Psychology in the Schools, Vol. 44(7), 2007
Published online in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/pits.20258

ERRORLESS LEARNING: REVIEW AND PRACTICAL APPLICATION FOR TEACHING CHILDREN WITH PERVASIVE DEVELOPMENTAL DISORDERS

MICHAEL M. MUELLER AND CHRISTINE M. PALKOVIC

Southern Behavioral Group

CYNTHIA S. MAYNARD

Georgia School of Professional Psychology at Argosy University

Errorless learning refers to a variety of discrimination learning techniques that eliminate or minimize responding to incorrect choices. This article describes experimental roots of errorless learning and applied errorless strategies. Specifically, previous research on stimulus fading, stimulus shaping, response prevention, delayed prompting, superimposition with stimulus fading, and superimposition with stimulus shaping are discussed. Educational applications for children with Pervasive Developmental Disorders (PDD) are presented for each technique so that school psychologists, educators, and teachers working with children with PDD can understand the underpinnings and practical applications of errorless techniques to use in skill acquisition programming in school settings. © 2007 Wiley Periodicals, Inc.

Errorless learning represents a set of teaching procedures designed to reduce incorrect responding as the student gains mastery over the work materials. Errorless learning procedures are used to teach children a variety of novel tasks that almost always involve choosing one response over a second response. Examples of common tasks include teaching a child to choose sight words, colors, numbers, shapes, pictures, or objects when given multiple stimuli from which to choose. Errorless learning procedures are derived from an experimental learning literature. As such, a discussion of this early literature is crucial to understanding the modern application to teaching children without errors.

Trial-and-Error and Errorless Learning Theory

People learn simple and conditional discriminations in various ways. Through the early 1960s, published discrimination learning tasks shared a common feature. That is, organisms learned through "trial and error." Trial-and-error learning has been defined differently through the years, but each definition relates to the same basic arrangement. In trial-and-error discrimination tasks, the correct choice (S+) and the incorrect choice (S−) are always presented together so that the arrangement of important stimulus materials is presented to the learner from the beginning and does not change throughout the learning task. For example, if a task is to teach a child to touch a red ball when presented with a red ball and a black block after hearing the verbal demand to "touch ball," the red ball and the black block are always paired together and the learner maintains the opportunity to select either one. Neither is larger or smaller, neither is farther or nearer to the learner, and no responses to either stimuli are blocked. The learner has equal opportunity to select the correct (S+) or the incorrect (S−) choice.

It was once believed that discrimination learning *required* the selection of both the correct (S+) and incorrect (S−) so the learner could experience the consequences of each (Hull, 1950; Spence, 1936). Because differential consequences are programmed for either selection, learning can occur from correct or incorrect responses. In the above example of a conditional discrimination task where a child can choose a red ball or a black block following the teacher's directive to "touch ball," a touch to the red ball would be followed by reinforcement. A touch to the black

Correspondence to: Michael M. Mueller, Ph.D., BCBA, Southern Behavioral Group, 1950 Spectrum Circle, Suite 400, Marietta, GA 30067. E-mail: mmueller@southernbahavioral.com

block would be followed by no programmed consequence (i.e., extinction; Sidman, Cresson, & Willson-Morris, 1974) or punitive or corrective consequences such as learning trials, aversive sounds (Mueller, Olmi, & Saunders, 2000), response cost (Pilgrim & Galizio, 1990), or time-outs/blackouts (Stromer, Mackay, McVay, & Fowler, 1998). The selection of the correct choice (S+) should continue in future trials because of reinforcer delivery. The selection of the incorrect choice (S−) should decrease because of the extinction schedule, aversive, or corrective consequences presented. Between the two selections and their associated consequences, children can learn which selection does and does not lead to reinforcement and persist with the response that does lead to reinforcement in the presence of its conditional and discriminative stimuli. Saunders and Williams (1998) described participants learning from either selection (i.e., selecting the correct choice [S+] and continuing to select the correct choice [S+] or selecting the incorrect choice [S−] and changing their response pattern to the correct choice [S+]) as "win–stay, lose–shift" (p. 207).

Trial-and-error learning necessarily creates an opportunity for "incorrect" or "error" selection during the learning trials as the learner experiences which selections lead to which consequences. More than 40 years of experimental research suggests that the selection of errors with some children can have several very serious and sometimes very dangerous side effects, including extinction induced aggression (Azrin, Hutchinson, & Hake, 1966; Kelly & Hake, 1970; Lerman, Iwata, & Wallace, 1999; Rilling & Caplan, 1973), different topographies of problem behavior (Ducharme, 2003), negative emotional responses (Weeks & Gaylord-Ross, 1981), and stimulus overselectivity (Lovaas, Koegel, & Schreibman, 1979). A type of discrimination learning that decreases the opportunity for incorrect choice (S−) selection and therefore reduces the negative effects of making incorrect choices is referred to as "errorless" learning. If error responses in some children can have untoward effects, this might be exacerbated in children with Pervasive Developmental Disorders (PDD) given their rigid adherence to rules, difficulty shifting behavioral repertoires, and learning superstitious behavioral chains (Green, 1996; Smith, 2001; Smith, Iwata, Goh, & Shore, 1995). If error production becomes part of the learned sequence in a novel task (e.g., touching the incorrect choice before shifting to the correct choice), the child might be more prone to experience the negative effects because the child could persist with a pattern of behavior that produces more errors.

Errorless learning is a bit of a misnomer. In fact, rarely are learning situations completely free of the opportunity to make errors. Therefore, errorless learning is not always totally error*less*, but, far more often than not, rather error *reducing*. In an evaluative paper of errorless procedures used to teach new skills to people with aphasic disorders (e.g., anomia), Fillingham, Hodgson, Sage, and Ralph (2003) defined errorless learning as "an approach whereby the task is manipulated to eliminate/reduce errors. Tasks are executed in such a way that the subject is unlikely to make errors" (p. 339). Even in this recent definition, the concept of error elimination is still included, because, as will be explained further, one typical characteristic of errorless learning procedures is the gradual increase of difficulty in the task that necessarily correlates with an increasing opportunity to make errors. The first published errorless learning description provides a wonderful example of this progression. There are six different general techniques that have been described in the literature, including stimulus fading, stimulus shaping, response prevention, delayed prompting, and superimposition with stimulus fading, and superimposition with stimulus shaping. The earliest published errorless procedure, and the procedure most commonly used in teaching children with PDD, is stimulus fading.

Research Review of Errorless Learning Techniques

Terrace (1963a), wrote the seminal research article of what we now call errorless learning in a classic demonstration of stimulus fading. Prior to the publication of Terrace's research, discrimination tasks were always conducted as trial and error, and the prevailing knowledge at the time was that selections to the correct choice (S+) and the S− (errors) were necessary components of discrimination learning (Hull, 1950; Spence, 1936). Terrace trained pigeons in a successive discrimination task in which key pecks to a red light (S+) were reinforced and those to a green light (S−) were not. Rather than presenting both the red and green lights together at their terminal intensity (so that errors could have been made by selecting the green light), stimulus control by the red light was established without the green light present. That is, the red light was presented and key pecks in its presence were reinforced. After consistent responding was established and key pecks were under discriminative control of the red light, the green light was very gradually faded both in duration and in the light's intensity beginning with a very brief and very dim presentation and ending with the duration and the intensity of both lights being identical. When the duration and intensity of the green light were presented in their terminal form, the pigeons did not peck the key in the presence of the green light, demonstrating for the first time discriminative control from an errorless learning training sequence. Initially, no errors were possible as control by the red light was established in the absence of the presentation of the green light. Gradually, the incorrect choice (S−) was added, creating an opportunity for errors to be made. This procedure, like so many others, was initially error eliminating and gradually moved to error reducing as the opportunity to respond to an S− increased over time.

Using identical or similar experimental arrangements for stimulus fading, other researchers used errorless procedures to train discriminations across multiple dimensions (e.g., shape and line orientation; Terrace, 1963b), reversing the function of the correct (S+) and incorrect (S−) choices (Sidman & Stoddard, 1966), and across a variety of organisms. Researchers next used similar procedures with human participants in experimental contexts and found similar outcomes with children of typical developmental trajectories (Everitt, 1977), those with mental retardation, (Duker, 1981), and those with autism (Schreibman, 1975). For example, in 1964, Moore and Goldiamond presented children with trial-and-error conditional discrimination tasks in which the sample was a triangle and the choices were an identical triangle (S+) and one rotated slightly in one direction or the other (S−). The children were unable to learn the discrimination through trial and error. Next, only the identical triangle (S+) was presented. Responding was established without the presence of the S− (rotated triangle) and then the very dim S− was gradually faded in by small increases in brightness until both choices were again presented at full intensity. All children in the study learned the discriminations under the stimulus fading procedure. In an interesting experimental variation, Moore and Goldiamond, at different points in the stimulus fading program for each child, reversed back to the baseline arrangement in which both choices were at full intensity. For all children, correct responding decreased until the stimulus fading program was put back in place. Beyond the simple demonstration of reversibility in children, they also demonstrated that for these children, the errors served as extinction trials and prolonged the number of trials required to make improvement on the simple matching task. The prolongation of the learning sequence is of obvious importance to those working with children with PDD due to the time constraints often imposed on early intervention instruction due to public school and parental financial considerations. The faster learning can occur, the more benefit it will have to those who provide the training to children with PDD.

The seminal work of researchers conducting errorless learning studies used stimulus fading techniques. Most early studies were basic experimental procedures aimed at understanding the results of errorless learning for its own sake and involved stimulus fading of an incorrect choice (S−) in one form or another (i.e., intensity, duration, etc.). Stimulus fading has also been successfully demonstrated with the dimension of size (Duffy & Wishart, 1987). Most fading methods share similar features. Each establishes responding to the correct choice (S+) in the absence of the incorrect choice (S−). Each very gradually presents a differing incorrect choice (S−) ending such that the correct (S+) and incorrect (S−) choices are eventually presented at the same size, shape, intensity, color, and so forth. Each also has used procedures whereby errors (although rare) result in a previous fading step presented on the next trial. It should be noted that each study described thus far relied on computerized or electronic experimental presentation and manipulanda such as light boards, light panels, microswitches, levers, or touch-sensitive screen technology.

Errorless procedures other than fading some dimension of the stimuli have been used with similarly successful results. Although stimulus fading is the most common, five other techniques have empirical support in experimental and applied settings. Another errorless learning procedure is stimulus shaping. Stimulus shaping initially presents different stimuli as choices (e.g., shapes, letters, colors, etc.) that are gradually changed over successive trials so that the terminal choices are physically different than their initial presentation (Lancioni & Smeets, 1986; Schilmoeller, Schilmoeller, Etzel, & LeBlanc, 1979). The systematic change of the physical properties of both the S+ and S− stimuli is the hallmark of a stimulus shaping procedure. Stimulus shaping has been used successfully to teach discriminations on computer and table-top arrangements. For example, Sidman and Stoddard (1966) initially taught children to select a circle when presented with a circle and an ellipse. Then, by making gradual physical changes over 43 different steps, the circle was transformed to a rectangle and the ellipses were shaped into circles. The circle that was initially the correct choice (S+) became a rectangle S+ and the ellipses that were initially the incorrect choices (S−s) were changed into circle S−s. This shaping procedure changed the circle from what was first a correct choice (S+) to an incorrect choice (S−) with no errors. Whereas stimulus fading makes gradual changes to some property of an incorrect (S−) stimulus, stimulus shaping makes complete physical transformation of both choices so that the original choices in a discrimination task may look nothing like the terminal choices of the discrimination task. Importantly, the initial choices should be those that are known to the child, so that correct responding begins immediately and continues as small changes shape the known stimuli to the terminal (initially unknown) choices.

Another method of errorless learning is called response prevention or graded choice. Storm and Robinson (1973) taught children errorless discriminations of colors while the terminal choices (i.e., the correct [S+] and the incorrect [S−] choices) were always present in their final forms. The required response was a bar press made under a colored shape illuminated on a flat panel in front of the child. During the errorless procedure, both colored shapes were presented. However, only the bar under the correct choice (S+) was initially available. So, although both visual choice stimuli were present on all trials, responding to the correct choice (S+) was established before responses could be made to the incorrect choice (S−). When responding to the correct choice (S+) was established and the bar under the incorrect choice (S−) was introduced, responding to the correct choice (S+) persisted with no, or very few, errors.

Another errorless method is delayed prompting. Delayed prompting is referred to differently by researchers, and variations of the procedure have been labeled constant prompting, delayed cue, delayed matching, or as time delay procedures (Alig-Cybriwsky, Wolery, & Gast, 1990; Coleman-Martin & Heller, 2004; Gibson & Schuster, 1992; Johnson, 1977; Keel & Gast, 1992; Koscinski & Gast, 1993; Mattingly & Bott, 1990; Venn et al., 1993). The basic application of delayed prompting (regardless of the procedural name) is the same, although there are, of course, slight procedural variations. Delayed prompting usually involves initially providing an immediate pointing, or other physical prompt, when stimuli are presented as choices. Next, there is a systematic increase in time from the presentation of choices to the onset of the prompt over successive trials. For example, in the initial delayed prompting procedure (called Delayed Cue by Touchette, 1971), children were presented with picture choices following a verbal demand. The experimenter pointed to the correct choice immediately upon choice presentation. Responding was accurate and the presentation of the choice stimuli gradually and increasingly preceded the presentation of the point prompt in a systematic manner. Finally, the children were selecting the correct choice before the prompt was delivered, or, essentially, without the prompt. Delayed prompting has been used to teach children individually (Johnson) or in group settings (Keel & Gast).

The final two errorless learning strategies are used less often and include superimposition with stimulus fading and superimposition with stimulus shaping (Lancioni & Smeets, 1986). Neither procedure makes physical changes to the choice stimuli such as typical stimulus fading or stimulus shaping procedures do. Instead, superimposed prompts are added and the changes (fading or shaping) that occur over time are to the superimposed prompts instead of the choice stimuli (Lancioni & Smeets; Terrace, 1963b). Both superimposition techniques combine a previously described technique (i.e., stimulus fading or stimulus shaping) with the superimposition of a cue that is later faded. For example, if a child is presented with two stimuli between which she cannot discriminate, either superimposition procedure would first superimpose two known stimuli over the unknown stimuli. After correct responding begins in the presence of the unknown stimuli, gradual changes through fading or shaping of the known (superimposed) stimuli are made so that the child eventually responds correctly to the initially unknown stimuli. An example would help illustrate the superimposition procedures.

In 1983, Schimek taught an 8-year-old to read printed words, write words, and to make auditory discriminations of words containing three digraphs (*th*, *ch*, and *sh*). The superimposition with stimulus fading technique that was used was unique in its table-top (vs. computerized) presentation and in its school setting. The printed word discrimination involved superimposing pictures of known objects that began with the same letters of the digraph printed on a card in front of the child. A large picture of a thumb was included with the *th* in the initial step of teaching the *th* digraph. Next, a hand with a regularly sized thumb was present on the card, and finally the card only contained the digraph *th*. Across all discrimination goals, the child eventually responded with 100% after 15 sessions. The physical properties of the digraph remained unchanged throughout the procedure. However, the size and shape of the superimposed thumb prompt included on the card was changed.

Using one or more of the methods described above, errorless learning enjoys demonstrated effectiveness across a wide variety of participants in numerous studies to improve a wide variety of goals. For example, errorless procedures have helped typical preschoolers (Venn et al., 1993), people with schizophrenia (Kern, Liberman, Kopelowicz, Mintz, & Green, 2002), Karsokoff's Syndrome (Komatsu, Mimura, Kato, Wakamatsu, & Kashima, 2000), memory impairments (Fillingham et al., 2003), brain injury (Wilson & Manly, 2003), Down Syndrome (Duffy & Wishart, 1987), mental retardation (Smeets, Lancioni, Striefel, & Willemsen, 1984), geriatric populations (Kessels & DeHaan, 2003), Herpes Simplex Encephalitis (Parkin, Hunkin, & Squires, 1998), Alzheimer's Disease (Clare, Wilson, Carter, Roth, & Hodges, 2002), and of course, children and adults with PDD (Schreibman, 1975).

Application of Errorless Learning in School Settings

Each of the six errorless learning procedural variations discussed above can be used to teach new discriminations to children with PDD in school settings. Children with PDD share common characteristics in learning that make errorless procedures sound choices for procedure selection. That is, rigid adherence to routines and response overselectivity and overgeneralization combined with problematic behaviors in response to failure or novel tasks can create situations in which limiting the incorrect response in a learning task is ideal (Green, 1996; Smith, 2001; Smith et al., 1995).

Errorless learning procedures can be used successfully without computerized assistance in a variety of formats and settings and across a very wide range of skills. In practice, it is very likely that more errorless procedures occur without computerized assistance for practical purposes when teaching children with PDD in schools and homes with discrete trial instruction.

Table 1 provides procedural names, descriptions, and examples of each of the six techniques discussed thus far. Below, practical examples and suggestions for each technique's usage are elaborated in more practical detail.

Cipani and Madigan (1986) provided a great application of stimulus fading (i.e., position fading) in a classroom setting to teach coin discrimination to children. First, a child was presented with a penny as the teacher said, "Touch penny." After the child was reliably responding to the penny in the presence of the verbal discriminative stimulus, another coin was added to the choice presentation but at an initially distant position. As trials continued, the S− coin was moved closer and closer to the penny until both were equidistant and randomly presented on either side of the sample. Cipani and Madigan then suggested using the same procedures with a dime, and as the number of coin discriminations increases, interspersing the coins as correct choices (S+s) so the child is eventually told to choose any coin from a group of different coins. As stimulus fading is most likely the easiest errorless learning technique to use, it can be applied most rapidly to current early intervention procedures such as discrete trial training in which trial-and-error techniques are being used. By beginning with only one choice in a wide variety of match tasks, a child can begin to respond to the only choice (i.e., the S+) available. By then slowly incorporating additional choices and moving them closer and closer to the correct choice, the child should continue to select the correct choice even in the presence of one of more distracters (S−s).

A practical table-top application of the stimulus shaping procedure to teach a child to discriminate letters would first involve the identification of two objects that are already discriminable to the child. Next, draw known objects onto a note card and begin the conditional discrimination task. Over successive trials, the known objects should be redrawn on new note cards so that they begin to change shape slowly to appear more and more like the terminal stimuli. The slow changes

Table 1
Procedural Names, Descriptions, and Examples for Six Errorless Learning Techniques

Procedure	Definition	Example
Stimulus Fading	Gradually increasing the dimension of the S−	Gradually increase the font of the S− until it is the same as the S+
Stimulus Shaping	Making physical changes to the S+ and S− over trials	Gradually change known letters into unknown letters over successive trials by changing their shape
Delayed Prompting	Gradually delay the onset of a prompt that identifies the S+	Provide initially immediate indication of the S+ and then gradually delay indication
Response Prevention	Physically preventing the learner from responding to the S−	Physically block responses to the S− until the learner responds independently to the S+
Superimposition with Fading	Superimposing physical prompts and using stimulus fading	Add pictures to accompany sight words cards and then gradually reduce the size until the pictures are no longer visible
Superimposition with Shaping	Superimposing physical prompts and using stimulus shaping	Teach a child to respond to known pictures in the presence of unknown sight words. Change the pictures gradually into the pictures of the unknown words

should allow the child to make the new discriminations as the stimuli change shape. The goal is to slowly transform pictures of known objects into pictures of the initially unknown objects. Although this technique may require a teacher to draw on his or her artistic skills, the child should ultimately demonstrate the novel discrimination if the steps gradually change a known object or item picture to an unknown object or item or picture.

Errorless learning using response prevention can very easily be used in many basic learning tasks in which physical choices are presented in front of a child and the child is asked to point, move, stack, combine line up, or touch in some way. For example, present a correct (S+) and incorrect (S−) stimulus to a child and tell the child to select the correct choice (S+; e.g., touch the boat). If the child begins to point to the incorrect choice (S−; e.g., a car), physically block the response and guide the child's hand to the correct choice (S+). Here, both choice stimuli are present from the onset of the task, but any responses aimed at the incorrect choice (S−) are physically prevented. Over successive trials or sessions, the blocking should occur less frequently as the child selects the correct choice without redirection.

A practical application of delayed prompting is to present choice stimuli in front of a child and tell him to choose the stimulus designated as the correct choice (S+) as the teacher points to the correct choice (S+). When the child is responding with the immediate prompt, the teacher systematically increases the delay by 0.5-s increments if the child does not choose the correct choice (S+) within the delay period. The delay should be increased slowly enough so that the child can respond before it is delivered. This technique also very easily lends itself to many common discrete trials tasks commonly presented with trial-and-error techniques such as three-step prompting. In a matching task in which the teacher presents two pictures (e.g., a truck and a bird) and the child is told to "show me the truck," the teacher upon saying the demand can immediately point to the truck. When the child begins to make a consistent response to the truck, the immediate pointing prompt can be systematically delayed.

Cipani and Madigan (1986) also described a very easy-to-do superimposition with stimulus fading procedure to teach children with autism to read an analog clock. First, the hour hand was covered in red, and red circles were placed next to the hour numbers on the clock. The minute hand was then covered in green, and small green circles were placed next to the minute numbers on the clock. The children were taught to read the red hand/red numbers and then the green hand/green numbers. The red and green cues were then systematically faded over time so that the hand length is what controlled the child's number-reading behavior. Another example from our own practice involved teaching a child to discriminate between the men's and women's bathroom signs. Through trial and error, the child only responded at chance levels when the demand, "Show me women's room," was delivered in the presence of both the men's and women's signs. The only difference in the black and white signs was the triangular skirt on the women's bathroom sign. In the superimposition with stimulus fading procedure, the skirt was colored green so that it was much different in appearance than the all-black signs. When responding to the women's sign was stable, the green in the skirt was systematically decreased until responding continued when both signs were again all black.

Superimposition with stimulus shaping was described by Etzel, LeBlanc, Schilmoeller, and Stella (1981). An application of their procedure to teach children sight words should begin with creating word cards containing words the child cannot currently read. The word cards should initially have the picture of the object on the card. For example, the card containing the word "CAT" would contain a picture of a cat. Over successive trials, the picture of the cat is slowly shaped and incorporated into part of the word. The picture of the cat could be shaped to resemble and then replace the letter "A". Eventually no trace of the picture remains on the card and the word itself controls sight word reading.

Although noncomputerized application of errorless learning procedures can be used, it is noteworthy that most of the studies described above were conducted using responses made on computers or by using some kind of computer technology. There are several advantages to computerized application. First, many computers can be purchased or outfitted after sale with touchscreen monitors. Touchscreen monitors allow the educator or teacher to precisely control fading steps in either direction. That is, small increments of intensity, size, position, shape, or color can be reliably controlled and presented to increase or decrease the S− stimuli dimensions in a well-controlled manner. Many touchscreens have multiple screen arrangements that allow for touch zones ("keys") to be arranged in a variety of experimental formats ranging from circular, nine-key arrangements to more traditional five key arrangements that use the center for sample presentation and the four corners for choice stimuli (Saunders & Williams, 1998). Mueller et al. (2000) arranged a touchscreen to present a sample key in the middle of the monitor. In a series of fading steps, the correct choice (S+) and then multiple incorrect (S−) stimuli were programmed to increase systematically from one to four choices. Touchscreen and computerized instructional aids can control many aspects of stimulus presentation not under strict control when an educator presents stimuli by hand. Some variables are simply under better control of with computerized delivery. These include but are not restricted to randomized or semirandomized positioning of choices, simultaneous presentation of stimuli, orienting responses to ensure a participant is oriented to the task, presentation of programmed reinforcer delivery, corrective consequences, and advancement or regression of a fading series. Each can be precisely controlled to decrease the potential for responses made under unwanted sources of stimulus control.

Discrimination learning often is a complex but necessary task for young children with PDD (Green, 1996; Smith, 2001). Because there are many potential negative side effects of incorrect responding when using trial and error procedures, teaching procedures can be arranged to eliminate or reduce responding to incorrect choice stimuli. The six different errorless learning procedures discussed above have been demonstrated to be effective in published research studies. Each procedure can be used to replace trial and error procedures that many children currently use in discrete trials, incidental teaching, natural environment teaching, and pivotal response training, as each of these methods use trials-type learning to teach discrimination tasks in one form or another. Future research should focus on determining if certain skills or tasks are better suited for these techniques, if verbal or motoric responses are better suited for either trial-and-error or errorless methods, whether either trial-and-error or errorless techniques lead to faster acquisition of a wide variety of novel tasks, and whether children with an existing history of learning through either method influences the learning rates of using another method.

References

Alig-Cybriwsky, C., Wolery, M., & Gast, D.L. (1990). Use of a constant time delay procedure in teaching preschoolers in a group format. Journal of Early Intervention, 14, 99–116.

Azrin, N.H., Hutchinson, R.R., & Hake, D.F. (1966). Extinction-induced aggression. Journal of the Experimental Analysis of Behavior, 9, 191–204.

Cipani, E., & Madigan, K. (1986). Errorless learning: Research application for "difficult to teach" children. Canadian Journal for Exceptional Children, 3, 39–43.

Clare, L., Wilson, B.A., Carter, G., Roth, I., & Hodges, J.R. (2002). Relearning face–name associations in early Alzheimer's Disease. Neuropsychology, 16, 538–547.

Coleman-Martin, M.B., & Heller, K.W. (2004). Using a modified constant prompt-delay procedure to teach spelling to students with physical disabilities. Journal of Applied Behavior Analysis, 37, 469–480.

Ducharme, J.M. (2003). "Errorless" rehabilitation: Strategies of proactive intervention for individuals with acquired brain injury and their children. Journal of Head Trauma Rehabilitation, 18, 88–104.

Duffy, L., & Wishart, J.G. (1987). A comparison of two procedures for teaching discrimination skills to Down's Syndrome and non-handicapped children. British Journal of Educational Psychology, 57, 265–278.

Duker, P.C. (1981). Prevention of incorrect responding for establishing instruction following behaviours. Journal of Mental Deficiency Research, 25, 25–32.

Etzel, B.C., LeBlanc, J.M., Schilmoeller, K.J., & Stella, M.E. (1981). Stimulus control procedures in the education of young children. In S.W. Bijou & R. Ruiz (Eds.), Behavior modification: Contribution to education (pp. 3–37). Hillsdale, NJ: Erlbaum.

Everitt, F.L. (1977). Role of errors in children's discrimination learning. Perceptual and Motor Skills, 45, 1187–1201.

Fillingham, J.K., Hodgson, C., Sage, K., & Ralph, M.A. (2003). The application of errorless learning to aphasic disorders: A review of theory and practice. Neuropsychological Rehabilitation, 13, 337–363.

Gibson, A.N., & Schuster, J.W. (1992). The use of simultaneous prompting for teaching expressive word recognition to preschool children. Topics in Early Childhood Special Education, 12, 247–267.

Green, G. (1996). Behavioral intervention for autism. In C. Maurice, G. Green, & S.C. Luce (Eds.), Behavioral interventions for young children with autism (pp. 29–42). Austin, TX: Pro-Ed.

Hull, C.L. (1950). Simple qualitative discrimination learning. Psychological Review, 57, 303–313.

Johnson, C.M. (1977). Errorless learning in a multihandicapped adolescent. Education and Treatment of Children, 1, 25–33.

Keel, M.C., & Gast, D.L. (1992). Small-group instruction for students with learning disabilities: Observational and incidental learning. Exceptional Children, 58, 357–368.

Kelly, J.F., & Hake, D.F. (1970). An extinction-induced increase in an aggressive response with humans. Journal of the Experimental Analysis of Behavior, 14, 153–164.

Kern, R.S., Liberman, R.P., Kopelowicz, A., Mintz, J., & Green, M.F. (2002). Application of errorless learning for improving work performance in persons with schizophrenia. The American Journal of Psychiatry, 159, 1921–1926.

Kessels, R.C., & De Haan, E.F. (2003). Mnemonic strategies in older people: A comparison of errorless and errorful learning. Age and Ageing, 32, 529–533.

Komatsu, S., Mimura, M., Kato, M., Wakamatsu, N., & Kashima, H. (2000). Errorless and effortful processes involved in the learning of face-name associations by patients with alcoholic Korsakoff's syndrome. Neuropsychological Rehabilitation, 10, 113–132.

Koscinski, S.T., & Gast, D.L. (1993). Use of constant time delay in teaching multiplication facts to students with learning disabilities. Journal of Learning Disabilities, 26, 533–544, 567.

Lancioni, G.E., & Smeets, P.M. (1986). Procedures and parameters of errorless discrimination training with developmentally impaired individuals. In N.R. Ellis & N.W. Bray (Eds.), International review of research in mental retardation (pp. 135–164). New York: Academic Press.

Lerman, D.C., Iwata, B.A., & Wallace, M.D. (1999). Side effects of extinction: Prevalence of bursting and aggression during the treatment of self-injurious behavior. Journal of Applied Behavior Analysis, 32, 1–8.

Lovaas, O.I., Koegel, R.L., & Schreibman, L. (1979). Stimulus overselectivity in autism: A review of research. Psychological Bulletin, 86, 1236–1254.

Mattingly, J.C., & Bott, D.A. (1990). Teaching multiplication facts to students with learning problems. Exceptional Children, 56, 438–449.

Moore, R., & Goldiamond, I. (1964). Errorless establishment of visual discrimination using fading procedures. Journal of the Experimental Analysis of Behavior, 7, 269–272.

Mueller, M.M., Olmi, D.J., & Saunders, K.J. (2000). Recombinative generalization of within-syllable units in prereading children. Journal of Applied Behavior Analysis, 33, 515–531.

Parkin, A.J., Hunkin, N.M., & Squires, E.J. (1998). Unlearning John Major: The use of errorless learning in the reacquisition of proper names following herpes simplex encephalitis. Cognitive Neuropsychology, 15, 361–375.

Pilgrim, C., & Galizio, M. (1990). Relations between baseline contingencies and equivalence probe performances. Journal of the Experimental Analysis of Behavior, 54, 213–224.

Rilling, M., & Caplan, H.J. (1973). Extinction-induced aggression during errorless discrimination learning. Journal of the Experimental Analysis of Behavior, 20, 85–92.

Saunders, K.J., & Williams, D. (1998). Stimulus control procedures. In K.A. Lattal & M. Perone (Eds.), Handbook of research methods in human operant behavior (pp. 193–228). New York: Plenum Press.

Schilmoeller, G.L., Schilmoeller, K.J., Etzel, B.C., & LeBlanc, J.M. (1979). Conditional discrimination after errorless and trial-and-error training. Journal of the Experimental Analysis of Behavior, 31, 405–420.

Schimek, N. (1983). Errorless discrimination training of digraphs with a learning-disabled student. School Psychology Review, 12, 101–105.

Schreibman, L. (1975). Effects of within-stimulus and extra-stimulus prompting on discrimination learning in autistic children. Journal of Applied Behavior Analysis, 8, 91–112.

Sidman, M., Cresson, O., Jr., & Willson-Morris, M. (1974). Acquisition of matching to sample via mediated transfer. Journal of the Experimental Analysis of Behavior, 22, 261–273.

Skinner, B.F. (1953). The science of human behavior. New York: Colliler-Macmillian.

Smith, R.G., Iwata, B.A., Goh, H., & Shore, B.A. (1995). Analysis of establishing operations for self-injury maintained by escape. Journal of Applied Behavior Analysis, 28, 515–535.

Smith, T. (2001). Discrete trial training in the treatment of autism. Focus on Autism and Other Developmental Disabilities, 16, 86–92.

Smeets, P.M., Lancioni, G.E., Striefel, S., & Willemsen, R.J. (1984). Training EMR children to solve missing minuend problems errorlessly: Acquisition, generalization, and maintenance. Analysis and Intervention in Developmental Disabilities, 4, 379–402.

Spence, W.W. (1936). The nature of discrimination learning with and without errors. Psychological Review, 43, 427–449.

Stromer, R., Mackay, H.A., McVay, A.A., & Fowler, T. (1998). Written lists as mediating stimuli in the matching-to-sample performances of individuals with mental retardation. Journal of Applied Behavior Analysis, 31, 1–19.

Storm, R.H., & Robinson, P.W. (1973). Application of a graded choice procedure to obtain errorless learning in children. Journal of the Experimental Analysis of Behavior, 20, 405–410.

Terrace, H.S. (1963a). Discrimination learning with and without "errors". Journal of the Experimental Analysis of Behavior, 6, 1–27.

Terrace, H.S. (1963b). Errorless transfer of a discrimination across two continua. Journal of the Experimental Analysis of Behavior, 6, 223–232.

Touchette, P.E. (1971). Transfer of stimulus control: Measuring the moment of transfer. Journal of the Experimental Analysis of Behavior, 15, 347–354.

Venn, M.L., Wolery, M., Werts, M.G., Morris, A., DeCesare, L.D., & Cuffs, M.S. (1993). Embedding instruction in art activities to teach preschoolers with disabilities to imitate their peers. Early Childhood Research Quarterly, 8, 277–294.

Weeks, M., & Gaylord-Ross, R. (1981). Task difficulty and aberrant behavior in severely handicapped students. Journal of Applied Behavior Analysis, 14, 449–463.

Wilson, F.C., & Manly, T. (2003). Sustained attention training and errorless learning facilitates self-care functioning in chronic ipsilesional neglect following sever traumatic brain injury. Neuropsychological Rehabilitation, 13, 537–548.