

AN EVALUATION OF TOILET-TRAINING PROCEDURES

BY

BRIAN D. GREER

Submitted to the graduate degree program in Applied Behavioral Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

---

Chairperson Pamela L. Neidert, Ph.D., BCBA-D

---

Claudia L. Dozier, Ph.D., BCBA-D

---

Derek D. Reed, Ph.D., BCBA-D

---

Edward R. Christophersen, Ph.D., ABPP, FAAP (Hon)

---

Barbara Thompson, Ph.D.

Date Proposed:  
7/12/2013

The Dissertation Committee for BRIAN D. GREER  
certifies that this is the approved version of the following dissertation:

AN EVALUATION OF TOILET-TRAINING PROCEDURES

---

Chairperson Pamela L. Neidert, Ph.D., BCBA-D

Date Approved:  
7/12/2013

## **Abstract**

Although systematic replications of Azrin and Foxx's (1971) procedures have proven extremely effective across a variety of populations and settings, the majority of behavioral toilet-training research has relied on complex multicomponent training packages. Therefore, little is known regarding the effectiveness of individual toilet-training components. In Study 1, we investigated the combined and individual effects of three commonly used components: (a) underwear, (b) a dense schedule of sits on the toilet, and (c) differential reinforcement. When all three components were combined, we observed overall improvements in toileting performance for five of six children. We observed overall improvements for two of four children exposed to only the underwear component. Overall improvements were not observed for any child exposed to only the dense-sit schedule component or to only the differential-reinforcement component. Study 2 was designed to determine whether training components that were ineffective when used in isolation add to treatment efficacy when combined with effective training procedures. The combination of the differential-reinforcement component with the underwear component produced no overall improvements in performance beyond gains observed when underwear was used alone.

## **Acknowledgements**

I would like to acknowledge all of those who helped me in completing this dissertation. I am particularly indebted to Dr. Pamela Neidert for her constant support and guidance, as well as to Dr. Claudia Dozier. It has been an absolute honor to work with both of you over the last five years. I am grateful to work everyday with an excellent group of behavior analysts of the highest caliber. I would like to thank all of my fellow graduate students for always allowing me to bounce ideas around and argue over seemingly minute study details. I would like to thank Dr. Derek Reed for his assistance with data analysis on this and other projects. Above all else, I would like to acknowledge my wife, Savanna. She is a source of never-ending encouragement and support.

## Table of Contents

Introduction.....	1
General Method.....	13
Study 1: Component Analysis of Commonly Used Toilet-Training Procedures.....	19
Subjects and Setting.....	19
Response Measurement and Interobserver Agreement.....	20
Procedure.....	23
Baseline.....	25
Toilet-Training Package.....	26
Underwear.....	27
Dense-Sit Schedule.....	28
Differential Reinforcement.....	28
Underwear.....	29
Dense-Sit Schedule (FT Sits).....	30
Differential Reinforcement.....	30
No-Treatment Probe.....	31
Data Analysis.....	33
Results and Discussion.....	35
Study 2: Determining Additive Effects of Differential Reinforcement.....	52
Subjects and Setting.....	53
Response Measurement and Interobserver Agreement.....	54
Procedure.....	54
Baseline.....	54

Underwear.....	55
Underwear + Differential Reinforcement.....	55
No-Treatment Probe.....	55
Results and Discussion.....	56
General Discussion.....	58
References.....	65
Appendices.....	69
Table 1.....	69
Figure 1.....	70
Figure 2.....	71
Figure 3.....	72
Figure 4.....	73
Figure 5.....	74
Figure 6.....	75
Figure 7.....	76
Figure 8.....	77
Table 2.....	78
Table 3.....	79
Table 4.....	80
Figure 9.....	81

## **An Evaluation of Toilet-Training Procedures**

Behavioral psychology (i.e., behavior analysis) has had a sizeable impact on toilet-training practices in the United States, and the work published by Azrin and Foxx (1971) is arguably the most eminent behavior-analytic research in toilet training. The effects of toilet-training procedures adapted from Azrin and Foxx's seminal work have been replicated countless times across multiple populations, and empirical investigations of what has become known as the "intensive-training approach" continue.

Despite its widespread influence, it is important to note that Azrin and Foxx (1971) was not the first toilet-training publication in behavior analysis. Rentfrow and Rentfrow (1969) reviewed a number of notable toilet-training studies prior to Azrin and Foxx's work. Early toilet-training procedures were based on operant (and sometimes respondent) conditioning and were aimed at improving the toileting skills of institutionalized children and adults with intellectual and developmental disabilities (IDD). The primary purpose of teaching toileting skills to institutionalized individuals was to improve the health conditions (e.g., decrease levels of dysentery and intestinal infections) in the institutions in which they lived (Rentfrow & Rentfrow, 1969). Early toilet-training studies were essentially management-minded interventions to improve health conditions, decrease soiled laundry, and decrease the amount of time staff members were required to care for residents. Toilet training also allowed institutionalized individuals the ability to engage in previously inaccessible and potentially more preferred and stimulating activities. However, this aspect of toilet training seemed to be of secondary importance to most early toilet-training researchers.

When direct measures of toileting performance were collected, early toilet-training research typically demonstrated improvements in toileting performance (e.g., Giles & Wolf,

1966); however, performance gains were often lost once treatment was removed, and long-term maintenance data were often excluded (Rentfrow & Rentfrow, 1969). Rentfrow and Rentfrow (1969) noted that researchers often attributed toileting gains observed in early toilet-training studies to environmental enrichment and improved staff-client ratios rather than to the toilet-training interventions themselves. Prior to Azrin and Foxx (1971), researchers had not developed a powerful set of toilet-training procedures able to exert strong and lasting improvements in toileting performance.

In their pioneering work, Azrin and Foxx (1971) taught nine institutionalized adults with IDD to remain dry (i.e., have fewer accidents) using a multicomponent training program that included: (a) fluid loading, in which beverages were provided every 30 min; (b) scheduled sits on the toilet every 30 min; (c) a differential reinforcement of other (DRO) behavior contingency for remaining dry; (d) an alarm that signaled when appropriate eliminations occurred; (e) praise and edible reinforcers for eliminating on the toilet; (f) access to a preferred chair when not sitting on the toilet; (g) an alarm that signaled the occurrence of an accident; (h) overcorrection following accidents; (i) timeout from fluids, edibles, social interaction, and the preferred chair following accidents; and (j) a prompting procedure used to teach undressing and dressing skills. Clearly, the authors sought to demonstrate the utility and power of the newly emerging field of applied behavior analysis, rather than unveil the behavioral processes by which toilet training occurs. The mean number of daily accidents for the group of individuals decreased immediately following implementation of the toilet-training program, and low-to-zero levels of accidents maintained for up to 140 days following the introduction of the toilet-training program.

The results of Azrin and Foxx (1971), as well as several systematic replications, indicated that immediate and long-term improvements in toilet-training performance could be achieved



using a multicomponent-training approach. Training components similar to those used by Azrin and Foxx (1971) have been combined in different ways to form comparable toilet-training programs, and these training “packages” have been used in the majority of subsequent toilet-training research (for a recent review, see Kroeger & Sorensen-Burnworth, 2009). Such toilet-training programs have shown to be extremely effective at decreasing the frequency of accidents and increasing appropriate skills (e.g., eliminating in the toilet, dressing oneself, independently requesting to sit on the toilet) in a relatively short period of time (Cicero & Pfadt, 2002). For this reason, the intensive-training approach has become the standard toilet-training intervention used with individuals diagnosed with IDD (Cicero & Pfadt, 2002).

Toilet-training procedures adapted from Azrin and Foxx’s (1971) seminal work have been replicated numerous times across multiple populations. Systematic replications of Azrin and Foxx (1971) have been successfully extended to a range of populations including older children with IDD (Azrin, Bugle, & O’Brien, 1971; Mahoney, Van Wagenen, & Meyerson, 1971; Sells-Love, Rinaldi, & McLaughlin, 2002), children with Autism Spectrum Disorder (Ando, 1977; Cicero & Pfadt, 2002; Kroeger & Sorensen, 2010; LeBlanc, Carr, Crossett, Bennett, & Detweiler, 2005), adolescents with psychological disorders (Boles, Roberts, & Vernberg, 2008; Friman & Vollmer, 1995), typically developing children (Azrin & Thienes, 1978; Foxx & Azrin, 1973; Hansen, 1979; Kimmel & Kimmel, 1970; Simon & Thompson, 2006), infants (Mahoney et al., 1971; Smeets, Lancioni, Ball, & Oliva, 1985), and older adults (Atkins & Mathews, 1997; Burgio et al., 1990; Schnelle et al., 1983; Spangler, Risley, & Bilyew, 1984). The initial extension to typically developing children (Foxx & Azrin, 1973) was so promising, a parent-friendly version of the intensive-training procedure was written by Azrin and Foxx (1974). Even parents have been taught to toilet train their children (Feldman et al., 1992).

However, future research should more closely examine the effectiveness of toilet-training procedures used with populations not often targeted in current toilet-training research. As highlighted in the review by Kroeger and Sorensen-Burnworth (2009), the vast majority of single-subject research on toilet training has been conducted with children and adults with IDD. Subject characteristics may dictate the need for alternative procedures or simply require comparable procedures be implemented at differing levels of intrusiveness. Future research with typically developing children may yield useful information regarding whether some considerations should be weighed more heavily when selecting toilet-training procedures for individuals without IDD.

Regardless of subject characteristics, most toilet-training research has prescribed that individuals sit on the toilet at scheduled times throughout the toilet-training process. Programmed toileting opportunities are designed to increase the likelihood that eliminations will occur while on the toilet without requiring individuals to engage in independent responses (i.e., self-initiate) to access the toilet. Caregivers can then provide reinforcement for appropriate eliminations more often than had they waited for the child to independently initiate a trip to the bathroom and then appropriately eliminate. Caregivers could presumably toilet train individuals more quickly by frequently placing individuals on the toilet (especially around times the child is likely to eliminate). However, there is no standard schedule by which caregivers should require individuals to sit, nor is there a standard duration of how long individuals should be required to remain on the toilet.

Some studies have used dense schedules of sits on the toilet. For example, Azrin and Foxx (1971) prompted subjects to sit on the toilet every 30 min for 20 min at a time. Foxx and Azrin (1973) prompted subjects to sit on the toilet every 10 min, and Azrin and Foxx (1974)

suggested caregivers sit their children every 15 min. Simon and Thompson (2006) prompted typically developing children to sit on the toilet every 30 min for 3 or 5 min. Other studies (especially those involving geriatric subjects) have used leaner sit schedules. For example, Atkins and Mathews (1997) demonstrated decreases in the amount (i.e., volume) of urine measured during accidents for two older adults with Alzheimer's disease when subjects sat on the toilet every hour. Unfortunately, a lack of experimental control in Atkins and Mathews prevents a determination of whether the prompted sits or some extraneous variable influenced subjects' toileting performance. Spangler, Risley, and Bilyew (1984) asked nursing home residents if they needed assistance using the toilet every 90 min and showed a decrease in accidents across residents. Schnelle et al. (1983) used a similar method with nursing home residents, but asked if assistance was needed every hour, and showed improvements in toileting performance across groups of residents. Burgio et al. (1990) prompted four residents to use the toilet four to seven times per day and demonstrated increases in the number of checks during which each subject was found to be dry. Given the wide range of sit schedules used in toilet-training research, future research could expose individuals to a range of sit schedules to identify both the optimal schedules and durations that facilitate toilet training.

In addition to prompting individuals to sit on the toilet according to a programmed schedule, most toilet-training studies have used a similar schedule to check individuals' undergarments. The rationale for undergarment checks is similar to the rationale for using scheduled sits: to increase the number of times caregivers can differentially respond to important toileting-related behaviors, presumably expediting the toilet-training process. Reinforcement is typically provided for being dry during a scheduled undergarment check (e.g., Azrin & Foxx, 1974), and the consequences arranged for accidents (having wet undergarments) have ranged

from telling children they are wet (e.g., Simon & Thompson, 2006) to spanking (e.g., Ando, 1977). Similar to the range of schedules used for programmed sits on the toilet, schedules used for undergarment checks have also varied. Azrin and Foxx (1971) described a one-hour check schedule; however, preferred items were delivered every 5 min in which the subject remained dry. Simon and Thompson (2006) checked children's undergarments every 15 min. Time between undergarment checks tend to be much longer when assisting older adults. For example, Schnelle et al. (1983) checked residents' undergarments every hour, and Spangler et al. (1984) conducted undergarment checks every 90 min. In addition to programmed sits on the toilet and scheduled undergarment checks, researchers have found that undergarment type (e.g., diapers, pull-ons, underwear, etc.) is a predictor of toilet-training success.

Two recent toilet-training publications examined how undergarment type affects toilet-training performance. Tarbox, Williams, and Friman (2004) demonstrated that diaper use increased the frequency of urinary accidents and decreased the frequency of eliminations on the toilet when compared to a no-diaper condition for one adult male with mental retardation. In a systematic replication, Simon and Thompson (2006) found that placing typically developing toddlers in underwear facilitated toilet training. Reductions in urinary accidents and increases in eliminations on the toilet were observed for two of five children when underwear was used. Interestingly, a comparison of diapers and pull-on training pants showed no improvement for any subject, and pull-ons were correlated with reduced performance for one subject. Collectively, the results of Tarbox et al. (2004) and Simon and Thompson (2006) indicate that the selection of undergarment type is important in toilet training, and that pull-on training pants may not facilitate acquisition of toileting skills.

Future investigations might evaluate the effects of other types of undergarments (e.g., cloth diapers, diapers that cool when in contact with liquid, etc.) on toileting performance. Also, an investigation of the behavioral mechanisms involved in the effectiveness of undergarment type might allow for better training procedures. That is, although placing children in underwear is procedurally an antecedent manipulation, underwear likely functions as a consequent (i.e., punishment) manipulation. For example, a child's accident is likely to be more salient to caregivers when the child is wearing underwear. Therefore, the latency with which caregivers change a child who has recently had an accident is likely decreased for children wearing underwear, and the changing process may be aversive for some children. Changes in toileting performance could result in children for whom the changing process is highly aversive.

Alternatively, underwear may facilitate toilet training via positive punishment as wetting oneself is likely aversive for some children. Capitalizing on the aversive properties of having an accident while wearing underwear, Giles and Wolf (1966) required children to remain in their soiled undergarments after having an accident. Unfortunately, the authors did not report when children wore diapers or underwear, making it impossible to determine whether requiring children to remain in soiled undergarments improved subjects' toileting performance. Therefore, although recent research has demonstrated that placing children in underwear is an effective method of toilet training some children (Simon & Thompson, 2006; Tarbox, Williams, & Friman, 2004), it remains unclear precisely how undergarment type affects toileting performance.

Differential reinforcement is perhaps the most common training component included across studies and typically involves the delivery of preferred edibles, drinks, leisure items, or praise. Often, a preference assessment is conducted to identify stimuli likely to function as

reinforcers for behaviors targeted by a toilet-training program. Simon and Thompson (2006) conducted a multiple stimulus without replacement (MSWO) preference assessment (DeLeon & Iwata, 1996) to select preferred items. Cicero and Pfadt (2002) conducted an informal preference assessment in which the authors assessed subject affect while interacting with items, the likelihood of compliance with demands when the item was used for compliance, and whether the subject requested access to the item to identify preferred items for three subjects. Boles, Roberts, and Vernberg (2008) used a reinforcer menu combined with tokens to treat one child's incontinence. Concerned with the possibility of large amounts of reinforcement delivery, Azrin and Foxx (1974) suggested that caregivers provide a variety of preferred beverages and foods to prevent reinforcer satiation and to serve as an establishing operation for voiding (i.e., fluid loading).

A close inspection of the reinforcement schedules typically programmed in toilet-training studies often reveals at least two types of reinforcement contingencies, and researchers do not often describe each reinforcement contingency explicitly. The first type of reinforcement contingency used in toilet-training research is differential reinforcement of alternative (DRA) behavior, most commonly arranged for appropriate eliminations. Azrin and Foxx (1971) provided an edible item and praise contingent on eliminations in the toilet. LeBlanc, Carr, Crossett, Bennett, and Detweiler (2005) provided access to highly preferred beverages, foods, and toys following eliminations in the toilet. Similarly, Simon and Thompson (2006) provided praise and a preferred edible item following appropriate eliminations.

Researchers have also implemented differential reinforcement of other (DRO) behavior for the nonoccurrence of accidents, which are typically detected at scheduled undergarment checks. In most cases, the stimuli (e.g., edible items, beverages, praise, etc.) delivered for

completion of the DRA contingency are also provided for completion of the DRO contingency. For example, in addition to providing edibles and praise for appropriate eliminations, Azrin and Foxx (1971) also delivered edibles and praise every 5 min in which subjects remained dry. Additionally, subjects received a drink every 30 min and maintained access to a preferred chair so long as accidents were not detected.

It is not clear whether providing the same stimuli for completion of the DRA and DRO components is best practice. Azrin and Foxx (1974) were clearly concerned with reinforcer satiation limiting toilet performance; however, the authors suggested that parents use the same stimuli for appropriate eliminations and for remaining dry at scheduled checks. Given this suggestion, reinforcement (although varied) would have likely been provided on a rather dense schedule, and reinforcer satiation could limit improvements in toileting performance. Therefore, it is possible that providing separate preferred items for the DRA and DRO contingencies would yield more rapid skill acquisition. Future research could evaluate this possibility. It is also possible that one contingency (DRA or DRO) is generally more influential in toilet training. Future research could also examine this possibility.

As previously mentioned, the majority of toilet-training studies using the intensive-training approach have focused on examining the effectiveness of complex, multicomponent toilet-training packages. The immediate and lasting improvements in toileting performance observed across a wide range of populations and settings is likely why multicomponent training packages have been evaluated extensively in the toilet-training literature. Unfortunately, the collective findings of these studies have offered limited information about the effectiveness of the individual procedures that comprise toilet-training programs. Some toilet-training components are likely to be more influential than other components. Other training components

may be ineffective, effective only when combined with other components, or even contraindicated. For example, if a dense schedule of prompting a child to sit on the toilet becomes aversive to the child, a certain degree of countercontrol might be expected. If the child comes to avoid scheduled sits on the toilet, the entire toilet-training process may become more difficult for caregivers and the child, and thus may prolong the toilet-training process. Therefore, it seems important to know the effects of individual toilet-training components that comprise toilet-training packages so that only those components that contribute to improving toileting performance are implemented, thereby improving effectiveness and efficiency.

Component analysis is one method by which researchers can identify those components that do and do not contribute to toilet training. Ward-Horner and Sturmey (2010) described two types of component analyses. In the “dropout method,” a training program is implemented with all components at full strength, and single components are subsequently withdrawn contingent on stable responding to determine the relative contribution of the removed component. In the “add-in method,” one component is introduced contingent on stable baseline performance, and subsequent components are introduced cumulatively to determine the additive value of each training component. For the purposes of conducting a component analysis of commonly used toilet-training procedures, the “add-in method” seems most appropriate. Once toileting-related skills are acquired, those skills may quickly generalize to natural sources of reinforcement, thereby complicating demonstrations of experimental control. Therefore, toileting performance may quickly improve and never diminish when using the “drop-out method,” preventing conclusions regarding whether the removed components were effective. However, the “add-in method” would allow researchers to examine toileting performance with the addition of each new component, and any improvements observed across children when particular components



are introduced would indicate an effective component. Additionally, the order of evaluated components should be counterbalanced across children to examine interactions between components and possible order effects (e.g., component B is more effective when added last).

A component analysis of commonly used toilet-training procedures would also require the inclusion of more subjects than are typically targeted in most single-subject research studies, as demonstrations of experimental control would be more safely attempted across children. That is, within-subject reversals in toileting performance may not occur reliably for all children. However, some researchers (Friman & Vollmer, 1995; Nordquist, 1971; Sells-Love et al., 2002; Taylor, Cipani, & Clardy, 1994) have found success using reversal designs to demonstrate experimental control of toilet-training procedures. However, anticipating reversals in toileting performance could be risky, as failing to reverse toileting performance would jeopardize experimental control. A more cautious approach to experimental design is to demonstrate experimental control across subjects while also attempting to demonstrate control within subject. Therefore, multiple baseline and reversal designs could be combined to yield a potentially powerful experimental design that demonstrates control across subjects while attempting to demonstrate control within subject. Additionally, elements of a multiple probe design could also be incorporated to improve the likelihood of demonstrating experimental control within subject by rapidly reversing toileting procedures to determine performance levels when treatment components are removed. Also, a multiple probe design would not require extended exposure to no-treatment conditions.

Regardless of the experimental design employed, a component analysis of commonly used toilet-training components would first require numerous demonstrations of the effectiveness of the combined toilet-training package prior to evaluations of its constituent components. That

is, evaluating the components of an ineffective toilet-training package makes little sense (unless there is reason to believe that the package could be improved by removing components that have a negative impact on toileting performance). The effectiveness of each component of the toilet-training package could be evaluated in isolation across children following repeated demonstrations of the effectiveness of the combined toilet-training package. Additional components could then be added, yielding new combinations of toilet-training components. Also, counterbalancing procedures could be used to evaluate interaction effects (e.g., components A + B yield better results than A + C or B + C) and possible order effects (e.g., one progression of added components proves more effective than other progressions). Ultimately, a better understanding of the effects of combined toilet-training components is needed to understand how complex toilet-training programs affect toileting performance. Additionally, a better understanding of the behavioral processes involved when effective toilet-training components are implemented could allow for a better behavioral understanding of toilet training as a process and inform the development of other more effective and efficient toilet-training programs.

Future research could conduct parametric analyses to “fine tune” toilet-training components to make ineffective components effective or determine a balance between effectiveness and efficiency by evaluating the “dose” at which components should be implemented. The integrity with which toilet-training procedures are implemented is also important in determining whether less-than-effective procedures are the result of an ineffective toilet-training program or inadequate implementation. The critical treatment-integrity value with which toilet-training components must be implemented for its effects to be clinically significant should also be evaluated by future research. More robust and efficient toilet-training programs

could be developed from the results from such evaluations. However, a component analysis of toilet-training procedures would identify those components worthy of further investigation.

The purpose of Study 1 was to evaluate the individual and combined effects of three commonly used toilet-training components: (a) underwear, (b) a dense-sit schedule, and (c) differential reinforcement on teaching toilet-training skills to typically developing children using the “add-in method” of component analysis. Study 2 was designed to further evaluate the combined effects of two toilet-training components to determine the additive effectiveness of an additional training component.

### **General Method**

The first classroom served toddlers between the ages of one to two-and-a-half years. The minimum teacher-child ratio in this classroom was 1:5 with a maximum child enrollment of 12. The second classroom served preschool children between the ages of two-and-a-half to four years. The minimum teacher-child ratio in the second classroom was 1:10 with a maximum child enrollment of 20. Both classrooms served children with and without IDD and operated Monday through Friday from 7:45 a.m. to 5:30 p.m. The daily operating time for each classroom was divided into three distinct classroom shifts: morning, midday, and afternoon. The morning and afternoon shifts were four hours in duration, and the midday shift was three hours in duration. Prior to and immediately following each shift, the classroom supervisors conducted 15-min feed-forward and feedback times in which important classroom and child information (including any changings in toileting programs) was discussed and reviewed with teachers on that shift. The third classroom served children with IDD between the ages of two-and-a-half to eight years. A 1:1 teacher-child ratio was used in this classroom. This classroom also operated Monday through Friday; however, the hours of operation were 9 a.m. to 3 p.m. Because of the

reduced operating hours, the third classroom only had two 4-hour classroom shifts: morning and afternoon. Classroom supervisors of the third classroom also conducted 30-min feed-forward and feedback times.

In each classroom, the toileting area was equipped with child-sized fixtures to promote toileting independence. Toilets and sinks were located within each child's reach. Paper towel and soap dispensers were also located within each child's reach with the use of a step stool in the toileting area. Toilet seat covers were used for smaller children to ensure proper positioning on the toilet. The toileting area in each classroom was physically adjacent to, but within the physical boundaries of, the general classroom area; however, a low barrier (operated by classroom teachers) prevented independent entry into the toileting area by children in the toddler classroom. This classroom modification was a safety requirement designed to prevent the unsupervised access of young children to the toileting area. In this classroom, classroom teachers permitted child access to the toileting area during scheduled checks of the child's undergarments and anytime the child independently said, or signed, "potty" or "bathroom." In the other two classrooms (serving older children or children with special needs), no barrier restricted children's access to the toileting area. Thus, children in these classrooms could move freely between the general classroom area and the toileting area.

Three-step prompting (also referred to as guided compliance or least-to-most prompting) was a general classroom procedure that was in place in each classroom prior to the start of the study, and its use continued throughout the study. Three-step prompting is a structured prompting procedure commonly used in early education classrooms and has been shown to be effective at teaching compliance to typically developing children (Wilder & Atwell, 2006), as well as children with developmental disabilities (Tarbox, Wallace, Penrod, & Tarbox, 2007).

The procedure consisted of a series of prompts arranged temporally in a least-to-most intrusive sequence. Instructions (vocal prompts) were restated with supplemental prompts until the child emitted the desired response. If compliance did not occur within a set period of time (e.g., 3 to 5 s) following a teacher's initial instruction, the instruction was restated with a model prompt. If compliance did not occur following the instruction and model, the instruction was repeated and physical guidance was used. Classroom teachers were instructed to provide praise if compliance occurred following the vocal or model prompt. Classroom teachers in the current study used three-step prompting as a general teaching procedure throughout the day to teach compliance. Three-step prompting was used to teach each child to say or sign "potty" or "bathroom" and to walk to the toileting area of the classroom. Additionally, three-step prompting was used to guide children through the general toileting process (e.g., pulling down pants, sitting down, wiping, standing up, pulling up pants, and washing hands).

Two multiple-stimulus-without-replacement (DeLeon & Iwata, 1996) preference assessments were conducted weekly with each subject to identify preferred edible and leisure items. One edible and one leisure preference assessment was conducted each week. Weekly edible preference assessments consisted of small food items (e.g., crackers, chips, and candies), and weekly leisure preference assessments consisted of small easily deliverable toys (e.g., toys that emitted noises or lights, dolls, small balls, necklaces, etc.). The items selected for inclusion in each child's preference assessment remained constant throughout the study. Occasionally, items were lost or misplaced, in which case duplicate items were used. If a duplicate item was not available, a similar item was substituted. Each preference assessment was conducted either in each child's classroom (in an area separate from other activities) or in a small therapy room.

The top two edible and top two leisure items from each child's preference assessment were selected for use for the upcoming week. Therefore, four items were selected for use each week, and those items were placed in the child's bin. Children's bins were labeled with each child's name and placed in the toileting area for quick access by the classroom teachers. Preferred items for that week remained stocked in each child's bin for the duration of the week. The contents of each child's bin changed each week and were dependent on the results of the weekly preference assessments. In rare cases, the same item was inadvertently used in the preference assessment arrays of more than one child. In these cases, two of the top three items were used if both children selected the same item as their first or second most preferred.

Subjects could gain access to the first two items selected during each of the most recent edible and leisure preference assessments throughout the following week according to the toileting condition in effect for each child. When edible items were delivered, one of each of the two edible items was provided following each response that met the delivery requirement outlined by the current toileting condition. Leisure items were also delivered when edible items were presented. Each of the two leisure items was provided for 30-s following each response that met the delivery requirement outlined by the current toileting condition. Two edible items and 30-s access to two leisure items were provided following each response that met criterion. This procedure for the delivery of preferred items remained constant across all conditions of the study. However, the contingencies arranged for preferred-item delivery varied across conditions.

Teachers in each classroom performed two general types of scheduled toileting routines with each subject: undergarment checks and scheduled sits on the toilet. Classroom teachers performed undergarment checks and scheduled sits on toilet throughout the day for each child. The only classroom activity that interfered with undergarment checks or scheduled sits on the

toilet was naptime. Sleeping was not disrupted for undergarment checks or scheduled sits.

However, children who were awake during naptime followed all toileting procedures. Children who slept during naptime were checked and prompted to sit on the toilet once they awoke.

To perform an undergarment check, classroom teachers approached the subject and provided a three-step prompt for the child to say “potty” or “bathroom” and walk to the toileting area. Younger children were then placed on a changing table used to assist with changing undergarments. Older children were permitted to stand, or a changing mat placed on the floor or the toileting area was used. Children using either the changing table or a changing mat were prompted to lie down on their backs. The child’s undergarments were then checked. Children who were found to be dry were permitted to wash their hands and return to the current classroom activity. Scheduled undergarment checks were typically completed within a few minutes if children were dry, such that children typically returned to the same activity in place prior to the undergarment check. Children who were wet or had soiled undergarments were changed into the same type of undergarment (e.g., children wearing diapers were changed into clean diapers) with minimal teacher attention. Teachers were instructed to tell the child “*child’s name*, you had an accident” and minimize the amount of attention that followed. Once the child was changed, the child was permitted to wash his or her hands and return to current classroom activity. If necessary, graduated guidance was used to prompt the children through the hand-washing routine. The toileting area, including the changing table or changing mat, was then sanitized for the next child. The cleaning of the toileting area was often completed as children washed their hands or immediately following the children returning to the current classroom activity. Undergarment checks were conducted every 30 min for each subject across all conditions of the study.

Classroom teachers also prompted the children to sit on the toilet at scheduled times throughout each day, and the first sit of the day immediately followed the check-in procedure conducted when each child first entered the classroom (typically, in the morning). To perform a prompted sit, classroom teachers approached the subject and provided a three-step prompt for the child to say “potty” or “bathroom” and walk to the toileting area. An undergarment check was performed in addition to a prompted sit on the toilet once the child entered the toileting area. Children were then prompted to pull down their pants and undergarments and to sit on the toilet. Children remained on the toilet for 3 min or until an appropriate elimination occurred, whichever came first. Classroom teachers provided attention (e.g., talked to, sung songs with) the child during the sit. If an appropriate elimination occurred at any time during the sit, classroom teachers provided descriptive praise (e.g., “I am so proud that you are dry!”), and the child was asked to stand up (i.e., escape from the toilet was provided contingent on appropriate eliminations). Classroom teachers then assisted in wiping the child. Regardless of whether the child appropriately eliminated, classroom teachers then used graduated guidance to prompt each child to wash his or her hands, and the child was permitted to return to current classroom activity. Classroom teachers prompted each subject to sit across all conditions (unless otherwise specified) of the study. However, the amount of time that elapsed between prompted sits depended on the toileting condition in effect for each child.

All independent requests (i.e., self-initiations) to sit on the toilet were honored throughout all conditions, and the above toileting procedure was then implemented. Classroom teachers praised self-initiations and immediately guided the child to the toileting area where the child was permitted to sit for 3 min. Regardless of the toileting condition in effect, self-initiations reset the prompted sit schedule. For example, if a child was prompted to sit on the toilet every 30 min



(per the current toileting condition) and a self-initiation occurred 20 min after the last prompted sit, the next scheduled sit would occur 30 min following the self-initiation. This was done to increase the likelihood of self-initiations, as prompting a child to return to the toileting area he or she recently exited could punish self-initiations if leaving preferred classroom activities to walk to the toileting area was aversive.

### **Study 1: Component Analysis of Commonly Used Toilet-Training Procedures**

This study was designed to examine the effectiveness of a comprehensive toilet-training program when component procedures were implemented in a combined training package and when components were implemented in isolation.

#### **Subjects and Setting**

Twenty incontinent children from a university-based early childcare center participated in Study 1. Subjects were selected from three of five classrooms the center operated. Classroom teachers or the classroom supervisor implemented all procedures in each condition of the study. Classroom teachers were either undergraduate practicum students or paid staff. Undergraduate students were enrolled in a practicum course designed to provide hands-on experience in early education and childcare. Paid staff were hired to cover shifts that either did not have sufficient teacher coverage (as required by Kansas childcare regulations) or that were less likely to provide a robust learning experience (e.g., the mid-shift time period in which children mostly ate and napped) for an undergraduate practicum student. Classroom supervisors were either graduate teaching assistants in Applied Behavioral Science or paid supervisors with extensive histories in early childhood education settings.

Subjects were selected for participation if parents expressed interest in help with toilet training, caregiver report suggested the presence of child-readiness skills (e.g., remaining dry for

more than two hours, demonstrating “interest” in the toilet, being able to sit for at least 3 min, etc.), and classroom teachers and supervisors recommended the child for training. At any given time in a given classroom, between 0 and 5 children were receiving toilet training. Table 1 lists each child’s name, age (in months) at the start of baseline, and diagnosis. All procedures were reviewed and approved by a university internal review board, and consent was attained from each subject’s legal guardian prior to baseline data collection.

### **Response Measurement and Interobserver Agreement**

Teachers and supervisors in each classroom collected primary data. Frequency data were collected on each subject’s number of urinary accidents, urinary appropriate eliminations, and self-initiations throughout each day. An *accident* was defined as urinating anywhere other than in the toilet. Accidents were recorded anytime the child was observed to be wet. Accidents were typically identified at scheduled undergarment checks that occurred every 30 min. Data were not collected on the first scheduled undergarment check of each day, as it was unclear when the child last sat on the toilet or was changed. An *appropriate elimination* was defined as urinating in the toilet. Appropriate eliminations were detectable when urine was observed to enter the toilet or when toilet paper was found to be wet after being wiped (following a sit on the toilet). A *self-initiation* was defined as the child independently manding for toilet access. Children typically said or signed “potty” or “bathroom” to gain access to the toilet; however, gestural mands were also recorded as self-initiations for the youngest children. In classrooms that did not have a barrier preventing the children from independently entering the toileting area, a self-initiation was also recorded if the child entered the toileting area and sat on the toilet independently.

Because the number of times a child eliminated (both appropriately and inappropriately) each day could vary (i.e., was not controlled) across days due to several variables including

amount of fluid intake and amount of time present in the classroom (i.e., differences in the opportunities to eliminate), an additional calculation was analyzed to interpret data on the frequency of appropriate eliminations and the frequency of accidents more easily. Each child's *daily percentage of appropriate eliminations* was computed by dividing the frequency of appropriate eliminations by the total number of eliminations (i.e., the sum of appropriate eliminations and accidents). Therefore, the percentage of appropriate eliminations reflects the relative number of urinations made in the toilet, and because the percentage is a metric of two primary dependent measures, the percentage of appropriate eliminations was the primary unit of analysis upon which inferences regarding the effect of training on performance were made.

Frequency data on both appropriate eliminations and accidents are not reported when presenting the percentage of appropriate eliminations, because changes in any two of these variables inherently implies a predictable (and quantifiable) change in the third. For example, if we observed that the percentage of appropriate eliminations for a child on a specific day was 50 percent and the frequency of accidents was 5, the frequency of appropriate eliminations would also be 5 ( $5 \text{ appropriate eliminations} / (5 \text{ appropriate eliminations} + 5 \text{ accidents})$ ). However, information from at least two of these measures is required to make specific inferences about the third measure. Only relative changes in the frequency of appropriate eliminations and accidents can be inferred if data are presented on the percentage of appropriate eliminations in isolation, and nothing can be said regarding the frequency of accidents if data are presented only on the frequency of appropriate eliminations and *visa versa*. As the frequency of accidents is commonly reported in toilet-training research, we chose to present data on the frequency of accidents in addition to data on the percentage of appropriate eliminations. Frequency data on self-initiations are also presented.

Interobserver agreement was assessed by having a second observer simultaneously yet independently collect data on the same measures of toileting performance as the primary data collector for all children on 15.8% (range, 8.1% to 30.5%) of undergarment checks and on 16.7% (range, 9.3% to 28.1%) of toileting opportunities for a combined average of 16.1% (range, 8.7% to 29.4%) of undergarment checks and toileting opportunities. The second observers were trained graduate students or trained undergraduate students enrolled in a research course (distinct from the undergraduate practicum course in which the classroom teachers were enrolled). Interobserver agreement coefficients were calculated for each measure of toileting performance by adding the number of agreements and dividing by the number of agreements plus disagreements and multiplying by 100. An agreement was scored if both the primary observer and secondary observer recorded the same information for a given category (e.g., both observers recorded that the child was dry) during the same toileting opportunity. Interobserver agreement coefficients averaged 97.2% (range, 92.1% to 100%) for undergarment check data, 92.6% (range, 74.2% to 100%) for toileting opportunity data, and 94.6% (range, 83.3% to 100%) for whether children self-initiated or were prompted to the toileting area.

The second observer also collected procedural integrity data by assessing teachers' implementation of each child's toilet-training protocol. Specifically, the second observer collected data on (a) the time at which each child was brought to the toileting area for an undergarment check or toileting opportunity and (b) whether teachers implemented the appropriate undergarment check or toileting opportunity on 15.8% (range, 8.1% to 30.5%) of opportunities. Correct implementation was assessed separately for each procedural integrity measure. To assess whether teachers performed undergarment checks and toileting opportunities at the appropriate time, correct implementation was defined as the teacher bringing the child to

the toileting area within 5 min of the scheduled time. Procedural integrity coefficients were calculated by adding the number of correct implementations and by then dividing by the number of correct and incorrect implementations. Procedural integrity coefficients averaged 90.5% (range, 78.0% to 98.1%) for teachers' implementation of undergarment checks and toileting opportunities at the appropriate time. Procedural integrity coefficients averaged 94.8% (range, 89.2% to 100%) for teachers' correct implementation of an undergarment check or sit.

In addition to the procedural integrity measures collected for all children, additional procedural integrity measures were collected for the last five subjects (Christy, Ernie, Gayle, Ivy, and Leah) included in Study 1. For these subjects, we also assessed teachers' use of the correct undergarment type (diaper or underwear) and whether teachers delivered access to each child's bin of highly preferred items only when appropriate for each opportunity to do so. For procedural integrity measures on teachers' correct delivery of each child's bin, each undergarment check contained two opportunities for teachers to deliver the bin (once for a self-initiation and once for having remained dry), and each sit contained three opportunities to deliver the bin (once for self-initiation, once for having remained dry, and once for having eliminated in the toilet). Incorrect bin delivery (i.e., errors of omission or commission) for any one of the opportunities to deliver the bin was scored as an error for all opportunities for a given undergarment check or sit. Additional procedural integrity data were collected for 11.9% (range, 8.3% to 18.0%) of opportunities, and procedural integrity coefficients averaged 97.7% (range, 91.7% to 95.1%) for teachers' use of the correct undergarment type and 90.5% (range, 82.2% to 96.3%) for teachers' correct delivery of each child's bin.

### **Procedure.**

Subjects were exposed initially to a set of baseline procedures in which some elements of common toilet-training programs were implemented (see description below). Baseline procedures were designed to mimic those procedures that served as the pre-intervention procedures already in effect in each child's classroom. Rather than use a no-treatment baseline, any gains in toileting performance observed during training phases would constitute improvements above and beyond those that would be expected from a less intensive toilet-training program. That is, improvements in toileting performance above those observed during baseline procedures constituted genuine and potentially more robust changes in performance, as at least as large performance gains would be expected as compared to no-treatment conditions.

Once stable levels of toileting performance were observed under baseline procedures, each subject was exposed to one of four possible training conditions. Subjects were exposed to either a comprehensive toilet-training program that was comprised of three individual component procedures, or subjects were exposed to one of the three training components subsumed in the comprehensive toilet-training program. Although single-subject data were collected and analyzed on a case-by-case basis, we also analyzed the effects of each training condition on levels of toileting performance across subjects to better determine the external validity of treatment gains produced by each training condition.

For several subjects, treatment gains produced under some initial treatment conditions were not sufficient. Additional toilet-training components were added to the existing toilet-training procedures if further improvements in toileting performance were possible (as determined by clinical judgment). For example, if sufficient gains in toileting performance were not observed after exposure to toilet-training component "A," toilet-training component "B" or "C" was added to component "A." If toileting performance under the newly combined toilet-

training arrangement (“A+B” or “A+C”) was insufficient, the remaining component (“C” or “B”) was then added to create the comprehensive “A+B+C” toilet-training program. The order of initial and subsequent component implementations were counterbalanced across subjects to minimize potential order effects and to better determine the additive contribution(s) of each component when combined with each remaining component. Additional toilet-training components were no longer added to a given subject’s toilet-training program once sufficient improvements in toileting performance were observed. This was done to minimize any ceiling effect that might preclude the interpretation of additive toilet-training gains made by the implementation of single training components. Parents were not instructed to implement the toilet-training procedures in effect in the classroom while at home. However, parents were informed of the current toilet-training procedures in effect with their child.

**Baseline.** Baseline procedures were designed to mimic a less-intrusive toilet-training program in which children were allowed to sit on the toilet at scheduled times and anytime the child self-initiated. Access to highly preferred items were also available for appropriate eliminations. Baseline procedures were not designed to constitute “no intervention.”

During baseline, children wore disposable diapers or pull-on training pants. Classroom teachers performed undergarment checks every 30 min. Praise was provided when children were found to be dry. Children found to be wet at anytime were changed into similar undergarments (i.e., a diaper or pull-on was used) with minimal attention. Children sat on the toilet every 90 min for 3 min or until an appropriate elimination occurred, whichever came first. Because undergarment checks were performed every 30 min with scheduled sits occurring every 90 min and undergarment checks were always performed just prior to sitting on the toilet, every third time the child was checked the child also sat on the toilet.

Access to the two highest preferred edible items and two highest preferred leisure items identified from the most recent preference assessment was provided contingent on appropriate eliminations during baseline. For one child (Aaron), only praise was delivered for appropriate eliminations during baseline, as baseline procedures in his classroom (the classroom serving children with IDD) differed from the baseline procedures used in the other classrooms. Aaron was the only child from this classroom. Bowel movements also resulted in access to the same preferred items; however, soiling was not the focus of this study. Preferred items were not delivered for any other response during baseline. All children were exposed to baseline procedures prior to their exposure to (at least) one of the following conditions.

**Toilet-training package.** A comprehensive toilet-training package was used to demonstrate general improvement in toileting performance and was designed such that potentially influential components could be isolated to test each component's unique effects on toileting performance. The combined toilet-training package was designed to improve all targeted measures of toileting performance and consisted of three modifications of baseline procedures: the substitution of underwear for diapers or pull-on training pants, a more dense schedule of sits on the toilet, and differential reinforcement of both self-initiations and remaining dry at undergarment checks. Six children (Lizzy, Aaron, Gayle, Jim, Ingrid, and Bethany) were exposed to the toilet-training package. The remaining children were exposed to only one of the three toilet-training components that, when combined, formed the toilet-training package. The children selected for inclusion in the toilet-training package as well as children selected for each of the remaining conditions were selected based on subject availability and experimenter convenience (i.e., satisfying our need for subjects in the combined group and in each of the component groups).



*Underwear.* First, children wore 100% cotton underwear during the toilet-training package instead of wearing diapers or pull-on training pants as in baseline. Children’s “plastic pants” (i.e., underwear-like briefs made of plastic with elastic waist and leg openings) were used (when available) to cover children’s underwear and minimize response effect required by the teachers when children had an accident. Parents of the children in the study purchased the underwear (as well as the diapers or pull-ons used in other conditions). Because parents purchased the underwear for their child, we did not specify the exact type of underwear they purchased. Therefore, two slightly different types of underwear were used. Most children wore thin cotton underwear, but some children occasionally wore cotton underwear that was slightly thicker. However, the children did not consistently wear only one type of underwear. The thicker underwear was slightly more absorbent than the thin underwear, but both types of underwear were much less absorbent than the diapers or pull-ons used in baseline and other conditions. However, we did not view the difference in absorbency between the thin and thicker underwear as being different enough to require parents only purchase one underwear type.

Regardless of the underwear type, children placed in underwear were changed more frequently than children who remained in diapers or pull-ons. Teachers were required (per State of Kansas childcare licensing regulations) to change all children as soon as children were detected to have had an accident. Because underwear is less absorbent than diapers or pull-ons, the clothing of children wearing underwear often became wet when the child had an accident. Therefore, teachers could more easily detect when a child had an accident when the child was wearing underwear. Children detected to have an accident across all conditions were changed as immediately as possible.

***Dense-sit schedule.*** The schedule of prompted sits on the toilet was also changed. Children were prompted to sit on the toilet every 30 min as compared to every 90 min in baseline. Children were still required to remain on the toilet for 3 min or until an appropriate elimination occurred, whichever came first. However, the schedule of prompted sits on the toilet was three times as dense as that used in baseline to increase the opportunity for appropriate eliminations. Despite creating a denser schedule of prompted sits on the toilet, children could potentially sit more frequently during baseline than during the toilet-training package. Self-initiations during the toilet-training package continued to result in a 3-min sit on the toilet. So, given the free-operant nature of self-initiations, the behavior of each child controlled (to some degree) the sit schedule in effect each day. Therefore, the number and distribution of sits on the toilet each child actually experienced was dependent on both the schedule of prompted sits as well as the frequency and distribution of self-initiations. However, each child sat at least every 30 min (except during nap times) during the toilet-training package. Also, children arrived at and departed from each classroom at different times, which also affected the actual number of sits on the toilet.

***Differential reinforcement.*** The third change in toilet-training procedures from baseline was differential reinforcement. The differential-reinforcement component of the toilet-training package was designed to improve each measure of toileting performance. Baseline procedures already targeted appropriate eliminations in that preferred edible and leisure items were delivered when children eliminated in the toilet. However, baseline reinforcement procedures did not directly target self-initiations or preventing accidents. Two preferred edible items and 30-s access to preferred leisure items selected from weekly preference assessments were delivered for self-initiations and when children were dry at undergarment checks in addition to appropriate

eliminations. Therefore, the differential-reinforcement component of the toilet-training package involved the addition of several differential-reinforcement contingencies. Differential reinforcement of alternative behavior (DRA) was used to reinforce self-initiations and appropriate eliminations, and differential reinforcement of other behavior (DRO) was used to reinforce the absence of accidents.

Each response that entered into a differential-reinforcement contingency (i.e., self-initiations and remaining dry at undergarment checks), in addition to appropriate eliminations targeted during baseline and during the toilet-training package, was measured and treated as a discrete event. Therefore, preferred items were delivered for each response, independent of the occurrence of other responses. In other words, preferred items were available for self-initiations, remaining dry at undergarment checks, and appropriate eliminations regardless of whether all, a portion of, or only one response occurred. Children could earn access to their bins of highly preferred items each time one of the above responses occurred. Therefore, children could access their bins of preferred items a maximum of three times during each trip to the toilet (i.e., access for self-initiating, access for remaining dry, and access for appropriately eliminating). However, if a teacher prompted the trip to the toileting area, the child was wet, and the child did not appropriately eliminate, bin access was denied. Access to preferred items did not interfere with the chain of responses involved in trips to the toileting area. When preferred items were delivered, children consumed edible items and interacted with preferred toys while being changed or while sitting on the toilet.

**Underwear.** The underwear condition was designed to isolate the individual effects on toileting performance of placing children in underwear. Although methodological differences exist between the studies, the inclusion of the underwear-component evaluation was designed to

be a systematic replication of Simon and Thompson (2006) and Tarbox et al. (2004). The underwear condition was the same as baseline, except children wore cotton underwear instead of diapers or pull-ons. Four children (Danny, Tammy, Sully, and Leah) were exposed initially to the underwear condition following baseline.

**Dense-sit schedule (FT sits).** The dense-sit schedule condition was designed to evaluate the individual effects on toileting performance of providing children more frequent opportunities to eliminate appropriately. This condition was the same as baseline, except children were prompted to sit on the toilet for 3 min every 30 min instead of every 90 min. Additionally, the dense-sit schedule was identical to the toilet-training package had the underwear and differential-reinforcement components been removed. Four children (Alton, Sebastian, Ernie, and Marge) were exposed initially to the dense-sit schedule condition following baseline.

**Differential reinforcement.** The differential-reinforcement condition was designed to evaluate the effects of arranging explicit reinforcement contingencies for toileting skills not directly targeted in baseline. This condition was the same as baseline, except that self-initiations and remaining dry at undergarment checks resulted in access to two preferred edible items and 30-s access to two leisure items. Appropriate eliminations continued to result in access to preferred items. Therefore, children could access their bins of preferred items a total of three times during each trip to the toilet (i.e., access for self-initiating, access for remaining dry, and access for appropriately eliminating). The differential-reinforcement condition was identical to the toilet-training package had the underwear and dense-sit schedule components been removed. Four children (Nancy, Blue, Christy, and Ivy) were exposed initially to the differential-reinforcement condition following baseline.

**No-treatment probe.** The no-treatment probe condition was designed to evaluate toileting performance when all treatment procedures (including those in baseline) were removed. No-treatment probes were conducted throughout each phase of children's toilet-training evaluations. Toileting performance measured during the no-treatment probe condition was used for two purposes: to quickly ascertain whether improved toileting performance was a function of the training procedures in place (or if stimulus generalization had occurred) and to evaluate toileting performance using the same procedures across time. Probes of baseline procedures were not used in place of the no-treatment probes, because we were primarily interested in determining toileting independence or each child's ability to maintain high toileting performance without teacher assistance. Probes, rather than phases, were used such that toileting training procedures could be rapidly reinstated if decrements in toileting performance were observed during no-treatment probes (indicating that previously observed high levels of toileting performance were likely a function of training procedures).

As noted above, baseline procedures included some elements (e.g., scheduled sits on the toilet, reinforcement of appropriate eliminations, praise for self-initiations, remaining dry, etc.) of toilet-training procedures commonly found in less-intensive toilet-training programs. These training procedures were removed during no-treatment probes. Children wore diapers or pull-ons, and undergarment checks occurred every 30 min. Scheduled sits on the toilet were removed during no-treatment probes (i.e., toilet access was permitted only if the child self-initiated). Also, preferred items were not delivered for any response during no-treatment probes. To ensure teachers did not inadvertently deliver preferred items, each child's bin of preferred items was removed from the toilet-training area. Praise delivered during baseline for self-initiations and for remaining dry at scheduled undergarment checks was also minimized during no-treatment

probes. Teachers were instructed to say “okay, let’s go to the potty” (or any similar statement) when self-initiations occurred and to say “you’re dry” (or any similar statement) when undergarments were found to be dry during undergarment checks. Escape from the toilet was still permitted contingent on appropriate eliminations during no-treatment probes. Therefore, other than escape from the toilet when a child appropriately eliminated, all toilet-training procedures used during baseline were removed during no-treatment probes.

The no-treatment probe condition was also created in an attempt to demonstrate an additional degree of experimental control. A combination of experimental designs was used to demonstrate that changes in toilet-training performance occurred when, and only when, changes in the independent variable (the toilet-training procedures in place) were implemented. Because within-subject experimental control is often difficult to achieve with skill acquisition experiments, our primary experimental design (i.e., multiple baseline across subjects design) was selected for its ability to demonstrate experimental control across subjects when reversals in performance are unlikely (as when teaching new skills that are likely to generalize to reinforcement contingencies outside experimenter control). Changes in toileting performance were sometimes delayed by several days following changes in toilet-training procedures (for an example, see data for Danny and Tammy in Figure 2), thereby limiting inferences made regarding the cause of behavior change. Reversals in toilet-training procedures were attempted with some children in addition to the multiple baseline design; however, these rarely assisted in contributing to experimental control (for an exception, see Gayle’s data in Figure 1).

One potential reason for our failure to reverse to baseline levels of performance following exposure to intervention procedures may have been due to functionally small differences between the less-intrusive toilet-training baseline procedures and each of the intervention

components when implemented in isolation or when combined. This interpretation also holds for why delayed effects were sometimes observed following implementation of toilet training. We attempted to demonstrate experimental control using rapidly implemented no-treatment conditions that were designed to approximate a multiple probe design. All toilet-training intervention procedures were removed during these no-treatment probes to enhance the procedural differences between intervention phases and when all intervention elements were removed. No-treatment probes were conducted throughout all toilet-training phases with children for whom it was used. Therefore, performance during no-treatment probes conducted during baseline phases also allowed for an evaluation of toilet-training performance when all toilet-training procedures were absent (during no-treatment probes) and when the low-intensity toilet-training baseline was implemented.

Unfortunately, no-treatment probes were created in response to questionable intervention effects and were not started until after some children had completed participation. Therefore, no-treatment probes were conducted with a subset of the children. Eleven children (Lizzy, Gayle, Jim, Ingrid, Bethany, Leah, Ernie, Christy, Ivy, Missy, and Jasmine) were exposed to the no-treatment probe condition.

### **Data Analysis**

Data collected on appropriate eliminations, accidents, and self-initiations were analyzed using a few different methods. First, visual inspection of each subject's data was used throughout the study to yield important information regarding when best to introduce selected training components. Components were introduced when stability in the majority of the dependent measures was observed. Additionally, component introduction was dependent on the

effects observed when other subjects were exposed to the same component (i.e., a multiple baseline was attempted).

Data collected on the effectiveness of each component were also analyzed by calculating the mean change in each measure of toileting performance (percentage of appropriate eliminations, frequency of accidents and self-initiations) from the previous phase. This was done according to the following formula:

$$\text{Mean Difference: } M_{\text{Treatment}} - M_{\text{Baseline}}$$

In the above calculation, M equals the phase mean of a given measure of toilet-training performance. Data collected on each measure of toileting performance were averaged for each phase of the study and were subtracted the previous phase mean from the subsequent phase mean. For example, if the mean percentage of appropriate eliminations during baseline was 10 percent, and the mean increased to 50 percent during implementation of the first component (i.e., during the phase subsequent to baseline), the mean change in appropriate eliminations was 40 percent (50 percent – 10 percent). This calculation was then completed for each dependent measure. Because training components were often added to existing components (e.g., A, A+B, A+B+C) when the effects of single components were insufficient at increasing overall toileting performance, the mean difference in each measure of toileting performance was calculated each time an additional component was added to each child's toileting program. In these cases, mean change was calculated as the change from previous phase data and not from baseline performance. This was done to better examine the additive contribution(s) of each component on toileting performance.

Standardized difference effect sizes (Faith, Allison, & Gorman, 1997) were also calculated to better quantify the additive effect of each toilet-training component on each



measure of toileting performance. This additional calculation was completed to better account for the amount of variability observed within phases. That is, the standardized difference effect size calculation accounts for both the change in mean levels of performance and variability.

Standardized difference effect sizes were calculated according to the following formula:

$$\text{Standardized Difference Effect Size} = \text{Absolute Value} [(M_{\text{Treatment}} - M_{\text{Baseline}}) / SD_{\text{Baseline}}]$$

In the above calculation, M equals the phase mean, and SD equals the standard deviation.

Similar to the calculation of the mean difference, standardized difference effect sizes were calculated for each measure of toileting performance when each training component was added to each child's toilet-training program. Standard difference effect sizes were also calculated as the change from previous phase data and not from baseline performance when additional training components were added to a toilet-training program for each child. For our purposes, the absolute value was removed from the standardized difference effect size formula, as negative values were important in some measures (e.g., reductions in accidents).

A combination of experimental designs was used in Study 1 to demonstrate both within- and across-subject experimental control of the arranged toilet-training procedures. A multiple baseline across subjects design was used as the primary experimental design, because acquisition of toileting skills occurred very quickly for some children, and returns to previous training phases did not always produce a reversal effect (i.e., newly acquired skills presumably came under the control of unprogrammed contingencies). Reversal data are presented for subjects for whom a reversal design was attempted and useful information was obtained regarding toileting skills under different procedures.

## **Results and Discussion**

Toilet-training data for children exposed initially to the toilet-training package following

baseline are depicted in Figure 1. Baseline performance for Lizzy showed a moderate number of accidents, a low-to-zero level of appropriate eliminations, and a variable but low level of self-initiations. The introduction of the toilet-training package did not produce an improvement in overall toilet-training skills for Lizzy, as the number of accidents remained consistent; however, an overall higher level of appropriate eliminations and self-initiations were observed after extended exposure to the toilet-training package. Data collected during Lizzy's sole no-treatment probe showed no change in toileting performance as compared to data collected during the toilet-training package. For Lizzy and all other children for whom the combined toilet-training package was insufficient, additional toilet-training procedures (e.g., fluid loading, enhanced differential-reinforcement procedures, toileting training alarms) were implemented to improve toileting performance. Data on those procedures are available from the author.

Baseline toilet-training data collected for Aaron showed a similar pattern of responding to Lizzy's baseline data. A low-to-moderate level of accidents, a low-to-zero level of appropriate eliminations, and zero self-initiations were observed during baseline for Aaron. Unlike toilet-training package data collected for Lizzy, a decrease in accidents and an increase in appropriate eliminations were observed during the toilet-training package for Aaron. Self-initiations remained low. No additional toilet-training components were implemented with Aaron.

Gayle's toileting performance was poor during the initial baseline. A moderate-to-high level of accidents, a moderate level of appropriate eliminations, and a low level of self-initiations were observed. Toileting performance was similar during no-treatment probes conducted during the initial baseline; however, decreased levels of appropriate eliminations were observed on days in which no-treatment probes were conducted. Overall improvements in Gayle's toileting performance were correlated with the introduction of the toilet-training package. A decrease in

accidents and an increase in appropriate eliminations were observed. There was no change in the level of self-initiations. Toileting performance remained consistent when baseline was reinstated, and a slight overall increase in self-initiations was observed. The no-treatment probe conducted during the second baseline phase produced a decrease in appropriate eliminations. Subsequent reintroduction of the toilet-training package produced additional improvements in the level of accidents and appropriate eliminations. However, self-initiations remained low. The final no-treatment probe produced a slight decrease in toileting performance; however, self-initiations increased. No additional toilet-training components were implemented with Gayle.

Jim's baseline toileting performance was initially poor, but increases in appropriate eliminations and self-initiations were observed over the course of baseline. Further increases in toileting performance were observed when the toilet-training package was introduced. Decreases in accidents and increases in appropriate eliminations were observed. Initially, the level of self-initiations was moderate (although responding was variable); however, the level of self-initiations decreased across the toilet-training package phase. Slight decreases in toileting performance were observed when no-treatment probes were conducted; however, increases in self-initiations were observed. Additional toilet-training procedures were implemented with Jim.

Baseline toileting performance for Ingrid was better than that observed for most of the children in the current study. A low-to-moderate level of accidents and moderate-to-high levels of appropriate eliminations and self-initiations were observed during baseline. Similar toileting performance was observed during the no-treatment probes. The introduction of the toilet-training package was correlated with a decrease in accidents and an increase in appropriate eliminations. Interestingly, self-initiations reduced to a low-to-zero level. Overall toileting performance declined when the final no-treatment probe was conducted. An increase in

accidents and a decrease in appropriate eliminations were observed. Similar to other children, an increase in self-initiations was recorded during Ingrid's final no-treatment probe. Additional toilet-training procedures were implemented with Ingrid.

The duration of exposure to baseline procedures was longer for Bethany. A variable but high level of accidents and low-to-zero levels of appropriate eliminations and self-initiations was observed during Bethany's baseline. Toileting performance was consistent during the no-treatment probes conducted during the baseline phase. Immediate improvements in toileting performance were observed for both accidents and appropriate eliminations when the toilet-training package was introduced. Bethany's self-initiations remained low. The final no-treatment probes were correlated with a slight decrease in toileting performance. However, an increase in self-initiations was also observed. Additional toilet-training components were not implemented with Bethany.

Overall, the toilet-training package produced decreases in accidents and increases in appropriate eliminations for the majority of children. Self-initiations did not increase during the toilet-training package phase for any child. Removal of all toilet-training components resulted in slight decreases in toileting performance for the majority of children when no-treatment probes were conducted. Interestingly, the no-treatment probe produced noticeable increases in self-initiations for most children once improvements in accidents and appropriate eliminations were established. For these children, self-initiations may have become a functional avoidance response, whereby self-initiating (and subsequently eliminating on the toilet) avoided the aversive properties of having an accident. Recall, there were no programmed sits on the toilet during no treatment probes. Therefore, independently requesting to sit (i.e., self-initiating) was the only way in which children could potentially avoid having an accident, aside from self-

restricting fluid intake. Fading procedures used to thin the programmed schedule of sits on the toilet may have interesting implications for when and how children learn to self-initiate.

Toilet-training data for children exposed initially to the underwear component of the toilet-training package following baseline are graphed in Figure 2. Danny was exposed to a reversal in toilet-training conditions. A low-to-moderate level of accidents and a low-to-zero level of self-initiations were observed during both of Danny's exposures to baseline. Appropriate eliminations remained at zero throughout Danny's initial baseline; however, a variable level of appropriate eliminations was observed during Danny's second baseline. Danny's initial exposure to the underwear component was correlated with delayed increases in appropriate eliminations. Danny's level of accidents remained constant; however, self-initiations increased initially during the underwear phase but decreased subsequently over time. Decreases in Danny's self-initiations were correlated with increases in appropriate eliminations. The second implementation of the underwear component was again correlated with delayed increases in appropriate eliminations. A delayed decrease in accidents and a moderate level of self-initiations was observed during the second underwear phase. Unfortunately, improvements in Danny's toileting performance are difficult to interpret, as delayed effects and two breaks from school preclude a definitive interpretation of the effect of underwear on toileting performance.

Additional toilet-training components were not implemented with Danny.

Tammy's baseline toileting performance was poor. A moderate level of accidents and low-to-zero levels of appropriate eliminations and self-initiations were recorded during Tammy's baseline. Changes in Tammy's toileting performance were not observed immediately following the introduction of the underwear component; however, a slight overall decrease in accidents was observed across the initial eight days of the underwear phase. During this time, appropriate

eliminations remained at zero, and self-initiations remained consistently low. For Tammy, the underwear component seemed to influence accidents before affecting the other toileting skills. Tammy's data suggest that she first learned to avoid having an accident by "holding it in" before learning how to eliminate in the toilet. Following the initial eight days of the underwear phase, the level of Tammy's appropriate eliminations increased, accidents further decreased, and self-initiations became more variable. Because further improvements in Tammy's toileting performance were possible, the differential-reinforcement component was added to Tammy's toilet-training program.

Consistently low toileting performance was observed for Sully during baseline. A moderately high level of accidents and near-zero levels of appropriate eliminations and self-initiations were observed during Sully's baseline. The introduction of the underwear component initially produced a slight decrease in accidents and increases in appropriate eliminations and self-initiations. After approximately the first week and a half of exposure to the underwear phase, substantial improvements in all of the measures of toileting performance were observed. Sully's accidents maintained at low levels and data collected on appropriate eliminations remained high. The differential-reinforcement component followed by the dense-sit schedule component was added to Sully's toilet-training program to improve the consistency of Sully's toileting performance.

Baseline toileting performance was similarly low for Leah. A variable low-to-moderate level of accidents and low levels of appropriate eliminations and self-initiations were observed for Leah during baseline. Three no-treatment probes were conducted during Leah's baseline phase, and toileting performance during the no-treatment probes was similar to her baseline toileting performance. The underwear component did not produce a change in any of the

toileting performance measures. Toileting data collected during Leah's final no-treatment probe indicated that the introduction of underwear did not facilitate toileting performance. The dense-sit schedule component followed by the differential-reinforcement component was added to Leah's toilet-training program to improve her overall low levels of toileting performance.

Overall, the introduction of the underwear component seemed to facilitate toilet-training performance for at least two of the four children with whom it was implemented. It is unclear whether the underwear component or some uncontrolled variable (such as potentially more intensive toilet-training procedures implemented during the summer breaks) affected toileting performance for Danny. The effects of underwear on toileting performance were somewhat delayed for Tammy and Sully. However, the lack of any improvement in Leah's toileting performance suggests that mere exposure to baseline procedures would have been unlikely to improve Tammy and Sully's toileting performance.

Toilet-training data for children exposed initially to the dense-sit schedule component of the toilet-training package following baseline are graphed in Figure 3. Baseline data collected for Alton showed a high level of accidents and low-to-zero levels of appropriate eliminations and self-initiations. Similarly poor toileting performance maintained when Aaron's sit schedule was changed from scheduled sits every 90 min to sits every 30 min.

A similar pattern of responding was observed for Sebastian. A moderate-to-high level of accidents and low-to-zero levels of appropriate eliminations and self-initiations were observed during Sebastian's baseline. Toileting performance remained low during the dense-sit schedule condition for Sebastian. The underwear component followed by the differential-reinforcement component was added to Sebastian's toilet-training program to improve Sebastian's toileting performance.

Ernie's baseline toileting performance was slightly higher than Alton and Sebastain's baseline performance. Although a moderate-to-high level of accidents was observed during Ernie's baseline, a moderate level of appropriate eliminations was observed. Ernie's self-initiations during baseline were variable. Two no-treatment probes were conducted during the baseline phase, and Ernie's performance during the no-treatment probes was similar to his baseline performance. The introduction of the dense-sit schedule was correlated with an increase in variability of accidents and an initial increase in variability of self-initiations. However, an overall increase in Ernie's appropriate eliminations was observed during the dense-sit schedule condition. Toileting data collected during Ernie's only no-treatment probe conducted during the dense-sit schedule phase produced mixed results. Although accidents were relatively low, appropriate eliminations decreased during the final no-treatment probe. Similar levels of self-initiations were observed. Because further improvements in Ernie's toileting performance were possible, the differential-reinforcement component followed by the underwear component was added to Ernie's toilet-training program.

Marge was exposed to an extended evaluation of the dense-sit schedule component to determine whether improvements in toileting performance are possible provided additional exposure to the dense-sit schedule. Marge's toileting performance was consistently poor throughout the evaluation. The level of accidents remained moderate, and low-to-zero levels of appropriate eliminations and self-initiations were observed through both baseline and both dense-sit schedule phases. The extended evaluation of the dense-sit schedule failed to produce increases in any of the toileting performance measures across a long period of time. To determine whether any modifications to Marge's toilet-training program would be able to improve Marge's toileting performance, a potentially more powerful change was made. The



underwear and differential-reinforcement components were combined and implemented in addition to the dense-sit schedule component rather than adding each toilet-training component sequentially.

In sum, the dense-sit schedule component of the toilet-training package did not produce overall improvements in toileting performance when implemented alone for any of the children. The results of Marge's extended evaluation also demonstrated that additional exposure to the dense-sit schedule component was equally ineffective. Also, Ernie's no-treatment probe data indicated similar toileting performance when using a dense-sit schedule and when all toilet-training components (including all prompted sits) were removed.

Toilet-training data for children exposed initially to the differential-reinforcement component of the toilet-training package following baseline are graphed in Figure 4. Baseline toileting data collected for Nancy showed moderate levels of accidents and appropriate eliminations and a variable level of self-initiations. The introduction of the differential-reinforcement component was not correlated with an overall improvement in toileting performance for Nancy. Accidents remained at a moderate level when the differential-reinforcement component was implemented, and overall decreasing trends in appropriate eliminations and self-initiations were observed. The underwear component followed by the dense-sit schedule component was added to Nancy's toilet-training program to improve low levels of toileting performance.

Variable but high levels of accidents and self-initiations were recorded during Blue's baseline. Blue's level of appropriate eliminations was also variable but decreased across baseline. An improvement in toileting performance was not observed for Blue with the introduction of the differential-reinforcement component. Levels of accidents and appropriate

eliminations remained moderate, and a decrease in self-initiations was observed when the differential-reinforcement component was implemented. Similar to the modifications made to Nancy's toilet-training program, the underwear component was added to Blue's toilet-training program followed by the dense-sit schedule component.

Baseline toileting data collected for Christy showed a variable moderate-to-high level of accidents and low-to-zero levels of accidents and self-initiations. Two no-treatment probes were conducted with Christy during the baseline phase. Christy's toileting performance when all treatment procedures were removed was similar to her baseline performance. The introduction of the differential-reinforcement component did not improve any of the toilet-training measures for Christy. Christy's toileting performance during the final no-treatment probe indicated that the differential-reinforcement condition was no better than the absence of toilet-training procedures. Christy's toilet-training program was modified first by adding the dense-sit schedule component then the underwear component.

A low-to-moderate level of accidents and low-to-zero levels of appropriate eliminations and self-initiations were recorded during Ivy's baseline. Ivy's toileting performance during the only no-treatment probe conducted during the baseline phase indicated no functional difference better the no treatment condition and baseline. An increase in the level of self-initiations was observed when the differential-reinforcement component was implemented for Ivy; however, levels of accidents and appropriate eliminations remained at baseline levels. Similar to Christy's data, Ivy's final no-treatment probe indicated no functional difference between the no treatment and the differential-reinforcement conditions. Ivy's toilet-training program was modified similar to Christy's by first adding the dense-sit schedule component then the underwear component.

The differential-reinforcement component of the toilet-training package failed to produce overall improvements in the toileting performance of any of the four children with whom it was implemented alone. Furthermore, Christy and Ivy's toileting performances during no-treatment probes were largely undifferentiated from their toileting performances during differential reinforcement. However, an increase in self-initiations was observed for Ivy. Unfortunately, this effect was not replicated with the other children, therefore limiting causal interpretations of whether differential reinforcement increased Ivy's self-initiations.

Prolonged exposure to baseline procedures improved toileting performance for some children. This effect can be observed to some degree with a closer examination of Jim's baseline data (Figure 1). Increases in appropriate eliminations and self-initiations are evident with additional exposure to baseline toileting procedures. Despite a change the in level of accidents across baseline, toileting performance seemed to improve simply due to extended exposure to the less-intensive toilet-training procedures included in baseline. Two children (Missy and Jasmine) demonstrated improvements in toileting performance during baseline. Figure 5 displays toileting data for Missy and Jasmine. Large improvements in all three measures of toileting performance were seen when exposure to baseline procedures was extended. Also, only partial reversals in toileting performance were observed for Missy and Jasmine when all toilet-training procedures were removed, indicating that toileting skills had likely come under other sources of control. The large overall improvements in Missy and Jasmine's toileting skills during baseline precluded evaluations of the toilet-training package or its components.

An alternate interpretation of the improvements seen in these children's baseline data is that of an influential but uncontrolled third variable. One possibility is that additional toilet-training procedures were implemented with children at home or outside the classroom setting.

Parents were not instructed to implement toilet-training procedures at home. However, we did not instruct parents not to attempt toilet training at home, and some parents reported that they attempted to implement some or all of the toilet-training procedures in place in the classroom with the children at home.

Another interpretation of improvements in toileting performance during baseline is subject maturation. As children age, the ease with which they become toilet trained is probably improved, as operant control of biological organs (e.g., the bladder) is likely to become more refined. Single subject research methodology is particularly susceptible to subject maturation as a threat to internal validity, especially when within-subject experimental control is unachievable. As noted above, the inclusion of the no-treatment probe was an attempt to demonstrate within-subject experimental control, and it proved useful in uncovering whether the teacher-implemented toilet-training procedures were maintaining improved toileting performance. In some cases, clear reversals (decrements) in toileting performance were observed when toilet-training procedures were removed indicating that the current toilet-training procedure(s) were maintaining improved performance. In other cases, clear reversals were not demonstrated. However, useful information regarding the variables responsible for maintenance of toileting skills was found. When improved toileting performance continued despite the removal of all toilet-training procedures (e.g., Missy and Jasmine's data on Figure 5), currently implemented toilet-training procedures were no longer contributing to improved toileting performance and, therefore, were unnecessary.

Although it appears that extended exposure to baseline procedures may have improved toileting performance for Jim, Missy, and Jasmine, mere exposure to baseline procedures was insufficient for other children. For example, Marge was exposed to a combination of baseline

and dense-sit schedule conditions for nearly twice as many school days as Missy and Jasmine, and improvements in toileting performance were not observed. Therefore, exposure to baseline procedures alone may have been sufficient for Jim, Missy, and Jasmine, but other children were unlikely to benefit from baseline procedures alone.

Summary component analysis data for each child exposed to the toilet-training package or one of the components of the toilet-training package following baseline are presented in Figure 6. The mean difference formula described above was used to calculate the average change (indicated by bar height) for each of the three measures of toileting performance from baseline during the first condition implemented following baseline. Standardized effect sizes were also calculated. Bars that appear with an asterisk indicate a standardized effect size above 1.0. Asterisks were not displayed for contraindicated effects (e.g., a large decrease in the frequency of self-initiations).

The toilet-training package produced improvements in appropriate eliminations and reductions in accidents for each of the six subjects with whom it was implemented. The effect of the toilet-training package on self-initiations was mixed. The effect sizes for children exposed to the toilet-training package were above 1.0 for appropriate eliminations for each child and above 1.0 for accidents of five of the six children. Self-initiation effect sizes were below 1.0 for all children exposed to the toilet-training package.

The underwear component also produced overall improvements in toileting performance for most children. Increases in appropriate eliminations, decreases in accidents, and slight increases in self-initiations were observed for three of four children exposed only to the underwear component. Effect sizes above 1.0 were calculated for two children across each of the three toilet-training measures.

The dense-sit schedule and differential-reinforcement components yielded mixed results across each of the toilet-training measures, and decreases in appropriate eliminations and increases accidents were observed for some children in each group. Effect sizes above 1.0 were found for appropriate eliminations for only one of the four children in each group. A large increase with an effect size above 1.0 was observed for one child in the differential-reinforcement component evaluation. Unfortunately, increases in self-initiations were not observed with the other children in either the dense-sit schedule component or differential reinforcement-component evaluation.

Mean difference data are re-presented in Figure 7 with the starting age (in months) of each child exposed to either the toilet-training package or a component of the toilet-training package following baseline. The toilet-training procedures (i.e., toilet-training package and the underwear component) found to be effective at increasing appropriate eliminations and decreasing accidents were effective regardless of children's age. Young children were just as likely as older children to show improvements in levels of appropriate eliminations and accidents, given that an effective intervention was used. In fact, the child with the largest decrease in accidents given any change in toileting procedures was one of the youngest children in the study.

Toilet-training procedures (i.e., dense-sit schedule and differential-reinforcement components) found to be ineffective at improving levels of appropriate eliminations and accidents were ineffective across children's ages. That is, older children were not more likely to become toilet trained with less-effective toilet-training procedures. Unfortunately, the effect of all toilet-training procedures on improving levels of self-initiations was mixed. Child age did not seem to influence changes in self-initiation levels either.

Mean difference data are again re-presented in Figure 8 and separated by each child's gender. As can be inferred from Figure 8, a roughly equal proportion of males and females were exposed to each of the toilet-training conditions following baseline, with all conditions filled with at least one male and one female. Although we observed a higher overall average percentage of appropriate eliminations following baseline with males, females produced a slightly lower overall average frequency of accidents following baseline. Overall average self-initiations remained unchanged for the majority of males and females. Therefore, the difference in combined toileting performances between males and females was largely negligible.

As noted above, additional toilet-training components were added sequentially when further improvements in toileting performance were possible. For example, Leah was exposed initially to the underwear component of the toilet-training package. However, Leah's toileting performance remained largely unchanged from baseline toileting performance. Therefore, an additional component (dense-sit schedule) of the toilet-training package was added to her existing toilet-training program, and the selection of which component to add was based on the counterbalancing of component implementation across children. The differential-reinforcement component was then added, because additional improvements in performance were possible. With the addition of the third component, the toilet-training program became the toilet-training package. Additional components were no longer added to children's toilet-training programs when adequate toileting performance was achieved.

Baseline averages of each measure of toileting performance as well as mean difference in toileting performance given the initial toilet-training program and when additional components were added (if necessary) to each child's toilet-training program is summarized in Table 2. Data that appear below each treatment header were calculated using the mean difference formula

above. Table 2 displays the additive effects on toileting performance of each toilet-training component introduced. Toilet-training procedures outside the scope of the study were used for some children. The labels “Other” and “N/A” identify when these procedures were used. Double dashes (--) were used to indicate when additional components were deemed unnecessary and therefore not introduced. Similar information is depicted in Table 3. Standardized difference effect sizes were calculated instead of mean differences for Table 3. The absolute value of the standardized difference effect size formula was removed from the calculations that appear in Table 3 to demonstrate the direction of change in each measure of toileting performance. Negative effect sizes appear in Table 3 and indicate that a reduction in the corresponding measure was observed.

Additional toilet-training components were unnecessary for three of the six children exposed initially to the toilet-training package. Additional toilet-training components were implemented with three of the four children exposed initially to the underwear component; however, a third component was implemented for only two children. Additional training components were implemented with all of the children exposed initially to the differential-reinforcement component or the dense-sit schedule component.

Interestingly, improvements in toileting performance were often observed when underwear was implemented as the second or third toilet-training component. Underwear was added as a second or third component for seven children; however, one child (Marge) was exposed simultaneously to the underwear and differential-reinforcement components. Of the six children who were exposed to only the underwear component as the second or third component added, four children showed increases in the mean difference of appropriate elimination, and the mean difference of accidents decreased for five children.



It is also interesting that the differential-reinforcement component was more likely to be effective when added to an existing toilet-training program, despite being generally ineffective when used alone. The differential-reinforcement component was added in isolation as a second or third component for five children. Of those five cases, the differential-reinforcement component was correlated with improvements in appropriate eliminations in all cases, and improvements in accidents were observed for four children.

Marge was exposed to the underwear and differential-reinforcement components simultaneously after the dense-sit schedule was shown ineffective. Very strong improvements in toileting performance were observed when underwear and differential reinforcement were combined. Large increases in levels of appropriate eliminations and self-initiations as well as near complete suppression of accidents were recorded when underwear and differential reinforcement were combined.

Some children (i.e., Tammy, Sully, and Sebastian) demonstrated improvements in toileting performance when differential reinforcement was added to an existing toilet-training program in which underwear was also used. The ineffectiveness of differential reinforcement when implemented alone as the first component in addition to the finding that differential reinforcement improvement toileting performance when added to children's toilet-training program in which underwear is used indicates that differential reinforcement may facilitate the effects of wearing underwear.

However, an alternate interpretation is that adding the differential-reinforcement component to toilet-training programs in which underwear is used is effective only so far as it permits additional exposure to the effective underwear component. Nancy and Blue's data help clarify the role of differential reinforcement when added to toilet-training programs in which

underwear is used. Solid improvements in appropriate eliminations and accidents were recorded only after underwear was added to the differential-reinforcement component. These data indicate that underwear alone improved toileting performance; however, it was unclear whether toileting performance would have been lower had the differential-reinforcement component been removed and the underwear component implemented alone. Therefore, it remains unclear what role (if any) the differential-reinforcement component played in promoting toileting performance when added to toilet-training programs in which underwear was used.

### **Study 2: Determining Additive Effects of Differential Reinforcement**

This study was designed to examine whether adding the differential-reinforcement component enhanced the effectiveness of the underwear component. In Study 1, the underwear component alone produced improvements in toileting performance for two of four children (see Figure 2). However, improvements were often delayed, as changes in toileting performance were not observed for several days following the introduction of underwear for some children (see Danny and Tammy's data in Figure 2).

In contrast, the differential-reinforcement component did not improve the toileting performance of any of the four children for whom it was implemented. However, the differential-reinforcement component was correlated with improvements in toileting performance when added as a second or third component to toilet-training programs in which underwear was also used (see Tables 2 and 3). Therefore, Study 2 was designed to determine whether differential reinforcement facilitates the effectiveness of underwear. Alternatively, adding the differential-reinforcement component may have allowed more time for underwear to exert control over toileting performance. In other words, initial low performance observed following the introduction of underwear might have reduced the averaged toileting performance

computed for the underwear effects in Tables 2 and 3, therefore artificially inflating the improved performance observed when differential-reinforcement was later added. Study 2 was designed to determine the additive potential of the differential-reinforcement component when used with underwear.

One method of examining the potential additive effects of differential reinforcement on underwear is to first assess toileting performance when the underwear component is implemented alone and again when differential reinforcement is added. If improvements in toileting performance are observed when differential reinforcement is added for children already wearing underwear, the differential-reinforcement component facilitates toileting performance. Furthermore, reductions in toileting performance observed when the differential-reinforcement component is removed (leaving only the underwear component intact) would further suggest an additive effect of differential reinforcement on underwear. However, steady improvements in toileting performance across underwear and underwear with differential reinforcement phases would suggest that the underwear component alone improved toileting performance and that differential reinforcement has no additive effect on underwear. Steady improvements across phases would also suggest that apparent improvements in summary measures (Tables 2 and 3) observed when differential reinforcement was combined with underwear in Study 1 were likely skewed due to a delayed underwear effect.

### **Subjects and Setting**

Four incontinent children from the same university-based early childcare center participated in Study 2. Subjects were selected for participation in Study 2 using the same selection procedures described in Study 1. Table 4 lists each child's name, age (in months) at the

start of baseline, and diagnosis. Consent was attained from each subject's legal guardian prior to baseline data collection.

### **Response Measurement and Interobserver Agreement**

All data-collection procedures were identical to those described in Study 1. Interobserver-agreement data were collected and analyzed for all children using the same procedures described in Study 1 on 17.4% (range, 16.1% to 19.1%) of undergarment checks and on 17.7% (range, 14.9% to 22.9%) of toileting opportunities for a combined average of 17.5% (range, 15.7% to 19.3%) of undergarment checks and toileting opportunities. Interobserver agreement coefficients averaged 97.8% (range, 96.2% to 99.2%) for undergarment check data, 98.0% (range, 95.7% to 100%) for toileting opportunity data, and 97.0% (range, 95.2% to 100%) for whether children self-initiated or were prompted to the toileting area.

Procedural integrity coefficients were calculated and analyzed for all children using the same procedures described in Study 1 on 17.4% (range, 16.1% to 19.1%) of opportunities. Procedural integrity coefficients averaged 89.9% (range, 81.4% to 93.5%) for teachers' implementation of undergarment checks and toileting opportunities at the appropriate time, 94.6% (range, 91.3% to 97.1%) for teachers' correct implementation of an undergarment check or sit, 99.7% (range, 99.0% to 100%) for teachers' use of the correct undergarment type, and 96.7% (range, 94.3% to 100%) for teachers' correct delivery of each child's bin.

### **Procedure**

The same university-approved procedures (including undergarment checks every 30 min and weekly MSWO preference assessments) described in the General Method and in Study 1 were used in Study 2.

**Baseline.** Baseline was identical to the baseline described in Study 1.

**Underwear.** The underwear condition was identical to the underwear condition described in Study 1.

**Underwear + differential reinforcement.** This condition combined the underwear condition with the differential-reinforcement condition described in Study 1. Children were placed in underwear, and access to highly preferred edible and leisure items (from weekly MSWO preference assessments) was arranged contingent on self-initiating and remaining dry at undergarment checks in addition to appropriately eliminating.

**No-treatment probe.** No-treatment probes were conducted as described in Study 1.

Similar to Study 1, we used a combination of experimental designs in Study 2 to demonstrate experimental control both across and within subject, including: (a) multiple baseline across subjects, (b) reversal, and (c) multiple probe. However, the critical comparison in Study 2 was not how toileting performance changed when underwear was implemented, rather we were interested in the additive effects of the differential-reinforcement component when combined with underwear. Therefore, implementation of the added differential-reinforcement component was staggered across children already exposed to the underwear component. If improvements in toileting performance were observed following the introduction of the differential-reinforcement component (following stability), the differential-reinforcement component was then removed from each child's toilet-training program, leaving only the underwear component. Researchers have used a similar reversal design within an overarching treatment evaluation to demonstrate experimental control (Iwata, Pace, Cowdery, & Miltenberger, 1994; Mason & Iwata, 1990; Smith, Iwata, Vollmer, & Zarcone, 1993). For example, Iwata, Pace, Cowdery, and Miltenberger (1994) demonstrated the ineffectiveness of sensory extinction on one subject's self-injury by showing that Jack's problem behavior decreased only when escape extinction was

implemented. Visual inspection of each child's toilet-training data was used as the primary means of data analysis.

## **Results and Discussion**

Toilet-training data for children who participated in Study 2 are depicted in Figure 9. Ellie's baseline data showed a high number of accidents with zero levels of appropriate eliminations and self-initiations. Similar performance was observed during her only no-treatment probe conducted in the baseline phase. Gradual improvements in Ellie's toileting performance were observed when the underwear component was implemented. A decrease in the frequency of accidents, and increased levels of appropriate eliminations were observed during the underwear phase. The level of self-initiations was low but variable. A decrease in Ellie's toileting performance was observed during her second no-treatment probe. Next, the differential-reinforcement component was added, and Ellie's toileting performance remained high; however, a decrease in the level of appropriate eliminations was observed when a third no-treatment probe was conducted. Ellie's toileting performance continued to remain high when the differential-reinforcement component was removed. Interestingly, an increase in self-initiations was also observed, suggesting a possible extinction burst. The level of self-initiations continued to increase during additional no-treatment probes. In fact, Ellie's overall toileting performance remained high during her final no-treatment probes. Therefore, we adjusted her toileting protocol such that she was placed back in underwear, and classroom teachers simply reminded Ellie every 90 min that she could sit on the toilet if needed. Ellie's toileting performance remained high during this last phase. No additional toilet-training components were implemented with Ellie.

Brandon's toileting performance remained poor across baseline, underwear, and

underwear with differential reinforcement phases, suggesting that neither underwear nor underwear with the differential-reinforcement component was sufficient to toilet train Brandon. As for children from Study 1 who were not toilet trained using the procedures described in this manuscript, additional toilet-training procedures were implemented to improve Brandon's toileting performance.

Hayes' baseline toileting performance was variable. A low-to-moderate frequency of accidents, a low-to-high level of appropriate eliminations, and a low-to-zero frequency of self-initiations were observed during Hayes' baseline. Hayes' no-treatment probe performance during the baseline phase was similar to Hayes' baseline toileting performance. Improvements in Hayes' toileting performance were observed when underwear was introduced. Teachers recorded a decrease in accidents and an increase in appropriate eliminations during the underwear phase. Self-initiations remained low. A slightly higher frequency of accidents and a lower percentage of appropriate eliminations were observed during the only no-treatment probe conducted during the underwear phase. The differential-reinforcement component was then added, and slight decrements in Hayes' toileting performance were observed. Hayes had more accidents and a lower percentage of appropriate eliminations when the differential-reinforcement component was added. A higher frequency of self-initiations was observed initially; however, self-initiations returned to previous levels across the phase. Additional procedures were implemented to improve Hayes' toileting performance.

Madelyn's baseline toileting performance was also poor. A low-to-moderate level of accidents and low levels of appropriate eliminations and self-initiations were observed during Madelyn's baseline. Madelyn's toileting performance remained largely unchanged when underwear was introduced; however, fewer overall accidents were observed. The added

differential-reinforcement component was correlated with a large increase in self-initiations. Unfortunately, accidents persisted and appropriate eliminations remained low. Next, the differential-reinforcement component was removed, and self-initiations returned to previous levels across the phase. In an attempt to increase Madelyn's number of daily urinations, teachers began offering Madelyn additional fluids (assorted reduced-sugar fruit juices per parental request) during each undergarment check (i.e., every 30 min) on day 44. Unfortunately, additional opportunities to consume fluids had no noticeable effect on Madelyn's toileting performance, as she often refused to drink or consumed the equivalent of a "sip." The differential-reinforcement component was then re-implemented, and previously observed higher frequencies of self-initiations were replicated. Additional procedures were implemented to improve Madelyn's toileting performance.

The addition of the differential-reinforcement component failed to improve overall toileting performance for any child in Study 2. However, an increase in the frequency of self-initiations was observed for one child (Madelyn). Unfortunately, Madelyn's increased requesting to sit on the toilet did not generalize to improvements in accidents or appropriate eliminations, raising questions as to whether it is valuable clinically to reinforce a child's saying "potty" if improvements are not observed in other measures of toileting performance. Ivy (Figure 4) demonstrated a similar pattern of responding in Study 1.

### **General Discussion**

In Study 1, we conducted a component analysis evaluating the combined and individual effects of three commonly used toilet-training procedures on levels of accidents, appropriate eliminations, and self-initiations. When underwear, a dense schedule of sits on the toilet, and differential-reinforcement components were combined, we observed overall improvements in



toileting performance for five (Aaron, Gayle, Jim, Ingrid, and Bethany) of six children. Overall improvements were observed for two (Tammy and Sully) of four children exposed initially to only the underwear component. Overall improvements in toileting performance were not observed for any of the four children exposed initially to the dense-sit schedule component alone. Likewise, none of the four children exposed initially to only the differential-reinforcement component demonstrated overall improvements in toileting performance. The toileting performance of two children (Missy and Jasmine) improved seemingly as a result of prolonged exposure to baseline contingencies, highlighting the presence of low-intensity toilet-training procedures found within the baseline condition. Additionally, we found no evidence that the results of Study 1 were due to the age of our subjects or their gender. Furthermore, the toilet-training package was as effective for our only child (Aaron) diagnosed with Autism Spectrum Disorder (ASD) as it was for other children, raising questions as to whether children with ASD are more difficult to toilet train as has been suggested by some researchers (e.g., Ando, 1977).

In Study 2, we further evaluated the additive effectiveness of the differential-reinforcement component when combined with underwear. Overall improvements in toileting performance were not observed for any of the four children when the differential-reinforcement component was combined with underwear. Because children in Study 2 were exposed initially to the underwear component following baseline, we effectively conducted a direct replication of the underwear evaluation in Study 1. Analyzing the Study 2 data in this manner, placing children in underwear produced an overall improvement in toileting performance for two (Ellie and Hayes) of four children.

Therefore, it appears that simply placing children in underwear is sufficient for some

children to improve toileting performance. Combining the results of Studies 1 and 2, eight children were exposed initially to the underwear component following baseline, and overall improvements in toileting performance were observed for four children for a 50% success rate. This percentage is similar to the success rate of underwear reported by Simon and Thompson (2006). Simon and Thompson (2006) found clear improvements in toileting performance for two (Anne and Beth) of five typically developing children for a 40% success rate. However, Darcy's data (Figure 1 in Simon & Thompson, 2006) showed a clear decrease in the frequency of accidents when underwear was introduced. We found similar patterns of responding for some children in which initial toileting performance following underwear introduction was not representative of later performance (i.e., a delay existed). Therefore, it is possible that had Darcy been exposed to additional days wearing underwear, larger improvements in toileting performance would have been observed (for striking similarity in responding, compare Darcy's data in Simon & Thompson, 2006 to Tammy's data in Figure 2 from Study 1). Had Darcy's data shown improvements in toileting performance, Simon and Thompson's (2006) success rate would have been 60%. Regardless, it appears from the current study and from the results of Simon and Thompson (2006) that the success rate of underwear is somewhere around 50%. Future research could determine more precisely the underwear success rate in both children with and without IDD by conducting similar procedures with additional subjects.

Although the underwear component was sufficient for some children in Studies 1 and 2 to become toilet training, it remains unclear whether underwear was a necessary component of the toilet-training package. Despite the differential-reinforcement component and dense-sit schedule component never being implemented simultaneously following baseline for any subject, both components were implemented sequentially with some children. Ivy, Christy, and Ernie were

exposed to the differential-reinforcement component and dense-sit schedule component sequentially, and overall toileting performance remained low when both components were implemented (see Tables 2 and 3). The lack of improvement in toileting performance when the differential-reinforcement component and dense-sit schedule component were implemented together suggests that underwear was both necessary and sufficient to improve toileting performance for children whose performance improved during the toilet-training package. However, the methodology of Study 1 does not permit an evaluation of why the underwear component was effective.

Informal observations conducted by classroom teachers and supervisors noted several potentially relevant stimulus changes that coincided with children wearing underwear. First, children wearing underwear often saturated their clothing (underwear, pants, and occasionally socks and shoes) when they had an accident while wearing underwear. Most children appeared uncomfortable (ceasing to move, walking with a strange gait, etc.) after having an accident while in underwear. The same children were not observed to saturate their clothing or appear uncomfortable when diapers or pull-ons were used. If the presence of saturated clothes is an aversive event, the act of saturating one's clothes (i.e., having an accident) could have been punished in children wearing underwear. So, it is possible that having an accident while wearing underwear serves as positive punishment for having an accident. Decreases in accidents and general improvements in toileting performance may have been due to positive punishment of accidents while wearing underwear.

Second, classroom teachers and supervisors were more likely to identify and respond more quickly to children in underwear that had an accident. At any point in which teachers identified that a child had an accident, the child was immediately prompted to the toileting area

to be changed. Children often resisted leaving the presumably preferred on-going classroom activities to go to the toileting area. For some children, teacher-delivered prompts to “walk to the potty” appeared aversive, and children wearing underwear were prompted to the toileting more quickly following an accident. The temporal contiguity between an aversive event (i.e., the removal of preferred activities) and the occurrence of an accident may have served as negative punishment for having an accident for children wearing underwear, therefore leading to a decrease in accidents and general improvements in toileting performance.

Third, classroom teachers and supervisors noticed that some children seemed to have strong preferences for specific pairs of underwear that were imprinted with preferred characters (e.g., action heroes, cartoon characters, etc.). Occasionally, children became emotional (crying, throwing tantrums, etc.) when their favorite pair of underwear was unavailable. Children who soiled the underwear that their parents supplied for the day were placed in underwear purchased by or donated to the classroom. It is possible that if children generally wear their most preferred underwear when available, losing preferred underwear following an accident and being required to wear non-preferred underwear could be aversive for some children. If children often lose preferred underwear contingent on having an accident, a decrease in behavior (i.e., accidents) leading to loss of the preferred underwear might occur. Negative punishment of accidents in the form of the removal of preferred underwear might have led to decreases in accidents and improvements in toileting performance for children wearing underwear.

Therefore, at least three possibilities exist regarding how the underwear component may have influenced toileting performance: (a) positive punishment of accidents via contingent wetting one’s clothing, (b) negative punishment of accidents via contingent removal of preferred activities, and (c) negative punishment via contingent removal of preferred underwear. It is also

possible that the underwear component operates by different processes across children. Future research should attempt to explain how underwear operates. More efficient and possibly more hygienic toilet-training procedures could be developed from such research.

One limitation of the current research is the ineffectiveness of the dense-sit schedule component. We did not observe systematic improvements in any of the three measures of toileting performance for any child for whom the dense-sit schedule was evaluated. It is likely that had we prompted children to sit more frequently than every 30 min, this condition would have been more effective. Future research could attempt to evaluate under what schedules of sits on the toilet children are likely to acquire toileting skills.

Another limitation is our apparent failure to either identify or effectively arrange reinforcers for some subjects. As discussed above, we performed two MSWO preference assessments each week with each child to identify highly preferred edible and leisure items for toilet training. Those items were then delivered according to each child's toilet-training condition in effect. We did not observe systematic improvements in any of the three measures of toileting performance for any of the children for whom the differential-reinforcement condition was implemented. A few possibilities for why this condition was ineffective include: (a) the absence of highly preferred items in the preference assessments, (b) preference assessments conducted too infrequently, (c) reinforcer satiation issues, and (d) fluctuations in the integrity with which teachers delivered preferred items. However, we found moderate success providing special individualized trips outside the classroom for appropriate eliminations and remaining dry at scheduled undergarment checks with children for whom the procedures described above were insufficient, suggesting that we simply did not identify effective reinforcers for some children.

In conclusion, toilet training is a particularly difficult endeavor, as toilet training involves the development of complex chains of responses including verbal responses (e.g., requesting toilet access) under stimulus control of inaccessible (i.e., private) events (e.g., a full bladder). Behavior analysis has done much to contribute to a science of toilet training, and many outside the field continue to use our procedures. However, more refined and practical toilet-training programs can be developed only through further research.

## References

- Ando, H. (1977). Training autistic children to urinate in the toilet through operant conditioning techniques. *Journal of Autism and Childhood Schizophrenia*, 7, 151-163.
- Azrin, N. H., Bugle, C., & O'Brien, F. (1971). Behavioral engineering: Two apparatuses for toilet training retarded children. *Journal of Applied Behavior Analysis*, 4, 249-253.
- Azrin, N. H., & Foxx, R. M. (1971). A rapid method of toilet training the institutionalized retarded. *Journal of Applied Behavior Analysis*, 4, 89-99.
- Azrin, N. H., & Foxx, R. M. (1974). *Toilet training in less than a day*. New York, New York: Simon and Schuster Inc.
- Azrin, N. H., & Thienes, P. M. (1978). Rapid elimination of enuresis by intensive learning without a conditioning apparatus. *Behavior Therapy*, 9, 342-354.
- Burgio, L. D., Engel, B. T., Hawkins, A., McCormick, K., Scheve, A., & Jones, L. T. (1990). A staff management system for maintaining improvements in continence with elderly nursing home residents. *Journal of Applied Behavior Analysis*, 23, 111-118.
- Christophersen, E. R. (2003). The case for evidence-based toilet training. *Archives of Pediatrics and Adolescent Medicine*, 157, 1153-1154.
- Cicero, F. R., & Pfadt (2002). Investigation of a reinforcement-based toilet training procedure for children with autism. *Research in Developmental Disabilities*, 23, 319-331.
- DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, 29, 519-533.
- Faith, M. S., Allison, D. B., & Gorman, B. S. (1997). Meta-analysis of single-case research. In R. D. Franklin, D. B. Allison, et al. (Eds.), *Design and analysis of single-case research* (pp. 245-277). Mahwah, NH: Erlbaum.

- Feldman, M. A., Case, L., Garrick, M., MacIntyre-Grande, W., Carnwell, J., & Sparks, B. (1992). Teaching child-care skills to mothers with developmental disabilities. *Journal of Applied Behavior Analysis, 25*, 205-215.
- Foxx, R. M., & Azrin, N. H. (1973). Dry pants: A rapid method of toilet training children. *Behaviour Research and Therapy, 11*, 435-442.
- Friman, P. C., & Vollmer, D. (1995). Successful use of the nocturnal urine alarm for diurnal enuresis. *Journal of Applied Behavior Analysis, 28*, 89-90.
- Giles, D. K., & Wolf, M. M. (1966). Toilet training institutionalized, severe retardates: An application of operant behavior modification techniques. *American Journal of Mental Deficiency, 70*, 766-780.
- Hansen, G. D. (1979). Enuresis control through fading, escape, and avoidance training. *Journal of Applied Behavior Analysis, 12*, 303-307.
- Iwata, B. A., Pace, G. M., Cowdery, G. E., & Miltenberger, R. G. (1994). What makes extinction work: An analysis of procedural form and function. *Journal of Applied Behavior Analysis, 27*, 131-144.
- Kimmel, H. D., & Kimmel, E. (1970). An instrumental conditioning method for the treatment of enuresis. *Journal of Behavior Therapy and Experimental Psychiatry, 1*, 121-123.
- Kroeger, K. A., & Sorensen-Burnworth, R. (2009). Toilet training individuals with autism and other developmental disabilities: A critical review. *Research in Autism Spectrum Disorders, 3*, 607-618.
- Kroeger, K., & Sorensen, R. (2010). A parent training model for toilet training children with autism. *Journal of Intellectual Disability Research, 54*, 556-567.



- LeBlanc, L. A., Carr, J. E., Crossett, S. E., Bennett, C. M., & Detweiler, D. D. (2005). Intensive outpatient behavioral treatment of primary urinary incontinence of children with autism. *Focus on Autism and Other Developmental Disabilities, 20*, 98-105.
- Mahoney, K., Van Wagenen, R. K., & Meyerson, L. (1971). Toilet training normal and retarded children. *Journal of Applied Behavior Analysis, 4*, 173-181.
- Mason, S. A., & Iwata, B. A. (1990). Artifactual effects sensory-integrative therapy on self-injurious behavior. *Journal of Applied Behavior Analysis, 23*, 361-370.
- Nordquist, V. M. (1971). The modification of a child's enuresis: Some response-response relationships. *Journal of Applied Behavior Analysis, 4*, 241-247.
- Rentfrow, R. K., & Rentfrow, D. K. (1969). Studies related to toilet training of the mentally retarded. *The American Journal of Occupational Therapy, 23*, 425-430.
- Schnelle, J. F., Traugher, B., Morgan, D. B., Embry, J. E., Binion, A. F., & Coleman, A. (1983). Management of geriatric incontinence in nursing homes. *Journal of Applied Behavior Analysis, 16*, 235-241.
- Sells-Love, D., Rinaldi, L. M., & McLaughlin, T. F. (2002). Toilet training an adolescent with severe mental retardation in the classroom: A case study. *Journal of Developmental and Physical Disabilities, 14*, 111-118.
- Simon, J. L., & Thompson, R. H. (2006). The effects of undergarment type on the urinary continence of toddlers. *Journal of Applied Behavior Analysis, 39*, 363-368.
- Smeets, P. M., Lancioni, G. E., Ball, T. S., & Oliva, D. S. (1985). Shaping self-initiated toileting in infants. *Journal of Applied Behavior Analysis, 18*, 303-308.

- Smith, R. G., Iwata, B. A., Vollmer, T. R., & Zarcone, J. R. (1993). Experimental analysis and treatment of multiply controlled self-injury. *Journal of Applied Behavior Analysis*, *26*, 183-196.
- Spangler, P. F., Risley, T. R., & Bilyew, D. D. (1984). The management of dehydration and incontinence in nonambulatory geriatric patients. *Journal of Applied Behavior Analysis*, *17*, 397-401.
- Tarbox, R. S. F., Wallace, M. D., Penrod, B., & Tarbox, J. (2007). Effects of three-step prompting on compliance with caregiver requests. *Journal of Applied Behavior Analysis*, *40*, 703-706.
- Tarbox, R. S. F., Williams, W. L., & Friman, P. C. (2004). Extended diaper wearing: Effects on continence in and out of the diaper. *Journal of Applied Behavior Analysis*, *37*, 97-100.
- Taylor, S., Cipani, E., & Clardy, A. (1994). A stimulus control technique for improving the efficacy of an established toilet training program. *Journal of Behavior Therapy & Experimental Psychiatry*, *25*, 155-159).
- Ward-Horner, J., & Sturmey, P. (2010). Component analyses using single-subject experimental designs: A review. *Journal of Applied Behavior Analysis*, *43*, 685-704.
- Wilder, D. A., & Atwell, J. (2006). Evaluation of a guided compliance procedure to reduce noncompliance among preschool children. *Behavioral Interventions*, *21*, 265-272.

Table 1

*Subject names, ages (in months), and diagnoses at Study 1 commencement.*

<b>Name</b>	<b>Age</b>	<b>Diagnosis</b>
Alton	21	None
Aaron	35	ASD
Bethany	20	None
Blue	28	None
Christy	24	None
Danny	25	None
Ernie	21	None
Gayle	39	None
Ingrid	21	None
Ivy	22	None
Jim	27	None
Jasmine	22	None
Leah	32	None
Lizzy	19	None
Missy	28	None
Marge	23	None
Nancy	32	None
Sully	28	None
Sebastian	24	None
Tammy	22	None

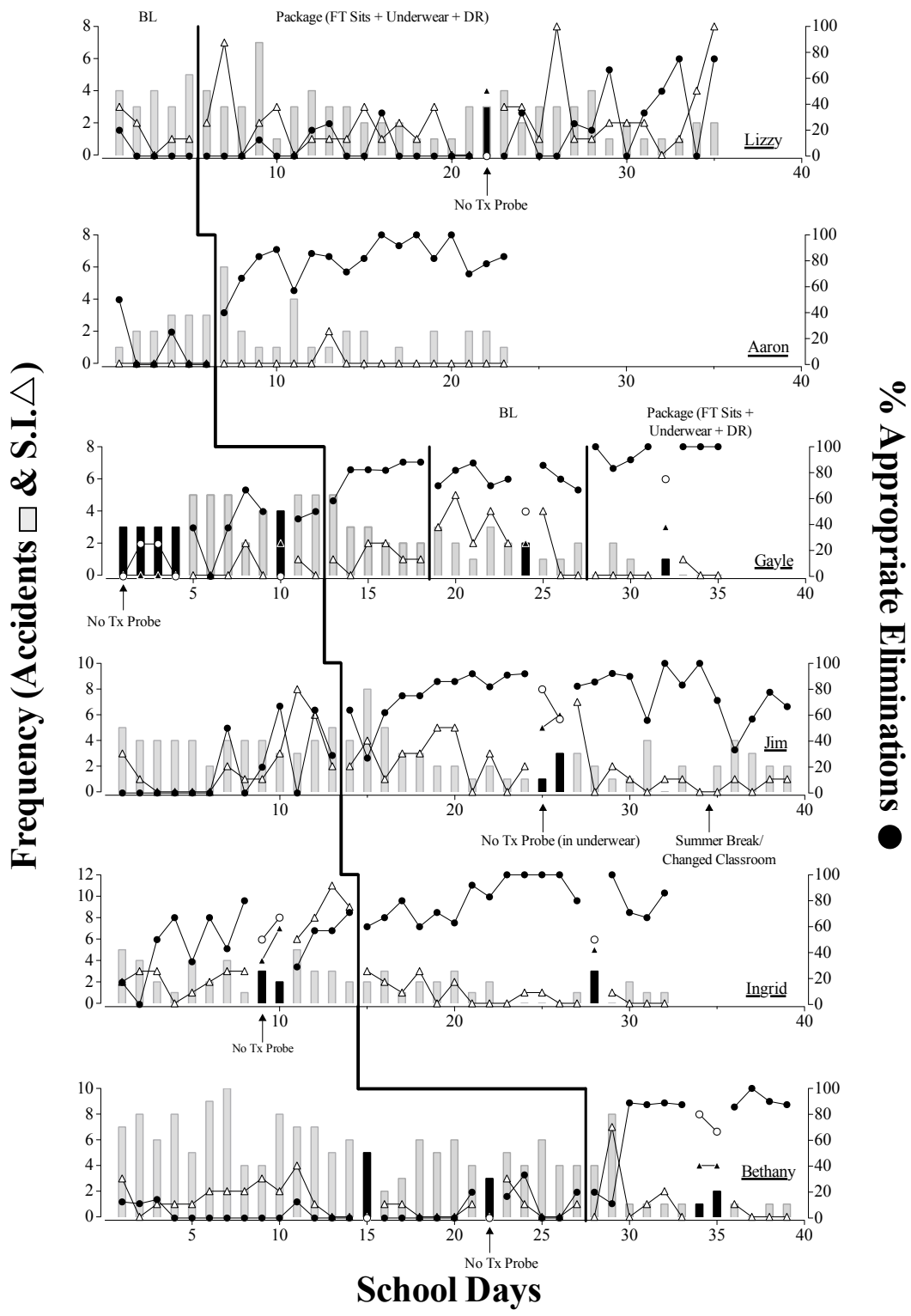


Figure 1. Component analysis data for children (Lizzy, Aaron, Gayle, Jim, Ingrid, and Bethany) exposed to the training package following baseline.

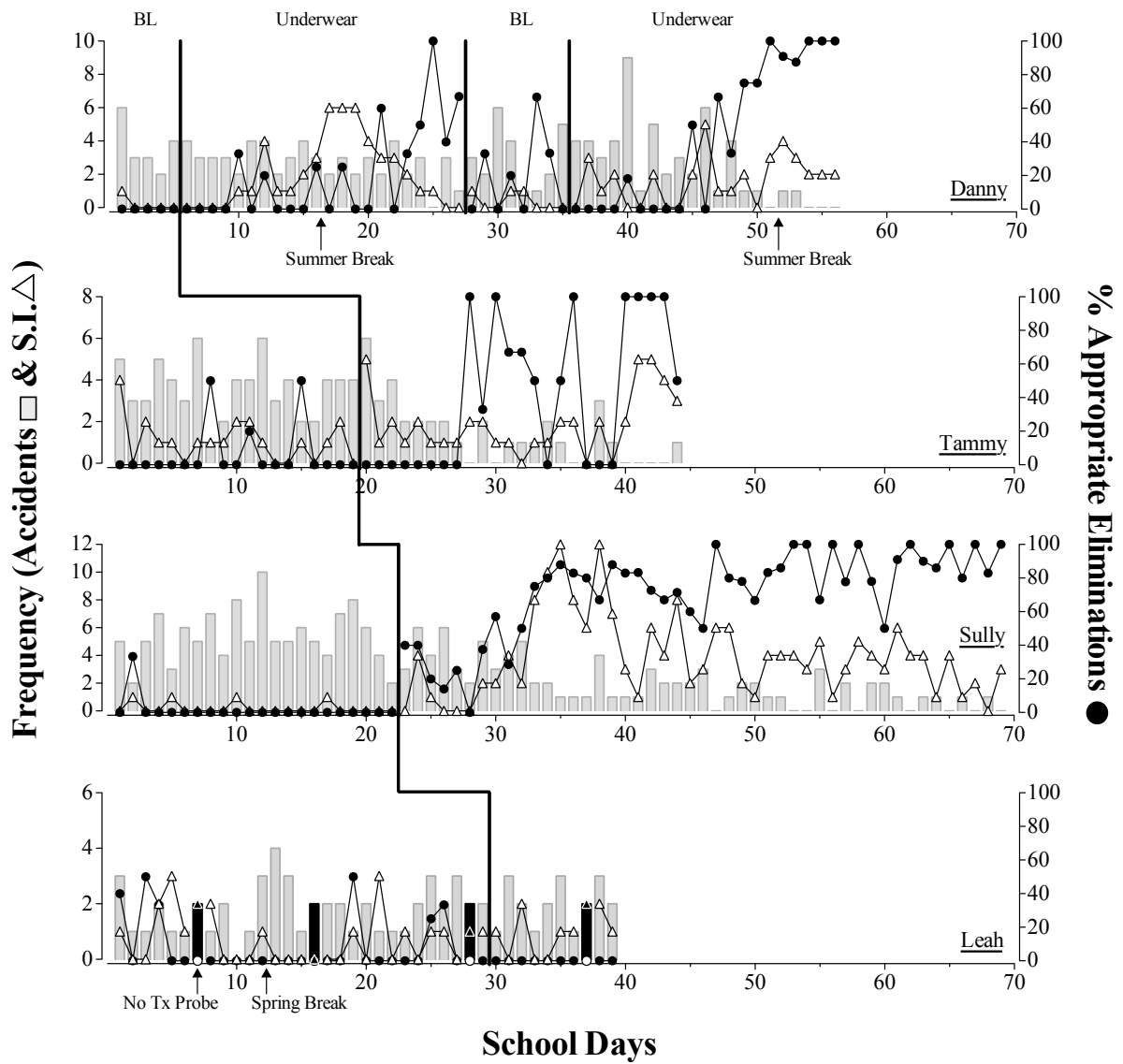


Figure 2. Component analysis data for children (Danny, Tammy, Sully, and Leah) exposed initially to the underwear component following baseline.

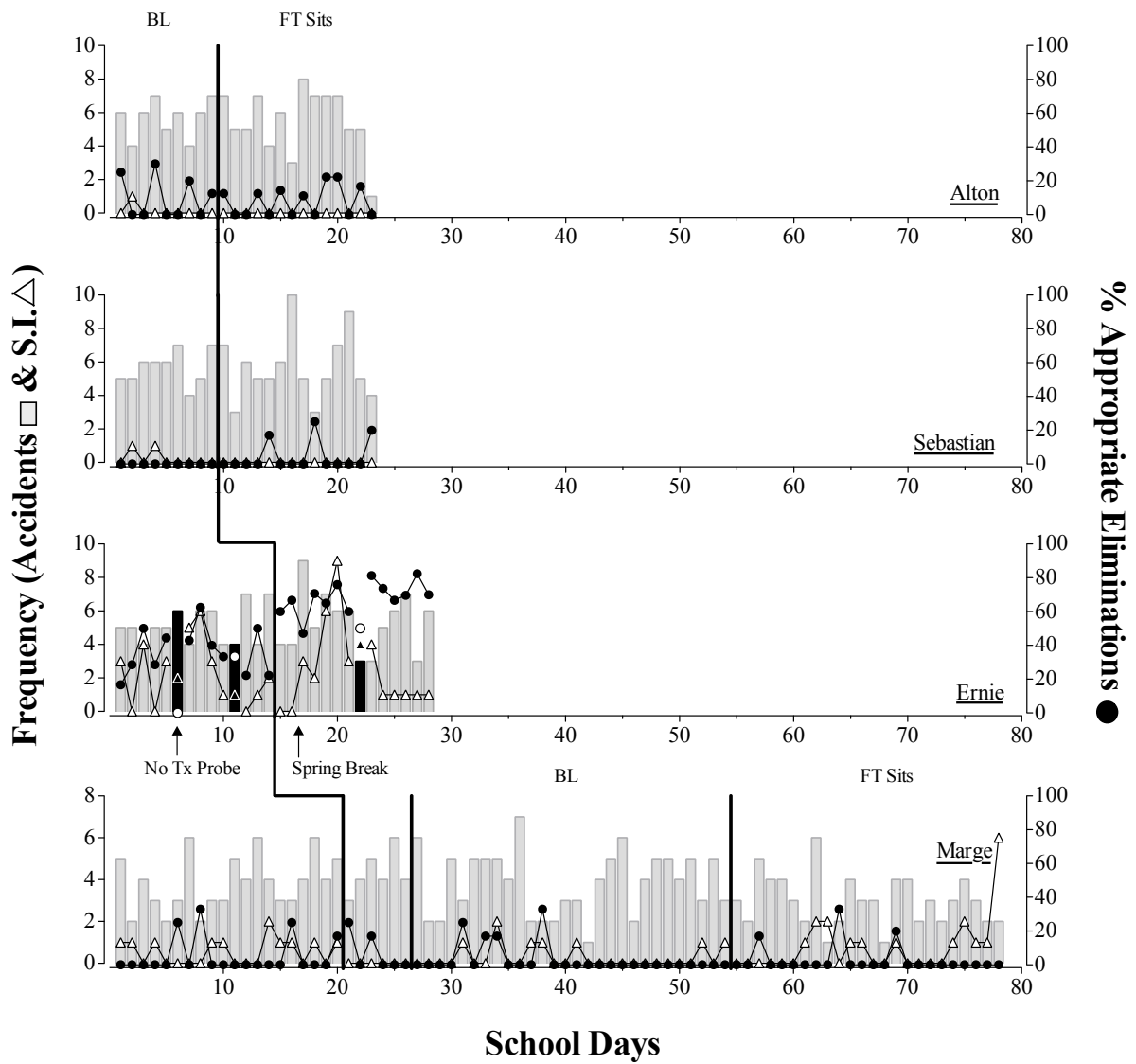


Figure 3. Component analysis data for children (Alton, Sebastian, Ernie, and Marge) exposed initially to the FT Sits component following baseline.

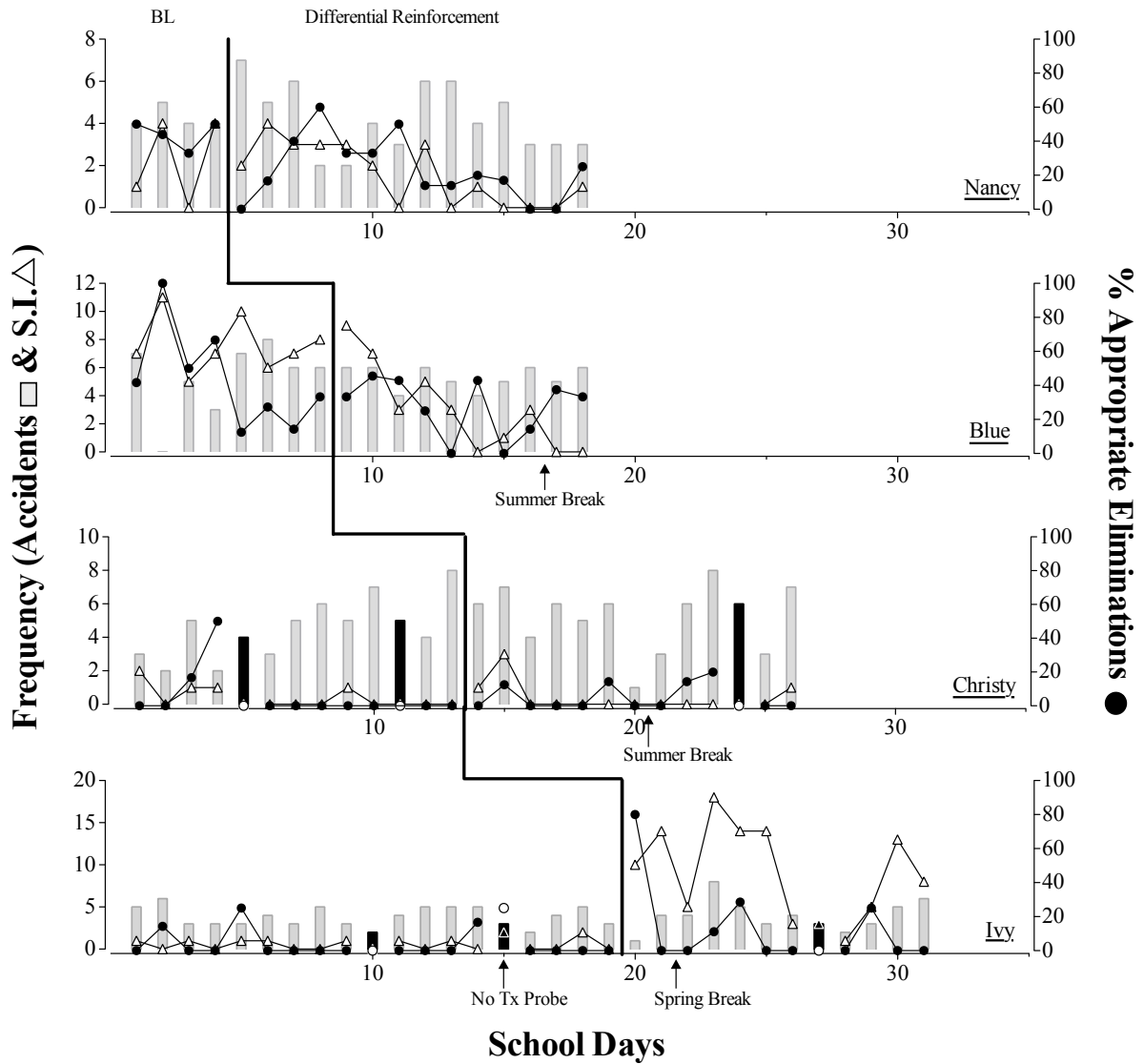


Figure 4. Component analysis data for children (Nancy, Blue, Christy, and Ivy) exposed initially to the differential-reinforcement component following baseline.

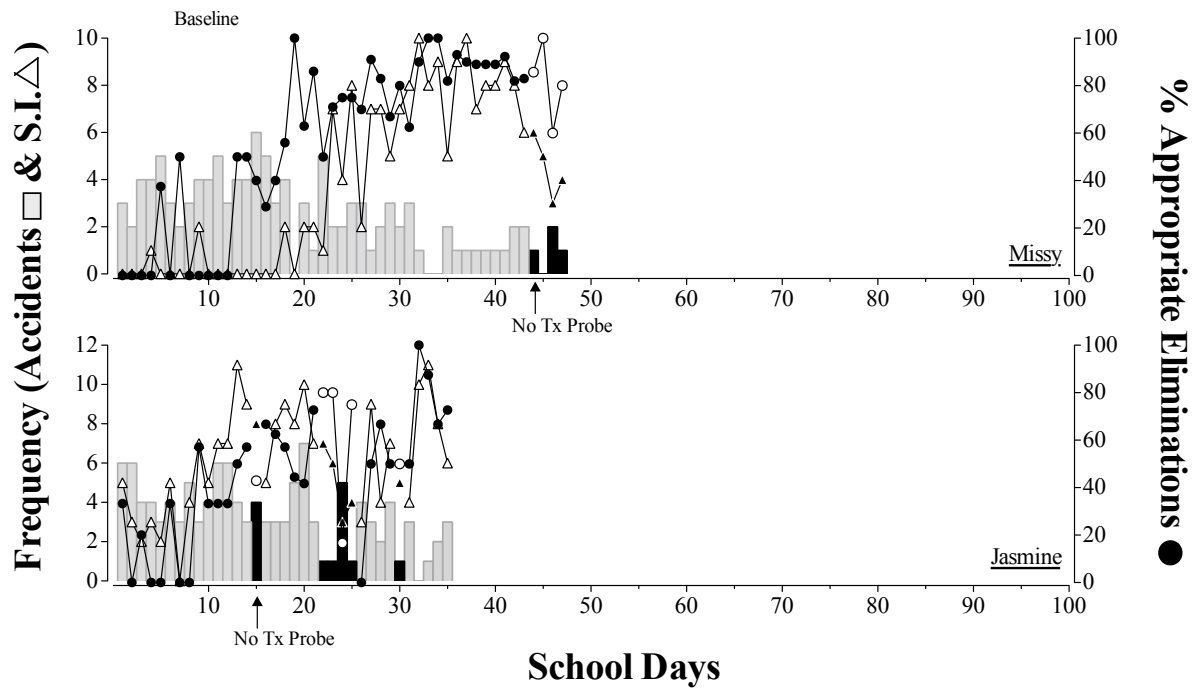


Figure 5. Toiling data for Missy and Jasmine demonstrating acquisition of toiling skills during baseline.



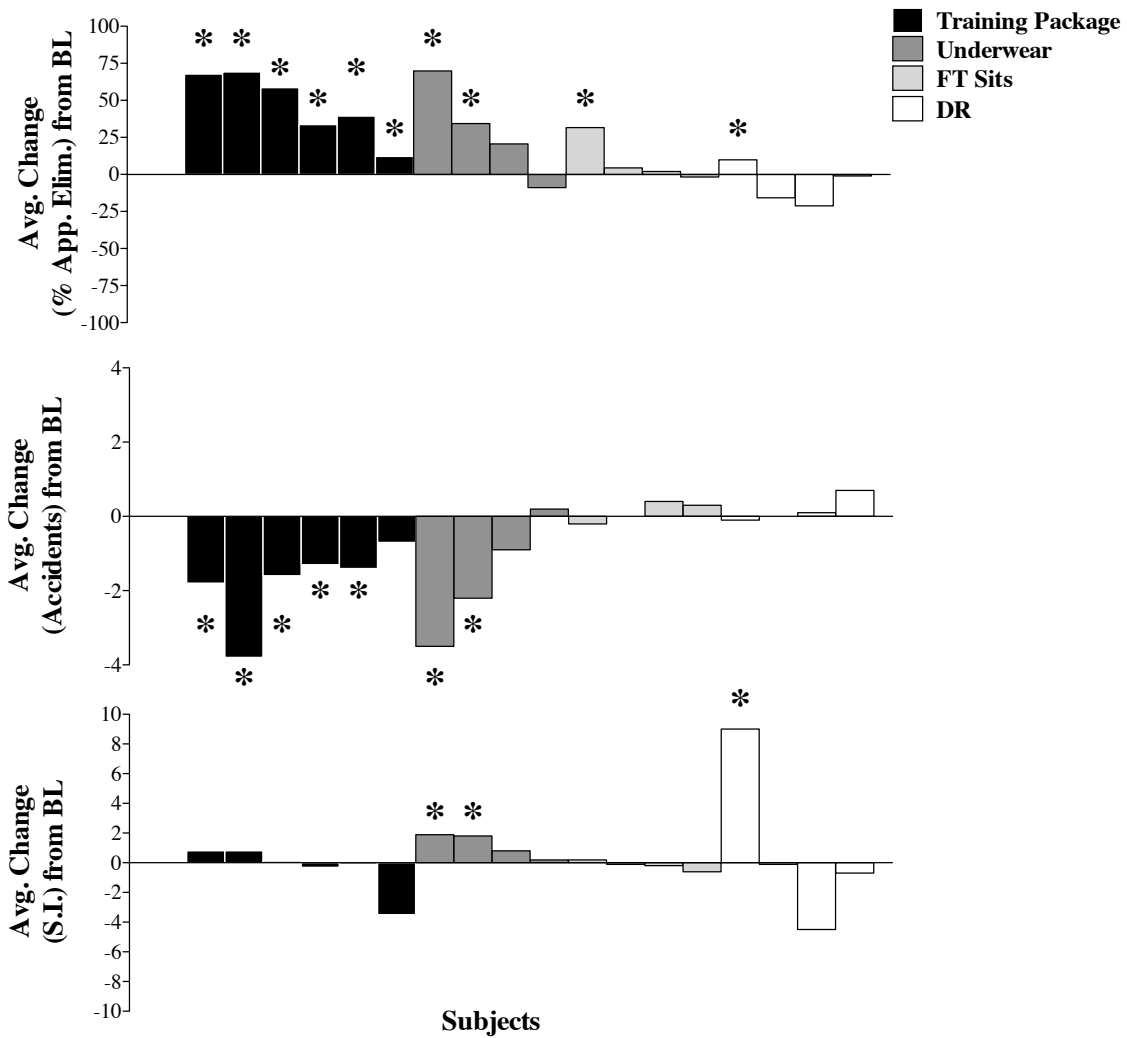


Figure 6. Summary component analysis data indicating average (mean) change in percentage of appropriate eliminations, accidents, and self-initiations for all subjects (except Missy and Jasmine) during the first phase of intervention following baseline. Asterisks above bars indicate effect sizes above 1.0. Data do not include days in which no-treatment probes were conducted.

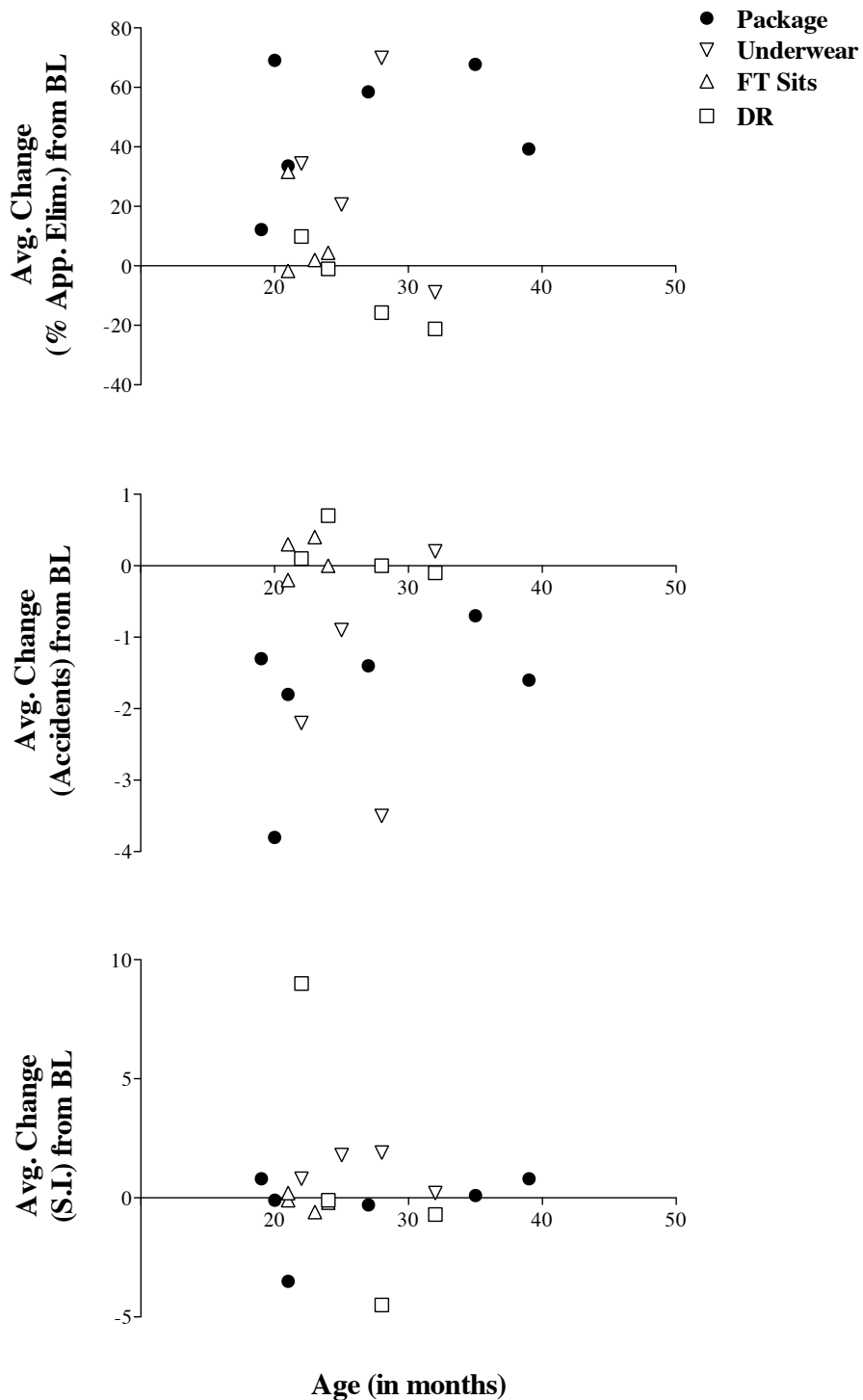


Figure 7. Mean difference in percentage of appropriate eliminations, frequency of accidents, and frequency of self-initiations from baseline during the first treatment implemented given subject age. Data do not include days in which no-treatment probes were conducted.

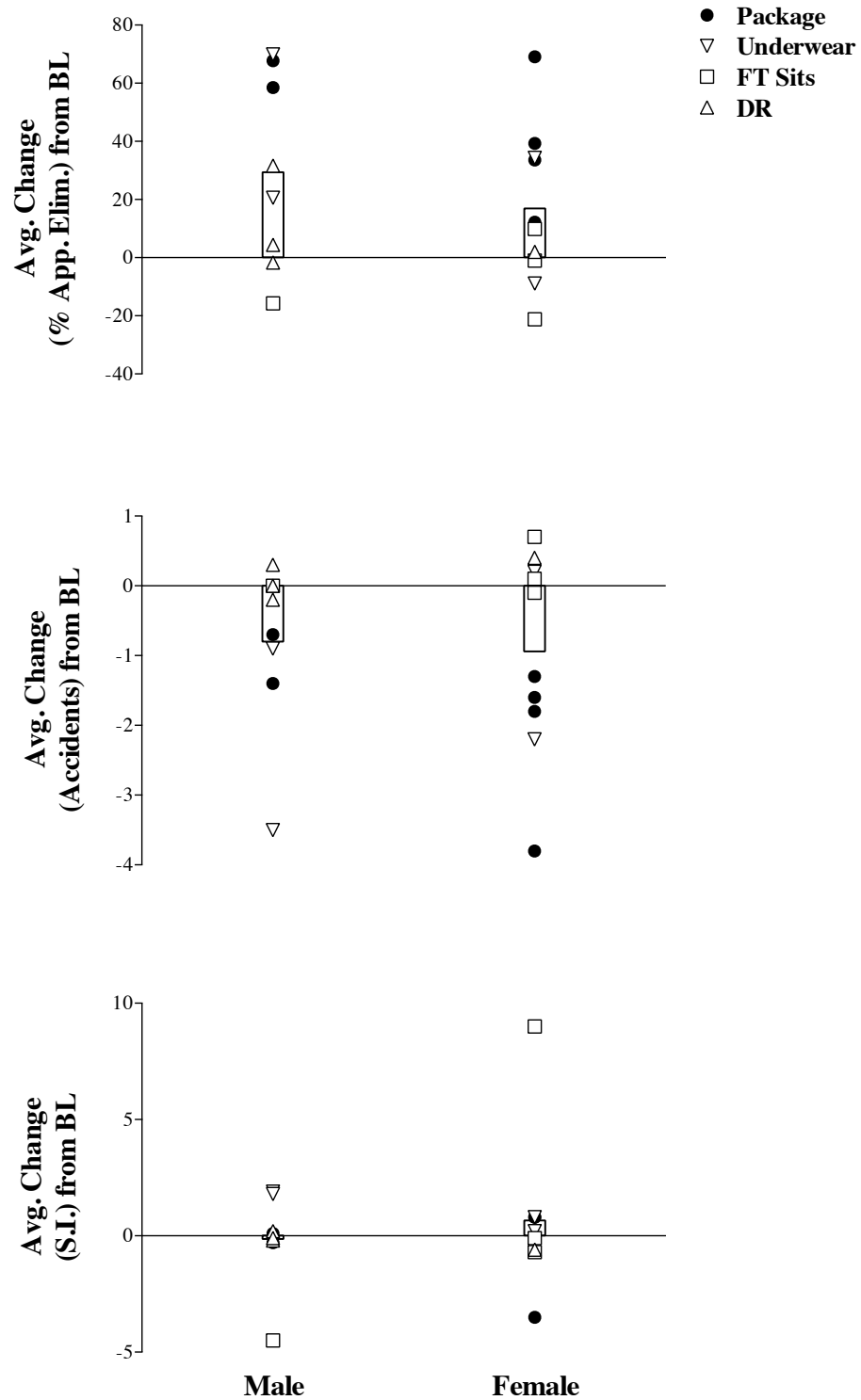


Figure 8. Mean difference in percentage of appropriate eliminations, frequency of accidents, and frequency of self-initiations from baseline during the first treatment implemented given subject gender. Data do not include days in which no-treatment probes were conducted.

Table 2

Mean baseline performance and mean difference in the percentage of appropriate eliminations (first change value), frequency of accidents (second change value), and frequency of self-initiations (third change value) for subjects exposed to the component analysis. Double dashes (--) indicate when additional treatments were no longer needed and therefore not implemented. "Other" indicates when training procedures outside the scope of the current study were implemented. Data do not include days in which no-treatment probes were conducted.

Subject	Baseline Average			Treatment	1 <sup>st</sup> Treatment			Treatment	2 <sup>nd</sup> Treatment Added			Treatment	3 <sup>rd</sup> Treatment Added		
	A.E.	Acc.	S.I.		A.E.	Acc.	S.I.		A.E.	Acc.	S.I.		A.E.	Acc.	S.I.
Jim	17.7	3.8	2.1	Package	58.5	-1.4	-0.3	Other	N/A	N/A	N/A	--	--	--	--
Aaron	12.5	2.3	0.0	Package	67.7	-0.7	0.1	--	--	--	--	--	--	--	--
Ingrid	47.6	3.0	4.3	Package	33.6	-1.8	-3.5	Other	N/A	N/A	N/A	--	--	--	--
Bethany	5.6	5.7	1.2	Package	69.1	-3.8	-0.1	--	--	--	--	--	--	--	--
Gayle	40.9	4.4	0.4	Package	39.3	-1.6	0.8	--	--	--	--	--	--	--	--
Lizzy	4.0	3.8	1.4	Package	12.2	-1.3	0.8	Other	N/A	N/A	N/A	--	--	--	--
Tammy	6.3	3.6	1.1	Underwear	34.3	-2.2	0.8	DR	59.3	-1.4	15.6	--	--	--	--
Sully	1.5	5.4	0.2	Underwear	69.9	-3.5	1.9	DR	18.1	-1.4	-0.6	FT Sits	7.1	-0.3	-0.6
Leah	8.9	1.8	0.7	Underwear	-8.9	0.2	0.2	FT Sits	0.0	0.0	-0.9	DR	10.0	0.5	1.2
Danny	0.0	3.6	0.2	Underwear	20.6	-0.9	1.8	--	--	--	--	--	--	--	--
Nancy	44.3	4.3	2.3	DR	-21.2	-0.1	-0.7	Underwear	59.8	-3.1	5.0	FT Sits	15.5	-0.8	-4.5
Blue	43.2	5.3	7.6	DR	-15.7	0.0	-4.5	Underwear	37.2	-2.6	-1.8	FT Sits	23.0	-2.0	-1.1
Ivy	3.3	4.0	0.5	DR	9.9	0.1	9.0	FT Sits	-13.2	2.2	-6.8	Underwear	4.9	-1.4	-1.8
Christy	6.1	4.5	0.5	DR	-1.0	0.7	-0.1	FT Sits	0.0	0.1	-0.2	Underwear	76.0	-4.2	0.6
Marge	5.0	3.9	0.6	FT Sits	2.0	0.4	-0.6	Underwear +DR	80.5	-3.8	6.3	Other	N/A	N/A	N/A
Sebastian	0.0	5.7	0.2	FT Sits	4.4	0.0	-0.2	Underwear	27.2	-3.2	0.5	DR	50.6	-2.1	-0.2
Ernie	36.8	5.2	2.3	FT Sits	31.6	0.3	0.2	DR	9.1	-3.1	-1.9	Underwear	-12.2	3.0	0.3
Alton	9.7	5.7	0.1	FT Sits	-1.7	-0.2	-0.1	Other	N/A	N/A	N/A	--	--	--	--

Table 3

Mean baseline performance and effect sizes for changes in the percentage of appropriate eliminations (first change value), frequency of accidents (second change value), and frequency of self-initiations (third change value) for subjects exposed to the component analysis. Double dashes (--) indicate when additional treatments were no longer needed and therefore not implemented. "Other" indicates when training procedures outside the scope of the current study were implemented. Errors (Err) are noted when previous phase performance was zero. Data do not include days in which no-treatment probes were conducted.

Subject	Baseline Average			Treatment	1 <sup>st</sup> Treatment			Treatment	2 <sup>nd</sup> Treatment Added			Treatment	3 <sup>rd</sup> Treatment Added		
	A.E.	Acc.	S.I.		A.E.	Acc.	S.I.		A.E.	Acc.	S.I.		A.E.	Acc.	S.I.
Jim	17.7	3.8	2.1	Package	2.2	-1.7	-1.2	Other	N/A	N/A	N/A	--	--	--	--
Aaron	12.5	2.3	0.0	Package	3.2	-0.8	ERR	--	--	--	--	--	--	--	--
Ingrid	47.6	3.0	4.3	Package	1.4	-1.2	-1.0	Other	N/A	N/A	N/A	--	--	--	--
Bethany	5.6	5.7	1.2	Package	7.5	-2.0	-0.1	--	--	--	--	--	--	--	--
Gayle	40.9	4.4	0.4	Package	1.9	-1.4	0.9	--	--	--	--	--	--	--	--
Lizzy	4.0	3.8	1.4	Package	1.4	-1.6	0.7	Other	N/A	N/A	N/A	--	--	--	--
Tammy	6.3	3.6	1.1	Underwear	2.1	-1.7	0.8	DR	1.4	-1.0	10.5	--	--	--	--
Sully	1.5	5.4	0.2	Underwear	9.8	-1.8	9.3	DR	0.7	-0.8	-0.8	FT Sits	0.6	-0.5	-0.5
Leah	8.9	1.8	0.7	Underwear	-0.5	0.2	0.2	FT Sits	ERR	0	-1.1	DR	ERR	0.5	ERR
Danny	0.0	3.6	0.2	Underwear	ERR	-0.6	4.1	--	--	--	--	--	--	--	--
Nancy	44.3	4.3	2.3	DR	-2.6	-0.1	-0.3	Underwear	3.2	-2.0	3.4	FT Sits	0.6	-0.6	-1.7
Blue	43.2	5.3	7.6	DR	-0.5	0.0	-2.3	Underwear	2.2	-3.2	-0.6	FT Sits	0.9	-0.8	-0.8
Ivy	3.3	4.0	0.5	DR	1.3	0.1	14.4	FT Sits	-0.5	1.2	-1.3	Underwear	ERR	-1.7	-1.0
Christy	6.1	4.5	0.5	DR	-0.1	0.3	-0.1	FT Sits	0.0	0.0	-0.2	Underwear	9.8	-3.5	1.2
Marge	5.0	3.9	0.6	FT Sits	0.2	0.4	-0.9	Underwear +DR	6.8	-3.7	ERR	Other	N/A	N/A	N/A
Sebastian	0.0	5.7	0.2	FT Sits	ERR	0.0	-0.5	Underwear	3.0	-1.6	ERR	DR	2.1	-1.6	-0.2
Ernie	36.8	5.2	2.3	FT Sits	2.3	0.3	0.1	DR	1.5	-1.8	-0.7	Underwear	-1.9	1.5	0.4
Alton	9.7	5.7	0.1	FT Sits	-0.1	-0.1	-0.3	Other	N/A	N/A	N/A	--	--	--	--

Table 4

*Subject names, ages (in months), and diagnoses at Study 2 commencement.*

<b>Name</b>	<b>Age (in months)</b>	<b>Diagnosis</b>
Brandon	20	None
Ellie	26	None
Hayes	26	None
Madelyn	25	None

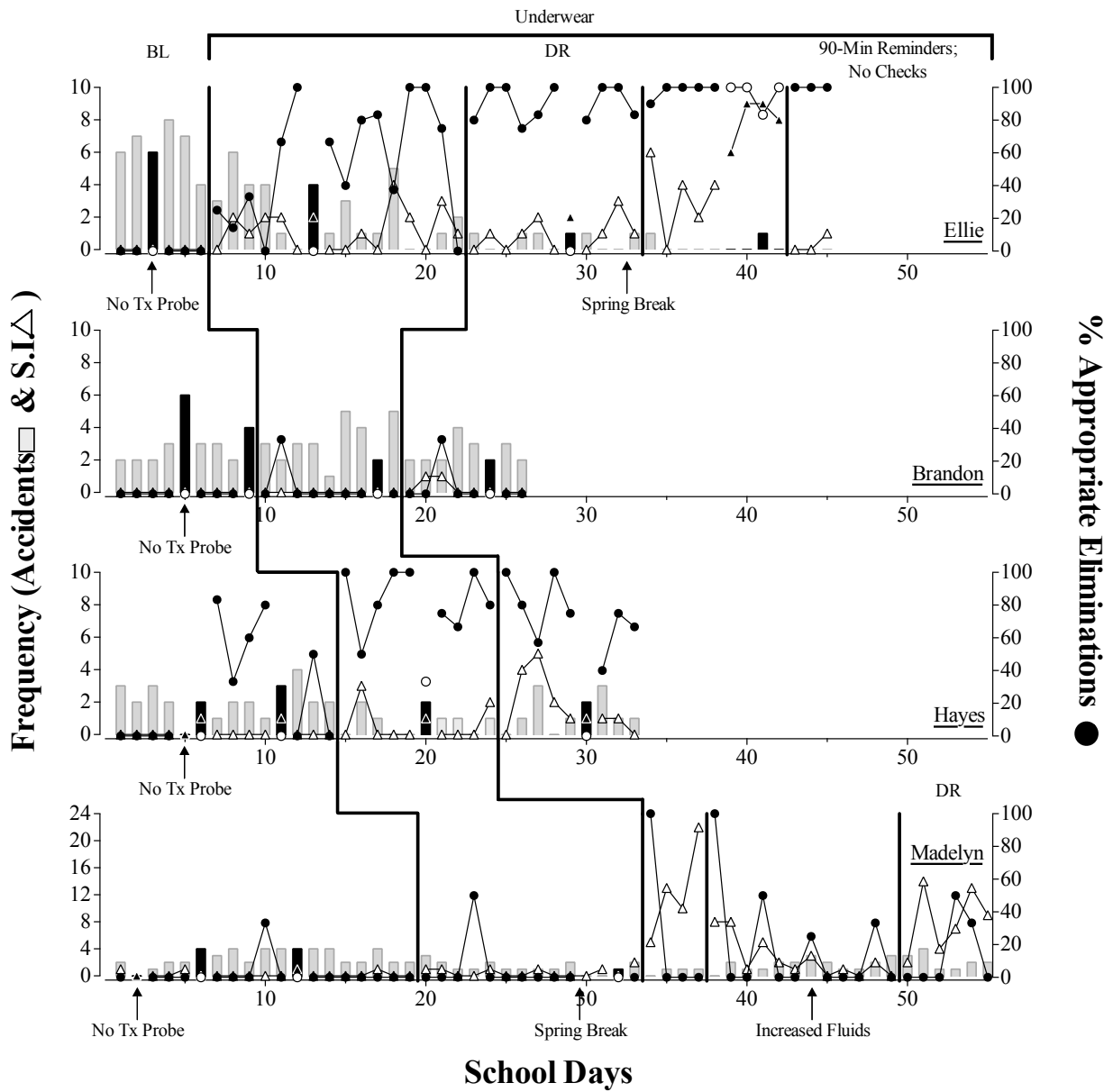


Figure 9. Toileting data for children (Ellie, Brandon, Hayes, and Madelyn) exposed initially to the underwear component followed by exposure to the underwear component combined with the differential-reinforcement component.