ashby, 1992 or 1993). GRT states that each observation of an exemplar can be represented as a point in a multidimensional perceptual space (Ashby & Townsend, 1986). Since repeated presentations of the same exemplar do not always lead to the same perceptual effect (Thurstone, 1927) the perceptual effects of each exemplar of a category are conveniently represented by a multivariate probability distribution. A category is thus represented as a probability mixture of the individual exemplar distributions (Ashby & Maddox, 1993). Here there arise two types of variation: within-exemplar and between-exemplar. Within-
exemplar variation is better known as perceptual noise and is determined by the extent to which
the psychological representation of an object varies on each presentation. Between-exemplar
variation is determined by the type of category structure and increases as the similarity between
pairs of category exemplars decreases. Using these variations, a category becomes a distribution
of exemplar points. A participant who is practiced on a particular categorical task partitions the
perceptual space into regions and associates a category label with each region. On each trial, the
learner determines in which region the exemplar representation falls and then emits the
associated response. The line or curve dividing the two response regions is called the decision
bound. Due to the generality of this model, many other models have been created which can be
mapped onto GRT in one way or another.

In attempting to provide evidence for a particular single-system model, researchers have
come up with results that supported one model and contradicted others. Since this had occurred
for all of the basic single-system models, it was thought that more complex multiple-system
models of categorization could better explain these results (Ashby et al., 1998).

Multiple System Models

COVIS

The COVIS (competition between verbal and implicit systems) model is based upon the
idea that there are two systems of categorization, verbal and implicit, that work in parallel
(Ashby et al., 1998). Here the implicit system is based upon GRT and is proposed to be
primarily supported by the basal ganglia. It is assumed that clusters of cells learn to respond to a
particular area in the distribution of one of the categories. The participant places the object in the
category that evokes the largest response from these clusters.

The verbal system is proposed to be primarily supported by the prefrontal cortex and
cingulate gyrus. When learning a category structure the verbal system first arbitrarily picks a
rule to categorize by. At this point all of the rules have a weight of zero. This rule is then
reinforced either positively or negatively depending on its success in categorizing the stimulus
presented. Through the combination of trial-by-trial feedback, the tendency of the person to
perseverate on each rule, and the anterior cingulate’s desire to pick a rule, the weight of each rule
is changed after each stimulus presentation. The response emitted by the verbal system, then, is
determined by the rule with the largest weight.
Each time a stimulus is presented both systems emit a deterministically decided response. Initially the verbal system learns faster than the implicit system since the implicit system has not yet learned to respond to the stimulus distributions. In the case where a rule strategy can be used that is almost as optimal as the implicit strategy, the bias towards verbal rules may cause the participant to rely on the verbal rule despite its inadequacy. Over time, the system that has the higher frequency of correct responses then becomes the dominant system in that particular categorization task.

ATRIUM

ATRIUM (Erickson & Kruschke, 1998) is also based upon the idea of having two parallel systems of categorization. In ATRIUM, though, one system is exemplar based and the other system is rule based. This model is an extension of ALCOVE, with ALCOVE representing the exemplar system. The final output of the ATRIUM model is a combination of the output from the verbal system and exemplar system. The degree to which each system contributes is determined by the output of its gate node.

Common to all multiple systems models is the distinction between an explicit, rule based system and a nonrule possibly implicit system. In an explicit categorization task, participants can verbally describe the rules that guided their decision as to category membership. Implicit categorization tasks, on the other hand, are those in which participants are unable to verbalize rules that delineate categories, but are still able to learn the rules. In implicit learning we are unaware of the decision processes used to make the categorical decision and instead just experience a ‘draw’ toward a particular category.

Generalization of Learning on Categorization Tasks

Explicit Category Learning

Research has shown that explicit category learning is generalizable between sets of stimuli within a rule (Haygood & Bourne, 1965; Bourne, 1970; Erickson & Kruschke, 1998). In 1970, Bourne conducted a study looking at different levels of rule complexity in stimuli and transfer of training for these rules. In his first experiment he had participants do a series of nine problems. In each problem the participant was told the relevant dimensions and had to categorize the object. The participant did this until they performed correctly on sixteen consecutive trials. The rule was the same for all nine problems for each participant, but the relevant dimensions varied for each problem. There were four groups of subjects with each
group being assigned to a particular rule, conjunctive, disjunctive, conditional, or biconditional. For each problem, the participants were significantly faster at learning the rule and all participants had learned the rule by the sixth problem.

On his second experiment, Bourne (1970) had participants first learn four different types of rules, conjunctive, disjunctive, conditional, and biconditional, by performing three consecutive categorization problems per rule. Then, the participants had to do more categorization problems with the same four rules but with varying relevant attributes until they had learned the rules. Significant interrule and interproblem transfer was obtained for all four of the rules.

In 1998, Erickson and Kruschke had participants perform a categorization task with rectangular stimuli that varied in the height of the rectangle and the horizontal position of an internal line segment. Most of the stimuli belonged to two categories that could be distinguished by a simple rule on one of the dimensions. The remaining two training stimuli were each members of their own categories and could be distinguished from the members of the two rule categories by integrating the values of both dimensions. Once training was complete, Erickson and Kruschke (1998) tested participants’ generalization by presenting them with stimuli not seen during training. During this phase of the experiment, no feedback was provided. In particular, Erickson and Kruschke were interested in how participants generalized what they had learned during training to stimuli that were similar to the exceptions. They found that even for novel stimuli that were more similar to the exception training instance than to the rule training instances, participants tended to make their classifications according to the rule. In other words, participants were able to generalize their rule-based training.

Implicit Category Learning

While there has been evidence of generalizability of explicit categories, there has been very little evidence of generalizability for implicit categories. Only one type of study has provided evidence of this, studies conducted on artificial grammar learning (e.g. Tunney & Altmann, 2001). Sequences of nonsense syllables were created using only letters from the first half of the alphabet. These sequences followed the rules of an artificial grammar created by the researcher. Participants were trained on sequences of these syllables by filling in sequences while being able to view other sequences. Once they became fairly accurate in filling in the sequences they were told that these sequences had been constructed using simple rules. The
participants were then shown novel sequences and asked to point out the sequences that followed the same rules as the training sequences. The participants were able to successfully categorize the novel stimuli in the appropriate category, indicating that they had learned the category. Since the participants were unaware of how they were able to classify the stimuli and since the letter strings used during training and test were different, it is safe to assume that transfer of training was due to the identical distributional properties of the grammar sets. In other words, the subjects implicitly learned the category structure.

Summary

From the discussion, the following should be evident to the reader:

1. The single versus multiple systems categorization debate has fueled much research.

2. Explicit rule-based task transfer of training has been demonstrated, supporting the idea that a single-system exemplar model (or a model that builds rules from exemplars) could be used.

3. Implicit task transfer of training has been demonstrated only in artificial grammar learning and not in a task requiring the integration of separable perceptual dimensions.

Therefore, transfer of training in an implicit task could add evidence for multiple system models because such a transfer could not be made using the memory of specific exemplars, something that is fundamental to all exemplar models.

The Current Study

Many experiments that have been conducted in order to investigate formal models of categorization use categories that consist of only a few exemplars. This creates a problem since most real-world categories are infinite in size and have continuous dimensions. Ashby and Gott (1988) created a categorization training technique in order to deal this. Here the researcher creates categories so that the values along each dimension are continuous and probabilistically distributed. On each trial an exemplar is sampled from this distribution and presented to the participant. This results in large, overlapping categories with multiple exemplars.

This technique was used in the current study in order to accurately represent categories by defining each of the categories as bivariate, normal distributions on continuously valued stimulus dimensions. In the learning task circle-line stimuli were used (Figure 1). The varying
dimensions were the size of the circle and the angle of the line. Participants were divided into three groups, with each group defined according to the type of category structure presented in the learning task. The groups were as follows: linear implicit, rules explicit, and quadratic implicit. The category structures for all three groups on the learning task are displayed in Figure 2. Each point in space represents the two relevant dimensions of the presented stimulus. The distributions are divided by their optimal bound. The optimal bound is defined as the equation that, if used, would result in the highest possible percent correct.

Figure. 1

In the linear implicit-trained group’s learning trials, both categories were correlated $+ .80$. In the example shown in Figure 2a, the best-fitting linear discriminate function ($y = x$) allows for 100% accuracy of classification, because the distributions only overlap by about 2%. For the quadratic implicit-trained group, the best-fitting quadratic discriminate function allows for 90% accuracy of classification due to more overlap of the distributions (Figure 2c). Since the stimuli in question have attributes that are not only separable but also non-commensurate, the task of learning to classify them is non-trivial. In fact, the task falls into the domain of implicit learning, because even though participants may categorize the stimuli with high accuracy, they cannot explain how they are exercising this expert knowledge. The decision rule is essentially non-verbalizable, and can be learned only by abandoning verbal rules and intuitively integrating the
For example, in Figure 2a the decision rule would sound something like this: ‘If the line orientation is greater than the size of the circle, the object is an A; otherwise, it’s a B.’ Because of the non-commensurate nature of the stimuli, this statement makes no logical sense.

Figure 2.

Stimulus distributions of the learning tasks for the (a) linear implicit-trained group (b) rule explicit-trained group and (c) quadratic implicit-trained group.

Following the nature of implicit learning, it is assumed that through trial-by-trial feedback the participants will learn to divide the stimuli into two predetermined categories (Figure 2a and c). In general, the optimal bound for two bivariate normal distributions is described by:
In the general case of a linear implicit condition, this bound becomes
\[ h(x, y) = b_1 x + b_2 y + c = 0. \]

In the general case of a quadratic implicit condition, the optimal bound becomes
\[ h(x, y) = a_1 x^2 + a_2 y^2 + a_3 x y + b_1 x + b_2 y + c = 0 \]
where \( a, b, \) and \( c \) are defined by the distribution parameters. Thus, the orientation of the optimal decision bound forced participants to integrate separable, non-commensurate dimensions in order to do well at the task.

In the rule explicit-trained group’s learning task, the same type and range of stimuli were used, but the stimuli were sampled from categories that did not require integration of the dimensions. Instead, they were created from categories oriented as in Figure 2b. The general optimal bound for the rules explicit condition is
\[ h(x, y) = (x-a_1)(y-a_2) = 0. \]
This allows a verbalizable rule to be successful on each trial (i.e. small circle and low angle or large circle and high angle belong to category A).

Figure 3.
After completing the learning task, all of the participants performed the same transfer task, a categorization task with a linear implicit category structure. Here the participants were presented with square-dot stimuli, where the height of the rectangle and the position of the dot (Figure 3) were the varying dimensions. In the transfer task, the category distribution is identical to the category distribution of the linear implicit-trained groups’ learning task (Figure 4). Here the decision rule sounds like: ‘If the dot orientation is greater than the height of the rectangle, the object is an A; otherwise, it’s a B.

Figure 4.

Stimulus distribution for the transfer task.

Predictions for the outcome of this study were formed according to the existing theories of multiple categorization systems. If implicit category learning can transfer, we would expect the linear implicit-trained group to learn faster on the transfer task than the rules explicit-trained group. Since not only did the rules explicit-trained group not learn the same category structure on both the learning and the transfer task, but even a different system was required for each of their tasks, we expect them to learn significantly slower on the transfer task than the linear implicit-trained group. We are not sure of what to expect for the quadratic implicit-trained group. If learning through the implicit system is due solely to the priming of that system, we would expect the quadratic implicit-trained group to perform as well on the transfer task as the
linear implicit-trained group. On the other hand, if the implicit system is learning the underlying category structure, we would expect the quadratic implicit-trained group to learn the transfer task slower than the linear implicit-trained group.

As the results of this study will demonstrate, participants who learn an underlying implicit category structure are able to generalize this learning to a new task with the same underlying category structure but different stimulus dimensions. This will be evidenced by the linear implicit-trained group learning faster on the transfer task than the rules explicit-trained group.

Method

Participants

Participants were 36 graduate and undergraduate students from the Psychology Department at Miami University, including 18 females and 18 males ranging between the ages of 21 and 36. All participants gave their time voluntarily and were not compensated.

Stimuli and Apparatus

Transfer of Training.

Stimuli for the learning task for both groups were circles bisected by straight lines (Fig. 1). The dimensions that varied were circle size and line orientation. Distribution parameters were chosen based on previous research (Ashby & Maddox, 1992; Thomas, 1998), and adhered to the following conventions (Table 1). Orientation units were such that 0 corresponded to a horizontal line pointing to the right and 200 corresponded to a line pointing straight up (angle = $\pi/2$). The size units were such that a circle with a radius of 300 filled the screen. Thus, sizes sampled over 300 units were discarded. Sizes below 30 units (1.4cm) were not used because these were deemed too small for reliable perception. To maintain sampling symmetry, no sampled orientation less than 30 or greater than 300 was used either. The probability of sampling a size or orientation outside of the restricted region was less than 2%.

The stimuli for the transfer task were rectangles varying in height and a dot varying in horizontal position (Fig. 3). The distribution of these stimuli retained the same underlying category structure as the learning task used in the implicit-trained condition (Fig. 2a). Distribution parameters were chosen based on their similarity to the circle-line stimuli, and adhered to the following conventions. The size units were such that a rectangle with a height of 620 filled the screen. Thus, sizes sampled over 620 units were discarded. Sizes below 30 units...
were not used because these were deemed too small for reliable perception. For the orientation of the dot, the farthest to the left the dot was able to go without placing it over the left edge of the rectangle was 30 units. The maximum value of the dot orientation was placed at 500 units. The probability of sampling a size or orientation outside of the restricted region was less than 2%.

All stimuli were presented using a Pentium-133 PC on a 15” MAG Innovision monitor with resolution set at 640x480. Observers were seated approximately 4 feet away from the screen in a dimly lit room. Stimuli subtended 3° of visual angle.

**Procedure**

A standard transfer of training experimental design was employed, with twelve subjects randomly assigned to one of three groups; the linear implicit-trained, the rule explicit-trained, or the quadratic implicit-trained. The participants were to complete a categorization learning task and a categorization transfer task.

For the learning and transfer task participants were presented with a computerized graphical categorization procedure. During the categorization task, the participant self-initiated trials by pressing the spacebar. The randomly generated stimulus appeared on the screen until either the subject responded or a time period of four seconds passed. The participant was asked to make a decision as to whether the presented object was a member of category A or category B. The keys ‘Z’ and ‘/’ on the keyboard were labeled ‘A’ and ‘B’ respectively for this purpose. Feedback was provided as to the correctness of a response, and in this manner the subject was expected to gradually learn to tell ‘A’-objects from ‘B’-objects. The probability of sampling from one or the other category distribution was .5, with each category being presented an equal number of times within a session. The distribution parameters are shown in Table 1.

Each session contained four blocks of 280 trials each, with a maximum of six sessions per task per subject, regardless of performance. The trials were organized into four blocks per session in order to encourage rest periods and decrease fatigue. Participants were told they could get up and stretch during these periods if preferred.
Table 1.

Learning Task, Linear Implicit-trained Group

\[ \mu_A = [135 \ 165] \]
\[ \Sigma_A = \begin{bmatrix} 2025 & 1984.5 \\ 1984.5 & 2025 \end{bmatrix} \]
\[ \mu_B = [165 \ 135] \]
\[ \Sigma_B = \begin{bmatrix} 2025 & 1984.5 \\ 1984.5 & 2025 \end{bmatrix} \]

Learning Task, Rules Explicit-trained Group

\[ \mu_A = [150 \ 150] \]
\[ \Sigma_A = \begin{bmatrix} 4900 & -4777 \\ -4777 & 4900 \end{bmatrix} \]
\[ \mu_B = [150 \ 150] \]
\[ \Sigma_B = \begin{bmatrix} 4900 & 4777 \\ 4777 & 4900 \end{bmatrix} \]

Learning Task, Quadratic Implicit-trained Group

\[ \mu_A = [135 \ 165] \]
\[ \Sigma_A = \begin{bmatrix} 2025 & 1984.5 \\ 1984.5 & 2025 \end{bmatrix} \]
\[ \mu_B = [165 \ 135] \]
\[ \Sigma_B = \begin{bmatrix} 2025 & 1984.5 \\ 1984.5 & 2025 \end{bmatrix} \]

Transfer Task, All Groups

\[ \mu_A = [135 \ 165] \]
\[ \Sigma_A = \begin{bmatrix} 2025 & 1984.5 \\ 1984.5 & 2025 \end{bmatrix} \]
\[ \mu_B = [165 \ 135] \]
\[ \Sigma_B = \begin{bmatrix} 2025 & 1984.5 \\ 1984.5 & 2025 \end{bmatrix} \]

Parameters of the stimulus distributions for all three learning groups and the transfer task.
Results

The dependent variable in each case is the number of blocks required to achieve 90% of optimal performance given the task. This measure was taken from earlier research using similar methods and stimuli (Thomas, 1998).

For reasons that will be explained shortly, the dependent variable was compared for the rules explicit-trained group and for the quadratic implicit-trained group on the learning task. The rules explicit-trained group took, on average, significantly more blocks to reach criterion on the learning task ($M_2 = 15.25$, $t = 8.001$, $p < .001$) than did the quadratic implicit-trained group ($M_3 = 5.58$).

To determine whether transfer of training had occurred, between-groups analysis was performed. In order to compare performance on the transfer task across the groups, participants from the linear implicit-trained group were rank-ordered on their learning task performance and matched with corresponding participants from the quadratic implicit-trained and rules explicit-trained group who had been similarly rank-ordered for training task performance. In all but three cases, the linear implicit-trained participants learned faster than the rule-trained group with one exception and two ties (in one case neither participant learned the training or the transfer task). In all but four cases, the linear implicit-trained participants learned faster than the quadratic implicit-trained participants with three exceptions and one tie. In all but four cases, the quadratic implicit-trained participants learned faster than the rules explicit-trained participants. A Wilcoxon test was conducted comparing the number of blocks to criterion on the transfer task for all of the groups. The results indicated a difference that approached significance between the linear implicit-trained and the rules explicit-trained group’s performance on the transfer task ($z = -1.892$, $p = .059$). No difference was found between the quadratic implicit-trained group and the linear implicit trained group ($z = -1.276$, $p = .202$) or between the quadratic implicit-trained group and the rules explicit-trained group ($z = -.044$, $p = .965$). A matched-pairs t-test was performed using the number of blocks to criterion on the transfer task. The mean number of blocks needed for the linear implicit-trained group ($M_1 = 10.42$) was significantly lower than the mean number of blocks for the rule-trained group ($M_2 = 13.92$), $t(11) = -2.259$, $p < .05$. A graphic comparison of participants’ performance within-groups and between-groups is shown in Figure 5.
Conclusion

This study provides preliminary support for the idea that categorization training can be generalized between different sets of implicitly defined stimuli that are composed using the same deep structure.

Figure 5.

Discussion

While none of the cognitive style assessments were significantly correlated with the dependent variables, most of them were large enough to give us confidence that significance will be attained with a larger sample size. As was mentioned earlier, we are currently conducting a
study with an even larger battery of cognitive style assessments and the same categorization tasks in order to further explore these relationships.

A post-hoc analysis of the quadratic distribution bound led to the discovery that a participant could use an explicit independent decision bound and still have obtained up to 89% correct. This bound is made up of four lines that form a box around the category B distribution. Because we are thus far unsure of the type of decision rules used by the participants we will be very cautious in interpreting the results of the quadratic linear-trained group until further analysis. We can, however, discuss the implications of each type of rule use.

It is often stated in previous literature (i.e. Ashby et al., 1998) that a person will opt for using an explicit verbal rule when they are able to do so and still succeed at the task. In the present case the participants would have used a box-shaped explicit verbal rule. Since we did not attempt to match the complexity of this distribution initially, we would now want to compare the complexity of this distribution to the complexity of the rules explicit-trained group’s learning distribution. Here we will use Boolean logic to decipher the complexity of the optimal verbal rule. For the rules explicit-trained group’s learning distribution, the rule is: ‘if x is less than 150 and y is less than 150 or if they are both greater than 150 the object belongs to category A’. This reduces to: \((a + b) \text{ or } (c + d)\). To calculate the Boolean complexity we simply count the literals, which gives us a value of four. The optimal rule for the quadratic implicit-trained group’s learning distribution reduces down to: \((a + b) \text{ and } (c + d)\). This also gives us a Boolean complexity of four. In calculating Boolean complexity, ‘or’ versus ‘and’ only matters when there is a disparity in parity (Feldman, 2000). Here the parity between the two rules is the same since there is the same number of literals in both rules, so we can presume that the two verbal rules are of equal Boolean complexity. In spite of this, the rules explicit-trained group took, on average, significantly more blocks to reach criterion on the learning task than did the quadratic implicit-trained group.

If the quadratic implicit-trained group had all used quadratic bounds we would not have any clear-cut predictions. If the implicit system learns simply by the priming of the neural areas, we would expect both implicit groups to perform equally well on the transfer task. If, on the other hand, the implicit system learns the underlying structure of the categories we would expect the quadratic implicit-trained group to take longer to learn on the transfer task than the linear implicit-trained group. While the neural areas would have been primed for the transfer task, the
quadratic implicit-trained group wouldn’t have the advantage of having learned the same underlying category distribution on both tasks. While not significant, this trend was seen in the data.

The third possibility, and the most likely of the three, is that some of the participants used an implicit decision bound and the rest used an explicit decision bound. If this is the case we will need to compare the dependent variable between both groups in order to be able to interpret the results.

As was demonstrated earlier, there are many models of categorization. This study provides support for the idea that people are able to generalize their learning of category structures to stimuli with different dimensions. While no formal model fitting has been done as of yet, the results are unable to be explained by some well-known existing models. There are a few different findings in our study that could be considered controversial in the view of several theories. While we acknowledge that more groups need to be run in order to cancel out any confounds, looking at how well our results are predicted by different categorization models is still warranted.

\textit{Stimulus Specific Models}

Obviously stimulus specific models like exemplar, prototype, and hybrids of these, do not fit our data. If a single system stimulus specific model applies, no transfer of training should obtain without the same stimulus. Since transfer of training did obtain, these systems must not have been involved. This is not to say that these categorization models are wrong, but that the system they represent may not be the only system employed in categorization.

\textit{COVIS}

In our scenario, the COVIS model predicts, theoretically, that the linear implicit-trained group would be able to categorize faster on the transfer task than on the learning task since the implicit system would have had many trials in which to build confidence. The rules group should take more trials to criterion on the transfer task than they had in the learning task since they did not have the verbal system’s weight advantage and, in fact, would have to switch from its high response confidence resulting from the previous task. This prediction was not seen. In fact, no trend at all was seen in the data: six participants were faster on the transfer task, five took longer, and one tied.

\textit{Future Directions}
More groups need to be added to the study in order to understand better its implications. Plans are in the making for at least three more groups, all of which will perform the same transfer task. One group will perform a non-categorization task using the same stimuli for their learning task. The second group will perform an explicit rule-based learning task with much simpler rules. The third group will perform a quadratic implicit based learning task where the participants will be unable to use an explicit rule to reach criterion.

Understanding the system(s) of categorization and creating models that predict their output is crucial to the field of cognitive science. This benefits us by providing a better understanding of how the general population perceives information. This understanding can thereby help to create rehabilitative programs and therapies for brain damaged populations with deficits in their ability to categorize objects.
References


