ORIGINAL ARTICLE

Using Behavioral Skills Training and Equivalence-Based Instruction to Teach Children Safe Responding to Dangerous Stimuli: A Proof of Concept

Antonia R. Giannakakos[1] · Jason C. Vladescu[2] · Kenneth F. Reeve[2] · April N. Kisamore[3] · Daniel M. Fienup[4] · Jacqueline N. Carrow[1]

Abstract
Over 3,000 children under 10 years old died in 2017 as a result of unintentional injuries (e.g. suffocation, drowning, firearm discharge, burns, and accidental poisonings). Some of these deaths likely occurred when children came across a dangerous item while unsupervised. In the current study two typically developing 4-year-olds were taught to engage in a safety response in the presence of a variety of dangerous stimuli. Behavioral skills training plus in-situ training was used to teach a three-step response in the presence of one dangerous stimulus. Following mastery of this response, equivalence-based instruction was used to create classes of dangerous (i.e., medicine bottle, a lighter, and a handgun) and nondangerous (i.e., container, a flash drive, and a hair dryer) stimuli. Results demonstrated that the addition of equivalence-based instruction to behavioral skills training plus in-situ training was effective at training a safety response across multiple dangerous stimuli.

Keywords Behavioral skills training · Equivalence-based instruction · In-situ training · Transfer of function · Safety skills

Over 3,000 children under 10 years old in the United States died in 2017 as a result of unintentional injuries (Centers for Disease Control [CDC], 2017). Specific unintentional injury causes include suffocation, drowning, firearm discharge, burns, and accidental poisonings. Some of these deaths likely occurred when children came across a dangerous item while unsupervised. For the top 10 leading causes of unintentional injury deaths, half were a result of contact with a dangerous stimulus (e.g., open water, swallowed objects, firearms, fires, and poison; CDC, 2017). Some of these deaths may be prevented if children are taught appropriate responses to dangerous stimuli at a young age.

✉ Antonia R. Giannakakos
  Antonia.giannakakos@mville.edu

✉ Jason C. Vladescu
  jvladescu@caldwell.edu

[1] Department of Special Education, Manhattanville College, 2900 Purchase Street, Purchase, NY 10577, USA

[2] Department of Applied Behavior Analysis, Caldwell University, 120 Bloomfield Avenue, Caldwell, NJ 07006, USA

[3] Hunter College, City University of New York, New York, NY, USA

[4] Columbia University, New York, NY, USA

The procedure most often used in the safety skills literature to teach safety responses is behavioral skills training (BST; Giannakakos, Vladescu, Kisamore, Reeve, & Fienup, 2019). BST involves (1) providing a vocal and/or written description of the target skill; (2) modeling correct implementation; (3) requiring the consumer to practice; and (4) providing feedback on performance (Parsons, Rollyson, & Reid, 2012; Rosales, Stone, & Rehfeldt, 2009). In regard to safety skills training, the response taught is typically a three-step response in which the consumer does not touch the item, leaves the area, and tells an adult. BST is often paired with in-situ training. In-situ training (IST) involves assessing whether the three-step safety response is emitted in the environment where the skill is expected while providing positive and corrective feedback as needed. Although BST plus IST is effective, the efficiency of these procedures may be limited. In the extant literature, participants were typically taught a safety response to a single danger (e.g., Lee, Vladescu, Reeve, Peterson, & Giannakakos, 2019; Jostad, Miltenberger, Kelso, & Knudson, 2008; Miltenberger et al., 2004). When training was provided specific to multiple dangers, training was provided specific to each stimulus (e.g., Rossi, Vladescu, Reeve, & Gross, 2017; Summers et al., 2011).

In an exception evaluating multiple dangers, Vanselow and Hanley (2014) designed a series of studies that evaluated the extent to which generalized responding occurred across three dangers following computerized BST. Four 5- and 6-year-old children were exposed to computerized BST for three dangers (lighters, poisons, strangers) and then IST for only one danger (lighters). Subsequent probes of the participant responding in the presence of poisons and abduction lures indicated that the safety response generalized to the poisons but did not generalize to a lure from a stranger. The limited generality of the safety responses following BST may have been due in part to the appropriate stimulus conditions not having been established during training.

When training a safety response, it is important for the response to occur under the appropriate stimulus conditions. One procedural modification that may increase the likelihood that appropriate stimulus control is established would be to incorporate nondangerous stimuli that share physical similarities with the dangerous stimuli into training. In the presence of the nondangerous stimuli a consumer would be expected to engage in a response other than the safety response (e.g., staying and playing). This would increase appropriate stimulus control, because it would increase the likelihood the safety response would be evoked only by those features specific to the dangerous stimuli and not those features shared by both. For example, a flash drive and a lighter may share several physical similarities such as color and shape, but only a lighter has a flint wheel.

Using BST plus IST could be a time-consuming process if a safety response to multiple dangers needs to be trained. A safety response has been shown to generalize across contexts (Lee et al., 2019), but it is unlikely the response will generalize to different dangers. Lee et al. (2019) evaluated the extent to which a safety response taught in the presence of a gun left in the open among the participants' toys would generalize to other contexts. Although the authors observed generalization to the novel contexts, they did not evaluate if the trained safety response would generalize to untrained dangers. Generalization could potentially be hindered by restricted stimulus control established over the safety response during BST plus IST. Following BST plus IST for only one danger the untrained dangerous stimuli have no history of differential reinforcement with regard to each other and therefore responses trained in the presence of only one are unlikely to be exhibited in the presence of the other untrained stimuli (Fields & Moss, 2008). However, the research suggests that combining BST plus IST with equivalence-based instruction (EBI) may provide a solution to the problem of training safety responses to multiple dangerous stimuli.

In an equivalence paradigm, a series of conditional relations are trained between stimuli; the structure of training results in the subsequent emergence of a series of untrained relations (Fields & Reeve, 2001). One potential outcome of EBI that has capacity for applied application is transfer of stimulus function. Transfer of function occurs when a response trained in the presence of one stimulus in a derived relation modifies the function of another stimulus or stimuli without direct training (Dymond & Rehfeldt, 2000). Transfer of function has been demonstrated in the equivalence literature (e.g., de Rose, McIlvane, Dube, & Stoddard, 1988; Dymond & Rehfeldt, 2000).

In one applied example, Taylor and O'Reilly (2000) evaluated if the combination of EBI and BST resulted in transfer of a grocery store task analysis to two untrained locations. Three individuals with autism spectrum disorder were exposed to EBI to establish three classes composed of supermarkets, shops, and restaurants. The members of each class were the spoken name of the location, the written name (e.g., supermarket), and a picture of the interior of each location. In the case of the supermarket, two additional pictures of novel supermarkets were used as exemplars (hereafter referred to as variants). BST was used to train the steps of grocery shopping in one supermarket. The authors found that the combination of EBI and BST resulted in all three participants correctly competing the steps involved in grocery shopping in the two untrained supermarkets without the need for direct training in those environments.

Given the findings in the EBI literature and those of Taylor and O'Reilly (2000), EBI and transfer of function may be ideally suited for application to safety skills training. EBI could be used to form functional classes of dangerous stimuli and establish separate dangers as being interchangeable for each other, and thus increase the number of dangerous stimuli

that exert control over the safety response. The formation of a dangerous equivalence class is of particular relevance, because a dangerous stimulus is not defined by its physical features alone, but by the consequences its misuse may produce. For example, a class of dangerous stimuli may include a lighter, a filled bathtub, and a handgun. All of these stimuli are physically disparate, but all could result in injury or death. The combination of EBI and BST plus IST, could result in responses trained to a single member of each class transferring to the other class members without additional instruction (Fields & Moss, 2008).

The current study had three purposes. First, we sought to evaluate whether BST could be used to teach a discriminated safety response, that is, an appropriate response in the presence of one dangerous stimulus and one nondangerous stimulus. Second, we evaluated the extent to which BST plus IST led to demonstration of the trained responses in the presence of untrained dangerous and nondangerous stimuli. Third, we employed EBI to establish two (dangerous and nondangerous), three-member classes of stimuli and evaluated whether transfer of function occurs across the equivalence class members.

Method

Participants

Two children with no clinical diagnoses or prior experience with safety response training participated in the current study. Participants were recruited via word of mouth from the local community. Informed consent was obtained from all individual participants included in the study. Jack was 4 years, 11 months old at the start of the study. He obtained a standard score of 113 (qualitive description: above average) on the Peabody Picture Vocabulary Test-Fourth Edition (PPVT-4; Dunn & Dunn, 2007) and a standard score of 87 (average) on the Expressive Vocabulary Test-Second Edition (EVT-2; Williams, 2007). Chrissy was 4 years, 11 months old at the start of the study and received standard scores of 96 (average) and 101 (average) on the PPVT-4 and the EVT-2, respectively. Both participants received ratings that indicated low levels or no problem behavior on the Disruptive Behavior Disorders Rating Scale (Friedman-Weieneth, Doctoroff, Harvey, & Goldstein, 2009) and the Home Situations Questionnaire (Barkley & Edelbrock, 1987) completed by at least one caregiver. Both participants demonstrated picture-to-object matching, object-to-picture matching, and three-step direction following prior to the start of the study. The experimenter taught participants to tact all experimental stimuli prior to the beginning of the evaluation. In addition, the experimenter conducted matching pretraining using stimuli not associated with the evaluation to provide participants with a learning history matching physically disparate stimuli.

Setting and Materials

All sessions for Jack took place in a room equipped with a one-way mirror at a private university. All sessions for Chrissy took place in a classroom at a church near her home. Remote viewing software was used to observe Chrissy during response assessments and IST. For sessions that required remote viewing, AtHome Video Streamer was installed on an iPad that was placed out of sight of the participant in the session room. The experimenter watched the live stream from a smart phone. The session room contained moderately preferred age appropriate toys, a table, and chairs. Participants were never in the session room without being trained or assessed. Tabletop instruction was conducted at a small table in a room that was not used for response assessment sessions. A video camera and the AtHome Video Streamer recorded sessions for data collection purposes.

During the response pretest, BST, post-BST plus IST, response posttest, and transfer of function sessions, materials included both dangerous and nondangerous stimuli. The dangerous stimuli comprised three exemplars (selected from a psychometric sort, described below) each of handguns, medicine bottles, and lighters. The nondangerous stimuli included three exemplars each of hair dryers, plastic containers filled with nondangerous items, and flash drives..

EBI materials were displayed in a three-ring binder. Trial sheets consisted of a sample stimulus displayed in the center of the top half of the page and two comparison stimuli displayed horizontally on the bottom half of the page. Trial sheets were printed on 21.59 cm X 27.94 cm white paper and presented horizontally in a clear sheet protector. The comparison stimuli were covered with a flap of paper, which was removed after the participant engaged in the observing response of touching the sample stimulus. A blank sheet of paper separated each trial sheet.

Psychometric Sort

The experimenter conducted a psychometric sort (Rosch, 1975) to identify different physically representative exemplars of each danger. The experimenter created 30 (10 for each handgun, lighter, and medicine) stimuli cards, each containing a picture of a dangerous stimulus. Ten students enrolled in graduate-level behavior analysis courses were asked to sort the stimuli cards from most to least representative of each experimental stimulus. The pictures used for the stimuli cards are available from the first author upon request. The most typical exemplar for each dangerous stimulus received a score of 1 and the least typical exemplar received a score of 10. The average position of each stimulus was calculated by adding

the scores and dividing by 10. For each dangerous stimulus, the experimenter selected the exemplars from each sort that received the highest average rating, lowest average rating, and an average rating closest to five. These specific stimuli were chosen to represent the highest, middle, and lowest end of each stimulus' perceptual characteristics and were used in the study. The stimuli with the highest and lowest average rating were used during training and the stimuli with an average rating closest to five were reserved for generalization testing. Pictures of the experimental stimuli are available in the Online Supporting Information.

Preassessments

Preference assessment and token economy Prior to the start of the study the experimenter conducted a 10-item picture-based multiple-stimulus-without-replacement preference assessment (MSWO; Carr, Nicolson, & Higbee, 2000). The top five ranked items were reserved for use as putative backup reinforcers during EBI. A token economy was created for each participant using the iReward application on an iPad Pro. Participants earned a token contingent on correct responses (both unprompted and prompted) and appropriate behaviors during preassessments (i.e., sitting in their seat, waiting for the next trial) and EBI sessions. After earning 10 tokens participants could exchange them for one of the backup reinforcers identified during the MSWO.

Card sorting task A card-sorting task was incorporated to assess whether participants demonstrated responding that suggested already established equivalence classes pertaining to the experimental stimuli (Arntzen, Norbom, & Fields, 2015). A detailed description of the task and its results can be found in the Online Supporting Information.

Dependent Variable, Interobserver Agreement, and Procedural Integrity

The primary dependent variable during the response pretest, BST, post BST plus IST, response posttest, and the transfer of function test was the safety response score. Responses in the presence of *dangerous* stimuli were scored based on a three-point system. Behaviors evaluated included touching the dangerous stimulus (defined as any contact between the participant's body or clothing and the dangerous stimulus), leaving the area (defined as the participants initiating to leave the room within 10 s of seeing the item; the participant must not have the item in hand), and telling an adult (defined as independently providing information about the presence of the dangerous stimulus to an adult within 30 s of leaving the room; using the correct name of the item). Safety responses were scored as follows: 3 = *did not touch the dangerous stimulus, left the room, and told an adult*; 2 = *did not touch the dangerous*

stimulus, left the room, did not tell an adult; 1 = *did not touch the dangerous stimulus, did not leave the room, did not tell an adult*; and 0 = *touched the dangerous stimulus.* Data were summarized as a response score per session.

In addition, the degree to which acquired safety responses were under appropriate stimulus control was evaluated by scoring participant responding in the presence of nondangerous stimuli. Responses in the presence of *nondangerous* stimuli were scored based on a two-point system: 2 = *stayed in the room and played (the participant may or may not have touched the stimulus)*; 1 = *left the area and/or told an adult about the nondangerous stimulus while touching it*; and 0 = *did not touch the item, left the area and told the experimenter about the presence of the nondangerous stimulus.* Data were summarized as a response score per session.

During EBI, the dependent variable was the percentage of unprompted correct responses. The experimenter collected data on unprompted correct responses, prompted correct responses, prompted incorrect responses, and unprompted incorrect responses during each session. Unprompted correct responses were defined as the participant touching the correct comparison stimulus within 5 s of the presentation of the corresponding sample stimulus. A prompted correct response was defined as the participant touching the correct comparison stimulus within 5 s of a model prompt. A prompted incorrect response was defined as the participant engaging in an error of omission (i.e., not engaging in a response) or commission (i.e., responding incorrectly) within 5 s of a model prompt. An unprompted incorrect response was defined as the participant engaging in an error of omission or commission within 5 s of the presentation of the sample stimulus. Self-corrections were scored as incorrect. The percentage of unprompted correct responses was calculated by dividing the number of unprompted correct responses by the number of trials in the session and multiplying by 100%.

The experimenter calculated trial-by-trial interobserver agreement data for a minimum of 28.5% (range: 28.5%–40%) of sessions for both participants during all phases of the study. An independent observer collected data on the dependent variable in vivo or from video. For response assessments (i.e., response pretest, post BST plus in-situ training sessions, the response posttest, and the transfer of function test), an independent observer scored data as either 0% (differing scores) or 100% (identical scores) per session. An agreement score was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100%.

For EBI, IOA data were calculated for a minimum of 30% (range: 30%–50%) of trials for all EBI test and training conditions. An independent observer collected data on participants' responding during EBI. An agreement score was calculated by dividing the number of agreements by the number

of agreements plus disagreements and multiplying by 100%. Mean IOA calculated for response assessment sessions was 100% for both participants. Mean IOA calculated for EBI sessions was 99% (range, 94%–100%) for Jack and 100% for Chrissy.

Procedural integrity (PI) data were collected on a minimum of 28.5% (range, 28.5%–50%) of sessions in each phase of the study. Mean PI calculated for response assessment and EBI sessions was above 96% (range: 86%–100%) for Jack and Chrissy. A secondary observer collected secondary PI data for 25% of response assessment sessions and EBI sessions. Mean PI IOA for response assessment sessions and EBI sessions was above 97% (range: 88%–100%) for both participants. See the Online Supporting Information for a full breakdown of IOA, PI, and PI IOA data.

Design and General Procedure

A nonconcurrent multiple baseline across-participants design (Harvey, May, & Kennedy, 2004) was used to evaluate the effectiveness of BST plus IST and EBI on participant demonstration of responses to experimental stimuli. A pretest/posttest design (Walker & Rehfeldt, 2012) was used to evaluate the effectiveness of EBI at forming classes of dangerous and nondangerous stimuli. For an experimental overview of the phases in the study see Table 1.

Response pretest Each response assessment session consisted of one trial with an intertrial interval of approximately 2 min. Sessions were conducted two to three times a week depending on participant schedules and each experimental meeting was 1 to 2 hr in length. Prior to each trial the experimenter placed a dangerous or nondangerous stimulus (hereafter referred to as an experimental stimulus) in the session room on a small table containing age-appropriate toys. These toys were those toys ranked in positions 6–10 during the MSWO. The experimental stimulus was placed within reach of the participant. Only one experimental stimulus was present during each trial. The order of presentation of experimental stimulus trials was arranged such that no more than two consecutive trials were conducted with a dangerous or nondangerous stimulus. Each experimental stimulus was presented at least once during the response pretest. At the start of the session the experimenter played with the participant outside the session room for 2 min. After this play period, the experimenter directed the participant to the session room and said, "I have to go do some work. You can come get me if you need me, and I'll come get you when I'm all done." The experimenter left the session room, closed the door, and observed the participant. No programmed consequences were provided for responses to dangerous and nondangerous stimuli during the response pretest regardless of response score. After 2 min if the participant did not touch an experimental stimulus or 10 s after the participant touched an

experimental stimulus, the experimenter reentered the room and told the participant it was time to go someplace else (e.g., "I'm done. Let's go play in the hallway."). If the participant engaged in any component of the safety response or responded in any way to the nondangerous stimulus, the experimenter made a neutral statement (e.g., "okay") and brought the participant to another area to play. If a participant left the session room to report the presence of a nonexperimental item (e.g., a block, toy), the experimenter made a neutral statement and brought the participant to another room to play.

Behavioral skills training The experimenter used BST to teach participants to differentially respond in the presence of one class member each from the dangerous stimulus class and the nondangerous stimulus class (i.e., handgun and hair dryer, lighter and flash drive, or medicine bottle and container; see Online Supporting Information). Two exemplars of each class member were used during training in an attempt to form a fully elaborated generalized equivalence class (Fields & Moss, 2008). The exemplars selected from each class for use during BST were counterbalanced across participants.

BST consisted of procedures adapted from Miltenberger et al. (2004). First, the experimenter provided information and instructions. Then, the experimenter presented each experimental stimulus one at a time to the participant. The experimenter presented a dangerous stimulus and said "This is a [stimulus]. It is very dangerous. If you find a [stimulus], don't touch the [stimulus], leave the room, and tell an adult." For a nondangerous stimulus, the experimenter said "This is a [stimulus]. It is not dangerous and safe to touch. If you find one you can keep playing and should not tell an adult about it." After providing instruction on all four stimuli (two dangerous, two nondangerous), the experimenter modeled the appropriate response while vocally describing each step of the respective response. Following the modeling of these steps the experimenter provided the participant with an opportunity to practice. The participant had the opportunity to rehearse the entire response while receiving positive and corrective feedback from the experimenter. The experimenter provided corrective feedback for any steps performed incorrectly. Corrective feedback consisted of the experimenter specifying what portion of the response the participant performed incorrectly. The experimenter then modeled the correct response again and had the participant engage in additional practice. Criterion for BST completion was when a participant demonstrated the entire response five consecutive times across each dangerous and nondangerous stimulus with no prompts or feedback from the experimenter. Unprompted correct responses

Table 1. Experimental Overview

Response Pretest	BST+ IST	Response Posttest	EBI Pretest	Train	Derived Relation Testing	EBI Posttest	Transfer of Function Test	Maintenance
Pretest- All dangerous and nondangerous stimuli								
	Dangerous stimulus Nondangerous stimulus							
		Posttest- All dangerous and nondangerous stimuli						
			BA AB CA AC BC CB					Trained Dangerous stimulus and nondangerous stimulus
				BA CA	AB AC BC CB			
						AB BA AC CA BC CB		
							All dangerous and nondangerous stimuli	

Note. A=Handgun/hair dryer, B=Medicine bottle/ container, C= Lighter/flash drive

during the practice portion of BST were consequated with praise and a token.

Post-BST and in-situ training Following BST, in-situ training (IST) was implemented. During IST sessions, the experimenter observed covertly and was not present in the session room. Contingent on a response score less than 3 (dangerous stimulus) or 2 (nondangerous stimulus) the experimenter entered the room and provided corrective feedback as per the example above. The experimenter then modeled the correct response and had the participant rehearse the response five times (Miltenberger et al., 2004). Positive and corrective feedback were provided throughout the rehearsals. This procedure was used for both the nondangerous and dangerous category responses. Following a response score of 3 in the presence of a dangerous stimulus, and a response score of 2 in the presence of a nondangerous stimulus, the experimenter provided the participant with positive vocal feedback. Post-BST and IST was considered complete when a participant correctly demonstrated the appropriate response across each trained experimental stimulus two times without requiring subsequent rehearsals.

Response posttest During the response posttest, the participant's responding in the presence of all experimental stimuli was assessed. The response posttest followed procedures described in the response pretest.

Review The experimenter conducted review sessions every 7 to 10 days during EBI to increase the likelihood that correct responses in the presence of the trained experimental stimuli would maintain while EBI was ongoing. Review sessions were identical to IST sessions.

Equivalence-based instruction general procedure The experimenter used a simple-to-complex training protocol (Imam, 2006) with a many-to-one training structure (Saunders, Drake, & Spradlin, 1999) to establish two (i.e., dangerous

and nondangerous), three-member equivalence classes. The dangerous class consisted of the following members: handguns, lighters, and medicine bottles. The nondangerous class consisted of the following members: hair dryers, flash drives, and containers. The experimental stimuli used during BST for each participant served as the nodal stimuli (i.e., the A stimuli) during EBI. For Jack the nodal stimuli during EBI were handgun and hairdryer and for Chrissy the nodal stimuli were lighter and flash drive. Two exemplars (hereafter referred to as variants) of each class member were taught and a third was reserved for generalization testing. A breakdown of all relations is available in the Online Supporting Information. For an overview of the assignment of stimuli during EBI for each participant, see Table 2.

EBI sessions were conducted 3 days a week and training sessions consisted of 16 trials. Three to four sessions were run during each experimental meeting with an intersession interval of 5 min. EBI trials were presented in a binder. The trial page included one sample stimulus and two comparison stimuli. During pretest, symmetry and equivalence probes, and the posttest no programmed consequences were delivered for unprompted correct and incorrect responses. A token and praise were delivered during each intertrial interval (ITI) for appropriate behavior, such as remaining seated. During the first session for each trained relation a 0-s prompt delay was used to provide a model prompt to the correct comparison stimulus. If the participant engaged in a prompted correct response the experimenter provided praise and delivered a token. If the participant engaged in a prompted incorrect response, the experimenter modeled the response again. Only praise was delivered during error correction trials. Following one session at 0-s prompt delay, the prompt delay was increased to 5 s.

During trials conducted with a 5-s prompt delay if the participant engaged in an unprompted correct response the experimenter provided praise and delivered a token. If the participant engaged in an unprompted incorrect response the experimenter provided a model prompt. After the participant engaged in a prompted correct response, then the experimenter represented the trial. This sequence was continued until the participant engaged in an unprompted correct response. Responses during error correction trials resulted in praise only.

If the participant engaged in two consecutive prompted incorrect responses, the experimenter required the participant to rehearse selection of the correct comparison stimulus five consecutive times and then represented the trial. This sequence continued until the participant engaged in a prompted correct response following the model prompt. This was included to prevent participants from practicing incorrect responses, which could inhibit acquisition of the trained relations.

Pretest. During pretest sessions, all possible relations were tested (i.e., BA, AB, CA, AC, CB, BC), including the generalization probes for each relation. All relations were presented in a 120-trial mixed trial block (i.e., all relations intermixed). Each relation was presented 16 times with an additional four trials for each relation presented using the generalization exemplars.

Training. Each training session consisted of 16 trials. Each sample was presented four times such that the correct comparison stimulus appeared equally on the right and left side across trials. Training for each relation continued until the participant demonstrated 100% unprompted correct responding for two consecutive sessions. First, the BA relation was trained across the two classes. For example, Jack was taught to match the lighter (sample) to the handgun (comparison) and the flash drive (sample) to the hair dryer (comparison).

The experimenter implemented feedback fading following mastery of a trained relation. Feedback fading was implemented during training of the BA and CA relations. Feedback fading consisted of providing praise and a token on a subset of trials (75%–50%–25%–0%; Fields, Reeve, Adams, & Verhave, 1991), and the size of this subset was systematically decreased until no praise was provided contingent on responding.

The AB symmetrical relation was assessed following mastery of the BA relation. If the participant did not score 94% or higher during the AB symmetrical relation probe, the experimenter repeated the training procedure for the BA relation, and then retested the AB relation. Following mastery of BA and AB relations, the CA relation was trained using procedures identical to those used to train the BA relation. Following mastery of the CA relation, the symmetrical AC relation was tested using procedures identical to the test of the AB symmetrical relation.

Equivalence probe. After a participant demonstrated mastery of the BA, AB, CA, and AC relations, we tested for

Table 2. Assignment of stimuli for EBI	Participant	A1	A2	B1	B2	C1	C2
	Jack	Hair dryer	Gun	Flash drive	Lighter	Container	Medicine
	Chrissy	Flash drive	Lighter	Container	Medicine	Hair dryer	Gun

Note: Class 1 was nondangerous and class 2 was dangerous. The A stimuli served as the nodal stimuli for both participants

the emergence of the BC and CB equivalence relations. Mastery of the equivalence relations was 87.5% or higher (at least 14 correct responses out of 16). If a participant did not meet this mastery criterion additional training of the BA and CA relations commenced.

Posttest. Following the test for equivalence, a 72-trial (Chrissy) or 120-trial (Jack) posttest of all relations (BA, AB, CA, AC, BC, and CB) including generalization stimuli was conducted using the same procedures as the pretest. Due to time constraints, we had to reduce the length of the EBI posttest for Chrissy. To do so, we removed the duplicate stimulus–stimulus pair from each relation. All generalization trials were included.

Transfer of function test The purpose of the transfer of function test was to determine whether the responses taught during BST plus IST in the presence of one member of the dangerous and nondangerous classes would be emitted in the presence of the other class members following EBI. For example, following BST plus IST with two exemplars of handguns and two exemplars of hair dryers, and EBI to create dangerous and nondangerous classes, participant responding was assessed in the presence of the remaining experimental stimuli (e.g., lighters, medicine bottles, flash drives, and containers). Procedures were identical to those used during the response pretest. A general praise statement was provided (e.g., "okay!") for all response scores.

Social Validity

Parents of the participants in the study as well as parents of 3- to 5-year-olds recruited via word of mouth from the local community were asked to rate their agreement to various questions related to safety response training on a five-point Likert-type scale (1 = no agreement; 5 = strong agreement). All caregivers indicated strong agreement with the goals of the study (see Table 3). Graduate students in ABA were asked to view pairs of video clips and indicate in which clips the participant demonstrated a safety response and to rate the acceptability of the safety response. Fifteen of the 17 respondents correctly selected the postintervention clip as depicting the most appropriate safety response on all six pairs. All respondents agreed that the safety response taught was appropriate, would prevent the participant from being harmed, and indicated they would teach the safety response to clients.

Results

Figure 1 displays the results in response to dangerous and nondangerous stimuli for Jack and Chrissy across the response pretest, post-BST plus IST sessions, the response posttest,

maintenance, and transfer of function test. For Jack the nodal stimuli used during EBI and trained during BST were gun and hair dryer. During the response pretest (Fig. 1, first panel) Jack either made contact with all dangerous stimuli or failed to leave the area and inform an adult about their presence. In the presence of all nondangerous stimuli (Fig. 1, second panel) Jack remained in the session room and played with the available toys. Following BST plus IST Jack demonstrated the appropriate safety response in the presence of the trained dangerous stimuli (i.e., guns) and the stay and play response during sessions with trained nondangerous stimuli (i.e., hair dryers). During the response posttest Jack did not demonstrate class-consistent responding in the presence of any untrained experimental stimuli, thus we implemented EBI.

Figure 2 displays the results for EBI for Jack. During the pre-EBI card-sorting task Jack incorrectly sorted the pictures of experimental stimuli, suggesting no preexisting class formation (the figure depicting sorting results is available in Online Supporting Information). During EBI Jack responded at below mastery during the pretest for both classes of stimuli and demonstrated mastery of the symmetrical relations following training of the baseline relations (figures including training of baseline relations is available in Online Supporting Information). During tests for equivalence, Jack demonstrated the emergence of all remaining relations (BC, CB). During the EBI posttest however Jack did not demonstrate mastery of all relations. It was hypothesized that the simple-to-complex training protocol disrupted the demonstration of the baseline relations during testing in a mixed-block. Therefore, the experimenter conducted remedial training of the BA and CA relations in the absence of the tests for symmetry and equivalence and then repeated the posttest. During the second administration of the posttest Jack demonstrated all trained and emergent relations at above 94% correct responding. Jack sorted all stimuli correctly during the post-EBI card-sorting task, providing further evidence of class formation. Responses taught during BST maintained at mastery levels during the maintenance probes conducted while EBI was ongoing. Following EBI, Jack demonstrated class-consistent responding in the presence of the remaining untrained dangerous stimuli (i.e., lighter and medicine), nondangerous stimuli (e.g., container and flash drive), and generalization variants during the transfer-of-function test.

For Chrissy, the nodal stimuli used during EBI and trained during BST were lighter and flash drive. During the response pretest (Fig. 1, third panel), Chrissy made contact with all dangerous stimuli. In the presence of nondangerous stimuli (Fig. 1, fourth panel) Chrissy initially did not engage in the desired response, however, on subsequent probes Chrissy remained in the session room and played. Following BST plus IST Chrissy demonstrated the appropriate safety response in the presence of the trained dangerous stimuli (i.e., lighters) and the stay and play response during sessions with trained

Table 3. Caregiver Social Validity Questionnaire Data

Component	Respondents					Average
	1	2	3	4	5	
It is important for preschool aged children to learn safety skills	5	5	5	5	5	5
It is important for my child to learn safety skills	5	5	5	5	5	5
I don't think my child will encounter dangers at this age	1	2	1	1	1	1.2
My child would behave safely if he/she encountered a potential danger	3	3	3	5	3	3.4
My child will become more fearful of the environment if he/she is exposed to dangers during this study	3	4	1	3	3	2.8

Caregiver Social Validity Open-ended Section

Danger	Respondent 1	Respondent 2	Respondent 3	Respondent 4	Respondent 5
Strangers	X	X	X		X
Guns		X	X	X	
Poison/ medicine		X	X	X	
Fire		X	X	X	
Knife			X	X	
Medical emergency				X	
Water safety				X	

Note. (top panel. 1= Do not agree, 3= Neutral, 5= Strongly agree. Medical emergency- one parent indicated they wanted their child to learn what to do if they were alone with a grandparent and that grandparent became sick. (bottom panel) An X indicates the respondent felt that learning a safety response related to that danger was important for their child.

nondangerous stimuli (i.e., flash drives). During the response posttest, Chrissy respond appropriately to all trained and generalization exemplars of the lighter and flash drive, but failed to demonstrate class-consistent responding as evidenced by engaging in the safety response in the presence of both untrained dangerous stimuli (i.e., guns and medicine) and untrained nondangerous stimuli (i.e., hairdryers and containers). Because Chrissy's pattern of responding suggested that appropriate stimulus control over the safety response had not been established for untrained stimuli, EBI was implemented.

Figure 2 displays the results for Chrissy during EBI. During the pre-EBI card-sorting task Chrissy incorrectly sorted the pictures of experimental stimuli., suggesting no preexisting class formation. During EBI Chrissy responded at below mastery during the pretest for both classes of stimuli and demonstrated mastery of the symmetrical relations following training of the baseline relations. During tests for equivalence, Chrissy demonstrated the emergence of all remaining relations (BC, CB). During the EBI posttest Chrissy demonstrated all trained and emergent relations at above mastery criterion. Chrissy sorted all stimuli correctly during the post-EBI card-sorting task, providing further evidence of class formation. Responding maintained at mastery levels during the maintenance probes conducted throughout EBI. During the transfer of function test following EBI, Chrissy engaged in the appropriate responses in the presence of the remaining untrained dangerous stimuli (i.e., gun and medicine), nondangerous stimuli (e.g., container and hairdryer) and generalization variants.

Discussion

Thousands of children die each year as a result of unintentional injuries (CDC, 2017). Identifying time-effective procedures for teaching safety responses to multiple dangerous stimuli is an important avenue for research. Because it is difficult to predict the types of dangers a child may encounter it is essential we identify procedures that reduce the need for direct training to each danger and subsequently the amount of instructional time needed to establish a safety repertoire. The current study provides a proof of concept for the use of EBI to supplement BST plus IST. The current study is a first step towards developing a more efficient procedure for teaching a safety response to multiple dangers.

The findings of the current study contribute to the safety literature in the following ways. First, the current study provides additional evidence that BST plus IST is an effective package for training a safety response to a single danger (e.g., Gatheridge et al., 2004; Gross, Miltenberger, Knudson, Bosch, & Bower Breitwieser, 2007; Himle, Miltenberger, Flessner, & Gatheridge, 2004; Jostad et al., 2008; Lee et al., 2019; Miltenberger et al., 2004). Second, it provides support for the use of EBI to establish physically disparate stimuli as members of the same class (e.g., Albright, Reeve, Reeve, & Kisamore, 2015; Fienup, Covey, & Critchfield, 2010).

In addition, to our knowledge this is the second study to train participants to respond differentially to both dangerous and nondangerous stimuli (Lee et al., 2019). By incorporating experimental stimuli with shared physical characteristics into

Fig. 1 Response scores for Jack and Chrissy in the presence of dangerous (first and third panels) and nondangerous stimuli (second and fourth panels) during pretest, post BST plus IST, posttest, review, and transfer of function. *Black-filled* shapes represent sessions with exemplar 1 of each stimulus, *gray-filled* shapes represent sessions with exemplar 2 of each stimulus, and *open-shapes* represent generalization sessions with untrained exemplars
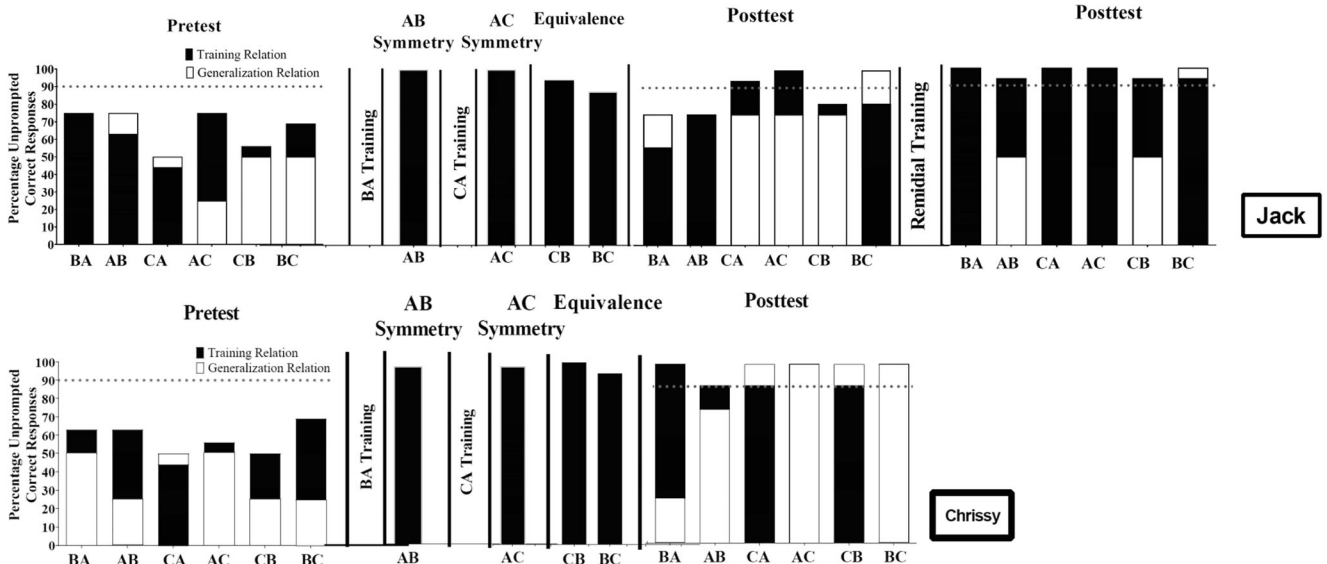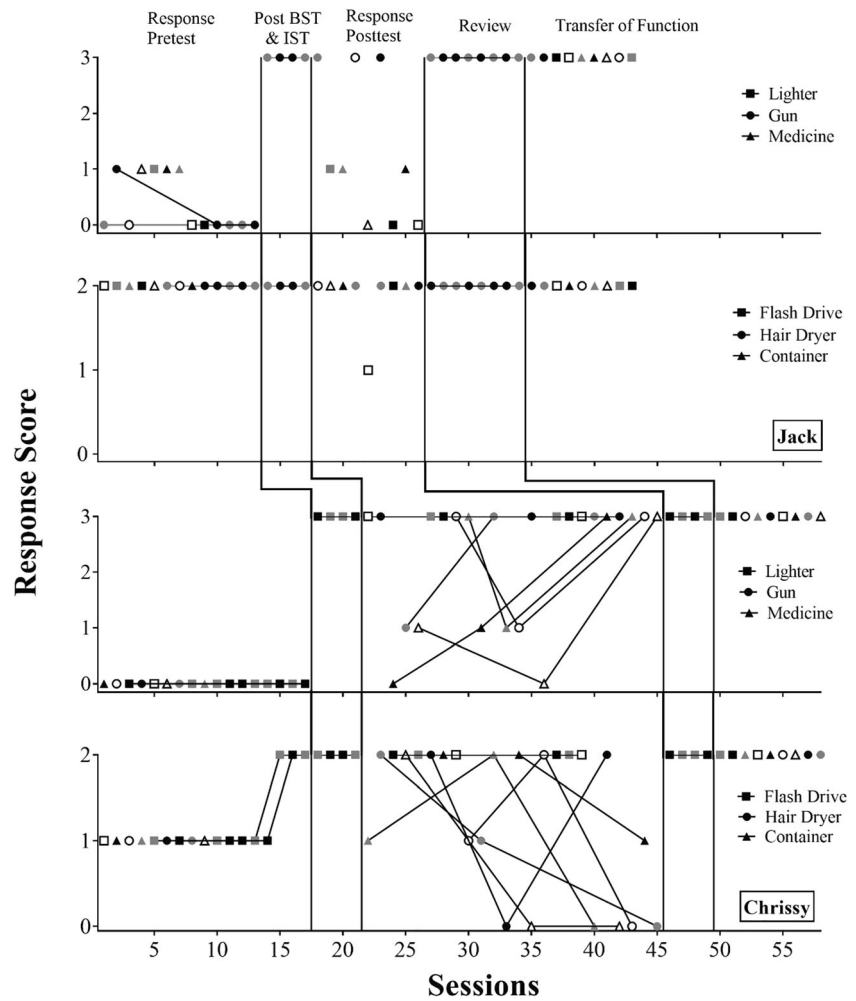

Fig. 2 Unprompted correct responses across EBI pretest, symmetry probes, and posttests for Jack and Chrissy. Black boxes represent scores on training relations. White boxes represent scores on generalization relations. An asterisk denotes that mastery criterion was met during feedback fading

training the current study was able to demonstrate that the safety response was evoked only by those features specific to the dangerous stimuli. Future research should continue to establish discriminated responding as part of safety instruction, as it is essential for consumers to engage in the safety response only under appropriate stimulus conditions. Engaging in the safety response under inappropriate conditions could weaken the strength of the response and impede long-term maintenance.

Furthermore, to our knowledge this is only the second study that has evaluated the extent to which generalization of a safety response to untrained dangers occurred following BST plus IST (Vanselow & Hanley, 2014). Similar to Vanselow and Hanley (2014), the current study found that the combination of BST plus IST for a single danger was not sufficient to produce generalized responding to untrained dangers for Chrissy and Jack. For these two participants, EBI effectively established classes of dangerous and nondangerous stimuli and produced class-consistent responding without the need for BST plus IST for each experimental stimulus. The successful transfer of function observed in this study provides further evidence that the training of a specific stimulus function prior to the implementation of EBI may facilitate transfer of function. However, this is only a circumspect conclusion because we did not conduct a comparison of different transfer of function arrangements. Future studies should compare the efficiency and merits of training a specific stimulus function prior to EBI to training a specific stimulus function following EBI to determine which is best when establishing safety responses to multiple stimuli.

This study extended the equivalence literature in several distinct ways. Of particular note, to our knowledge this is the first applied study to establish a fully elaborated generalized equivalence class (Fields & Moss, 2008) by including the most representative and least representative variants of each member during training. EBI created a shared reinforcement history between these variants and established each as a member of a secondary perceptual class. By incorporating the most and the least representative member of each class during training, we were able to establish an additional median variant stimulus as a member of each perceptual class without the need for direct training during EBI. These variants possessed physical characteristics that fell between the two extremes of the most and least representative stimuli. It is probable that the most and least representative variants of each class member form the boundaries of each perceptual class. It follows that any other stimuli that fall between these boundaries will also be members of the perceptual classes (Fields & Reeve, 2001).

However, unlike in previous studies (e.g., Fields & Moss, 2008) the members of our perceptual classes were not created through digital manipulation and therefore the differences between them are not readily quantifiable. Rather we used a psychometric sort (Rosch, 1975) to assign each of the 10 potential class members a ranking along a continuum. Although this sorting procedure provided a systematic method for selecting the boundary stimuli of each perceptual class, to our knowledge, such a procedure has yet to be evaluated empirically in the extent literature. Because we only evaluated the generality of class-consistent responding in the presence of the midpoint variant we cannot be certain of the size of the perceptual classes that were established in this study. It is possible that each class only contains the three members included in this study. Future researchers should consider conducting a more extensive evaluation of generalization when establishing generalized equivalence classes.

One important consideration when evaluating the results of the current study is that the design was intentionally overengineered to include several components needed to demonstrate experimental control. For example, to verify that generalized responding to untrained dangers had not occurred following BST plus IST the current study included response assessments with every exemplar of every dangerous and nondangerous stimulus. However, in a practical application the response posttest phase may potentially be unnecessary. In addition, following training of each relation during EBI several feedback fading steps were implemented. Feedback fading was included to ensure participants were gradually exposed to decreasing rates of reinforcement prior to the tests for symmetry. It is possible that all steps of the feedback fading procedure may not be required. Although we did not record sessions length these experimental components resulted in the procedures being quite lengthy. Future research is necessary to determine what components can be eliminated from the current experimental arrangement, before the true efficiency of EBI to supplement BST plus IST can be evaluated. Finally, in the current study EBI was conducted using tabletop materials, although this format was the most feasible for study purposes, several studies have used electronic formats that may be desirable for practical applications (Albright et al., 2015; Fienup & Critchfield, 2011).

The current study has several limitations. First, the study included only two participants. Additional direct replications are needed to support the findings of this study and to increase the external validity. Second, the context in which the participant encountered a dangerous stimulus was not varied during the study. Although we chose to keep these contextual variables constant so as to evaluate the generality of the safety response independent from other variables that may affect generalization (e.g., motivation, social contingencies), it should be acknowledged that the safety response taught in this study may not transfer to different, more natural contexts. The conditions under which a child might encounter a dangerous stimulus could be quite complex. For example, a child who is left unattended with a medicine bottle after having observed the parent ingest the medication may be more likely to do the same. A child who has observed a lighter used to light the

candles on a birthday cake may be more likely to play with the lighter if it is left out near candles. However, a recent study found that when a safety response to a gun was taught in a single context, identical to the one used in this study, participants responding generalized from the trained context to other contexts that may represent a range of situations in which a child may encounter a dangerous stimulus (Lee et al., 2019). Though promising, additional research is needed to determine the exact mechanisms behind and limitations of this observed generality.

Third, only two comparison stimuli were included during EBI. Including only two comparison stimuli could lead to falsely concluding that an equivalence class has formed (Sidman, 1987). Sidman (1987) suggested that with only two comparison stimuli, a score as high as 75% is possible without the participant actually acquiring the two classes. To reduce the risk of these types of errors, Sidman suggested setting a mastery criterion of 90% across all relations. Although there is the potential that the two-comparison format lead to participants in the present study forming one class and a null class, it seems unlikely given our inclusion of the mastery criterion suggested by Sidman and participant demonstration of class-consistent responding during the transfer of function test. Furthermore, participants also demonstrated class-consistent responding during the post-EBI card-sorting task (Arntzen et al., 2015). To our knowledge this is the first study to use a card-sorting task to assess the formation of stimulus classes with young children.

One additional avenue for future research might be to evaluate the extent to which new members can be added to the established classes and the subsequent maintenance of those classes. Although previous research has provided evidence for the expansion of equivalence classes (e.g., Saunders et al., 1999), research has yet to evaluate the extent to which the expanded classes maintain. Camargo and Haydu (2015) evaluated the extent to which classes of varying sizes maintained and found that six-member classes were more likely to maintain at a 6-week follow up than three-member classes. Although this provides preliminary evidence that larger classes may result in better maintenance, further evaluation is needed to determine whether there is a break point at which class size negatively affects maintenance of that class.

A final limitation is that the graduate students that completed our social validity measure on the outcomes of the current study were students at the same university where the study was conducted. It is possible that the educational backgrounds of those students may have influenced their responses.

This study provides initial evidence to support the use of EBI to supplement BST plus IST to teach safety skills to young children and demonstrates the formation of a generalized equivalence class of dangerous and nondangerous stimuli. Because we cannot predict what dangers a child may encounter, it seems reasonable to equip children with the training

to respond appropriately to many dangers. This method may prepare children to respond to a variety of dangers without the need for individual danger training. Such training could help prevent some of the thousands of accidental deaths and injuries of children that occur every year.

Compliance with Ethical Standards

Ethical Approval 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 Declaration of Helsinki and its later amendments or comparable ethical standards.

References

Albright, L., Reeve, K. F., Reeve, S. A., & Kisamore, A. N. (2015). Teaching statistical variability with equivalence-based instruction. *Journal of Applied Behavioral Analysis, 48*, 883–894. https://doi.org/10.1002/jaba.249.

Arntzen, E., Norbom, A., & Fields, L. (2015). Sorting: An alternative measure of class formation? *The Psychological Record, 65*, 615–625. https://doi.org/10.1007/s40732-015-0132-5.

Barkley, R. A., & Edelbrock, C. (1987). Assessing situational variation in children's problem behaviors: The Home and School Situations Questionnaires. In R. Prinz (Ed.), *Advances in behavioral assessment of children and families* (pp. 157–176). Greenwich, CT: JAI Press.

Camargo, J., & Haydu, V. B. (2015). Class size effects on the establishment, maintenance, and recovery of equivalence relations. *Behavior Analytic Research & Practice, 15*, 36–49. https://doi.org/10.1037/h0101068.

Carr, J. E., Nicolson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple-stimulus preference assessment in naturalistic context. *Journal of Applied Behavior Analysis, 33*, 353–357. https://doi.org/10.1901/jaba.2000.33-353.

Centers for Disease Control and Prevention, National Center for Injury Prevention and Control (CDC). (2017). Web-based Injury Statistics Query and Reporting System (WISQARS). Retrieved from http://www.cdc.gov/injury/wisqars

De Rose, J. C. McIlvane, W. J., Dube, W. V., & Stoddard, L. T. (1988). Stimulus class formation and functional equivalence in moderately retarded individuals' conditional discrimination. *Behavioral Processes, 17*, 167–175. doi:https://doi.org/10.1016/0376-6357(88)90033-2

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). San Antonio, TX: Pearson Education.

Dymond, S., & Rehfeldt, R. A. (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst, 23*, 239–254. https://doi.org/10.1007/bf03392013.

Fields, L., & Moss, P. (2008). Formation of partially and fully elaborated generalized equivalence classes. *Journal of Experimental Analysis of Behavior, 90*, 135–168. https://doi.org/10.1901/jeab.2008.90-135.

Fields, L., & Reeve, K. F. (2001). A methodological integration of generalized equivalence classes, natural categories, and cross-modal perception. *The Psychological Record, 51*, 67–87. https://doi.org/10.1007/bf03395387.

Fields, L., Reeve, K. F., Adams, B. J., & Verhave, T. (1991). Stimulus generalization and equivalence classes: A model for natural categories. *Journal of Experimental Analysis of Behavior, 55*, 305–312. https://doi.org/10.1901/jeab.1991.55-305.

Fienup, D. M., Covey, D. P., & Critchfield, T. S. (2010). Teaching brain-behavior relationships economically with stimulus equivalence technology. *Journal of Applied Behavior Analysis, 43*, 19–33. https://doi.org/10.1901/jaba.2010.43-19.

Fienup, D. M., & Critchfield, T. S. (2011). Tanty of equivalence-based programmed instruction: Efficacy and efficiency in a college classroom. *Journal of Applied Behavior Analysis, 44*, 435–540. https://doi.org/10.1901/jaba.2011.44-435.

Friedman-Weieneth, J. L., Doctoroff, G. L., Harvey, E. A., & Goldstein, L. H. (2009). The Disruptive Behavior Rating Scale-Parent Version (DBRS-PV): Factor analytic structure and validity among young preschool children. *Journal of Attention Disorders, 13*, 42–55. https://doi.org/10.1177/1087054708322991.

Gatheridge, B. J., Miltenberger, R. G., Huneke, D. F., Satterlund, M. J., Matten, A. R., Johnson, B. M., & Flessner, C. A. (2004). Comparison of two programs to teach firearm injury prevention skills to 6- and 7-year-old children. *Pediatrics, 114*, 294–299. https://doi.org/10.1542/peds.2003-0635.

Giannakakos, A. R., Vladescu, J. C., Kisamore, A. N., Reeve, K. F., & Fienup, D. M. (2019). A review of the literature on safety response training. *Journal of Behavioral Education.* Advance online publication. https://doi.org/10.1007/s10864-019-09347-4.

Gross, A., Miltenberger, R., Knudson, P., Bosch, A., & Bower Breitwieser, C. B. (2007). Preliminary evaluation of a parent training program to prevent gun play. *Journal of Applied Behavior Analysis, 40*, 691–695. https://doi.org/10.1901/jaba.2007.691-695.

Harvey, M. T., May, M. E., & Kennedy, C. H. (2004). Nonconcurrent multiple baseline designs and the evaluation of educational systems. *Journal of Behavioral Education, 13*, 267–276. https://doi.org/10.1023/B:JOBE.0000044735.51022.5d.

Himle, M. B., Miltenberger, R. G., Flessner, C., & Gatheridge, B. (2004). Teaching safety skills to children to prevent gun play. *Journal of Applied Behavior Analysis, 37*, 1–9. https://doi.org/10.1901/jaba.2004.37-1.

Imam, A. A. (2006). Experimental control of nodality via equal presentation of conditional discriminations in different equivalence protocols under speed and no-speed conditions. *Journal of Experimental Analysis of Behavior, 85*, 107–124. https://doi.org/10.1901/jeab.2006.58-04.

Jostad, C. M., Miltenberger, R. G., Kelso, P., & Knudson, P. (2008). Peer tutoring to prevent firearm play: Acquisition, generalization, and long-term maintenance of safety skills. *Journal of Applied Behavior Analysis, 41*, 117–123. https://doi.org/10.1901/jaba.2008.41-117.

Lee, N., Vladescu, J. C., Reeve, K. F., Peterson, K. M., & Giannakakos, A. R. (2019). Effects of behavioral skills training on the stimulus control of gun safety responding. *Journal of Behavioral Education, 28*, 187–203. https://doi.org/10.1007/s10864-018-9309-8.

Miltenberger, R. G., Flessner, C., Gathridge, B., Johnson, B., Satterlund, M., & Egmo, K. (2004). Evaluation of behavioral skills training to prevent gun play in children. *Journal of Applied Behavior Analysis, 37*, 513–516. https://doi.org/10.1901/jaba.2004.37-513.

Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: a guide for practitioners. *Behavior Analysis and Practice, 5*, 2–11. https://doi.org/10.1007/BF03391819.

Rosales, R., Stone, K., & Rehfeldt, R. A. (2009). The effects of behavioral skills training on implementation of the Picture Exchange Communication System. *Journal of Applied Behavior Analysis, 42*, 541–549. https://doi.org/10.1901/jaba.2009.42-541.

Rosch, E. (1975). Cognitive representation of semantic categories. *Journal of Experimental Psychology, 104*, 192–233. https://doi.org/10.1037/0096-3445.104.3.192.

Rossi, M. R., Vladescu, J. C., Reeve, K. F., & Gross, A. C. (2017). Teaching safety responding to children with autism spectrum disorder. *Education & Treatment of Children, 40*, 187–208. https://doi.org/10.1353/etc.2017.0009.

Saunders, R. R., Drake, K. M., & Spradlin, J. E. (1999). Equivalence class establishment, expansion, and modification in preschool children. *Journal of Experimental Analysis of Behavior, 71*, 195–214. https://doi.org/10.1901/jeab.1999.71-195.

Sidman, M. (1987). Two choices are not enough. *Behavior Analysis, 22*, 11–18.

Summers, J., Tarbox, J., Findel-Pyles, R. S., Wilke, A. E., Bergstrom, R., & Williams, W. L. (2011). Teaching two household safety skills to children with autism. *Research in Autism Spectrum Disorder, 5*, 629–632. https://doi.org/10.1016/j.rasd.2010.07.008.

Taylor, I., & O'Reilly, M. F. (2000). Generalization of supermarket shopping skills for individuals with mild intellectual disabilities using stimulus equivalence training. *The Psychological Record, 50*, 49–62. https://doi.org/10.1007/BF03395342.

Vanselow, N. R., & Hanley, G. (2014). An evaluation of computerized behavior skills training to teach safety skills to young children. *Journal of Applied Behavior Analysis, 47*, 51–69. https://doi.org/10.1002/jaba/105.

Walker, B. D., & Rehfeldt, R. A. (2012). An evaluation of the stimulus equivalence paradigm to teach single-subject design to distance education students via Blackboard. *Journal of Applied Behavior Analysis, 45*, 329–344. https://doi.org/10.1901/jaba.2012.45-329.

Williams, K. T. (2007). *Expressive vocabulary test* (2nd ed.). San Antonio, TX: Pearson Education.