Adventitious Reinforcement of Maladaptive Stimulus Control Interferes with Learning

Kathryn J. Saunders¹ · Kathleen Hine¹ · Yusuke Hayashi² · Dean C. Williams¹

Published online: 17 June 2016

© Association for Behavior Analysis International 2016

Abstract Persistent error patterns sometimes develop when teaching new discriminations. These patterns can be adventitiously reinforced, especially during long periods of chance-level responding (including baseline). Such behaviors can interfere with learning a new discrimination. They can also disrupt already learned discriminations, if they re-emerge during teaching procedures that generate errors. We present an example of this process. Our goal was to teach a boy with intellectual disabilities to touch one of two shapes on a computer screen (in technical terms, a simple simultaneous discrimination). We used a size-fading procedure. The correct stimulus was at full size, and the incorrect-stimulus size increased in increments of 10 %. Performance was nearly error free up to and including 60 % of full size. In a probe session with the incorrect stimulus at full size, however, accuracy plummeted. Also, a pattern of switching between choices, which apparently had been established in classroom instruction, re-emerged. The switching pattern interfered with already-learned discriminations. Despite having previously mastered a fading step with the incorrect stimulus up to 60 %, we were unable to maintain consistently high accuracy beyond 20 % of full size. We refined the teaching program such that fading was done in smaller steps (5 %), and decisions to “step back” to a smaller incorrect stimulus were made after every 5—instead of 20—

trials. Errors were rare, switching behavior stopped, and he mastered the discrimination. This is a practical example of the importance of designing instruction that prevents adventitious reinforcement of maladaptive discriminated response patterns by reducing errors during acquisition.

Keywords Teaching discriminations · Stimulus control · Superstitious behavior adventitious reinforcement · Resurgence · Programmed instruction · Developmental disabilities

It can be perplexing when a child begins making errors on a previously mastered task. Even the best of us sometimes conclude that, “He can do it, he’s just not trying.”

A key characteristic of an effective instructional program is that prerequisite skills are identified and taught *before* attempts to teach the more complex skills that they comprise. When instruction is well sequenced, errors are kept to a minimum. Deleterious effects of errors have long been recognized. Some problems are obvious: if errors persist, the student is not learning. Moreover, a learner who encounters many errors might discontinue the task entirely or become irritated or aggressive (e.g., DePaepe, Shores, Jack, & Denny, 1996; Gunter, Denny, Jack, Shores, & Nelson, 1993; Gunter, Denny, & Venn, 2000; Holland, 1960; Jack et al., 1996; Moore & Edwards, 2003; Weeks & Gaylord-Ross, 1981). A primary reason for these effects is that errors are signals for nonreinforcement or extinction, and they decrease the overall rate of reinforcement.

Less well recognized is that errors can have persistent, detrimental effects on learning and performance, even if the exact conditions that originally produced the errors have been modified. In other words, errors are not neutral events—they can make subsequent teaching more difficult. This characteristic

✉ Kathryn J. Saunders
ksaunders@ku.edu

Dean C. Williams
deanwms@ku.edu

¹ University of Kansas, Life Span Institute, Lawrence, KS, USA

² Division of Social Sciences and Education, Pennsylvania State University, Hazleton, 76 University Drive, Hazleton, PA 18202, USA

has profound implications for teaching. It is this characteristic of errors that we illustrate with this case study.

Errors are behaviors under stimulus control. Like any behavior, they can be reinforced, that is, made more likely. A study by Stoddard and Sidman (1971) provides an example that is especially striking because the subjects, four rhesus monkeys, were not learning a new task. They were performing a previously learned skill: selecting a circle from an eight-stimulus display containing the circle and seven ellipses (the ellipses were identical to one another). Ellipses are slightly flattened circles; they were the same width as the circle, but slightly shorter in height.

This was basic research, and the next condition was specifically designed to produce errors. The monkeys were exposed to a discrimination that literally was impossible. The display showed eight identical circles. On each trial, one stimulus was randomly designated as correct. That is, food was delivered on approximately one out of eight trials, regardless of the stimulus selected. To varying degrees, the monkeys developed position biases. That is, a large proportion of responses were controlled by a few of the eight key positions. The procedures were thus similar to those used in Skinner's classic demonstration of adventitious reinforcement ("Superstition in the pigeon," 1948), in which a reinforcer was delivered periodically regardless of what the subject was doing.

After the impossible discrimination, the monkeys were re-exposed to the previously mastered circle-ellipse discrimination. At this point, accuracy was *lower than it was before the impossible discrimination*, and the monkeys exhibited a mixture of two sources of stimulus control. On some trials—more than would be expected by chance—the circle controlled selections (a correct response), but on others, positions controlled selections. The position-controlled responses had presumably been reinforced adventitiously during the impossible-discrimination training, and this undesired stimulus control persisted when the previously mastered task was reinstated.

Stoddard and Sidman (1967) reported similar results in a study of children with intellectual disabilities (ID). One group was taught the circle-ellipse discrimination with size fading. The height of the incorrect stimuli (ellipses) was much lower than the circle at first and became more similar to the height of the circle over trials. The other group was taught with small differences between the circle and the ellipse from the beginning of training (i.e., no fading). Children taught with size fading made fewer errors, ultimately maintained higher accuracy, and discriminated smaller differences between the size of the correct and incorrect shapes, than children who began training with small differences between the circle and ellipse. Unlike in the 1971 study, the discrimination that produced errors was not literally impossible, just extremely difficult.

The procedures used with the children had another feature that appears to have produced undesired stimulus control—a

correction procedure that allowed the child to continue selecting one of the eight choice stimuli until he chose the correct key. That is, the trial did not advance until the child made a correct response (under the assumption that this teaches the child what to do, instead of what not to do). In addition to the forms of undesired stimulus control seen in the study described above, however, the correction procedures engendered "adventitiously reinforced sequences" (Sidman & Stoddard, 1967 p. 12). That is, the contingencies established a superstitious chain of key selections rather than a single key selection.

Several other published studies have analyzed laboratory, discrete-trial performance in terms of undesired stimulus control (e.g., Doughty & Saunders, 2009; McIlvane, Kledaras, Callahan, & Dube, 2002). The source of undesired stimulus control (i.e., errors) is typically a position within a stimulus array (position bias) or, in teaching conditional discriminations, a particular comparison stimulus (regardless of the sample stimulus; a stimulus bias). We will use the term "maladaptive" because the behavior interferes with the development of new "adaptive" behavior that would produce a higher rate of reinforcement. In the studies discussed thus far, maladaptive stimulus control was defined by persistent selection of the incorrect stimulus shape (stimulus bias), or the persistent selection of a position regardless of the stimulus shape (position bias).

This case study extends this framework for understanding persistent errors to another form of undesired stimulus control—one that may be especially likely to develop in classroom instruction. Specifically, in simple discrimination sessions that engendered many errors and no learning, we observed a maladaptive selection response in many trials: Immediately after touching a stimulus and regardless of whether the touch was correct or incorrect, the participant moved his hand toward the location of the other stimulus (even though both stimuli disappeared when one was selected). That is, if he initially touched the left stimulus, he immediately moved his hand across the screen toward the right side. This "crossover" was not interrupted by the immediate delivery of the consequences for the initial touch.

Our interpretation is that this pattern of responding may have been adventitiously reinforced in the classroom when the participant encountered a discrimination that he had not learned. Anecdotally, he had also exhibited this maladaptive crossover topography in pretests with tabletop tasks, suggesting that it may have developed during classroom teaching experiences.

Worth noting is that we did not initially intend to study the development of maladaptive stimulus control. The participant was recruited for a treatment study on the reduction of aberrant behavior that was occurring during one-to-one instruction in his classroom. The goal was to remove stimulus control over aggression and improve instructional programming such that he performed an academically relevant task that produced high rates of reinforcement without teacher prompting. We

would then gradually reinstate classroom conditions while maintaining the adaptive behavior. We selected a simple simultaneous discrimination task in hopes that he could learn it independent of teacher prompting.

Upon observing the maladaptive stimulus control, we turned our attention to several questions. First, was there a predictable relation between errors on the simple discrimination task and the occurrence of the crossover topography? Second, once the switching had been established within the context of the task that we were teaching, was it subsequently exhibited when the participant was exposed to a previously mastered discrimination (i.e., was it maladaptive)? Third, could we ultimately eliminate the maladaptive topography?

In addressing these issues, we extended Stoddard and Sidman's (1967, 1971) work in several ways. Stoddard and Sidman (1967, 1971) used a laboratory task specifically designed to be difficult and generate errors (i.e., a circle-ellipse discrimination). We used a much simpler task that met an educational objective. The discrimination was between two stimuli that differed in both color and shape. Thus, we extend the demonstration of adventitiously reinforced stimulus control to teaching procedures for an academically relevant discrimination.

Method

Participant

EC was a 14-year-old boy with ID. He was recruited for a treatment study on the reduction of aberrant behavior, including frequent outbursts of screaming during one-to-one instruction. In addition to ID, clinical records indicated that EC was diagnosed with autism, mood disorder, attention deficit hyperactivity disorder, and intermittent explosive disorder. He did not achieve a basal score on the Peabody Picture Vocabulary Test, 4th Edition (which was administered in both English and Spanish, because EC came from a Spanish-speaking home). His medications at the time of this study included Adderall (30 mg), Risperdal (2 mg), and Depakote (1500 mg).

Apparatus and Setting

To minimize distractions, training occurred in a small, sparsely furnished room. EC sat in front of a touch screen mounted on the wall. Small edible rewards were dispensed into a cup that was mounted below and to the left side of the screen. All sessions were videotaped; the camera was mounted in the ceiling. For safety monitoring, one staff remained in the room with EC, but interaction was kept to a minimum.

Procedures

Computerized Simple-Discrimination Task In the terminal version of the task, two stimuli—a blue square and a yellow circle—were displayed in the bottom two corners of the screen on each trial. Touching the blue square (the S+) produced a reinforcer. The S+ was on each side equally often, and the location varied quasi-randomly, with the restriction that it was not presented in the same location on more than three consecutive trials. Sessions had 20 trials.

Correct responses immediately produced tones and the delivery of an edible. Incorrect responses produced a different tone and a black screen. Responses resulted in the immediate disappearance of the stimuli. The screen was blank for 3 s between trials. Touching the blank screen extended the time before the next trial began, to prevent the inadvertent reinforcement of blank screen touches via the appearance of the stimuli immediately after a screen touch.

Training The initial sessions displayed a single stimulus (i.e., no incorrect stimulus) on each trial. The purpose was to prepare the participant for discrimination training by (a) ensuring that he did not touch the screen unless the stimulus was present, (b) ensuring that his screen touches cleanly activated the touch screen, and (c) providing experience retrieving the reinforcer from the cup.

As baseline, we first presented the terminal version of the task, described above. Although correct responses were reinforced, we expected these “trial-and-error” procedures to be ineffective. Given two sessions at chance level, with no increase in accuracy over the total of 40 trials, we implemented a teaching procedure—size fading.

Size fading is an “errorless” discrimination training procedure in which the incorrect stimulus (the S−) starts out so small that responding to it is highly unlikely, and selection of the correct stimulus (S+) is reinforced. The size of the S− is gradually increased as long as accuracy is high. If errors occur, the size of the S− is reduced to a previously mastered level. Size fading is not always the best choice for discrimination training. Criterion-related prompts (as in stimulus-shaping procedures; Schilmoeller, Schilmoeller, Etzel, & LeBlanc, 1979) are generally recommended for teaching discriminations, but such procedures require an easily or already mastered discrimination as a start point. For example, in Schilmoeller et al., the stimuli were an apple and a witch's hat at the beginning of teaching. The stimuli in this easily learned discrimination gradually morphed into two abstract shapes. For EC, the target discrimination differed in both color and shape. Although we did not test alternatives, it seemed unlikely that we would find an “easier” discrimination to use as a start point.

In the first step of the size-fading procedure, the S− was 10 % of its full size—nearly invisible. Given high accuracy in

a step, the size of the S– increased, in the next session, in increments of 10 %. The criterion for moving to the next fading step was a session score of at least 95 % correct or 90 % if there were no errors in the last 10 trials (sessions had 20 trials). If accuracy did not meet criterion, but was at least 75 %, that teaching step was repeated in the next session (i.e., the S– size stayed the same). If session accuracy was below 75 %, we returned to the previous step (the S– size decreased by 10 %).

We followed the fading protocol up to the point at which the S– was 60 % of terminal size. Because EC's accuracy did not fall below 95 % across the six size-fading sessions, we made a fateful decision—we presented a probe session with differential reinforcement in which the S– was at full size (i.e., the session was the same as those presented before size fading began). As will be described in more detail later, accuracy fell to chance levels, and the participant's response topography changed such that, in many trials, he quickly moved his finger away from the stimulus first touched, toward the other stimulus. We refer to these as “crossover responses.”

Observations of Screen Touches The computer software registered, and recorded as correct or incorrect, the first screen touch for each trial. The crossovers, however, had to be recorded by observers. All sessions were videotaped from an angle that showed the computer screen and the participant's hand position. Screen crossovers were recorded when, following a registered response to the touch screen, one or both of EC's hands crossed the midline of the screen. Note that EC did not actually touch the alternate stimulus, as it disappeared immediately when a registered response occurred.

Interobserver Agreement Two independent observers reviewed video of 15 20-trial sessions chosen randomly from throughout the study and scored the occurrence and nonoccurrence of crossovers. Eleven trials were not scored because the response was obscured in the video. Observers agreed on 285/289 (98.6 %) trials.

Additional Method and Results

The bars in Fig. 1 indicate the size of the S–, expressed at percentage of full size. The circles show the percentage of trials with a correct selection; squares show the percentage of trials with crossovers. In the two baseline sessions in which the S– was at full size, accuracy was around chance levels, and crossovers occurred on 70–80 % of trials. In the first fading session, the S– was presented at 10 % of its full size, and very high accuracy was maintained as the S– size increased across sessions up to 60 % of full size (session 8). So far, it is a virtually flawless example of size-fading procedures!

Fig. 1 Grey bars indicate the size of the S–, as a percentage of the full size, circles show discrimination accuracy, and squares show the percentage of trials with a “crossover”

The next session (session 9), however, was a probe with the S– at full size. Reinforcement contingencies were the same as in fading sessions. Accuracy dropped to chance levels. In addition, crossovers occurred on most trials and regardless of whether the first response (i.e., the response that produced feedback) was correct or incorrect. In fact, in the first two sessions and the probe session (sessions 1, 2, and 9), 47 % of the crossovers occurred following a correct response. Thus, crossovers were not EC's reaction to the consequences for errors.

Reducing the S– to 50 % of its full size for two sessions reinstated the high accuracy and low number of crossover responses shown before the probe, but accuracy dropped to near chance levels and crossovers increased when we reinstated the 60 % S– in session 12. We again reduced the S– to 50 % but did not recover high accuracy across two sessions (13 and 14). Thus, EC was now making many errors at an S– level (50 %) that he had previously performed only with high accuracy.

To re-establish consistently high accuracy, we presented a session with no S– (session 15) followed by another attempt at the fading sequence. High accuracy and low levels of crossovers occurred as the S– was reintroduced at 10 % then increased to 20 and 30 % on sessions 16–19, but accuracy decreased and crossovers increased on subsequent sessions at 30 and 20 % S– levels. Again, he had not shown high error and crossover responding at these S– values before the probe at full size.

We made five additional attempts at the fading sequence that are not shown graphically. Similar to sessions 14–20, these sequences began with no S–, and the S– size increased by 10 % across sessions until accuracy fell below 75 %. When this low accuracy occurred, the S– size was decreased to 0 % and another fading attempt started. Across five attempts, EC never met criterion to advance beyond the S– size of 30 %. Thus, EC did not perform discriminations that he had demonstrated previously.

It is important to note that, in practice, we do not recommend allowing this much exposure to procedures that generate many errors. We violated good practice purely for

experimental purposes—to replicate the disruptive effect of errors. Fortunately, the problem was readily solved with improved instructional programming.

We next made two changes in the instructional procedures meant to decrease errors. First, after the first two sessions, the S[−] size was changed in smaller steps—5 instead of 10 %. Put another way, we doubled the number of fading steps. Second, from a size of 10 % upward, accuracy was evaluated every five trials, instead of every 20 trials. If all five trials were correct, we increased the S[−] size by 5 % for the next five trials. If four of five trials were correct (80 %), the S[−] size stayed the same. If accuracy was below 80 %, the S[−] size decreased by 5 % for the next five trials. That is, when the procedures were generating errors on more than 20 % of the trials, the fading step would be discontinued after only 5 trials instead of 20 trials. That is, we returned to an easier step after many fewer errors than the 20-trial procedure.

Figure 2 shows the results of the revised fading sequence for each 20-trial session. Note that, with a step size of 5 %, EC could proceed through a maximum of four S[−] sizes per 20-trial session. Session accuracy fell below 80 % only once in the entire 14-session training sequence, and thus, EC met criterion in nearly the minimal number of trials. Along with the increases in accuracy, crossovers decreased. In the final two sessions, with the S[−] at full size, EC maintained perfect accuracy across two sessions, and there were no crossovers.

Discussion

In the course of teaching an adolescent with ID a simple discrimination—selecting one of two stimuli that differed in both color and shape presented on a computer screen—we demonstrated persistent, deleterious effects of errors. Specifically, the participant showed high accuracy and a single, discrete touch of the correct stimulus, throughout the first six steps of a size-

fading procedure—up to the point at which the S[−] was 60 % of the size of the S⁺.

When we presented a one-session probe with the S[−] at full size, however, accuracy fell to chance levels. In addition, the participant quickly moved his finger from the stimulus that he touched first, toward the location of the other stimulus, on many trials. This response topography occurred even though the first touch resulted in the disappearance of both stimuli and regardless of whether the first touch was correct or incorrect.

In numerous subsequent attempts to conduct the fading procedure, we were unable to increase the S[−] to larger than 30 % of the terminal size without engendering many errors and the crossover response. That is, after the probe session procedures induced errors, the participant demonstrated low accuracy and the maladaptive topography with S[−] sizes at which accuracy had previously been virtually perfect.

This finding—that maladaptive stimulus control generated by a history of errors can interfere with previously learned tasks—is consistent with the findings of Stoddard and Sidman (1967, 1971). The early studies differed from ours in several ways. Stoddard and Sidman's simple discrimination (a circle vs. an ellipse) was more difficult. They generated the maladaptive topography by presenting an impossible discrimination (both shapes were the same). Finally, errors involved a position bias—a typical side effect of discrete-trial instruction. A well-accepted interpretation of the acquisition of stimulus control by position is that it may be reinforced on about half of the trials in a two-choice task.

We interpret the crossover response similarly, except that we suspect that it was firmly established before the participant began our study. We observed it in “tabletop” testing conducted before the study and in the first two computerized sessions—before the fading procedure began.

A plausible account is that, in attempting to reinforce a specific response occurring within a stream of behavior, well-meaning teachers reinforced other behaviors in the stream. This scenario is much as described by Skinner in the classic Superstition in the pigeon. In that study, a reinforcer was delivered periodically without regard to the pigeon's behavior. As Skinner wrote, “The bird happens to be executing some response” as the reinforcer is delivered and “tends to repeat the response” (Skinner, 1948 p. 169). Because the response increases in frequency, it tends to further co-occur with reinforcer delivery and may become the predominant response in the context. Moreover, once established, it may remain in the subject's repertoire even following a history of extinction and acquisition of another response through direct teaching.

Along with adventitious reinforcement, the process of resurgence may have operated in the present case. *Resurgence* refers to the reoccurrence of a response (i.e., crossovers) that had been replaced by a new response (i.e., stimulus selections) when the new response was no longer reinforced consistently (see Lattal & St. Peter Pipkin, 2007, for a review and

Fig. 2 Grey bars indicate the maximum size of the S[−] in the 20-trial session, as a percentage of the full size, circles show discrimination accuracy, and squares show the percentage of trials with a “crossover”

discussion of relevance to application). Technically, our demonstration may fit the concept of “reinstatement,” in which the resurgent response is controlled by specific stimulus conditions. In this case, frequent errors constitute the stimulus conditions that controlled crossovers. Crossovers had dropped out during the initially successful fading procedures, but the large number of errors in the probe session (session 9, Fig. 1) reinstated crossovers.

Regardless of whether or not EC’s response originated prior to our study, its persistence, and interference with learning, underscores a major practical problem in conducting instruction that involves a selection response: the difficulty of precisely defining the response. When the computer decides whether to deliver a reinforcer, there is no variation in response definition across trials. Thus, it can be beneficial to take the decision as to whether the response is correct out of the hands of a human. This may not be possible in many classrooms, however. For tabletop tasks, stimulus presentation could be engineered to force a more discrete response. For example, the learner could be required to push the correct stimulus toward the experimenter. Or, the learner might be required to place a coin on the chosen stimulus. We have used a method in which the learner touches the stimulus through a hole in a Plexiglas overlay, which makes the response more discriminable. These response definitions also make it more difficult to rapidly change the response. Although we have attributed the final success in teaching the simple discrimination to improved instructional programming, it is also likely that the consistency of the response definition allowed by computerization was beneficial.

These considerations extend to practices that may be common in tabletop instruction. For example, how does a teacher react to a response that is quickly “self-corrected?” We can envision at least two forms of undesired stimulus control that could be established if the self-corrected response results in a reinforcer. One is that a sequence of two responses could be reinforced as a unit; this appears to describe the outcome of the present study. The second is that, when the first response in the sequence is incorrect, the participant might quickly perceive that the teacher is not reaching for a reinforcer and change to the correct response. In this study, crossovers occurred after both correct and incorrect responses and thus were not cued by the computerized feedback.

The data suggest that, like some well-meaning teachers, we were using bad teaching procedures. The participant ultimately acquired the simple discrimination under a teaching procedure that incorporated a more stringent “step-back” criterion. In the new teaching sequence, we evaluated accuracy after blocks of 5 trials, rather than after sessions of 20 trials. This method greatly reduced the number of trials presented in teaching steps (i.e., S– sizes) that generated maladaptive stimulus control.

In addition to their implications for teaching, these findings have implications for designing research on the acquisition of stimulus control: the baseline period is not benign. That is, although long baselines are good for demonstrating experimental control, the errors generated can increase adventitious reinforcement of error patterns that may disrupt future acquisition of the desired stimulus control. This concern forms part of Horner and Baer’s (1978) rationale for the development of the multiple-probe design, which reduces the number of, and increases the time between, baseline sessions.

Horner and Baer (1978) pointed out (in different terms) that maladaptive stimulus control topographies can develop during baseline sessions. And, if so, the subsequent teaching procedures would have to be powerful enough to both eliminate the maladaptive behavior and establish the new behavior. This, in turn, may result in slower acquisition of the new behavior than might occur otherwise—as suggested in the early work of Stoddard and Sidman (1967, 1971) and by the present results. In addition, and paradoxically, this slower acquisition might suggest to the researcher that, to better demonstrate experimental control, the baseline should have been longer—because immediacy of effect is one marker of control by the independent variable (Cooper, Heron, & Heward, 1987).

These data support the observations of Sidman and Stoddard (1967): “Once the procedure has generated errors, response patterns incompatible with correct responses may be reinforced and hinder the child in learning” (p. 10). Here, adventitious reinforcement presumably established the crossover pattern. The initial steps of the size-fading procedure produced stimulus control by the larger stimulus and nearly 100 % accuracy and reinforcement on every trial. When the S– was again introduced at full size, the participant no longer demonstrated the stimulus control relation involving the larger stimulus. The crossover responding was reinstated, and it persisted even at S– sizes previously associated with accurate responding. Under these conditions, the crossovers were again adventitiously reinforced. The revised teaching program reduced the negative impact of errors by decreasing their likelihood and by stepping back to a previously acquired discrimination after fewer errors, which decreased adventitious reinforcement of the switching pattern.

We conclude by noting two issues of particular importance to practitioners. First, when teaching discriminations, strictly maintaining the definition of a correct response is critical. This is something we all know, of course, but it can be extremely challenging to implement and maintain when “in the thick of things.” Ideally, the definition should allow little latitude in the therapist’s determination of response initiation and completion. In our task, the response definition was completely objective—the computer detected a screen touch. A second, related, issue is that the first response should determine whether or not a reinforcer is delivered. Allowing ambiguous “self-

correction” responses creates the potential for adventitious reinforcement of problematic response topographies.

Acknowledgments This research was funded by the National Institute of Child Health and Human Development grant no. P01HD055456 and center grants P30 HD002528 and P30 DC005803 to the University of Kansas. We appreciate the support of the Parsons State Hospital and Training Center. We also thank Carol Cummings, Carlos Sanchez, and Sarah Hall for their assistance in collecting and analyzing data. Address correspondence to Kate Saunders (ksaunders@ku.edu) or Dean Williams (deanwms@ku.edu).

Compliance with Ethical Standards

Conflict of Interest This research was funded by grant no. P01HD055456 from the National Institute of Child Health and Human Development to the University of Kansas.

Ethical Approval/Human Rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This work was approved by the University of Kansas Institutional Review Board.

Animal Welfare This article does not contain any studies with animals performed by any of the authors.

Informed Consent Written informed consent and behavioral/verbal assent were obtained from all individual participants included in the study. Because the participant was not capable of giving written informed consent, the parents or legal guardian provided written consent, and the participant provided verbal and behavioral assent.

References

- Cooper, J., Heron, T. E., & Heward, W. L. (1987). *Applied behavior analysis*. Columbus: Merrill.
- DePaepe, P. A., Shores, R. E., Jack, S. L., & Denny, R. K. (1996). Effects of task difficulty on the disruptive and on-task behavior of students with severe behavior disorders. *Behavioral Disorders, 21*, 216–225.
- Doughty, A. H., & Saunders, K. J. (2009). Decreasing errors in reading-related matching to sample using a delayed-sample procedure. *Journal of Applied Behavior Analysis, 42*, 717–721.
- Gunter, P. L., Denny, R. K., Jack, S. L., Shores, R. E., & Nelson, C. M. (1993). Classroom management strategies: are they ecological contexts for coercion. *Behavioral Disorders, 18*, 265–274.
- Gunter, P. L., Denny, R. K., & Venn, M. L. (2000). Modification of instructional materials and procedures for curricular success of students with emotional and behavioral disorders. *Preventing School Failure: Alternative Education for Children and Youth, 44*, 116–121.
- Holland, J. G. (1960). Teaching machines: an application of principles from the laboratory. *Journal of the Experimental Analysis of Behavior, 3*, 275–287.
- Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: a variation on the multiple baseline. *Journal of Applied Behavior Analysis, 11*, 189–196.
- Jack, S. L., Shores, R. E., Denny, R. K., Gunter, P. L., DeBriere, T., & DePaepe, P. A. (1996). An analysis of the relationship of teachers’ reported use of classroom management strategies on types of classroom interactions. *Journal of Behavioral Education, 6*, 67–87.
- Lattal, K., & St. Peter Pipkin, C. (2007). Resurgence of previously reinforced responding: research and application. *The Behavior Analyst Today, 10*, 254–266.
- McIlvane, W. J., Kledaras, J. B., Callahan, T. C., & Dube, W. V. (2002). High-probability stimulus control topographies with delayed S+ onset in a simultaneous discrimination procedure. *Journal of the Experimental Analysis of Behavior, 77*, 189–198.
- Moore, J. W., & Edwards, R. P. (2003). An analysis of aversive stimuli in classroom demand contexts. *Journal of Applied Behavior Analysis, 36*, 339–348.
- 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.
- 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.
- Skinner, B. F. (1948). Superstition in the pigeon. *Journal of Experimental Psychology, 38*, 168–172.
- Stoddard, L. T., & Sidman, M. (1967). The effects of errors on children’s performance on a circle-ellipse discrimination. *Journal of the Experimental Analysis of Behavior, 10*, 261–270.
- Stoddard, L. T., & Sidman, M. (1971). The removal and restoration of stimulus control. *Journal of the Experimental Analysis of Behavior, 16*, 143–154.
- Weeks, M., & Gaylord-Ross, R. (1981). Task difficulty and aberrant behavior in severely handicapped students. *Journal of Applied Behavior Analysis, 14*, 449–463.