Generalization of categorization skills in language learning disabled children

Janet Elizabeth McLellan

Follow this and additional works at: https://scholarworks.umt.edu/etd

Let us know how access to this document benefits you.

Recommended Citation


This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
Generalization of Categorization Skills in
Language Learning Disabled Children

By
Janet Elizabeth McLellan
B.Sc., Trent University, 1984

Presented in Partial fulfillment of the requirements
for the degree of
Master of Arts
University of Montana
1990

Approved by

Chairman, Board of Examiners

Dean, Graduate School

Date

May 9, 1990

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Language learning disabled children (LLD) have been shown to be deficient in the acquisition of categorization skills which is hypothesized to lead to deficits in semantic knowledge, memory, word retrieval, word usage and comprehension. As a result, training of categorization skills has been used to remediate these deficits, however, the efficacy of this training to increase semantic skills has not been examined. If categorization tasks are to be used to remediate a wide variety of disorders, generalization of categorization across behaviors must occur. A first step in determining whether training in categorization generalizes across behaviors is to examine whether generalization of categorization occurs from trained to untrained items within categories and then to untrained categories.

This study used a single-subject multiple baseline design to investigate whether such generalization occurs for two eight-year-old LLD students. While one subject demonstrated generalization of training to untrained categories, the other did not. Both subjects demonstrated within category generalization. Implications for therapy and future research are discussed.
Acknowledgements

My sincere thanks to the members of my committee: Randy Weirather, Paul Silverman, Barbara Bain, Beverely Reynolds and Julie Waughbaugh whose guidance and advice were invaluable. Special thanks to Marty Crossman, the resource room teacher who allowed me into her classroom, and the parents and children without whom the study would not have been possible.
# Table of Contents

Abstract...................................................................................................................ii  
Acknowledgements.............................................................................................iii  
List of Figures.......................................................................................................vi  
Introduction.............................................................................................................1  
\hspace{1em} Development of Normal Categorization........................................3  
\hspace{1em} Evidence of Categorization Deficits in LLD Children .................6  
\hspace{1em} Training and Generalization of Categorization as a Strategy to Increase  
\hspace{1em} Recall .................................................................9  
\hspace{1em} Summary .............................................................11  
Hypothesis.............................................................................................................12  
Materials and Procedures..................................................................................13  
\hspace{1em} Subjects ..................................................................................13  
\hspace{1em} Training and Testing Environment ...........................................17  
\hspace{1em} Pre and Post Test Measures ..................................................18  
\hspace{1em} Materials ..................................................................................18  
\hspace{1em} Entry Tasks ........................................................................20  
Design .............................................................................................................22  
\hspace{1em} Baseline ..............................................................................24  
\hspace{1em} Training ............................................................................25  
\hspace{1em} Acquisition and Generalization Probes ................................27  
Reliability .........................................................................................................28  
Analysis ............................................................................................................29  
Results ...............................................................................................................30  
\hspace{1em} Pretest ..............................................................................30  
\hspace{1em} Baseline ..............................................................................31  
\hspace{1em} Acquisition and Generalization of Categorization ....................33  
\hspace{1em} Subject One .........................................................................33  
\hspace{1em} Subject Two .........................................................................36  
\hspace{1em} Summary ............................................................................37  
\hspace{1em} Post test ..............................................................................37  

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Data From Subject One Category Pair A</td>
<td>32</td>
</tr>
<tr>
<td>2. Data From Subject One Category Pair B</td>
<td>32</td>
</tr>
<tr>
<td>3. Data From Subject One Category Pair C</td>
<td>33</td>
</tr>
<tr>
<td>4. Data From Subject Two Category Pair A</td>
<td>34</td>
</tr>
<tr>
<td>5. Data From Subject Two Category Pair B</td>
<td>35</td>
</tr>
<tr>
<td>6. Data From Subject Two Category Pair C</td>
<td>35</td>
</tr>
</tbody>
</table>
Introduction

Learning disabled children are those who experience problems in the classroom but have no apparent physical, intellectual, or environmental disorder (Cavalier, 1980). Estimates are that two-thirds of these children have reading and spelling problems, one-half have arithmetic problems, one-half have handwriting problems, one-half have receptive language problems, one-quarter have expressive language problems, and one-quarter have conceptual confusions (Cavalier, 1980). Those with language difficulties are referred to as the learning/language disabled.

Many learning/language disabled (LLD) children are deficient in the acquisition of semantics or word meaning (Bauer, 1979; Hoskins, 1983; Swanson, 1986; Wigg and Semel, 1980). More specifically, LLD children often have difficulty retrieving words; organizing definitions and explanations along logical, hierarchical lines; comprehending abstract or unusual word meanings and comprehending idioms and metaphors (Hoskins, 1983; Wigg and Semel, 1980). These comprehension errors seem to be related to difficulty in recognizing the relationship between categories of words (Hoskins, 1983).

The ability to semantically categorize words has been viewed as essential in the development of word meaning, memory, and conceptual skills (Bannatyne, 1971; Hoskins, 1983; Kuczaj, 1982; Nelson, 1979; Wiig and Semel, 1980). Categorization is defined as the ability to abstract common properties of objects and create a conceptual construct or a class which can then be extended to include all objects possessing these common properties (Lavetelli, 1973).
To remediate the semantic difficulties of LLD children, several authors have suggested using categorization or sorting tasks (Hoskins, 1983; Wigg and Semel, 1980). These tasks include a variety of activities such as sorting pictures on the basis of semantic category membership, identifying an item which does not belong to a semantic class, naming the semantic category to which an item belongs, or enumerating a class or field (Wiig and Semel, 1980). Sorting and categorizing tasks have been hypothesized to aid in memory (Cremack, 1983; Ceci, 1984; Paker, Freston and Drew, 1975), word retrieval (Wiig and Semel, 1980), comprehension and word usage (Hoskins, 1983; Wiig and Semel, 1980).

Although training in categorization has been suggested for use with LLD children, little research has focused on the effectiveness of using categorization tasks in treatment. For categorization training to be considered effective, generalization of treatment effects should be observed (Hughes, 1985). Generalization has been defined as "the occurrence of relevant behavior under different non-training conditions (i.e. across subjects, people, behaviors, and/or time) without the scheduling of the same events in those conditions as had been scheduled in the training conditions" (Stokes and Baer, 1977, p.350).

At least three types of generalization have been defined: response, stimulus, and maintenance generalization (Hughes, 1985). Stimulus generalization occurs when a trained response is elicited by stimuli not involved in the original training situation (Hedge, 1985; Hughes, 1985). These stimuli may include verbal and visual antecedents, persons, and physical settings. Response generalization occurs when training of one behavior results in the acquisition of a new but related behavior, and maintenance occurs when a trained behavior is exhibited over time (Hughes, 1985). None of the preceding types of generalization have been systematically examined with young LLD children and the training of categorization skills.
The present study is designed to examine response generalization of trained categorization skills in LLD children. LLD children who are deficient in categorization skills were taught to sort words into categories such as "time concepts" and "feelings". Their ability to categorize words into two untrained categories was examined to determine if response generalization occurred. Normal development of categorization skills, evidence for categorization deficits in LLD children, and training and generalization of categorization will be discussed in the following sections.

Development of Normal Categorization

In adults, knowledge within semantic memory appears to be organized in a hierarchical system composed of semantic fields (or categories) which are groups of related words (Collins and Quillian, 1969; Kuczaj, 1982; Reed, 1982; Smith, Shoben and Rips, 1974; Wiig and Semel, 1980). Semantic memory is the store for meaningful information (Smith, 1978). This network of semantic information provides an efficient system for storing and retrieving information (Hoskins, 1983). The ability to group words into categories increases as the semantic system develops with increasing age (Cermack, 1978; Daehler and Bukatoko, 1985; Lange, 1978; Nelson, 1978; Reed, 1982).

Several theories have been hypothesized to explain the development of conceptual categories in children. These theories can be divided into three groups: feature abstraction theories, functional theories and prototype theories (Daehler and Bukatko, 1985). Feature abstraction theorists suggest that children abstract the features of objects which are labeled for them by others. These features are initially based on the perceptual aspects of stimuli, but with increased development, children are able to derive more abstract features.
Development involves adding components or features to the concept until the child's knowledge resembles adult knowledge (Clark, 1973). Functional theorists suggest that children identify the relationship between objects based on function rather than perceptual similarities. These functional relations are used as the basis for conceptual groupings (Nelson, 1978). Prototype theorists suggest that children form idealized mental representations, or prototypes, of concepts as they come in contact with different objects and events in the world.

Presently, prototype theories appear to be most accepted as they explain why objects can be categorized together without specific perceptual or functional features linking them together (Daehler and Bukatko, 1985). Rosch (1973a,b) suggested categories are composed of a core meaning surrounded by other category members of decreasing similarity so that some instances of a category are better exemplars than others. The boundaries of categories are not always clear at the periphery. For example, "tomatoes" and "apples" are both fruit but "apples" are better examples of this category than are "tomatoes". Adults quickly judge which are good or poor exemplars of a category with a great deal of agreement (Rosch, 1973a). This finding was also supported by Battig and Montague (1969) who tabulated the frequency with which objects were named by adults in response to a category name. In addition, Hupp and Mervis (1982) determined that training with good exemplars compared to poor exemplars of categories resulted in improved acquisition and generalization of category names for mentally retarded children. Thus the exemplars chosen to represent each category should be considered in studies of categorization.

Acquisition of semantic relations is important in developing word meaning and in the efficient storage and retrieval of information. The acquisition of a semantic system in children involves learning individual word meanings and the relationships between words...
(Hoskins, 1983; Kuczaj, 1982; Nelson, 1978; Wiig and Semel, 1980). The relationships between words or semantic relations include opposites, hyponymy (i.e., superordinate-subordinate relationship), part-whole relationships, collectives (i.e. collection of objects or people and their members), and synonymy (Kuczaj, 1982). The acquisition of semantic relations, particularly that of hyponymy, structures the semantic system to form the adult hierarchical semantic system (Kuczaj, 1982).

Research indicates that skill in categorization is gradually gained during the preschool years, starting with basic level categories and proceeding to superordinate and subordinate categories. Researchers agree that development of the hierarchical semantic system is essentially complete by middle childhood, although estimates of the exact age of acquisition range from five (Daehler and Bukatko, 1985) to eight years of age (Nelson, 1978). Further refinements in the number and reliability of the categories appear to be made after this age (Winters and Brzoska, 1976).

Studies of word association (e.g. Emerson and Gekoski, 1976) suggest semantic memory becomes more conventionally organized within a hierarchical system after the age of seven or eight years (Nelson, 1978). Younger children tend to respond to word association tasks with words that may serve as cues in the formation of phrases or sentences (syntagmatic association). At the age of seven or eight years, children begin to respond with words that are related by class (paradigmatic association). For example, when young children are asked to associate a word with "dog", they are more likely to respond "bone" while older children are more likely to respond "cat".

Rosch, Mervis, Gray, Johnson and Boyes-Braem (1976) further investigated the nature of category development. They defined three levels of categories: basic, subordinate and
superordinate. The authors stated the basic level is the most fundamental as it is the most cognitively efficient. Rosch et al. indicated that at the basic level categories are most distinctive from one another. Members of a superordinate category share only a few attributes among each other (for example, members of the category "animal" have a wide variation with only a few attributes which define category membership). Objects within a basic category share attributes which are common to most members (for example, all cats have four legs and whiskers, and this commonality forms their membership in the category of cats). Subordinate categories contain objects with specific defining characteristics but with many attributes which overlap with other categories (for example, wild and domestic cats both have all the characteristics of cats but their habits, markings, etc. define their membership with either the wild or domestic category). They predicted children should learn basic level categories most easily. Investigations by Rosch et al. and Mervis and Crisafi (1982) confirmed this order of acquisition. They found that 2 to 3 year-old children categorized accurately at the basic level. Four to five year-olds categorized at the basic and superordinate levels while 5 to 6 year-olds could categorize at all three levels.

Evidence of Categorization Deficits in LLD Children

The results of research have indicated at least some LLD children have deficits in categorization skills. Specifically, LLD children failed to use categorization as a strategy for recalling information (Cremack, 1983; Freston and Drew, 1974; Lorsbach, 1982; Parker, Freston and Drew, 1975; Torgenson, 1977). The semantic memory of LLD children appeared not to be as structured as that of their peers (Baker et al., 1987; Bauer, 1979; Dallago and Moely, 1980; Swanson 1986). In addition, LLD children were found to process semantic information more slowly than normal children (Lorsbach, 1982). They
also had difficulty extracting the semantic features of words and were less adept at recognizing relationships between words (Ceci, 1984; Cremack, 1983; Lorsbach and Gray, 1985). Finally, LLD children may have less knowledge about why words belong to certain categories than do normal children (Harris 1979; Scott, Greenfield and Sterental, 1986).

Hoskins (1983) found the organization of semantic information in memory affects the efficiency with which information was remembered and retrieved. Furthermore, studies which investigated the memory of LLD children have consistently found the LLD children remembered fewer items from a list and evidenced less clustering in the items recalled than did non-LLD children (Bauer, 1979; Ceci, 1984; Cremack, 1983; Dallago and Moely, 1980; Freston, 1974; Parker, Freston and Drew, 1975; Swanson, 1986; Torgenson, 1977; Torgenson, 1980).

The cause of these memory deficits is not clear. Some researchers present evidence that the structure of semantic memory (organization of words in semantic memory) is deficient (Swanson, 1986). Others present evidence that the processes for encoding or retrieving the knowledge from semantic memory are deficient (Ceci, 1984; Cremack, 1983; Lorsbach, 1982 and Lorsbach and Gray, 1985). Others suggest that both structure and processes are deficient (Baker et al., 1987). This diversity of opinion may be due to the heterogeneous nature of the learning disabled population, the differing ages of subjects, and the widely varying methodologies (e.g. differing tasks, differing measures, etc.) used in these investigations. The presence of processing deficits does not preclude structural deficits or vice versa. Both processing and structural deficits may be present in one child or processing deficits may be present in some children while structural deficits are present in others. Baker et al. (1987) suggested that deficits in either area may influence the other. For example, inadequate encoding processes may impede semantic acquisition and eventually, impede semantic structure.
Deficits in category knowledge of LLD children have been found by some investigators but not by others. Lorsbach (1982) examined the speed and accuracy of categorization in LLD and non-LLD 9-11 year old children. The children named the category to which a picture of an object belonged after a choice of two categories was provided. He found LLD children differed in speed rather than accuracy of categorization.

Other investigators studying younger children have found significant differences in the accuracy of categorization in LLD children. Scot, Greenfield and Sterental (1986) investigated two kinds of category knowledge, extensional and intentional knowledge, in six, seven and eight year-old LLD and non-LLD children. Extensional knowledge referred to exemplars that were considered members of a category while intentional knowledge referred to knowledge of properties that define category membership. They found that LLD children had deficient knowledge bases particularly in the intentional knowledge about differences between category members.

Harris (1979) examined the ability of 60 normally achieving and LLD seven and nine year old boys to sort 20 common objects into four groups (foods, toys, school supplies and cooking utensils). Significant differences were noted between the LLD group and the normal group with regard to the number of items they correctly categorized and in the types of reasons they gave for their categorization. Some nine year old LLD children based groupings on perceptual reasons unlike their same-age normal peers. Some seven year-old LLD children based categories on figural or graphic representations while others placed all the items in one pile or gave numerous contradictory reasons for the categories formed. Learning disabled children also tended to give more semantically empty reasons (e.g. "they're all things"), possibly reflecting less well developed categorical knowledge.
Previous research shows that at least some LLD are deficient in categorization skills but the demonstration of categorization deficits may be task dependent. That is, consideration must be given to the difficulty of the task (e.g. Lorsbach, 1982). On easy tasks, deficits may not be demonstrated by LLD children. The categorization deficits of LLD children are apparent in sorting tasks. In addition, these tasks appear to be useful in teaching categorization to LLD children as they are easily adapted to the therapy situation. However, their effectiveness needs to be empirically documented.

Training and Generalization of Categorization as a Strategy to Increase Recall

Categorization exercises have been hypothesized to benefit word retrieval, conceptualization, comprehension, and word usage as well as recall. Bannatyne (1971), Hoskins (1983) and Wiig and Semel (1980) suggest that if categorization skills are weak, comprehension and word usage will also be impaired. Wiig and Semel (1980) suggest language remediation of LLD children should be directed toward developing a category structure in which to organize new information. This category structure may be developed by various sorting and categorization tasks (Wiig and Semel, 1980). Sorting and classification tasks have found their way into clinical practice. Several language programs contain tasks for the development of categories (e.g. Semel, 1975; Plourde, 1985). Williams (1986) emphasized basic categorization and classification skills in her training program to improve reading comprehension.

The training of categorization skills in LLD children has been investigated as a strategy to increase recall (Cremack, 1983; Dallago and Moely, 1980; Swanson, 1986; Torgenson,
Murphy and Ivey, 1979) but not as a strategy to increase semantic skills. Swanson (1986) and Cremack (1983) found that instructions orienting LLD children's attention to the semantic content of words increased their recall. Torgenson et al. (1979) and Dallago and Moely (1980) found instructing LLD children to sort items by semantic categories increased the number of items recalled from a list of items. The findings of Dallago and Moely (1980) and Torgenson et al. (1979) imply that LLD children can be trained with a sorting task to categorize, which can be beneficial to recalling words. However, these researchers have not examined the accuracy of categorization before and after training.

Recently, investigators have come to recognize the importance of generalization of strategic behaviors such as categorization. Generalization is essential since the training of each appropriate behavior in every situation is not feasible nor desirable. Gelzheiser (1984) trained LLD adolescents to categorize as a means of improving recall of a list of words and examined whether these skills generalized to recall of prose material (response generalization). To train categorization skills, Gelzheiser (1984) used a method of direct instruction to promote generalization in LLD adolescents. This method included training skills to a level of mastery, informing the student of the usefulness of the procedure, giving the student feedback on the level of performance, giving incentives for correct performance, fading, using a variety of training material, and increasing task difficulty. The LLD children were taught to follow four rules: sort before studying, study by group, name the groups, and cluster to recall. The author found that trained LLD children used categorization as a strategy to improve recall of word lists, and recall of word lists was improved for trained groups but not the control group. This training generalized to recall of prose passages. Gelzheiser's (1984) study demonstrated LLD children can be trained to categorize words to increase recall and this training generalizes to untrained material. These researchers were concerned with the use of categorization as a strategy to increase recall
and as a result did not examine the accuracy of categorization before and after training. In addition, the researchers did not examine which categories were being formed by the subjects in either the list or prose condition.

Wacker and Greenebaum (1984) presented evidence that response generalization of a sorting task occurs in moderately and severely retarded persons. Generalization to a different shape and a different dimension (color) was tested. Subjects were first trained to sort circles and squares with either nonverbal or verbal training. Nonverbal training involved modeling of correct response, feedback on accuracy of response, correction trials for incorrect responses, and social reinforcement for correct responses. In addition to these components, the verbal training required the subjects to name the shape before sorting. After verbal training, all clients acquired the target behavior and generalized their performance to a novel shape as well as to an untrained dimension. However, none of the subjects successfully generalized the sorting behavior after nonverbal training. Verbal training facilitated generalization while nonverbal training did not. This study demonstrated that sorting skills may generalize for mentally retarded persons. The ability to generalize after training of a sorting task has not been demonstrated for LLD children.

Summary

LLD children have been shown to be deficient in the acquisition of categorization skills. This deficit may lead to deficits in semantic knowledge, affecting memory, word retrieval, word usage and comprehension. Some authors (Wiig and Semel, 1980) have suggested using categorization tasks in remediating these difficulties. Research has shown that training LLD adolescents in categorization with sorting tasks improved recall of word lists and this strategy generalized to other material such as prose. However, training
categorization to improve semantic skills has not been studied with LLD children. If categorization tasks are to be used to remediate comprehension, word usage, word retrieval and conceptualization deficits, generalization across many behaviors and situations must occur. Generalization of sorting behavior as a result of training has been demonstrated for mentally retarded children but has not been studied with LLD children. A first step in determining whether training in categorization generalizes across behaviors is to examine whether generalization of categorization occurs from trained to untrained items within trained categories and then to untrained categories.

Research Question

Will the effects of training in categorization of words generalize to untrained exemplars within trained categories and to untrained categories for learning-language disabled children?
Material and Procedures

Subjects

Two subjects were recruited through the special services division of the school districts of Bangor and Brewer, Maine. Subject selection criteria were provided to school district Special Education Directors (see below). A letter describing the study and requesting permission for the child to participate in the study was distributed to each parent. The name and home phone number was given to the investigator and the parent was then contacted by telephone. Information about the following criteria for admission to this study was obtained through the school records or from a report from the teacher. The criteria were:

1) identification as learning disabled by the school district with no other handicap present as identified by school records within the last year (See Appendix A for the State of Maine definition of Learning Disability).

2) a performance or full scale IQ within ± one standard deviation of the mean on the Weschler Scale of Intelligence - Revised administered by the school district personnel within the last year.

3) passage of a hearing screening test administered by school personnel within the last year.
4) placement in the first or second grade and an age between 8;0-8;11.

5) identification by a licensed Speech Language Pathologist as having difficulties with language within the last year.

6) identification as a native speaker of English.

Once a subject was identified by the above criteria, the investigator administered the following tests to determine if the child had deficits in categorization skills and yet understood the task of sorting. These tasks are described later in the chapter. They included:

1) the word classes subtest of the Clinical Evaluation of Language Fundamentals-Revised (CELF-R) (Wiig, Semel and Secord, 1987) with a score of no more than the tenth percentile.

2) an entry sorting task with a criterion performance of no more than 60% correct.

3) an entry perceptual card sorting task (i.e. on the basis of color) to demonstrate that the subject could indeed sort.

Thirteen potential subjects were approached. Five of these subject's parents did not respond to the initial letter of invitation to participate. Three parents denied permission for their children to participate in the study. Two of the remaining subjects did not meet entry requirements. One of these subjects scored in the average range on the Word Classes and Relations Subtest on the CELF-R and the other subject received an unacceptably high
score of 75% correct on the entry task. The remaining two subjects met all of the entry requirements.

For the three subjects failing the CELF-R word classes subtest, the following additional information was collected: treatment history, socio-economic status, academic performance, psychological reports, medical reports and information of previous speech/language testing.

All subjects were from a lower middle class background. The first subject was referred to special education after she was required to repeat her kindergarten year. She was referred for distractibility, a slow pace of working and poor writing skills. She was identified as Learning Disabled a year later on the basis of range and scatter on her WISC-R scores. She had received academic support and speech/language services since that time. At the time of this study, she was working on organizational skills, staying on task, fine motor skills, following directions, active listening and vocabulary.

Most recent testing indicated that the first subject had a verbal IQ of 91, a performance IQ of 78 and a full scale IQ of 84 on the WISC-R. Areas of significant weakness included vocabulary, expressive language skills, verbal reasoning, object assembly, visual memory and motor coordination (poor scores on motor coordination accounted primarily for the low performance IQ scores: She scored well above average on block design). Her Bender Gestalt scores were within normal limits. Language testing indicated average understanding of basic concepts and grammar, however, receptive/expressive vocabulary and expressive grammar were below average. On the Test of Auditory Comprehension she scored at the 30th percentile, on the Boehm Test of Concepts she scored at the 50th percentile, on the Peabody Picture Vocabulary Test she scored at the ninth percentile and
on the *Test of Language Development* she scored at the fifth percentile for expressive and receptive vocabulary, the ninth for expressive grammar and the 50th percentile for receptive grammar. Achievement tests indicated she had a reading level of 2.4 grades which was above her current grade level, and math skills at grade level.

The second subject was referred for special education after she was in first grade for reasons of difficulty in reading, math, oral language, understanding directions and distractibility. She was classified as LD at that time on the basis of scatter and range on her *WISC-R* scores and she has been receiving academic support and speech/language services since that time. She repeated first grade. At the time of this study, she was working on organizational skills, following directions, visual-motor integration, vocabulary, inferential comprehension, listening for details, relating stories in an organized sequence, and identification of absurdities.

Most recent testing indicated that this subject had a verbal IQ of 86, a performance IQ of 91 and a full scale IQ of 89. Particular areas of weakness included comprehension, verbal reasoning, practical judgement and block design. No visual perceptual dysfunction was noted on the Bender-Gestalt. Language testing indicated difficulty in understanding vocabulary and in processing directions. She scored below average on the *Test of Language Development* for processing information and following oral directions. She scored at the ninth percentile on the *Peabody picture Vocabulary Test* and below average on the *Token Test For Children, Parts one, three, and four*. Achievement testing indicated she had reading skills above grade level (3.3 grades) and math skills at grade level.

A third subject was administered the *CELF-R* and failed the categorization portion of this test but passed the entry task for the study. Even though this subject did not complete therapy in this study, descriptive information will be provided here because his
performance will be discussed later. Subject three was referred to special education for reading, math, and oral language when he was in kindergarten. He received speech/language therapy at the preschool level, primarily for articulation. He has received academic support and continued speech/language services since entering school. At the time of this study, he was working on following directions, auditory memory, visual motor skills, listening for details, identification of absurdities and production of /s/ and /z/.

Most recent testing indicated he had a verbal IQ of 94, a performance IQ of 95 and a full scale IQ of 93. He demonstrated strengths in vocabulary and picture completion and weaknesses in identification of similarities and coding. The Bender-Gestalt was significant for possible visual perceptual problems. Language testing indicated average receptive grammar and vocabulary skills but below average expressive skills. On the Test of Auditory Comprehension of Language he scored at the 90th percentile and on the Peabody Picture Vocabulary Test he scored at the 55th percentile. Achievement tests indicated reading skills well below grade level (at the time of this study, he continued to have difficulty identifying letters) but math skills at grade level.

Training and Testing Environment

All therapy and baseline procedures were carried out in the resource room at the child's school. Pre- and post-test measures were given after school at the child's home due to the length of time needed to administer these tasks. In all cases testing was carried out in a quiet environment, with good lighting and with the examiner seated next to the child at the table. All distractions were minimized. Parents were not present in the examination room.
Pre and Post Test Measures

A profile of each subject's language abilities was obtained prior to the study by the examiner administering the CELF-R. This test was readministered at the end of therapy to determine if changes in overall language functioning occurred during the time of training. In scoring of the CELF-R the 90% confidence intervals were determined using the procedures outlined in the manual. The difference between the pre and post-test scores is likely due to chance when their confidence intervals overlap. Therefore, scores were considered significantly different when their confidence intervals did not overlap.

Materials

Six categories were used in this study. For each category 11 stimulus words were selected (See appendix B). One of the 11 words was randomly selected to represent each category. This word was printed on an index card with an iconic symbol representing its meaning (see appendix C). The semantic categories were then randomly paired. Ten distractor words were also randomly assigned to each category pair (see appendix B). Therefore, the stimulus items consisted of 1) 3 pairs of 20 words (ten words from each category), 2) 10 distractor words assigned to each category pair (30 distractor words in all), 3) six words printed on index cards to represent each category and a blank card to represent the category of "distractor". A blank card was used to represent the distractor items as no single word would be representative of this group of words.

All of the vocabulary chosen was included on the compiled vocabulary lists by Buckingham and Dolch (1936) at the kindergarten level except for the following words:
shy, yell, shout, whisper, inch, limp and hike. These words were at the grade two level and were included to make up the ten words necessary in each category. Buckingham and Dolch compiled their vocabulary list from several different sources. These sources included a free-association task (Buckingham and Dolch, 1936), the most common words spoken by children up to age six (as cited by Buckingham and Dolch, 1936; Horn, 1925; and the International Kindergarten Union, 1928), the most common words written by children (as cited by Buckingham and Dolch, 1936; Jones, 1915; Bauer, 1916; Tidyman, 1921; Studley and Ware, 1914; and Payne and Garrison, 1931), and reading vocabularies (as cited by Buckingham and Dolch, 1936; Gates, 1926 and Thorndike, 1931). Those listed at the kindergarten level were considered by Buckingham and Dolch as "undoubtedly known to first grade children" (p.20).

Five adults were asked to match these stimuli to the index cards as described below in the baseline section to help establish the validity of these stimuli. The adults sorted the words in agreement with the examiner 94% of the time. Where disagreements occurred, the word was usually categorized as a distractor by the subject and as one of the category words by the examiner. Six of the distractor words were categorized as a category word by one adult. No one word was categorized in disagreement with the examiner by more than one adult except for "kind", "before" and "front" where two people disagreed with the examiner. Single disagreements occurred on the following words: soft, apple, full, late, past, short, across, tired, cry, whisper, shy, call, long, step, because, back, left, around, climb, roll, think and man.
Entry Tasks

Before the administration of the baseline probes, a test trial was given in which the child sorted two sets of 10 cards (10 red and 10 blue cards). The child was presented with the mixed deck of cards. The child was instructed in the following way: "Some of these picture cards belong together. Put them into two groups so that the ones that go together are in the same pile". The child was shown two boxes on the table where the cards could be placed. All children that could not sort at this basic level with at least 80% accuracy were excluded from the study as their understanding of the task of sorting would be uncertain.

To ensure the child had not already acquired adequate categorization skills, the word class subtest of the CELF-R was administered. Wüg, Semel and Secord (1987) developed this language battery to assess the language skills of children. The word classes subtest evaluates the ability to perceive the associative relationships between word concepts. They included this subtest in their battery because they noted LLD students often encounter difficulty in accurate and broad understanding of specific categories of words. During this subtest students are presented with four words auditorily. They must state the two words which "go together best". For example, if given the words "far", "near", "big" and "late", they would be required to say "far" and "near" belong together. Children doing poorly on this subtest are likely to have difficulties categorizing. Only children with scores below the 10th percentile for their age were considered suitable as subjects. Children scoring at this level would typically be considered disordered in this area.

In screening the categorization skills of the subjects, it was necessary to determine at what level of task difficulty the subjects would fail to categorize. Matching of cards with
pictures or colors on the basis of category membership is one such task. A more difficult task would be matching of spoken words to pictured category prototypes. It was the intent of this study that the entry screening task and the subsequent training task be the same. However, this was not the case and several types of entry tasks were given since the presence of a categorization deficit varied considerably as a function of the task chosen to display it.

The first task chosen was a nonverbal task using concrete categories. In this task, subjects were required to sort 20 pictures of common objects into two categories such as "animals" and "toys". Three pairs of categories were used. Two boxes were placed in front of the subject. Each box contained a prototype picture of one member of a category pair. The subjects were presented with a mixed deck of cards and instructed to sort the cards so that the ones that go together were in the same box. This task was administered to ten normally achieving first graders between the ages of six and eight. All these subjects were able to sort these pictures at or above the 90% level. The three LLD subjects who passed all other criteria were then administered this task. All subjects performed at a high level on this task (above 95%).

The number of categories was then increased to three to determine the effect on the subjects' ability to categorize. This had little effect on the subjects' level of categorization and all LLD subjects could preform at or above the 95% level. Subjects' ability to sort pictures into subcategories such as "fruits" and "vegetables" was then assessed for three pairs of subcategories. This reduced the subjects' performance slightly but performance remained above the 80% level. These tasks were then presented auditorily; again, all subjects' performances remained above the 80% level.
The final task devised is described in the baseline procedures below and the materials are described in the materials section above. The subjects were required to match abstract words with a prototype word on an index card. Distractor items were also included to increase the difficulty of the task. All children who could sort more than 60% correctly were excluded from the study. A score of less than 60% was considered a sufficient level of failure by the investigator so that improvement to an 80% level or above, which was used as an achievement criterion in this study, would be significant.

Design

A multiple baseline design across subjects and behaviors was used to assess the effects of treatment. In a multiple baseline design, treatment is systematically applied first to one behavior, subject, or setting and then to the next behavior, subject or setting while baseline is extended for untrained behaviors, settings or subjects. Experimental control is demonstrated by the replication of treatment effects on other behaviors, settings or subjects (McReynolds and Kearns, 1983). In this study, a multiple baseline design across behaviors was used to assess the generalization of sorting behavior to untrained categories (response generalization). The multiple baseline design across subjects was used to maintain experimental control. The start of therapy was staggered across subjects so that improvement in the ability to categorize could not simply be attributed to increased exposure to the materials, classroom activities or the passage of time.

The study had three phases: baseline, treatment and generalization. At the beginning of each session of the study, the subjects sorted three groups of 30 stimulus words (20 category words and 10 distractor words) described previously. The number of correctly
categorized items was recorded and served as the dependent measure throughout the study. However, the status of each category pair as a baseline, training or generalization probe changed throughout the study.

After a stable baseline was established for all category pairs, which were scored together as a single sorting task, treatment was applied to one-half of the items of the first category pair (five items from each category and five distractor items). Acquisition of the trained items was observed by the subject's ability to sort this first half of the items of the first category pair at the beginning of the session. Within category generalization was assessed by observing the subject's ability to sort the second half of the items of the category pair, and across category generalization was assessed by observing the subject's performance on the category pairs which were not yet trained. If the subject failed to generalize the trained categorization skills to the untrained items within the trained category pair, the second half of the items was trained while training of the first half of the items stopped. Acquisition of those items of the second half of the category pair was then assessed by observing the subject's ability to categorize them at the beginning of each session. Generalization to the other categories was assessed by continuing to observe the subject's performance on those items at the beginning of each session.

If, subsequent to training of the first category pair, response generalization failed to occur to the second category pair, training of the first half of those items was started. Within category generalization was measured by the subject's performance on the second half of the items in the second category pair, and across category generalization was measured by the subjects' performance on the third category pair. If generalization failed to occur after training of the first half of the items, the second half of the items was trained. The final pair of categories was trained only if generalization failed to occur after training of the second pair. This was to ensure the the third category was indeed trainable. A more
detailed chronology of treatment can be found in appendix D. Use of this design allowed assessment of distance of generalization (i.e. within or across category generalization), and the amount of training necessary for generalization to occur if it occurs (i.e. training of one or two categories).

Baseline

Each subject's ability to categorize the three groups of 30 words (20 category words and 10 distractor words each) described in the materials section was assessed. For each group, the experimenter placed three index cards in front of the child. A word representative of each of the two categories used in the sorting task was printed on two of the index cards respectively. A symbol of the word (see appendix C) was also placed on the card for the benefit of those subjects who were unable to read. The third card was blank and represented the distractor words. The child was told "We are going to practice putting words into categories or groups. I am going to say some words. Some of the words will belong best with ___(e.g. morning), others will belong best with ___(e.g. far), others won't belong with either ___(morning) or ___(far). When I say a word that belongs with ___(morning), I want you to say ___(morning). When I say a word that belongs with ___(far), I want you to say ___(far). When I say a word that doesn't belong with either, I want you to say 'blank'". The instructions were repeated prior to each session. All items which the child assigned to its matching category were scored as correct. Items were scored as incorrect when the child said they belonged with the other category or the distractors. Distractors categorized as belonging to one of the other categories were also scored as incorrect.
The order of presentation of the three groups of words was randomized for each baseline session. In addition, the order of words within each group was random for each session. Verbal, noncontingent reinforcement was used to maintain the child's involvement in the task. Four baseline sessions were conducted for the first subject and seven for the second child. If the baseline for any child was unstable, it was extended until a stable or declining baseline was reached.

Training

A verbal method of training categorization described by Wiig and Semel (1980) and elements of direct instruction described by Gelzheiser (1982) were used in training. Both verbal training and direct instruction have been demonstrated to promote generalization (Gelzheiser, 1984; Wacker and Greenebaum, 1984).

During the first training session, the uses of categorization (see appendix E for specific instructions) were explained to the subject. The child was then presented with three index cards: one card with the word and picture representative of the first category of the category pair, one with a word and picture representative of the second category, and a blank card representative of the distractor words. First the child was asked to describe the referent of one of the words on the cards. The child described its critical features (see appendix F for the salient characteristics used in training). If the child failed to give the salient characteristics, the examiner tried to elicit these through questions (see appendix G). If the child continued to be unable to give the salient characteristics of the item, the experimenter provided them for the child.
The same procedure was used for the card from the other category. The experimenter then asked how the first and second cards were different. If the child failed to provide an adequate explanation, the experimenter gave the child the information. The difference was described by contrasting the critical features of the two categories. For example, if the categories were "feelings" and "ways to communicate", the examiner might say "This is a word telling how people feel but this is a word describing ways we tell each other things, so they belong to different groups." The child was then auditorily presented with a word from one of the two categories and asked if it belonged with the first, second or blank card. They were then asked why they chose a particular card and the similarities/differences between the word and the item on each card representing a category. If the child did not correctly identify which category the word belonged to, the examiner stated the critical features of the word and the critical features of the item on the index card for each category. The child then categorized the word. If the child responded incorrectly, the examiner said "No, that's not a ___, that's a____" and told the child the salient features which determined the category it belonged to. This procedure continued until all the words were categorized. Verbal feedback ("That's right! It belongs there" or "No, That belongs in the other pile") was provided on all responses. Verbal reinforcement was given for correct responses such as "Good job", "Great" and "That's right" etc. A more detailed description of treatment procedures is given in appendix H.

The words were matched to cards two times each session. The same task was administered for at least two subsequent training sessions. After these two sessions, or when the child could correctly identify the salient characteristics of each category and describe the differences between categories for each word for one session (two matches), they were no longer asked to verbalize these facts unless a word was incorrectly categorized. Training continued until a 90% accuracy level was reached for two consecutive sessions on the acquisition probes (see below) or until the end of the seventh training session.
Acquisition and Generalization Probes

Acquisition of trained categories and generalization to untrained categories were examined with generalization and acquisition probes administered at the start of every training session. These probes were identical to the baseline probes (three groups of 30 words described in the materials section) and were administered and scored in exactly the same way. An increase in the percentage of trained items that were correctly categorized to 90% indicated the subject had acquired the skill for the items trained. An increase in the percentage of untrained items in the trained categories to 80% indicated this skill had generalized within the trained categories. If generalization to the second half of the exemplars within trained categories failed to occur by the seventh training session, the second half of the items from that category was trained using the same training procedure described above.

The remaining two groups of 30 words were used to assess generalization from trained to untrained categories. An increase in the percentage of correctly categorized items within these categories to 80% correct indicated generalization from trained to untrained category pairs. If such generalization failed to occur by the ninth training session, training of one of the two untrained category pairs was initiated and the remaining pair was used to assess generalization to untrained categories. If generalization still failed to occur, the final category pair was trained.
Reliability

Reliability of the therapy procedures was examined by having an independent judge listen to three audio taped treatment sessions for each subject. All sessions had been audio rather than video taped in order to minimize intrusiveness in the classroom environment. The sessions scored by the independent judge were randomly selected by that judge from the tape recordings of all the sessions. The judge was then given a detailed list of the procedures such as that in appendix H. The independent observer listened to the tapes and recorded "+" when a stated procedure was presented by the examiner to the subject. The judge recorded the absence or inaccurate presentation "-" when the examiner did not follow the procedure as stated. Specifically, the accuracy of feedback, delivery of reinforcement, accuracy of stimulus presentation, appropriate use of probe questions, and accurate statement of the category names and salient features were examined. A sample score form can be found in appendix I. The number of positive (+) scores was divided by the total number of scores (+ and -) and the result was multiplied by 100 to get the percentage of times the examiner followed the stated procedures.
Analysis

Percentages of correct categorization for the acquisition and generalization probes were graphed for each pair of categories across both subjects in the following format:

Acquisition of trained categories and generalization to untrained categories was determined through visual inspection of the graphs. If an increase was noted in the percentage of correct categorization of trained items to 90% correct for two consecutive sessions, then acquisition skills were inferred. If the percentage of correct categorization of untrained exemplars of untrained category pairs was observed to increase to 80% or above, response generalization of this skill was inferred.
Results

Pretest

Three subjects were administered the CELF-R and found to be below average in categorization skills. All three subjects received below-average scores for receptive language (see Appendix J). Subjects one and two scored in the low-average range for expressive language skills. Expressive and receptive scores were markedly different for the second subject; this was apparently due to a high score in the "Sentence Assembly" subtest for subject two. The third subject scored in the first percentile for expressive language. All three subjects were below average in overall language scores. The results of the CELF-R pretest are summarized in appendix J.

On the initial entry task, subject one correctly categorized 57% of category pair A, 40% of category pair B and 44% of category pair C. Subject two correctly categorized 53% of category pair A, 53% of category pair B and 56% of category pair C. Subject three correctly categorized 56% of category pair A, 76% of category pair B and 80% of category pair C. As Subject three scored above the 60% level for two of the three category pairs, further intervention was not completed.

Subject three’s pattern of response to the task differed significantly from the other two subjects. He clearly demonstrated association of words in one or both of the categories for all category pairs. For example, he categorized 80% of the items correctly in the "ways to
communicate" category while he categorized only 20% of the items correctly in the "feelings" category. In total, he categorized 67% of the category items correctly while Subjects one and two categorized only 30% and 24% respectively. As there were only three categories to which an item could belong, the latter scores were judged to be at chance level. Subject three incorrectly categorized seven of the thirty distractor words as category words. Four of the category words were categorized incorrectly as belonging to the other category. Neither of these types of errors were committed by the other two subjects. Their errors all resulted from including category items as distractor items. Therefore, subject three's errors differed not only quantitatively but also qualitatively, suggesting this subject was using a different approach or strategy to the task.

**Baseline**

Four baseline sessions were administered for the first subject and seven for the second subject. An increase in performance was noted from the first to the second sessions for the first subject but declined for the next two sessions. Therapy was begun after this session. The highest level of correct categorization was noted on category pair A and the lowest on category pair B.

An increase in the second subject's performance was noted until the fourth session for category pairs A and B, however, performance declined after this and therapy was started after the seventh session. Baseline for category pair C remained relatively stable until the end of the fourth session and then declined slightly. While baseline scores for category pairs A and B rose above 60% (70% and 63%, respectively), scores for category pair C did not rise above 56%. Data across the course of baseline and therapy for subject one is shown in figures one, two and three and for subject two in figures four, five and six.
Figure 1
Data from "Subject 1, Category Pair A"

Figure 2
Data from "Subject 1, Category Pair B"
Acquisition and Generalization of Categorization

Subject One

Subject one demonstrated learning of the trained items. She generalized this training to untrained items within trained categories but not to untrained categories. Training of category pair A took place over six sessions. The subject quickly acquired the trained items and scored 100% on categorization of trained items in the seventh session overall. The training generalized to untrained items. Criterion for generalization was met in the seventh session; however, ability to categorize untrained items declined quickly after reaching criterion. Training was continued to determine if ability to categorize the untrained items would stabilize. By session ten, the ability to categorize these items had risen to 87%. Generalization to category pairs B and C did not occur. The percentage of items correctly categorized did not rise above baseline levels during the training of category pair A.

Training of category pair B took a greater number of sessions. Training of this pair was stopped for a two week period after the fourth session because of Christmas break. Criterion for acquisition of trained items was not reached until the 6th session of training.
Criterion for generalization to untrained items was actually met before this in the third session. Training of category pair B did not generalize to category pair C. Levels of performance on this category did not rise above baseline levels. Level of correct categorization for category pair A was maintained, although this category pair was not trained at this time.

Acquisition of the trained items of category pair C was achieved by the third session after training was started. Generalization to untrained items also occurred quickly. Ability to categorize both trained and untrained items in category pairs A and B remained high although training of these items was not taking place at this time. Therapy was concluded after the third session of training category pair C. Data is displayed in figures one, two and three.

Figure 4

Data from "Subject 2 Category Pair A"

![Graph showing percentage correct over sessions for trained and untrained items.](image-url)
Figure 5

Data from "Subject 2 Category Pair B"

Figure 6

Data from "Subject 2 Category Pair C"

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Subject Two

Acquisition of category pair A did not occur as rapidly for subject two as for subject one. During verbal training of this subject, she was able to correctly describe the properties which placed items in each category. This ability was not demonstrated in the acquisition probes until later, when nonverbal training was implemented. While ability to categorize untrained items within this category increased, it did not meet the criterion of 80% before the seventh training session, therefore, training of these items was carried out for three sessions (sessions 14, 15 and 16). The training generalized to category pair B as a level of mastery on category pair A was achieved. While correct categorization increased for category pair C, it did not reach the 80% criterion. Therefore, category pair B was trained to determine if correct categorization of the untrained category pair C would increase without training. This was the case. As the subject achieved a level of mastery for category pair B, correct categorization of items in category pair C increased to the 84% level. Again this subject did not acquire the trained items as quickly as the first subject and demonstrated more variability in correct categorization of these items. Therapy was concluded after the seventh session of training category pair B.

At the end of training, the subject was asked why she categorized items in category pair C as she did to determine what her basis of categorization was. She responded that she placed items with "under" when "you can go over, or in or out like you can go under". She described words she placed with "walk" as "things you do with your legs." Therefore, it appeared that the subject had abstracted a common feature from the presented items to determine category membership that was similar to the examiner's. Data for subject two can be seen in figures four, five and six.
Summary

Despite the similarity of these two subjects in their major areas of language difficulty and school performance, their ability to generalize this categorization task to untrained categories was different. The first subject did not demonstrate generalization to untrained categories but did generalize to untrained items within trained categories. She demonstrated acquisition of trained items and she maintained this level of performance on both trained and untrained items when the category was not being trained. The second subject demonstrated generalization to untrained categories. However, her acquisition of the first trained category pair was slower than for the first subject. While the second half of the items were trained in this category because criterion was not reach in the proscribed time, they were increasing and may have reached criterion if they had not been trained. Both category pair B and C reached criterion for generalization without training. It took this subject longer to reach criterion for generalization for category pair C than for B. This may be because this subject found this category pair C more difficult. Performance on category pair C did not rise during the baseline period and was slightly lower for category pair C than for category pair B during baseline (56% vs 63%).

Post-test

The CELF-R was readministered to determine the effects of training on language abilities. Post-test subtest, composite and overall scores for both subjects either increased or remained the same. Post-test scores are summarized in appendix J. This increase in scores may be due to a testing effect that occurs when readministration of the test takes place after a short period of time (Wiig, Semel and Secord, 1987). To examine whether the increase in scores from pretest to post-test was significant, 90% confidence intervals were
determined according to the procedures outlined in the *CELF-R*. Because of the possibility of testing effects, only those pre and post-test scores whose confidence intervals did not overlap or coincide were considered as significantly different. Wiig, Semel and Secord (1987) recommend this procedure in comparing subtests. Using this comparative procedure, none of the subtest, composite or overall scores significantly differed from pretest to post-test.

**Reliability**

Results obtained by the independent judge indicated that the examiner followed stated procedures 97% of the time. In total, there were eight disagreements between the observer and stated procedures. All disagreements occurred during nonverbal training. The examiner failed to positively reinforce the child for a correct response in 6 out of 118 opportunities. The examiner also failed to ask the child why he categorized as he did when he categorized an item incorrectly in one out of seven opportunities and failed to have the child recategorize an item one time in six opportunities. The examiner also failed to state the differences between items when the child failed to do so in one out of two opportunities. However, similarities between items were stated.
Discussion

The results of this study indicate generalization of categorization skills occurred for the two learning disabled students studied; however, the extent of generalization varied between the two. One subject generalized trained categorization ability to untrained categories while the other subject generalized her training only to untrained exemplars of trained categories. While scores on the CELF-R increased for both subjects, these increases were not demonstrated to be significant at the 90% level. However, the overall Receptive Language Score for subject two was significant at the 80% level.

The categorization task was effective in teaching these skills. This was particularly evident in the case of the first subject who showed an increase in the ability to categorize the words in each category pair only when training of that pair was introduced. The second subject demonstrated increased ability to categorize all three category pairs when a level of mastery was achieved in the first trained category.

Considerable variation was noted in the subjects' performance on acquisition probes. While some variability was expected, several factors may have affected the acquisition of trained items. The first subject learned how to categorize the third category pair more quickly than the second category pair. The slower learning of the second category pair may have been due to the Christmas break and/or boredom with the task. Task boredom must be considered as a factor affecting acquisition of items. The second subject, for example, actually improved after a break in therapy at Christmas time. Having a greater variety of tasks and categories during therapy would alleviate some of this problem and
increase motivation to learn the task. Subjects were also sick on and off during the course of therapy with colds. This may account for some of the variability in performance noted. For example, the second subject complained of a cold and was coughing and sneezing during sessions 19, 20 and 21. During these sessions, performance declined slightly.

The second subject's performance on the acquisition probes appeared to improve when nonverbal training was introduced. The reason for this was not clear. The increased similarity between training and probe conditions with nonverbal training may have been a factor in the improved scores.

Why one subject generalized training to untrained categories and the other did not is not clear. At the beginning of therapy, both subjects had difficulty drawing associations between words. Both seemed to have difficulty abstracting common features as shown by their placement of the majority of items in the distractor category. This is consistent with other findings that some LLD children have difficulty extracting semantic features of words and are less adept at recognizing relationships between words (Ceci, 1984; Cremack, 1983; Lorsbach and Grey, 1985). It is possible that subject two could categorize more proficiently at the beginning of therapy and, therefore, generalized the skill more readily. The second subject showed increases in performance over a greater number of sessions in the baseline condition. However, the highest level of accuracy in categorizing items was only slightly greater for the second subject. The highest level of accuracy achieved by the first and second subjects during baseline for pair A was 66% and 70%, for category pair B was 46% and 63% and for category pair C, 53% and 56%, respectively. Therefore, this does not seem an entirely satisfactory explanation.
Generalization did not appear to be related to general levels of language competence as measured by the *CELF-R*. The subject with poorer overall language scores generalized the skill while the other subject did not. In addition, mastery was achieved by the first subject across all three category pairs but this did not automatically lead to generalization of those skills. Mastery appeared to be closely related to generalization for the second subject. Therefore, mastery of categorization skills may be a necessary but not sufficient condition for generalization.

When the results of this study were reported to the resource room teacher and the speech/language therapist, they stated they were not surprised by the results. They said the first subject had difficulty generalizing learning in other areas at school. The possibility exists that certain students have more difficulty generalizing all skills. These students may have difficulty not only drawing association between words but also between events, situations, places, etc. If this is the case, ways to aid generalization in all areas need to be found.

For logistical reasons, two aspects of direct training discussed by Gelzheiser (1984) were not used in this study. These include using a variety of materials and increasing task difficulty. Adding these to the training may increase the client's ability to generalize training to other aspects of language. In addition, increasing the number of categories trained and using a variety of activities that could directly aid in classroom work may also help increase generalization.

As suggested by Scot, Greenfield and Sterental (1986), there may actually be two skills involved in categorization: 1) the ability to abstract common features and 2) the ability to extend these similarities to other items. If two different skills are involved, then
generalization of categorization skills to untrained categories might not occur for the first subject even with training of additional categories using a variety of tasks. Direct practice in the area of weakness may be effective in teaching either one or both of these skills. In the case of the first subject, practice in extracting common features of several items may help generalize her skills to other categories. Whether this type of training would have encouraged generalization needs to be further investigated.

The possibility that training received in the classroom or by the speech/language therapist which encouraged generalization to untrained categories for the second subject cannot be completely ruled out. However, this seems unlikely as both subjects had the same resource room teacher and speech/language therapist and were working on similar goals. In addition, generalization appeared closely related to mastery of categorization in trained categories. Regardless of the effective components of training, subject two demonstrated the ability to generalize categorization skills. The specific components of training which lead to generalization of this skill need to be further investigated.

Another finding of this study was that categories using abstract concepts were necessary to challenge the subjects. The original task was changed in several ways in order to challenge the subjects. Originally, the task involved placing pictures from two superordinate categories such as toys and furniture into two groups. In a first attempt to make the task more difficult, the number of categories was increased to three. This did not have an effect on the subjects' ability to categorize. Next, the subjects were required to categorize members of subordinate rather than superordinate categories. This appeared to be a slightly more difficult task for the subjects but they continued to perform at a high level (above 80%). The task was changed from a visual to an auditory task. Changing from a visual to auditory task using the concrete words was not sufficient to challenge these subjects. The two subjects appeared able to categorize items that were concrete or
picturable such as animals and toys presented visually or auditorily with a high degree of accuracy.

Two final adjustments were made to the task in order to challenge the subjects. First, more abstract words were used and, second, a distractor category was added. While the subjects were able to categorize concrete or picturable objects with a high degree of accuracy, they had difficulty with abstract associations. In addition, the inclusion of a distractor category may have caused the subjects to fail.

By including this category, several different approaches to the task were possible. The subject could involve either or both inclusionary and exclusionary cognitive processes to categorize each item. The subject had the opportunity to assign an item to a category because of matching attributes or because the item did not fit somewhere else. Which of these two strategies caused the subjects to fail this task can not be determined from this study. Determining if the subjects would fail to categorize concrete items with a distractor category or categorize abstract items without the distractor category would be interesting. Future research should be conducted on the strategies used by LLD and normal subjects in completing different kinds of categorization tasks.

While further research needs to be completed to determine the validity of the categories used in this study and the ability of normal children to complete this task, the preliminary data presented here suggest adults can and do draw abstract associations in the presence of the distractor category. In addition, one same aged LLD child (subject three) drew these associations, suggesting this skill is developmentally appropriate for children of this age. LLD children are known to have particular difficulty with abstract definitions (Wiig and Semel, 1980). These two subjects did not appear to draw associations between items at the beginning of therapy. This may in part explain why some LLD children have difficulty with abstract definitions. Inability to see these associations may be related to their
difficulties in forming these concepts. Conversely, incomplete conceptualization of abstract concepts may interfere with their ability to form categories with these words. Another issue to be considered is flexibility of categorization. The ability to see several associations between words would be important in developing multiple definitions of words. This issue has not been examined in the literature.

It is interesting to speculate why one of the three LLD children identified by the CELF-R as having difficulty with categorization was able to complete the categorization task while the other two were not. While three subjects are not a normative sample, it is interesting to note that the two subjects who failed the categorization task had specific difficulties with vocabulary while the one that did not fail had a strength in vocabulary skills as shown by his scores on the WISC-R and the PPVT. While subject three received the same score as the other two subjects on the CELF-R, he successfully passed the entry task for this study. He responded to the task by making both quantitatively and qualitatively different errors from the other two subjects. This suggests he used a different strategy or approach to the task. Perhaps these strategies allowed him to learn vocabulary more readily. It is possible that his reportedly poor auditory memory influenced his performance on the CELF-R word classes subtest which relies heavily on auditory memory. The relationship between vocabulary learning, word usage, word retrieval and categorization has still not been clearly described. Further research to elucidate these relationships will go a long way to adding to our understanding of how children learn or fail to learn language.

In conclusion, assumptions that training is generalizing to untrained categories can not be made. If categorization tasks are to be effective in structuring and organizing semantic memory, this kind of generalization is necessary as not all possible associations between
words can been taught to clients. Testing of this kind of generalization during therapy is important. Several different techniques may be useful in facilitating generalization. For example, varying the task, providing direct practice for extracting abstract features and providing additional trained categories may be tried. Further research is needed to determine which techniques will be effective in facilitating generalization of categorization skills. The strategies used by LLD and normal children need to be examined if a full understanding of children's ability to categorize is to be reached. In addition, more complex word associations need to be examined as LLD children mature. Training of more concrete associations may not generalize to more abstract associations. Understanding of these associations may be influential in the LLD child's understanding of abstract and multiple definitions of words. The results of this study suggest some relationship between semantic language skills and categorization may exist but the nature of this relationship is unclear. The influence of categorization skills on semantic language skills needs to be further studied.
References


Appendix A: State of Maine Definition of Learning Disability

"1. The student exhibits a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which manifests itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations, and the disorder adversely affects the student's educational performance.

2. There is a severe discrepancy of approximately 1.5 standard deviations between the student's achievement and ability as determined by individual assessment of intelligence and academic achievement in one or more of the following areas: oral expression, listening comprehension, mathematical calculation or mathematical reasoning.

3. The discrepancy is not primarily the result of visual, hearing, or motor handicaps: mental retardation; emotional disturbance; or environmental, cultural, or economic disadvantage." (State of Maine, Department of Educational and Cultural Services, Special Education Regulations and Guidelines).
## Appendix B: List of Stimuli

### Category Pair A

<table>
<thead>
<tr>
<th>Feelings</th>
<th>Ways of Communicating</th>
<th>Distractors</th>
</tr>
</thead>
<tbody>
<tr>
<td>happy</td>
<td>talk (exemplars of categories)</td>
<td>soft</td>
</tr>
<tr>
<td>surprised</td>
<td>say</td>
<td>apple</td>
</tr>
<tr>
<td>love</td>
<td>call</td>
<td>clean</td>
</tr>
<tr>
<td>kind</td>
<td>tell</td>
<td>shiny</td>
</tr>
<tr>
<td>angry</td>
<td>ask</td>
<td></td>
</tr>
<tr>
<td>sad</td>
<td>yell</td>
<td>full</td>
</tr>
<tr>
<td>hate</td>
<td>cry</td>
<td>plant</td>
</tr>
<tr>
<td>shy</td>
<td>scream</td>
<td>big</td>
</tr>
<tr>
<td>afraid</td>
<td>shout</td>
<td>bread</td>
</tr>
<tr>
<td>tired</td>
<td>whisper</td>
<td>pepper</td>
</tr>
<tr>
<td>funny</td>
<td>speak</td>
<td>dirty</td>
</tr>
</tbody>
</table>

### Category Pair B

<table>
<thead>
<tr>
<th>Time Concepts</th>
<th>Distance Concepts</th>
<th>Distractors</th>
</tr>
</thead>
<tbody>
<tr>
<td>morning</td>
<td>far (exemplars of categories)</td>
<td>empty</td>
</tr>
<tr>
<td>minute</td>
<td>near</td>
<td>because</td>
</tr>
<tr>
<td>seconds</td>
<td>inch</td>
<td>telephone</td>
</tr>
<tr>
<td>before</td>
<td>mile</td>
<td>dog</td>
</tr>
<tr>
<td>while</td>
<td>close</td>
<td>fat</td>
</tr>
<tr>
<td>noon</td>
<td>short</td>
<td>cut</td>
</tr>
<tr>
<td>night</td>
<td>across</td>
<td>sing</td>
</tr>
<tr>
<td>late</td>
<td>long</td>
<td>green</td>
</tr>
<tr>
<td>day</td>
<td>step</td>
<td>word</td>
</tr>
<tr>
<td>past</td>
<td>foot</td>
<td>build</td>
</tr>
<tr>
<td>tomorrow</td>
<td>yard</td>
<td></td>
</tr>
</tbody>
</table>
### Category Pair C

<table>
<thead>
<tr>
<th>Spatial Concepts</th>
<th>Ways of Moving</th>
<th>Distractors</th>
</tr>
</thead>
<tbody>
<tr>
<td>under</td>
<td><em>walk</em> (exemplars of categories)</td>
<td></td>
</tr>
<tr>
<td>up</td>
<td>limp</td>
<td>think</td>
</tr>
<tr>
<td>back</td>
<td>climb</td>
<td>man</td>
</tr>
<tr>
<td>down</td>
<td>run</td>
<td>sweet</td>
</tr>
<tr>
<td>left</td>
<td>hop</td>
<td>nose</td>
</tr>
<tr>
<td>through</td>
<td>jump</td>
<td>brown</td>
</tr>
<tr>
<td>behind</td>
<td>fall</td>
<td>rocket</td>
</tr>
<tr>
<td>around</td>
<td>skip</td>
<td>waterfall</td>
</tr>
<tr>
<td>beside</td>
<td>roll</td>
<td>break</td>
</tr>
<tr>
<td>front</td>
<td>crawl</td>
<td>cheap</td>
</tr>
<tr>
<td>between</td>
<td>hike</td>
<td>baby</td>
</tr>
</tbody>
</table>
Appendix C: Index Card Stimuli

happy
talk
Far
Morning
Walk
Under
Appendix D: Treatment Chronology

Coding of Stimuli for Use in Treatment Map

Groups of stimuli (note: each group also has an additional 2 target stimuli on index cards)

A. (n=30)
1. Ways to communicate words (10)
   a. a random set (5)
   b. remaining set (5)
2. Feeling words (10)
   a. a random set (5)
   b. remaining set (5)
3. Distractors (10)
   a. a random set (5)
   b. remaining set (5)

B. (n=30)
1. Time concept words (10)
   a. a random set (5)
   b. remaining set (5)
2. Distance concept words (10)
   a. a random set (5)
   b. remaining set (5)
3. Distractors (10)
   a. a random set (5)
   b. remaining set (5)

C. (n=30)
1. Prepositions (10)
   a. a random set (5)
   b. remaining set (5)
2. Ways to move (10)
   a. a random set (5)
   b. remaining set (5)
3. Distractors (10)
   a. a random set (5)
   b. remaining set (5)
Treatment Map
(codes refer to the sets of stimuli identified above)

Baseline sessions: A,B,C

Treatment sessions:

1. Therapy: A1a, A2a, A3a
   Generalization Probes
   within category: A1b, A2b, A3b
   across category: B, C
   Acquisition Probes: A1a, A2a, A3a

Criteria for Continuation:
Repeat session until one of the following criteria is met:
   a) end of session 6
   b) 80% correct on within category probe
   c) 80% correct on across category probe
   d) If b or c are not met, continue sessions:

7. Therapy: A1b, A2b, A3b
   Generalization Probes
   across category: B, C
   Acquisition Probes: A1b, A2b, A3b

9. Repeat session 7
10. Repeat session 7
11. Therapy: B1a, B2a, B3a
   Generalization Probes
   within category: B1b, B2b, B3b
   across category: C
   Acquisition Probes: B1a, B2a, B3a

Criteria for Continuation:
Repeat session 10 until one of the following criteria is met:
   a) end of session 16
   b) 80% correct on the within category probe
   c) 80% correct on the between category probe
   d) If b or c are not met, continue session:

17. Therapy: B1b, B2b, B3b
   Generalization Probes
   across category: C
   Acquisition Probes: B1b, B2b, B3b
18. Repeat session 17
19. Repeat session 17
20. Therapy: C1a, C2a, C3a
   Generalization Probes
   within category: C1b, C2b, C3b
   Acquisition Probes: C1a, C2a, C3a
   Criteria for Continuation:
   Repeat session 20 until one of the criteria are met:
   a) end of session 25
   b) 80% correct on within category category probe
   c) If criteria b is not met continue:
27. Therapy: C1b, C2b, C3b
   Acquisition Probe: C1b, C2b, C3b
28. Repeat session 27
29. Repeat session 27
Appendix E: Uses of Categorization

"In the world there are many things that belong together. These are things that are alike in some way. Sometimes it is important to know what things go together and what things do not. For example, it can help you to learn new words. Someone tells you a 'smiff' is a vegetable. You may not know what a smiff is but you know it is something that you eat and that it grows in the ground. These are two things that are alike about all vegetables. It can also help you to remember things. If you try to remember what you have in your room, you can try to think about all the clothes you have, then all the toys you have, then all the furniture you have. So, you see, learning which things go together can be a big help. What we're going to do now is learn about how things can go together."
Appendix F: Critical Features

Feelings

a word describing feelings people have

Ways to Communicate Words

a word saying how you can say something to somebody

Time Concepts

a word saying when something happened or telling about time

Distance Concepts

a word saying how far or how close something is or telling about distance

Spatial Concepts

a word saying where something is

Ways to Move

a word saying how you can move from one place to another
Appendix G: Elicitation Questions

When somebody says that word what are they telling you about?

What does that word mean?

Is the word telling you about ______ or _______?
Appendix H: Treatment Procedure

1) Word on index cards is presented to child who is requested to give its salient features.

   Failure: Examiner elicits salient characteristics with probe questions.
   Failure: Examiner gives salient characteristics.

2) The second word card is presented to the child and the child is requested to give its salient characteristics.

   Failure: Examiner elicits salient characteristics with probe questions.
   Failure: Examiner gives salient characteristics.

3) Child is required to state the differences between the two words.

   Failure: Examiner gives the difference between the two words in terms of their salient characteristics.

4) A third word is presented auditorily from one of the two categories and the child is requested to categorize it.

5) The child is asked why they categorized it as they did and to state the differences/similarities between this word and the first two words on the index cards.

   Failure: Examiner gives the differences/similarities.

6) Child is asked whether the word belongs with the first or second word presented or neither, if their first categorization was incorrect.

   Failure: Examiner gives the category names and salient features that determine category membership.

7) Examiner provides feedback on the accuracy of the match and verbal reinforcement for every correct match.

8) Repeat 4) through 7) for all additional words.

After the child correctly verbalized the salient features of the word for two consecutive sessions, they were not required to verbalize this information (i.e. steps 1-3 and 5 were
omitted), unless they sorted a word incorrectly. The following procedure was used in this case:

1) The child sorts a word incorrectly.

2) Examiner asks the child why they sorted as they did and requires the child to state the differences/similarities between this word and the words on the index cards.

Failure: examiner gives the similarities/differences.

3) Examiner asks the child which category the word belongs in.

Failure: Examiner restates the salient characteristics.

4) Child categorizes card.

Failure: Examiner gives the category and salient characteristics of the card.

5) Examiner provides feedback as to the accuracy of the match.
Appendix I: Score Form For Reliability Check

Instructions: Score (+) when presented as in outlined procedure, score (-) when procedure differs and note how it deviates from outlined procedure. All procedures are to be scored for each opportunity for occurrence in the treatment session.

1. Instructions are provided at the beginning of the session.

Training Procedure When Verbalization Required:

2. Salient characteristics of the first word requested:
   a) If the child fails 2. salient characteristics are elicited through probe questions:
   b) If child fails a. salient characteristics are stated by the examiner:

3. Salient characteristics of the second word requested:
   a) If child fails 3. salient characteristics are elicited through probe questions:
   b) If child fails a, salient characteristics are stated by the examiner:

4. Differences between the first and second words are requested:
   a) If child fails 4. differences are stated by the examiner:

5. Categorization of a third, verbally presented word from either category is requested by the examiner:
   a) If correct on 5. the child is reinforced:
   b) Regardless of correctness, child is required to state why they categorized as they did:
   c) If child does not provide differences and similarities in response to b, he is asked to do so:
   d) If child fails 5.c, examiner elicits them through probe questions:
   e) If child fails 5.d, examiner states differences and/or similarities:
   f) If child fails 5, he now attempts to recategorize or is asked to do so:
      i) If correct on f, child is reinforced:
      ii) If incorrect on f, child is corrected and the salient characteristics are given by the examiner:
6. Score each subsequent word according to all scoring dimensions of #5:

Training Procedure When Verbalization Not Required

5. Categorization of a third, verbally presented word from either category is requested by the examiner:

   a) If correct on 5. child is reinforced:

   b) If incorrect:

      i) Child is required to state why he categorized as he did:

      ii) If he does not provide differences/similarities in response to i, he is asked to do so:

      iii) If correct differences/similarities aren't provided by the child, they are stated by the examiner:

      iv) Child attempts to recategorize or is asked to do so:

      v) If correct the child is reinforced:

      iv) If incorrect a second time, salient features of category membership are restated by the examiner:

      vii) Child attempts to recategorize a third time or is asked to do so:

      viii) If correct, child is reinforced:

      ix) If still incorrect, child is corrected and salient characteristics are restated:

6. Recycle through item 5 for subsequent words.

The first item 5 is operable only for words
1. presented during the first two training sessions.
2. whose category membership the child had not justified by stating salient characteristics and differences twice in one session.

The examiner will identify the reliability tapes as suitable for the first or second scoring procedure for items 5 and 6.
Appendix J: Pretest and Post-test Results of CELF-R

### Pretest

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Standard Score</th>
<th>Confidence Interval</th>
<th>Percentile Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>1  2  3</td>
<td>1  2  3</td>
<td>1  2  3</td>
</tr>
<tr>
<td>Oral Directions</td>
<td>8* 4 5</td>
<td>6-10, 2-6, 3-7</td>
<td>25  2  5</td>
</tr>
<tr>
<td>Word Classes</td>
<td>6  5  6</td>
<td>4-8, 3-7, 4-8</td>
<td>9  5  9</td>
</tr>
<tr>
<td>Semantic Relationships</td>
<td>7  7  5</td>
<td>5-9, 5-9, 3-7</td>
<td>16  16 5</td>
</tr>
<tr>
<td><strong>Receptive Language Score</strong></td>
<td>80** 70 70</td>
<td>72-88, 62-78, 62-78</td>
<td>9  2  2</td>
</tr>
<tr>
<td>Formulated Sentences</td>
<td>10  7  3</td>
<td>8-12, 5-9, 1-5</td>
<td>50  16 1</td>
</tr>
<tr>
<td>Recalling Sentences</td>
<td>6  8  6</td>
<td>4-8, 6-10, 4-8</td>
<td>9  25 9</td>
</tr>
<tr>
<td>Sentence Assembly</td>
<td>8  10 6</td>
<td>6-10, 8-12, 4-8</td>
<td>25  50 9</td>
</tr>
<tr>
<td><strong>Expressive Language Score</strong></td>
<td>87 88 67</td>
<td>79-95, 80-96, 59-75</td>
<td>22  21 1</td>
</tr>
<tr>
<td>Total Language Score</td>
<td>82  77 66</td>
<td>76-88, 71-83, 60-72</td>
<td>12  6 1</td>
</tr>
</tbody>
</table>

* mean=10; standard deviation=3
** mean=100; standard deviation=15
## Post-test

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Standard Score</th>
<th>Confidence Interval</th>
<th>Percentile Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Oral Direction</td>
<td>8</td>
<td>7</td>
<td>6-10</td>
</tr>
<tr>
<td>Word Classes</td>
<td>8</td>
<td>7</td>
<td>6-10</td>
</tr>
<tr>
<td>Semantic Relationships</td>
<td>10</td>
<td>9</td>
<td>8-10</td>
</tr>
<tr>
<td><strong>Receptive Language Score</strong></td>
<td>91</td>
<td>85</td>
<td>83-98</td>
</tr>
<tr>
<td>Formulated Sentences</td>
<td>12</td>
<td>9</td>
<td>10-14</td>
</tr>
<tr>
<td>Recalling Sentences</td>
<td>6</td>
<td>8</td>
<td>4-8</td>
</tr>
<tr>
<td>Sentence Assembly</td>
<td>10</td>
<td>10</td>
<td>8-12</td>
</tr>
<tr>
<td><strong>Expressive Language Score</strong></td>
<td>95</td>
<td>93</td>
<td>87-103</td>
</tr>
<tr>
<td><strong>Total Language Score</strong></td>
<td>92</td>
<td>88</td>
<td>84-98</td>
</tr>
</tbody>
</table>