JASH
1984, Vol. 9, No. 4, 287-295



# Responding in the Presence of Nontrained Stimuli: Implications of Generalization Error Patterns

Robert H. Horner, G. Thomas Bellamy, and Geoffrey T. Colvin.

**Author Information**
Robert H. Horner, Ph.D., Assistant Professor, Special Education and Rehabilitation, University of Oregon.
G. Thomas Bellamy, Ph.D., Associate Professor, Special Education and Rehabilitation, University of Oregon.
Geoffrey T. Colvin, Ph.D., Behavior Specialist, Central Services Offices, Casper, Wyoming.

**Article Descriptors**
generalization; stimulus control; error patterns.

*The technology of applied behavior analysis has documented impressive advances in procedures for changing behavior. Behavior change is often functional, however, only if it endures over time and is performed in nontrained situations. One approach to building a technology that meets these standards has been systematic analysis of generalization "successes." The present paper extends this approach through an analysis of generalization failures (i.e., the patterns of generalization errors). A format is provided for categorizing the errors made in generalization situations, and linking these errors back to stimulus characteristics present during training. Stimulus control serves as the pivotal concept that directs this error analysis, and provides the bridge between infrahuman research on generalization and the broader applied issue of obtaining performance across nontrained situations. Implications of the error analysis are defined for building a technology of generalization, and for developing an effective research methodology for studying generalized responding in applied settings.*

Recent reviews of generalization research have synthesized the progress toward building a technology and theory of applied generalization. These reviews have relied on a careful analysis of interventions that produce responding under nontrained stimulus conditions (Drabman, Hammer, & Rosenbaum, 1979; Sanders & James, 1983; Stokes & Baer, 1977). The products of these reviews have been taxonomies of effective intervention strategies and operational *classes of generalization*. The utility of these contributions is now evident in the attention being given to research and intervention focused on generalization.

Progress in science and technology, however, results not just from review of successes, but also from systematic analysis of problems and errors (Kuhn, 1962). The purpose of this paper is to provide an analysis of generalization error patterns to supplement existing reviews of generalization successes. Problems and errors of generalization are those instances in which no behavioral gains are evident in nontreatment situations even after excellent gains have been achieved within the treatment setting. The underlying concept that allows delineation of generalization error patterns, and functional solutions for avoiding the error patterns, is stimulus control. Error patterns identified from a stimulus control perspective have direct implications both for a technology of generalized responding and for the research methodology used to study generalization in applied settings. Within this context, the term *generalized responding* includes any situation in which newly acquired responses are performed in the presence of nontrained stimuli. The responses may be the result of stimulus control, stimulus generalization, or a variety of other phenomena. The term generalized responding should not be confused with *response generalization*, with which it shares only semantic similarity.

## The Applied Problem of Generalization

Behaviors that are modified and skills that are taught under specific stimulus conditions all too often fail to occur in other appropriate conditions, occur too frequently in inappropriate conditions, and do not maintain over time regardless of the conditions (Becker, Engelmann, & Thomas, 1975; Stokes & Baer, 1977). The applied problem is complex: Behavioral and educational gains must be translated into lifestyle changes that are functional across the variety of stimulus conditions imposed by the natural environment. Stated succinctly, the applied problem is to deliver interventions that reliably and efficiently result in the acquisition of adaptive behaviors that endure over time, are performed across the full range of appropriate stimulus conditions, and are not performed across inappropriate stimulus conditions.

It is tempting to label this entire problem *generalization*, and press the search for a single intervention that produces the needed behavior change. An alternative approach is provided by emphasizing the concept of *stimulus control*. Stimulus control refers to a functional relationship between presentation of an antecedent stimulus and change in the probability of a response (Terrace, 1966). From a stimulus control perspective, the applied problem is not one of generalizing behavior, but of bringing adaptive responses under the control of appropriate stimuli. Viewed in this way, the problem of generalization becomes many different problems. If a behavior does not occur in a novel setting it may be because stimuli in that setting do not exert sufficient stimulus control over the response. If conditioned behaviors occur in inappropriate situations it may be because training procedures have brought these behaviors under control of inappropriate stimuli or have not restricted the range of controlling stimuli.

Indeed, stimulus control provides the foundation for much of the existing behavioral technology. The presentation and fading of antecedent stimuli (e.g., $S^D_1$ and $S^D_2$), use of correction procedures and application of differential consequences are all procedures that build stimulus control (Snell, 1983; Terrace, 1966). Instruction is defined as complete, or successful, when a student's responding is under control of appropriate stimuli. The stimulus "t" should serve as the $S^D$ for the response /t/; the stimulus "say cup" should be the $S^D$ for the response "cup"; a "walk light" at a street corner should be the $S^D$ for crossing the street; and a verbal request from a teacher normally should be the $S^D$ for a compliant response. The ability of teachers and behavior managers to bring adaptive responses under stimulus control is the strength of our current teaching technology. It is by extending this attention to stimulus control beyond training settings that a technology for teaching enduring behavior across novel stimulus conditions will occur. One key to such an extension is shifting our training focus from simple Stimulus $\longrightarrow$ Response relationships to building control by stimulus classes.

For behaviors to occur under appropriate, nontrained stimulus conditions the learner must treat the new conditions as if they were similar to conditions encountered during training. This involves bringing a target behavior under control of a group or class of stimuli that are similar in specific ways across all appropriate situations. A *stimulus class* is a set of stimuli that share certain characteristics. All stimuli (or stimulus conditions) that contain this set of stimulus characteristics are members of the class. All stimuli that contain none, or only some, of the characteristics are not members of the stimulus class (Becker et al., 1975). From a stimulus control perspective, the applied problem of generalization translates to defining the stimulus class that should control a target behavior, and intervening so that only those stimulus characteristics common to all members of the stimulus class control responding (c.f., Becker et al., 1975; and Horner, Sprague, & Wilcox, 1982 for more detailed discussions).

The concept of stimulus class has been useful in both experimental (Honig & Urcuioli, 1981) and applied (Guess, Keogh, & Sailor, 1978; Patterson, 1974) research. It is a functional concept in part because it encourages teachers and behavior managers to define the full range of stimulus conditions that should control responding, and in part because it leads to a separation of stimuli (and stimulus characteristics) into two groups: those that are relevant and should control the target behavior; and those that are irrelevant and should have no control over responding. A comprehensive behavioral technology should produce responding across any or all members of a stimulus class, should define procedures for bringing those responses under control of the relevant stimuli common to all members of the target stimulus class, and should ensure that irrelevant stimuli do not exert stimulus control.

When the applied problem of generalization is viewed as a problem of stimulus control, attention necessarily shifts from *where* target behaviors occur to *what* class of stimuli do or do not control them. This shift provides the logical foundation for defining generalization error patterns and the intervention procedures for avoiding and remediating such patterns.

## Error Patterns

When a learned response is performed in a nontrained, but inappropriate situation, or fails to be performed in a nontrained, appropriate situation, a *generalization error* has occurred. When generalization errors occur, questions should arise as to what controlling stimulus variables affect the errors, and how training could have avoided the errors. Adoption of this approach focuses attention on the *patterns* of errors after instruction, and the implications of these error patterns for applied intervention and research. With a few notable exceptions (e.g., Carnine & Becker, 1982; Rincover & Koegel, 1975; Sidman & Stoddard, 1967; Stella & Etzel, 1983), existing research has done little to analyze the nature of generalization errors. Those studies that have addressed generalization errors indicate at least four important error patterns associated with responding in nontrained stimulus conditions. With each case the error patterns can be attributed to specific variables in the training and/or performance settings. Each of these error patterns is described briefly below, followed by a discussion of intervention procedures for avoiding such errors, and research procedures for furthering our analysis of responding in nontrained stimulus conditions.

**Irrelevant stimuli control the target response.** If training results in a target response coming under control of irrelevant stimuli, a complex error pattern results in which the response occurs in appropriate stimulus conditions that contain the irrelevant stimuli (Rincover & Koegel, 1975); but also occurs in inappropriate conditions that contain the irrelevant stimuli (Anderson & Spradlin, 1980; Sidman & Willson-Morris, 1974); and does not occur in appropriate conditions that do *not* contain the irrelevant stimuli (Rincover & Koegel, 1975).

Rincover and Koegel (1975) documented this pattern with children labeled autistic who were unintentionally taught that irrelevant furniture or trainer cues were $S^D$s for compliance with simple verbal commands. Only when these irrelevant stimuli were presented simultaneously with the verbal command did responding occur. Similarly, students with handicaps who learn that the discriminative stimulus for crossing a street is initiation of crossing by another person (e.g., wait until someone else begins to cross, then "go") exhibit an error caused by the response being under control of an irrelevant stimulus (e.g., another person beginning to cross).

Engelmann and Carnine (1982) have described several situations in which the behavior of students with handicaps has come under control of irrelevant stimuli. They attribute this control to training procedures that systematically pair responding in the presence of irrelevant stimuli with reinforcement. A stimulus comes to control a response when the response is differentially reinforced over multiple trials in the presence of that stimulus. If irrelevant stimuli are paired consistently with reinforcement, the student may learn to treat irrelevant stimuli as if they are relevant. This error pattern can be avoided if multiple training examples are used that include both positive instances (where the response is appropriate) and negative instances (where the response is inappropriate). The positive examples should sample the range of relevant stimulus variation the student will encounter, and avoid consistent presentation of any irrelevant stimuli (Becker et al., 1975).

**Irrelevant stimuli control irrelevant responses.** A second pattern of errors develops when, following successful training, the student does not perform the target behavior under novel conditions because irrelevant stimuli in the novel conditions exert more powerful control over irrelevant responses than the relevant stimuli exert over the target response. This may result from training procedures that produce insufficient stimulus control, or because the learner has a prior history with the irrelevant stimuli that has resulted in a very powerful, but inappropriate, stimulus control relationship. It should be noted that these pre-existing, irrelevant Stimulus $\longrightarrow$ Response relationships, while unacceptable from a teacher or parent perspective, are often very "functional" (i.e., produce reinforcers) from the learner's perspective. Our use of the terms *irrelevant stimuli* and *irrelevant responses* are drawn from the teacher/parent perspective. The very point we wish to emphasize, however, is that adaptive responding does not occur in some nontrained stimulus conditions because stimuli defined as irrelevant by the teacher are considered relevant by the learner.

Horner and Budd (in press) have described the error pattern of irrelevant stimuli controlling irrelevant responses with an 11-year-old, nonverbal student who persistently grabbed and yelled to obtain things. The student was taught to use hand signs to ask for a variety of items for which he typically grabbed. Even though he acquired the ability to use signs, he continued to grab and yell until training procedures were employed in those settings and times typically associated with grabbing and yelling. Once the student began signing at natural times during the day, grabbing and yelling responses ceased to occur. The authors suggest that the use of signs may not have occurred across all parts of the student's day because grabbing and yelling were under stronger stimulus control in those situations than sign use.

Berler, Gross, and Drabman (1982) have identified a similar example of irrelevant stimulus control with three boys with mild handicaps. The subjects were successfully trained to use several social skills in a training setting. Training improvement did not, however, affect the students' social behavior in a "free play" setting that contained the same discriminative stimuli for appropriate social skills. It is possible that the subjects had learned inappropriate social behaviors during their previous years of free play sessions, and that the resulting irrelevant stimulus control over these inappropriate behaviors was too strong to afford an opportunity for the newly acquired social skills.

The applied problem of developing responses that are performed across nontrained situations is often discussed as if a student enters these nontrained situations with no prior history. In fact, behaviors such as street crossing, purchasing items in stores, sharing-instead-of-aggressing-against-peers are not performed under totally new conditions. The student may enter these situations having already learned from years of experience that an inappropriate behavior pattern is functional. If a technology of instruction is to deal effectively with performance across classes of nontrained situations, consideration must be given to procedures that identify and address the prior learning that students bring to training and generalization settings.

At a minimum, training should occur with a set of positive and negative examples that include those irrelevant stimuli that control irrelevant responses. This tactic will result in more errors during the initial phase of training, but the student will learn that the problem stimulus is irrelevant, and be more likely to perform correctly in nontrained situations presented after training. When the irrelevant control is very strong, it is possible that generalization of an adaptive response will not occur unless specific procedures are implemented to break the existing irrelevant (competing) control (Horner & Budd, in press; McDonnell, Horner, & Williams, 1984).

**Restricted stimulus control.** Restricted stimulus control exists when a response that should be under the control of multiple relevant stimuli or multiple characteristics of a relevant stimulus is only controlled by a subset of those stimuli (Carnine & Becker, 1982; Carr & Kologinsky, 1983; Coon, Vogelsberg, & Williams, 1981; Engelmann & Carnine, 1982). If all members of a stimulus class contain a group of relevant characteristics, and a learner treats only some of those characteristics as relevant, s/he will make errors by responding to those examples outside the stimulus class that contain only a subset of the relevant stimulus characteristics.

An example of restricted stimulus control was encountered by the authors on a vocational task requiring placement of a small resistor on a circuit board. The response (place the resistor in the proper holes with the proper orientation) needed to be controlled by several stimulus characteristics (i.e., different colored stripes on the resistor). A worker who was severely retarded had learned to place a resistor with red, blue, and brown stripes in a particular set of holes. After reaching criterion on one training resistor possessing red, blue, and brown stripes, he consistently placed a part with red, green, and orange stripes, and another with red, brown, and black stripes into those same holes on the board. His error was treating all resistors (red, blue, brown; red, green, orange; and red, brown, black) as if they were the same. His responses were controlled only by red stripes when, in fact, correct responses required attention to all three colors.

This error pattern may also describe the generalization errors reported by Stella and Etzel (1983). They trained nonhandicapped preschoolers to "count musical notes." A quarter note was to get a count of one; a half note a count of two; a dotted half note a count of three; and a whole note a count of four. One error pattern exhibited by subjects was to give each note a count of one. Correct counting required simultaneous attention to multiple relevant characteristics, i.e., color of note, presence of stem, presence of dot. It is possible that the error pattern was due to subjects only responding to a subset of the relevant characteristics.

To avoid the development of restricted stimulus control, training examples should be chosen so some examples contain all the relevant characteristics, and some contain subsets of the relevant characteristics. With this complete set of positive and negative instances, correct performance (i.e., responding only in the presence of all relevant characteristics) can be differentially reinforced (Engelmann & Carnine, 1982).

**Limited variations of the target response.** Responding across nontrained situations in the natural environment often requires that the target response occur with minor variations in topography and intensity across different stimulus situations. For example, there are many ways of throwing a ball, opening a door, or manipulating a toothbrush. The stimulus demands of a particular situation dictate which of the response variations will be most effective. Part of responding in a generalized manner involves selecting the correct response variation in novel stimulus situations. One pattern of errors that limits adaptive responding is the use of the wrong response variation when a new stimulus condition is presented.

Engelmann and Carnine (1982) suggest that the use of inappropriate variations of target behaviors is often the result of selecting inappropriate training examples. Students with severe retardation, for example, were trained by Sprague and Horner (1984) to activate vending machines by pushing the correct button. When these students were presented with new vending machines that required a pull response, their consistent error was to push the activation lever. A similar group of students were taught to use screwdrivers. When screwdriver training was only done with a large, philips screwdriver that allowed success with a "low precision" response, the students performed repeated precision errors when given a task requiring the use of a small, slotted screwdriver (Colvin & Horner, 1983; O'Connell, 1981). Student performance improved when they received training with examples that included different types and sizes of screwdrivers that sampled the range of response demands experienced in nontrained situations.

A similar response pattern was demonstrated by Garcia, Baer, and Firestone (1971) in their analysis of experimental control over imitation responses within and across topographical boundaries. After being trained to imitate motor or vocal responses, subjects demonstrated generalized imitation to nontrained stimuli within the response modality they were trained in (motor or verbal). Subjects did not imitate in nontrained situations requiring application of the *generalized imitation* skill to a response in the nontrained modality. It is possible that this error pattern was due to the limited variations (i.e., single modality) of the target response performed during training.

## Implications of Error Patterns for a Technology of Generalization in Applied Settings

A comprehensive intervention technology does not exist that systematically prevents the error patterns described above. There are, however, many controlled studies and applied recommendations addressing pieces of the needed technology (Baer, 1981; Drabman et al., 1979; Engelmann & Carnine, 1982; Stokes & Baer, 1977). The current need is to define procedures for avoiding generalization error patterns and to translate these procedures into intervention guidelines for teachers and behavior managers. Four areas exist in which intervention guidelines are needed: (a) definition of instructional objectives, (b) selection of teaching examples, (c) specification of procedures used during instruction, and (d) analysis of performance settings. The following discussion focuses on the implications of the error pattern analysis for each of these points of intervention.

### Defining Instructional Objectives

Programming for generalized responding begins by defining the range of stimulus situations across which the response is expected to occur (i.e., the stimulus class). Instructional objectives should be written to define where the target response will and will not be performed (Baer, 1981; Becker et al., 1975).

Once the stimulus class is defined, the teacher or behavior manager should examine variations in the target responses that are required for responding across the full range of stimulus conditions represented by the stimulus class. This is of major importance for objectives related to motor skills (i.e., dressing, vocational behaviors, eating). Appropriate selection of teaching examples requires information about both the situations in which the target response should occur and the different ways the response should be performed (Horner et al., 1982).

### Selecting Teaching Examples

Existing literature documents the importance of selecting appropriate teaching examples to achieve generalized responding (Horner & McDonald, 1982; Hupp & Mervis, 1981; Sprague & Horner, 1984). While there continues to be an array of unanswered questions, the central message is clear: Generalized responding is taught by presenting a set of teaching examples that sample the range of relevant stimulus and response variation represented in the target stimulus class.

Baer (1981) conveys this message by recommending, "Think through the structure of the generalized behavior change that you want to produce, to see what its components are. Be sure to represent those components in the list of examples, and try to maximize their representation as early as seems practical" (pp. 20–21). Engelmann and Carnine (1982) provide a similar message and add a recommendation to select negative examples that teach the learner when *not* to respond.

If the learner is to respond correctly in new situations, s/he must be under control of relevant stimuli that are similar across trained and nontrained situations. To ensure that the learner attends to, and is under control of, the relevant stimuli, a set of examples should be selected that are the same only in that they contain these relevant stimuli. Irrelevant stimuli should be part of each example but different irrelevant stimuli should be associated with different examples and no irrelevant stimulus should be contained in all examples. This will both result in the learner coming under stimulus control of the correct stimuli, and avoid the error pattern of irrelevant stimuli controlling the target response.

In addition, the set of examples should teach the learner just how much variation in the target response is acceptable. The set of teaching examples should sample the range of acceptable response variation, as well as stimulus variation. In most applied situations a set of teaching examples that meet these criteria can be obtained from the natural environment. With some functional skills, however, the selection of an appropriate range of teaching examples may be difficult. When this occurs the necessary variation may be achieved via adding *simulation* teaching examples to a restricted group of natural setting teaching examples (c.f., Horner, McDonnell, & Bellamy, 1984). The critical variable is presentation of teaching examples that sample the full range of stimulus and response variation the learner will encounter after instruction. Following these procedures will avoid the development of errors due to both restricted stimulus control and limited response variations.

### Specifying Instructional Procedures

Procedures for delivering instruction are a strength of our current technology, and are well detailed in a variety of texts (Becker, Engelmann, & Thomas, 1975; Sailor & Guess, 1983; Snell, 1983). The present analysis of generalized responding is consistent with the vast majority of these existing teaching procedures. Only two additions are suggested from the analysis of error patterns: present multiple training examples within individual sessions; and use a conservative criterion for terminating training. Individual teaching sessions should be characterized by multiple trials with different examples. It is tempting to use a single, simple example to promote rapid learner success. While this approach leads to success in training, it can easily result in major error patterns during generalization trials. The most common errors for this approach are related to irrelevant stimuli controlling the target response (Horner & McDonald, 1982) and use of a limited range of response variations (Colvin & Horner, 1983).

The second suggestion is to use a stringent criterion for terminating instruction. Far too little attention has been given to the relationship between criteria for terminating training and the generalization and maintenance of responding thereafter (Liberty, Haring, & Martin, 1981). Strong stimulus control relationships, however, are more likely to avoid error patterns in which irrelevant stimuli control irrelevant responses.

### Analyzing the Performance Environment

The fourth intervention area in need of review is the prior analysis of performance environments. Errors related to "irrelevant stimuli controlling irrelevant responses" point to the need for teachers to have a clear understanding of the controlling variables in the performance setting prior to intervention. A pretraining analysis should identify the relevant stimuli within performance settings that need to control responding, and irrelevant stimuli that have a high potential for developing inappropriate stimulus control. If a student is already performing inappropriately across performance settings, the analysis should identify the inappropriate response and the stimuli controlling it. An analysis of this type should lead to the selection of a range of training examples, some of which include the identified irrelevant stimuli. This approach would decrease the control by irrelevant stimuli in nontrained situations.

## Implications of the Error Analysis for Applied Research on Generalization

The pattern of errors exhibited by learners in nontrained situations has a number of implications for research. Specific recommendations are presented for the definition of research questions, construction of dependent variables, description of stimulus conditions, documentation of experimental control, and emphasis on social validity.

### Definition of Research Questions

Research questions focused on generalized responding will be clarified by a language system that allows discrimination of the stimulus control variables under analysis. This requires improved precision in the definition of research variables and a re-evaluation of how the terms *stimulus generalization*, *transfer*, *maintenance*, and *stimulus control* are used in describing performance under nontrained conditions.

Behavior does not generalize, stimulus control generalizes. Stimulus generalization refers to the control exerted by a previously neutral stimulus over a target response as a function of bringing that response under control of a different stimulus (Pavlov, 1927; Terrace, 1966). If, prior to training, neither Stimulus$_1$ ($S_1$) nor Stimulus$_2$ ($S_2$) control a target response, and if after training with $S_1$, $S_2$ also controls the response, the control developed with $S_1$ has generalized to $S_2$. The response did not generalize to $S_2$; rather the stimulus control exerted by $S_1$ generalized to $S_2$. The important characteristic of this process for applied researchers is recognition that it is stimulus control that has generalized. Generalization does not refer to the simple occurrence of a response in a novel setting, but the control of that response by nontrained stimuli. This definition of stimulus generalization is consistent with that used by Skinner (1953) and early operant researchers (Guttman & Kalish, 1956; Jenkins, 1965). It emphasizes the simultaneous development of stimulus control by trained and nontrained stimuli as a function of the similarity of those stimuli along relevant characteristics. The definition has significant advantages. It has been used for decades by operant researchers, and therefore brings a precise language system and a rich body of analysis to the applied problem. It is consistent with stimulus control research in applied settings, and it has direct implications for the development of effective instructional procedures.

This definition of stimulus generalization also carries two significant constraints for applied researchers. The first is the difficulty of identifying and measuring those stimuli that control responding in applied settings. Where laboratory researchers have focused on simple responses (i.e., key pecks) controlled by single characteristics (i.e., color) of individual stimuli (i.e., disc), the applied researcher is faced with a "behavior stream" (Schoenfeld & Farmer, 1970) that involves complex responses (i.e., tantrumming) performed in stimulus environments that make identification of the precise discriminative stimulus (much less the controlling stimulus characteristics) extremely difficult. It is not surprising, therefore, to find that applied researchers have redefined generalization by focusing on performance in nontrained conditions rather than control by nontrained stimuli (Stokes & Baer, 1977).

The second constraint is that the restricted definition of generalization addresses only a part of the applied problem. Stimulus generalization does not necessarily address those situations where training is conducted in one room with a training stimulus and testing occurs with that same stimulus in a different room. Correct performance in the second room may not be controlled by novel stimuli, but by the same stimuli used in training. Hence it may not be an example of stimulus generalization but simply an example of stimulus control (Rincover & Koegel, 1975). Similarly, responding that continues to be under the same stimulus control over many trials, or over an extended passage of time (maintenance), is not an example of control by a novel stimulus (i.e., not an example of stimulus generalization), yet it is a part of the applied problem. If the applied problem is addressed from a stimulus control perspective, then the more narrow definition of stimulus generalization must be accompanied by a recognition that the applied problem represents several behavioral processes, and will require a response that involves manipulation of several independent variables.

The fact that generalization errors form specific patterns suggests a need for greater precision in the way research questions are defined, and the way responding in nontrained situations is described. Toward this end, more precise analysis is needed of the multiple processes affecting responding in nontrained situations. Transfer and maintenance, like generalization, refer to stimulus control, not behavior. Behavior does not transfer or maintain, rather the stimulus-response relationship (i.e., stimulus control) transfers or maintains. The stimulus-response *relationships*, therefore, should serve as the focus of intervention strategies and research analysis.

**Selection of Dependent Variable Measures**

Applied interventions are typically designed to produce behavior change across a class of appropriate stimulus conditions. Consequently, applied research focused on generalization should use performance across a class of appropriate stimulus conditions as a dependent variable. The measurement of such a dependent variable should document stimulus control by relevant stimuli, and allow assessment of the four generalization error patterns.

It is tempting to measure generalization by showing improved performance in one new setting or with a small set of randomly selected nontrained stimuli. Demonstrations of this type, however, do not define the range of controlling stimuli in the nontrained conditions, do not exclude the possibility that irrelevant stimuli are controlling the response, and as such may lead to inappropriate conclusions (Perkins, 1965). If, however, the dependent variable is measured by performance across a systematically selected range of stimuli, a mistaken analysis is unlikely.

One procedure for measuring responding under nontrained stimulus conditions is to begin by defining the stimulus class across which responding is desired, and building a dependent measure through systematic selection of nontrained examples from this class. *Systematic* selection involves defining a set of examples that: (a) sample the range of relevant stimulus variation within the stimulus class; (b) control presentation of irrelevant stimuli across examples; (c) sample the range of response variations (topographical demands) required by the stimulus class; and (d) include negative examples (stimulus conditions that should not control responding) that are maximally similar to positive examples. Several examples of studies that have used this approach now exist (Anderson & Spradlin, 1980; Horner & McDonald, 1982; Horner, Williams, & Steveley, 1984; McDonnell et al., 1984; Sprague & Horner, 1984).

**Description of Stimulus Conditions**

An important methodological consideration for generalization research is the need for precise descriptions of independent variables (Johnston & Pennypacker, 1980; Peterson, Homer, & Wonderlich, 1982). Interpretation of generalization research requires detailed documentation of the stimuli used during training as well as those procedures used to measure the dependent variable (Anderson & Spradlin, 1980; Hupp & Mervis, 1981). The generalization error patterns discussed earlier index multiple relationships between the stimuli used during training and stimulus control in nontrained situations. Consequently, any study attempting to document generalized responding bears a burden of defining the similarities and differences between training and nontraining situations. This is a difficult but necessary task for applied researchers who manipulate variables in complex stimulus settings.

**Documentation of Experimental Control**

Generalization is a consistent topic of analysis in applied research (Hayes, Rincover, & Solnick, 1980). Yet there are few examples in which experimental control of generalization is documented (Kendall, 1981; Robinson & Swanton, 1980; Rusch & Kazdin, 1981). This may be due to the way dependent variables have been defined, or, as Kendall (1981) suggests, to inherent constraints in single-subject designs. Regardless of the reason, practitioners and researchers need demonstrations of functional relationships between independent variables and generalized responding. The use of dependent variables that assess performance across nontrained stimulus conditions plus exploration of creative research design options should lead to group designs and single-subject designs that document experimental control over generalized responding.

**Documentation of Social Validity**

Applied research addressing generalized responding bears a particular burden of documenting social as well as experimental validity (Kazdin, 1977; Voeltz & Evans, 1983). The purpose of teaching generalized responses in applied settings is so the learner will be able to use the acquired behavior(s) across the range of situations that arise as part of his/her daily routine. While experimental analysis of the multiple situations and behavioral variations that occur across a day may be unwieldy, less formal documentation that the behavior pattern under analysis is functional for the learner can occur. Subjective measures (Koegel, 1983) and less controlled documentation that the learner is applying the targeted behavior (Wolf, 1978) should be considered as supplements to more experimentally rigorous measurement procedures.

## Summary

The experimental and applied analyses of behavior have served us well in documenting procedures for behavior change. For this technology to meet today's applied challenges, however, increased attention must be focused on the extent to which behavior change extends to the range of stimulus conditions characteristic of natural environments. Documentation and analysis of various techniques that successfully produce generalized responding have proven useful. A logical extension of this analysis is available by focusing on the generalization error patterns that learners perform after training is completed. A taxonomy of these error patterns has been presented that examines the applied problem of generalization within a stimulus control logic. The results of this examination have implications for instructional procedures and the research methodology used to analyze responding in the presence of nontrained stimuli.

## References

Anderson, S. R., & Spradlin, J. E. (1980). The generalized effects of productive labeling training involving common object classes. *Journal of the Association for the Severely Handicapped, 5*(2), 143–157.

Baer, D. M. (1981). *How to plan for generalization.* Lawrence, KS: H & H Enterprises.

Becker, W., Engelmann, S., & Thomas, D. (1975). *Teaching 2: Cognitive learning and instruction.* Chicago: Science Research Associates.

Berler, E. S., Gross, A. M., & Drabman, R. S. (1982). Social skills training with children: Proceed with caution. *Journal of Applied Behavior Analysis, 15,* 41–54.

Carnine, D. W., & Becker, W. C. (1982). Theory of instruction: Generalization issues. *Educational Psychology, 2,* 249-262.

Carr, E. G., & Kologinsky, E. (1983). Acquisition of sign language by autistic children II: Spontaneity and generalization effects. *Journal of Applied Behavior Analysis, 16,* 297–314.

Colvin, G. T., & Horner, R. H. (1983). Experimental analysis of generalization: An evaluation of a general case program for teaching motor skills to severely handicapped learners. In D. Hogg & P. Mittler (Eds.), *Advances in mental handicap research: Volume 2. Aspects of competence in severely mentally handicapped people* (pp. 309–345). Chichester, England: John Wiley.

Coon, M. E., Vogelsberg, R. T., & Williams, W. (1981). Effects of classroom public transportation instruction on generalization to the natural environment. *Journal of the Association for the Severely Handicapped, 6*(2), 46–53.

Drabman, R. S., Hammer, D., & Rosenbaum, M. S. (1979). Assessing generalization in behavior modification with children: The generalization map. *Behavioral Assessment, 1,* 203–219.

Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and applications.* New York: Irvington Publishers.

Garcia, E., Baer, D. M., & Firestone, I. (1971). The development of generalized imitation within topographically determined boundaries. *Journal of Applied Behavior Analysis, 4,* 101–112.

Guess, D., Keogh, W., & Sailor, W. (1978). Generalization of speech and language behavior: Measurement and training tactics. In R. Schiefelbusch (Ed.), *Bases of language intervention* (Vol. I, pp. 373–397). Baltimore: University Park Press.

Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology, 51,* 79–88.

Hayes, S. C., Rincover, A., & Solnick, (1980). The technical drift of applied behavior analysis. *Journal of Applied Behavior Analysis, 13,* 275–285.

Honig, W. K., & Urcuioli, P. J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior, 36*(3), 405–445.

Horner, R. H., & Budd, C. M. (in press). Teaching manual sign language to a nonverbal student: Generalization of sign use and collateral reduction of maladaptive behavior. *Education and Training of the Mentally Retarded.*

Horner, R. H., & McDonald, R. S. (1982). Comparison of single instance and general case instruction in teaching a generalized vocational skill. *Journal of the Association for the Severely Handicapped, 7*(3), 7–20.

Horner, R. H., McDonnell, J. J., & Bellamy, G. T. (1984). *Teaching generalized behaviors: General case instruction in simulation and community settings.* Unpublished manuscript, University of Oregon.

Horner, R. H., Sprague, J., & Wilcox, B. (1982). Constructing general case programs for community activities. In B. Wilcox & G. T. Bellamy, *Design of high school programs for severely handicapped students* (pp. 61–98). Baltimore: Paul H. Brookes.

Horner, R. H., Williams, J. A., & Steveley, J. D. (1984). *Acquisition of generalized telephone use by students with severe mental retardation.* Manuscript submitted for publication.

Hupp, S. C., & Mervis, C. B. (1981). Development of generalized concepts by severely handicapped students. *Journal of the Association for the Severely Handicapped, 6*(1), 14–21.

Jenkins, H. M. (1965). Generalization gradients and the concept of inhibition. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 55–61). Stanford: Stanford University Press.

Johnston, J. M., & Pennypacker, H. S. (1980). *Strategies and tactics of human behavioral research.* New Jersey: Lawrence Erlbaum Associates.

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification, 1,* 427–452.

Kendall, P. C. (1981). Assessing generalization and the single-subject strategies. *Behavior Modification, 5*, 307-319.

Koegel, R. L. (1983). *Pivotal response related to generalization.* Paper presented at the American Psychological Association Conference, Anaheim, CA.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

Liberty, K. A., Haring, N. G., & Martin, M. M. (1981). Teaching new skills to the severely handicapped. *Journal of the Association for the Severely Handicapped, 6*(1), 5-13.

McDonnell, J. J., Horner, R. H., & Williams, J. A. (1984). Comparison of three strategies for teaching generalized grocery purchasing to high school students with severe handicaps. *Journal of the Association for Persons with Severe Handicaps, 9*(2), 123-133.

O'Connell, J. M. (1981). *Acquisition of functional and nonfunctional response topographies by moderately and severely handicapped students performing a vocational task.* Unpublished masters thesis, Division of Special Education and Rehabilitation, University of Oregon.

Patterson, G. R. (1974). A basis for identifying stimuli which control behaviors in natural settings. *Child Development, 45*, 900-911.

Pavlov, I. P. (1927). *Conditioned reflexes.* Translation by G. V. Anrep, London: Oxford University Press.

Perkins, C. C. (1965). A conceptual scheme for studies of stimulus generalization. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 38-54). Stanford: Stanford University Press.

Peterson, L., Homer, A. L., & Wonderlich, S. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis, 15*, 477-492.

Rincover, A., & Koegel, R. (1975). Setting generality and stimulus control in autistic children. *Journal of Applied Behavior Analysis, 8*, 235-246.

Robinson, V., & Swanton, C. (1980). The generalization of behavioral teacher training. *Review of Educational Research, 50*, 486-498.

Rusch, F. R., & Kazdin, A. E. (1981). Toward a methodology of withdrawal design for the assessment of response maintenance. *Journal of Applied Behavior Analysis, 14*, 131-140.

Sailor, W., & Guess, D. (1983). *Severely handicapped students: An instructional design.* Boston: Houghton Mifflin.

Sanders, M. R., & James, J. E. (1983). The modification of parent behavior: A review of generalization and maintenance. In M. Hersen & A. S. Bellack (Eds.), *Behavior Modification* (Vol. 7, No. 1). Beverly Hills: Sage.

Schoenfeld, W. N., & Farmer, J. (1970). Reinforcement schedules and the "Behavior Stream." In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules* (pp. 215-247). New York: Appleton-Century-Crofts.

Sidman, M., & Stoddard, L. T. (1967). The effectiveness of fading in programming a simultaneous form discrimination for retarded children. *Journal of the Experimental Analysis of Behavior, 10*, 3-15.

Sidman, M., & Willson-Morris, M. (1974). Testing for reading comprehension: A brief report on stimulus control. *Journal of Applied Behavior Analysis, 7*, 327-332.

Skinner, B. F. (1953). *Science and human behavior.* New York: Macmillan.

Snell, M. E. (Ed.). (1983). *Systematic instruction of the moderately and severely handicapped* (2nd ed.). Columbus, OH: Merrill.

Sprague, J. R., & Horner, R. H. (1984). The effects of single instance, multiple instance and general case training on generalized vending machine use by moderately and severely handicapped students. *Journal of Applied Behavior Analysis, 17*, 273-278.

Stella, M. E., & Etzel, B. (1983). Effects of criterion-level probing on demonstrating newly acquired discriminative behavior. *Journal of the Experimental Analysis of Behavior, 39*, 479-498.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Terrace, H. (1966). Stimulus control. In W. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 271-344). New York: Appleton-Century-Crofts.

Voeltz, L. M., & Evans, I. M. (1983). Educational validity: Procedures to evaluate outcomes in programs for severely handicapped learners. *Journal of the Association for the Severely Handicapped, 8*(1), 3-15.

Wolf, M. M. (1978). Social validity: The case for subjective measurement, or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203-214.

## Author Notes

The authors express thanks to Shawn M. Boles, Julie A. Williams and Bruce Wald for their comments on early drafts of this paper.

The activity which is the subject of this report was supported in whole or in part by the U.S. Department of Education, Contract 300-82-0362. However, the opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement by the Department should be inferred.