

# Deviations from Prescribed Prompting Procedures: Implications for Treatment Integrity

Laura L. Grow · James E. Carr · Kristin V. Gunby ·  
Shaireen M. Charania · Lucita Gonsalves ·  
Inas A. Ktaech · April N. Kisamore

Published online: 14 April 2009  
© Springer Science+Business Media, LLC 2009

**Abstract** The acquisition of new skills may be hindered when teaching procedures vary from previously validated approaches or contain errors. In the present study, we compared the acquisition and maintenance of response chains taught using a perfectly implemented system of least prompts and a multiple verbal prompts procedure (i.e., addition of multiple verbal prompts and failure to follow through with more intrusive prompts). Four children, aged 6–9, participated in the study. An adapted alternating treatments design was used to compare the effectiveness and efficiency of learning during the system of least prompts and the alternative system of least prompts. Results were consistent with those obtained in previous studies in that the perfectly implemented and alternative prompting procedures were effective in teaching new skills for all participants. However, the perfectly implemented treatment required fewer trials to mastery for 4 of the 5 evaluations. Response chains taught under the multiple verbal prompts condition had poorer maintenance for 2 of the 5 evaluations. The results of the current study suggest that deviations from empirically identified teaching procedures may reduce the speed with which new skills are acquired.

**Keywords** Behavioral acquisition · System of least prompts · Treatment integrity

## Introduction

Response prompting is an effective strategy for teaching skills to children with autism, among other disabilities (MacDuff et al. 2001). Response prompts (e.g.,

---

L. L. Grow (✉) · J. E. Carr · A. N. Kisamore  
Department of Psychology, Western Michigan University, Kalamazoo, MI 49008-5439, USA  
e-mail: laura.l.grow@wmich.edu

K. V. Gunby · S. M. Charania · L. Gonsalves · I. A. Ktaech  
Kinark Child and Family Services, Markham, ON, Canada

gestures, physical guidance) are designed to produce more accurate responding during the initial teaching period. Prompts are then systematically removed to transfer stimulus control to the discriminative stimuli that should ideally occasion correct responses. One prompt fading method is the *system of least prompts* (SLP) which provides increasing assistance until the correct response occurs (Horner and Keilitz 1975).

The SLP is comprised of at least four procedural components (Doyle et al. 1988; Wolery and Gast 1984). First, at least two prompts should be selected and arranged in a least-to-most intrusive fashion. Verbal, model, and physical prompts are most commonly used in the SLP. Second, the response interval, or time allotted to respond between prompts, should be chosen. The most commonly reported response interval is 5 s (Doyle et al.); however, the response interval should ideally be selected based on the amount of time needed to complete the task. Another consideration is the amount of time the learner typically takes to respond to instructions. Generally speaking, the response interval should be similar to the amount of time that a learner typically requires to emit a response. For example, if a learner takes approximately 10 s to respond to instructions, a response interval of 3 s would be unreasonable. Third, consequences for both correct and incorrect responses should be selected. Differential reinforcement is commonly used in conjunction with SLP such that independent responses result in highly preferred consequences and prompted responses result in less preferred consequences.

Although, several authors have recommended “optimal” procedural parameters for the SLP, some studies have evaluated the extent to which planned variations are likely to lead to more efficient acquisition (e.g., Steege et al. 1987; West and Billingsley 2005). For example, West and Billingsley compared the traditional SLP to a revised SLP. The traditional system included: (a) verbal, gestural, and physical prompts, (b) a 5- to 10-s response interval, depending on the task, (c) a discriminative stimulus accompanied by the gestural and physical prompts, and (d) praise and access to a small food item for correct responses. The procedures were similar for the revised system except that the discriminative stimulus was omitted during response prompts. Both variations were effective in teaching individuals with moderate to severe mental retardation to pour a beverage, spray water on plants, and retrieve a backpack. However, the revised SLP system was more efficient in transferring stimulus control from therapist-delivered prompts to the discriminative stimulus. This study suggests that the omission of the discriminative stimulus during response prompts results in greater efficiency during SLP. However, some tasks may require the discriminative stimulus to be repeated if correct responses fail to occur. For example, auditory-visual conditional discriminations (e.g., receptive identification) require additional presentations of the auditory conditional stimulus (e.g., “Which one is blue?”) because it is impossible to complete the task without it.

Steege et al. (1987) conducted another study that evaluated the effectiveness and efficiency of variations of the SLP. Two types of SLP procedures were evaluated. The “traditional” method consisted of issuing increasing amounts of assistance until the learner emitted the correct response. The “prescriptive” method involved starting with the level of prompt that was required to occasion correct responding in

the preceding trial. For example, if a correct response was emitted following the gestural prompt, the following trial was initiated by the presentation of the discriminative stimulus followed by the gestural prompt rather than the verbal prompt. Both traditional and prescriptive methods resulted in skill acquisition; although, the prescriptive method was more efficient in terms of trials to mastery and duration of training. Taken together, the results of these studies suggest that certain variations in the SLP alter its efficiency. However, other variations might alter its effectiveness and efficiency, especially if the variations can be considered errors of treatment integrity.

Repp et al. (1981) assessed the types of prompts that therapists naturally provided during instruction and evaluated the likelihood of compliance following each type of prompt. The authors found that single verbal prompts were the most common form of prompt; however, this prompt was the least likely to evoke correct responding. One implication of the study is that behavior-change agents may continue to provide multiple verbal prompts despite the less-than-ideal outcomes that may result from their use.

Consistently implementing a prompting procedure with perfect integrity may be less likely in natural settings such as classrooms. Thus Wolery and Shuster (1997), noted that additional research should be conducted to evaluate instructional methods with less-than-ideal treatment integrity for learners with developmental disabilities. In school settings, teachers may provide additional verbal prompts in the form of coaxing (e.g., “Come on, you can do it”) instead of providing more intrusive prompts as indicated by the prompted procedure. Multiple verbal prompts may be issued during instruction with individuals because verbal prompts require the least effort of response prompts and are portable unlike model and physical prompts. Therefore, the purpose of the current investigation was to evaluate treatment integrity failures during the SLP. More specifically, the SLP was compared to *multiple verbal prompts* (MVP) which included additional verbal prompts, omitted the opportunity to respond to the model prompt, and excluded the physical prompt.

## Method

### Participants and Setting

Four children who had been previously diagnosed with autism participated in the study. Max, Jason, Anton, and Tony were 6, 7, 9, and 7 years old, respectively, at the beginning of the study. Inclusion in the study required that participants (a) had a previous diagnosis of a developmental disability (e.g., autism spectrum disorder or mental retardation), (b) had a generalized motor imitation repertoire, (c) exhibited little or no severe problem behavior (e.g., aggressive or self-injurious behavior), (d) tolerated physical contact, and (e) had a history of responding appropriately to vocal instructions. The participants were considered tolerant of physical contact if they did not engage in behaviors that indicated stress such as physically resisting contact or crying following light physical contact from the experimenter (e.g., pat on the back, touch their arm).

**Table 1** Results of the Gilliam autism rating scale (GARS), peabody picture vocabulary test III (PPVT-III), and expressive vocabulary test (EVT)

	GARS		Standard score	
	Autism quotient	Probability of autism	PPVT-III	EVT
Max (6 years 5 months)	63	Very low	63	55
Jason (7 years 10 months)	77	Low	73	49
Anton (9 years 3 months)	103	Average	40	40
Tony (7 years 2 months)	119	Above average	73	58

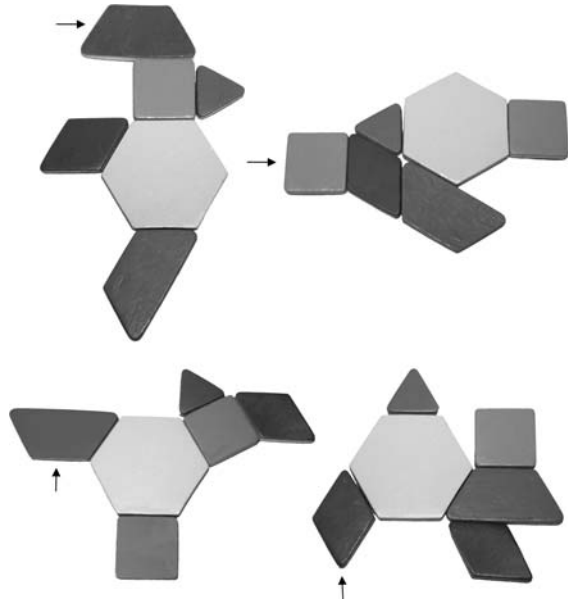
Table 1 shows the results of several standardized assessments that were administered to each participant or one of his parents. All standardized assessments were conducted as part of the study and occurred prior to the experimental analysis. The purpose of the standardized assessments was to document some of the key characteristics of participants at the time of the study. The Gilliam Autism Rating Scale (GARS), which was completed by a parent, assesses social behavior, stereotypy, communication, and developmental history (Gilliam 1995). The GARS autism quotient, a global measure of the domains, and corresponding probability of an autism diagnosis are presented in Table 1 and ranged from very low to above average. The Peabody Picture Vocabulary Test-Third Edition (PPVT-III) provides a measure of receptive language skills. The forms of PPVT-III are reliable with each other and scores are stable up to 1 month. The PPVT-III correlates well with other common measures of vocabulary and moderately well with verbal repertoires (Dunn and Dunn 1997). The standard scores across participants ranged from 40 to 73. The Expressive Vocabulary Test (EVT) measures expressive language skills. There is a high degree of internal and test-retest reliability in the measurement of EVT scores. The EVT is moderately correlated with other measures of cognitive ability (e.g., Wechsler Intelligence Scale for Children) and the PPVT-III (Williams 1997). The standard scores across participants ranged from 40 to 58. Overall, the language assessments indicated that participants' skills were weaker than might be expected from their same-aged peers; however, the severity of the deficit varied across participants.

Teaching trials were conducted in the corner of a classroom for Max, Jason, and Anton and in the school's cafeteria for Tony. The experimenter sat next to the participant at a table during teaching trials. Two to eight teaching trials were conducted 3–5 times per week for each participant. Trials ranged from 15 s to 1 min 30 s in duration across experimental conditions.

## Tasks

Four arbitrary, 5-step response chains were designed for the study (see Fig. 1). The participant was presented with a magnetic board and a starter geometric piece. The starter piece was different with respect to either color and shape or position and orientation across response chains. The arrows in Fig. 1 indicate the starter piece associated each response chain. Each response chain required the participants to

**Fig. 1** Graphical depiction of the four stimulus sets. Clockwise from the top left, stimulus-set names are “Frooodle”, “Zemic”, “Monig”, and “Delib”. Arrows indicate the starter piece in each set



place five geometric, magnetic pieces on a magnetic board to create a unique configuration. The task difficulty was held constant across tasks by using similar pieces to minimize response effort as a factor of acquisition. Task difficulty was held constant by using similar materials which included the same number of steps for each response chain.

#### Dependent Variables, Data Collection, Interobserver Agreement, and Treatment Integrity

The primary dependent variable was the percentage of correct responses during teaching trials. For each step in the 5-step response chain, an observer scored the participant's behavior as correct, prompted, or a non-response. A correct response was scored if the participant completed a given step without experimenter-delivered prompts (i.e., during the independent opportunity in the prompting sequence). A prompted response was scored if the experimenter delivered prompts during a step and a correct response followed. No response was scored when the participant did not display the target behavior within 5 s of the discriminative stimulus (e.g., “Make a frooodle”) during baseline or during the MVP condition when the experimenter completed the step for the participant. The percentage of correct responses during the trial was calculated by dividing the number of correct responses by the number of opportunities to respond (i.e., five). The secondary dependent variable was the number of trials required to meet the mastery criterion (i.e., three consecutive trials with 100% correct responding) for each response chain during both prompting procedures.

Data were collected by a second, independent observer for each dependent variable across participants. An agreement was defined as both observers scoring a

correct, prompted or no response. A disagreement was scored if the observers scored a response differently from one another. Point-by-point agreement was calculated by dividing the number of agreements for a trial by the number of agreements and disagreements and multiplying by 100%. Point-by-point agreement was assessed for 40% of sessions and averaged 99% (range, 80–100%) across experimental conditions for Max. Point-by-point agreement was assessed for 36% of sessions and averaged 100% across experimental sessions for Jason. Point-by-point agreement was assessed for 37% of sessions and averaged 100% across experimental conditions for Anton. Point-by-point agreement was assessed for 69% of sessions and averaged 99% (range, 80–100%) across experimental conditions for Tony's first evaluation. Point-by-point agreement was assessed for 34% of sessions and averaged 99% (range, 80–100%) across experimental conditions for Tony's second evaluation.

Treatment integrity was assessed across experimental conditions for each participant. An independent observer recorded correct implementation of (a) prompts, (b) the order in which prompts should have occurred, (c) the response interval between prompts, and (d) consequences for correct and incorrect responses. The specific steps that were required for proper implementation of each experimental condition depended on the experimental condition in place and the participants' behavior. Percentage of treatment integrity during a trial was calculated by dividing the number of properly implemented steps by the total number of steps and multiplying by 100%.

Treatment integrity checks were conducted for 40% of sessions and averaged 98% (range, 80–100%) across experimental conditions for Max. Treatment integrity checks were conducted for 36% of sessions and averaged 99% (range, 80–100%) across experimental sessions for Jason. Treatment integrity checks were conducted for 37% of sessions and averaged 99% (range, 93–100%) across experimental conditions for Anton. Treatment integrity checks were conducted for 69% of sessions and averaged 99% (range, 93–100%) across experimental conditions for Tony's first evaluation. Treatment integrity checks were conducted for 34% of sessions and averaged 98% (range, 87–100%) across experimental conditions for Tony's second evaluation.

## Procedures

### *Preference Assessments*

Prior to the study, parents or teachers were interviewed to identify preferred toys and activities to evaluate in a paired-stimulus preference assessment, which was subsequently conducted to identify potential reinforcers to use in the study (Fisher et al. 1992). The top four items that were selected for each participant during the paired-stimulus preference assessment were used during a brief multiple-stimulus (without replacement) preference assessment (MSWO; De Leon and Iwata 1996) prior to each teaching trial. The purpose of the brief MSWO assessment was to identify the programmed consequence used in the subsequent teaching trial.

### Trial Format

A trial consisted of (a) the presentation of a discriminative stimulus (e.g., “Make a froodle” and the magnetic board with a starter piece), (b) an independent opportunity to respond, (c) the delivery of prompts as necessary to complete the step, and (d) the presentation of a reinforcer, if appropriate. This process was repeating until all steps in the response chain were completed. The vocal discriminative stimulus was only presented prior to the first step and omitted during subsequent steps. The other aforementioned components were included in subsequent steps in the response chain. Participants were given an opportunity to respond independently during each step in the response chain. The response chain consisted of placing five geometric, magnetic shapes on a board in a configuration that was associated with the specific discriminative stimulus delivered to the participant. Response chains were taught using total task presentation in which the teaching trial was initiated with the first step in the chain and instruction continued until the final step was completed. See Fig. 2 for a graphical depiction of the antecedents for one step in the response chain presented across experimental conditions.

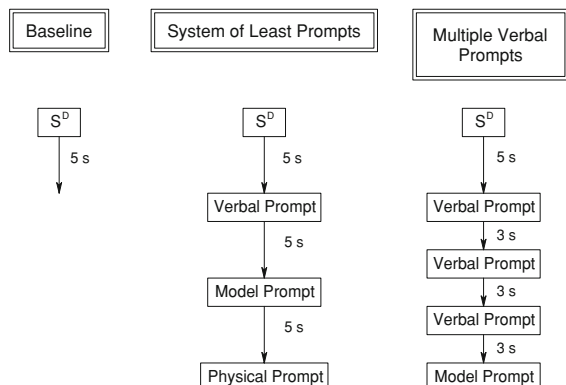
### Baseline

Baseline trials were conducted prior to teaching trials for each participant. Baseline consisted of presenting materials that were appropriate to the specific task (i.e., magnetic board with starter piece and appropriate geometric shapes to complete the chain). The experimenter presented the discriminative stimulus to the participant and allowed 5 s for the participant to start the response chain. A trial was terminated after the first error or 5 s with no responding. Regardless of the participant’s performance, prompts were not delivered.

### System of Least Prompts

The procedural parameters recommended by Wolery and Gast (1984) were employed during the SLP condition. The experimenter presented the discriminative

**Fig. 2** Antecedents presented in each condition for a single step of the response chain



stimulus to the participant and allowed 5 s for the participant to start the response chain. If the participant completed the first step in the chain independently, the experimenter delivered enthusiastic praise with physical attention (e.g., tickles, pat on the back). If the participant responded incorrectly or did not respond within 5 s, the experimenter delivered a verbal prompt (e.g., “Put the yellow piece on.”). If the participant completed the step following the verbal prompt, the experimenter delivered praise only. If the participant responded incorrectly or did not respond within 5 s, the experimenter delivered a model prompt (i.e., “Put the yellow piece on like this” while simultaneously modeling the step). If the participant completed the step following the model prompt, the experimenter delivered praise only. If the participant responded incorrectly or did not respond within 5 s, the experimenter delivered a physical prompt (i.e., “Put the yellow piece on” while using hand-over-hand guidance). No attention followed steps completed following physical guidance. The SLP was implemented for each step in the response chain until the task was completed. If the participant completed the entire response chain without experimenter prompts, the experimenter allowed the participant 30 s of access to a toy.

### *Multiple Verbal Prompts*

Procedures in the MVP condition were identical to those employed in the SLP condition with the following exceptions: (a) the response interval was 5 s following the discriminative stimulus and 3 s following subsequent prompts, (b) three verbal prompts were included in the prompting hierarchy, and (c) the experimenter completed the step for the participant. The first verbal prompt during the MVP condition was identical to the verbal prompt in the SLP condition (e.g., “Put the blue piece on”). The subsequent two verbal prompts included information from the first verbal prompt along with a coaxing statement (e.g., “You can do it” or “I know you can”). The final prompt involved the experimenter completing the step for the participant (e.g., “You should have put the piece on like this” while simultaneously placing the piece on for the participant). The purpose of this condition was to evaluate the combined effects of issuing repeated verbal prompts, not providing the opportunity to respond after the model prompt, and omitting the physical prompt during the SLP. This specific variation of the SLP was based on clinical experience related to common prompts provided by teachers. As mentioned previously, teachers may be less likely to provide more instructive prompts (e.g., model and physical prompts) than to issue additional verbal prompts. Therefore, this variation of SLP might be expected to occur in classroom settings during educational programming.

### Experimental Design

The effects of the SLP and MVP conditions were evaluated using an adapted alternating treatments design (Sindelar et al. 1985). An adapted alternating treatments design is suited for evaluations that compare treatments that are intended to increase and maintain behavior. Rather than exposing the same behavior

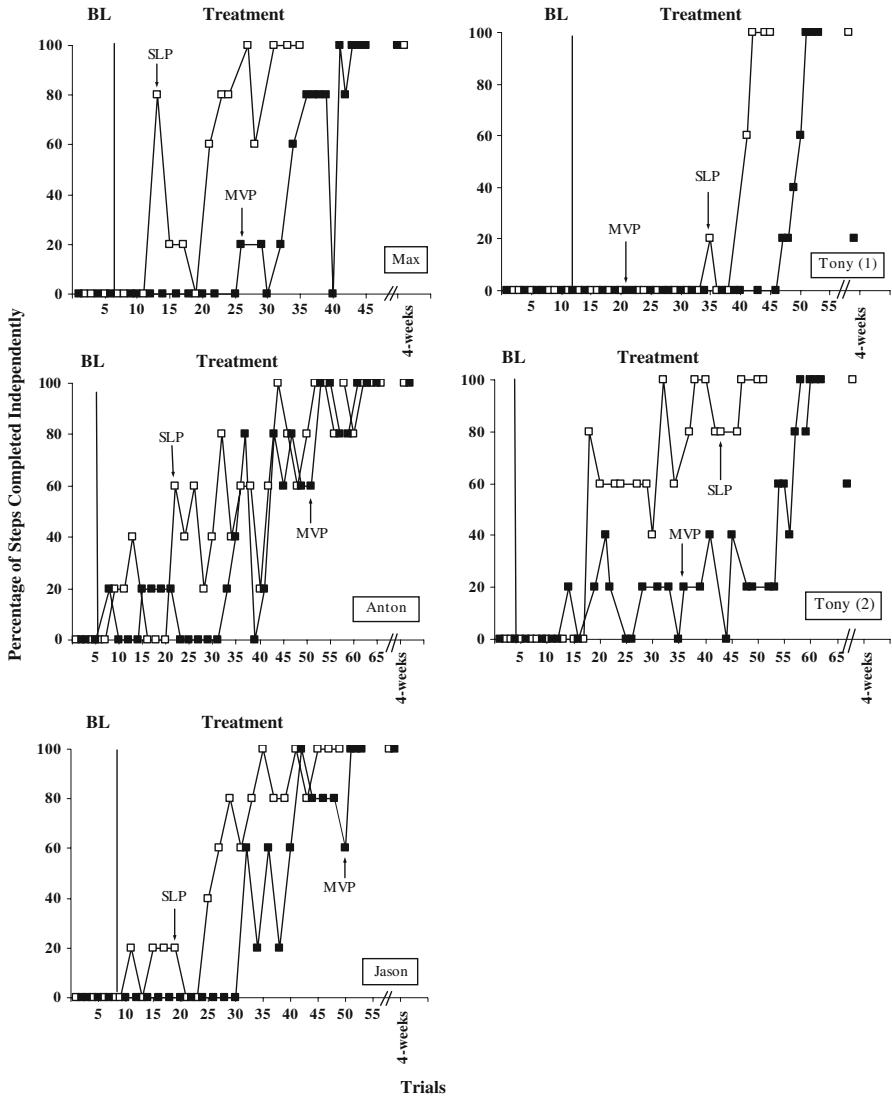
to two or more independent variables (as with a standard alternating treatments design), two or more behaviors that are deemed equivalent are each exposed to different independent variable manipulations. Each participant was taught two tasks with the exception of Tony who was taught four tasks. Baseline trials were conducted to determine the level of accuracy for each task prior to implementing teaching trials. One task was taught using the traditional SLP and another task was taught using the MVP condition for each participant. Max, Jason, and Anton were taught the “Froodle” using the SLP condition and the “Zemick” using the MVP condition. The tasks were counterbalanced for Tony during the first evaluation to demonstrate that any differences in the acquisition of the response chains was due to the differences in the teaching conditions rather than some other variable (e.g., task difficulty). That is, Tony was taught the “Zemick” using the SLP condition and the “Froodle” using the MVP condition. For Tony’s second evaluation, the “Monig” was taught using the SLP condition and the “Delib” was taught using the MVP condition.

### Follow-up Procedure

A 4 week follow-up trial was conducted for all response chains across all participants. The follow-up procedure was identical to the baseline trial conditions described earlier in the paper.

## Results

The percentage of steps completed independently during the study for each participant is presented in Fig. 3. None of the participants completed any of the steps independently for either response chain during baseline. Both SLP and MVP conditions produced acquisition of the response chain for Max and Tony (first evaluation; top panel, Fig. 3). Although both procedures were effective, the response chain associated with the SLP condition was acquired more quickly than the response chain associated with the MVP condition for both Max and Tony. Both response chains were maintained at 100% for Max during the follow-up. For Tony (evaluation 1), the response chain taught using the SLP condition was completed independently during the follow-up; whereas the response chain taught under the MVP condition was completed 20% independently. The middle panel depicts the results of the study for Anton and the second evaluation for Tony, both of whom acquired the response chains regardless of the prompting procedure. The same number of teaching trials was required for Anton to learn the response chains under SLP and MVP conditions. Both response chains were maintained at 100% for Anton. For Tony (evaluation 2), the response chain taught using the SLP condition was completed independently during the follow-up; whereas the response chain taught under the MVP condition was completed 60% independently. The bottom panel shows the results of Jason’s evaluation. Although both procedures were effective, the response chain associated with the SLP condition was acquired slightly quicker than the response chain associated with the MVP condition. For all

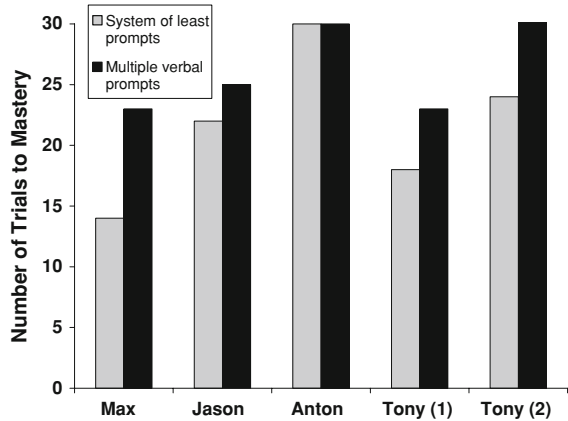


**Fig. 3** Percentage of steps completed independently during baseline (BL), system of least prompts (SLP), and multiple verbal prompts (MVP) conditions

participants, response chains were acquired in both the SLP and MVP conditions. During 4 of the 5 evaluations, participants acquired the response chain under the SLP condition before the response chain using the MVP condition.

Figure 4 displays the number of trials that were required to meet the mastery criterion (i.e., three consecutive trials with 100% independent responding) for both SLP and MVP conditions for each participant. Overall, fewer trials were required under the SLP condition to teach the response chains for 4 of the 5 evaluations. Both procedures were equally effective and efficient for one participant (Anton). An

**Fig. 4** Number of teaching trials required to the meet mastery criterion for tasks taught during system of least prompts and multiple verbal prompts conditions



average of 22 teaching trials was necessary to teach the response chain using the SLP (range, 14–30). An average of 27 teaching trials were required to teach the response chain using the MVP condition (range, 23–33). Overall, both teaching procedures were effective in teaching the response chain, but the SLP condition produced the most efficient learning for the majority of the evaluations.

## Discussion

The results of this study add to the body of literature that examines the impact of altering previously validated teaching procedures during behavioral acquisition programs. Skills taught using the MVP method were associated with less efficient acquisition during 4 of the 5 evaluations and poorer maintenance in 2 of the 5 evaluations. The counterbalancing of tasks for Tony's first evaluation adds strength to the internal validity of the study. That is, 3 of 4 participants learned the response chain associated with the traditional SLP more quickly regardless of the task. The results suggest that the difference in the efficiency with which tasks were learned was related to the differences between the traditional SLP and MVP conditions rather than the task itself.

Although all participants learned the response chains regardless of the prompting condition, additional trials were required to teach the response chain under the MVP condition. The severity of the impact of the change in the prompting procedure varied across participants. That is, the number of additional trials that were needed to meet the mastery criterion ranged in value from 3 to 9 training trials across evaluations. The average length of a teaching trial was 1 min (range, 30 s to 1 min 45 s) in the present study resulting in an approximate 3–9 min additional training time that was required to meet the mastery criterion using the MVP condition.

Although extensive additional teaching trials were not needed during the MVP condition for the participants in the present study, a considerable impact in the efficiency of learning across a long period of time may occur if changes in prompting procedures maintain during instruction over time. In natural settings such as schools, several areas of functioning (e.g., discrimination, attending, social and

daily living skills) are addressed by targeting a variety of behaviors. If changes in prescribed teaching procedures occur across multiple skill sets over an academic school year, a considerable amount of teaching time may be lost and fewer academic targets may be taught.

Variations in teaching procedures in natural settings can be conceptualized as treatment integrity failures. The results of the study highlight the importance of assessing how treatment integrity failures affect response to behavioral interventions. Although treatment integrity has been described as a critical feature of applied behavior-analytic methodology, it is rarely reported in experimental reports (Peterson et al. 1982). Consequently, it is largely unknown how certain kinds of treatment integrity errors (and their frequencies) correlate with behavioral outcomes. One of the ways researchers have compensated for this lack of information is by conducting studies in which procedural errors were intentionally introduced and comparing their effects with those of intact treatments.

Two studies have been conducted that evaluate the effects of treatment integrity failures on performance during academic instruction (Holcombe et al. 1994; Noell et al. 2002). Holcombe and colleagues evaluated the effects of high and low treatment integrity during receptive and expressive identification tasks using a constant time-delay procedure. Noell and colleagues (2002) extended the Holcombe et al. (1994) study by examining the impact of varying levels of treatment integrity during mathematics instruction. In these previous evaluations, target behaviors were acquired even when prompts were omitted during a proportion of teaching trials; however, omitting prompts during teaching trials produced less efficient response acquisition than the standard teaching procedure. Similarly, in the present study, response chains were acquired regardless of the prompting procedure. However, the MVP condition which was associated with the addition of unnecessary verbal prompts and the omission of an opportunity to respond after the model prompt. Taken together, the results of these studies suggest that modifications to antecedent prompts may not compromise the effectiveness of a teaching procedure but may reduce its efficiency. Children participating in early intervention programs should benefit from teaching procedures that produce rapid skill acquisition. Compromised treatment integrity on antecedent prompts may ultimately reduce opportunities to learn new skills.

For Max and Tony (both evaluations), very little acquisition occurred until the response chain associated with the SLP was acquired. Slow acquisition may be due to the nature of the MVP condition or the similarity in response chains across experimental conditions. By equating the response chains across conditions, we may have suppressed acquisition during the MVP condition. Examination of the raw data does not support this conclusion for either participant. That is, a shift from experimenter-completion prompts to verbal prompts did occur before the response chain associated with the SLP condition was acquired for Max and Tony. Regardless, the response chain taught using the SLP condition was acquired first. Moreover, the similarity between the response chains provides a more conservative demonstration of the impact of treatment integrity failures during the SLP.

Although the MVP condition was effective in teaching the participants to complete the response chain, the utility of this condition as a viable teaching

procedure is questionable. The participants in this study may not be entirely representative of other children with developmental disabilities. Max and Jason had experience with early and intensive behavioral intervention, had relatively fewer deficits in language skills as shown by the PPVT-III and EVT, and engaged in little or no problem behavior. Moreover, after a number of exposures to the experimental conditions, Max, Jason, and Tony began to audibly self-prompt during trials prior to scheduled experimenter-delivered prompts. Learners who lack appropriate attending skills when presented with auditory discriminative stimuli (e.g., instructions) or display noncompliance may not benefit from instruction that utilizes the MVP method. In addition, learners who do not have prerequisite attending repertoires or display noncompliance may not benefit from verbal or experimenter-completion prompts which may result in greater detriments in efficiency. The SLP and MVP conditions were equally effective and efficient to teach the response chains for Anton. One explanation is that neither condition is the best teaching procedure for acquiring new skills for Anton. Anton required the most trials to meet the mastery criterion during the SLP and the second most to meet the criterion for the MVP condition which leaves room for the possibility that another teaching procedure (e.g., errorless prompting) may have been more efficient.

While the MVP condition was not detrimental in producing skill acquisition, the condition as a teaching procedure may be contraindicated for children who are noncompliant or require extensive training to acquire academic skills. Children who engage in noncompliant behavior during academic tasks may not benefit from teaching that is similar to the MVP condition. During the MVP condition, the participant was not required to complete any of the steps in the chain (i.e., the physical prompt was omitted and no opportunity was provided to respond after the model prompt) during teaching trials. One treatment for noncompliant behavior is guided compliance, which involves physical guidance, often during the initial stages of treatment (Wilder et al. 2006). Children with noncompliance who are exposed to MVP may be less likely to respond during trials. The MVP condition may maintain or exacerbate noncompliant behavior as a result of the omission of the physical prompt and the opportunity to respond after the model prompt.

The errors introduced in the MVP condition may be particularly detrimental to slower learners. Generally speaking, prompts are added during teaching to increase the likelihood that a correct response will occur. The MVP condition utilized three verbal prompts which may not occasion correct responses during initial teaching trials for slower learners. If verbal prompts are not effective in facilitating appropriate responding, a considerable number of errors may occur on the part of the learner. High errors in classroom settings may result in reduced access to high-quality reinforcers which may make instructing the learner more difficult in the future. An additional limitation of the MVP condition as a teaching procedure is that the number of academic goals that can be met over time may decrease.

Several limitations of the study exist and warrant discussion. First, some participants included in the study had skills that may have attenuated potential detrimental effects of the MVP condition. Skills included self-prompting, appropriate attending, and compliance. Future research could address this issue by including participants with a broader range of preexisting skills to determine if

errors during SLP are more likely with less advanced learners. Second, the MVP condition contains several intervention components (i.e., additional verbal prompts, removal of the opportunity to respond after the model prompts, omission of the physical prompt). Since multiple components were used during the MVP condition, the specific components that reduced the efficiency of the SLP condition remain unknown. Future research might focus on examining the separate and combined effects of the MVP components to determine their relative contribution to the reduced efficiency. Third, the tasks taught in the study lack social validity given the arbitrary nature of the response chains. Future investigations might include tasks that children frequently come in contact with during instruction at school (e.g., daily living skills, matching to sample tasks). Additional research should determine whether the acquisition of tasks that are commonly presented during instruction is negatively impacted by the MVP method. Finally, these findings may not represent results that would be obtained with children from other populations (e.g., typically developing children, children with learning disorders).

The results of the study add to the current literature on treatment integrity during behavioral interventions by again showing that deviations from standard teaching procedures may alter the efficiency of learning. Treatment agents in the natural environment may sometimes drift from prescribed behavioral interventions. If and how deviations from teaching procedures produce changes in behavioral outcomes is an important and growing area of research. Information obtained from these studies may guide how treatment agents are trained to implement behavioral interventions and how treatment integrity is assessed over time.

## References

- De Leon, 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. doi:[10.1901/jaba.1996.29-519](https://doi.org/10.1901/jaba.1996.29-519).
- Doyle, P. M., Wolery, M., Ault, M. J., & Gast, D. L. (1988). System of least prompts: A literature review of procedural parameters. *The Journal of the Association for Persons with Severe Handicaps*, 13, 28–40.
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.
- Fisher, W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, L. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis*, 25, 491–498. doi:[10.1901/jaba.1992.25-491](https://doi.org/10.1901/jaba.1992.25-491).
- Gilliam, J. (1995). *Gilliam autism rating scale examiner's manual*. Austin, TX: Pro-Ed.
- Holcombe, A., Wolery, M., & Snyder, E. (1994). Effects of two levels of procedural fidelity with constant time delay on children's learning. *Journal of Behavioral Education*, 4, 49–73. doi:[10.1007/BF01560509](https://doi.org/10.1007/BF01560509).
- Horner, R. D., & Keilitz, I. (1975). Training mentally retarded adolescents to brush their teeth. *Journal of Applied Behavior Analysis*, 8, 301–309. doi:[10.1901/jaba.1975.8-301](https://doi.org/10.1901/jaba.1975.8-301).
- MacDuff, G. S., Krantz, P. J., & McClannahan, L. E. (2001). Prompts and prompt-fading strategies for people with autism. In C. Maurice, G. Green, & R. M. Foxx (Eds.), *Making a difference: Behavioral intervention for autism* (pp. 37–50). Austin, TX: PRO-ED.
- Noell, G. H., Gresham, F. M., & Gansle, K. A. (2002). Does treatment integrity matter: A preliminary investigation of instruction implementation and mathematics instruction. *Journal of Behavioral Education*, 11, 51–67. doi:[10.1023/A:1014385321849](https://doi.org/10.1023/A:1014385321849).

- Peterson, L., Homer, A. L., & Wonderlich, S. A. (1982). The integrity of independent variables in behavior analysis. *Journal of Applied Behavior Analysis, 15*, 477–492.
- Repp, A. C., Barton, L. E., & Brulle, A. R. (1981). Correspondence between effectiveness and staff use of instructions for severely retarded persons. *Applied Research in Mental Retardation, 2*, 237–245. doi: [10.1016/0270-3092\(81\)90017-5](https://doi.org/10.1016/0270-3092(81)90017-5).
- Sindelar, P. T., Rosenberg, M. S., & Wilson, R. J. (1985). An adapted alternating treatments design for instructional research. *Education & Treatment of Children, 8*, 67–76.
- Steege, M., Wacker, D., & McMahon, C. (1987). Evaluation of the effectiveness and efficiency of two stimulus prompt strategies with severely handicapped students. *Journal of Applied Behavior Analysis, 18*, 293–299. doi: [10.1901/jaba.1987.20-293](https://doi.org/10.1901/jaba.1987.20-293).
- West, E. A., & Billingsley, F. (2005). Improving the system of least prompts: A comparison of procedural variations. *Education and Training in Developmental Disabilities, 40*, 131–144.
- Wilder, D. A., Atwell, J., & Wine, B. (2006). The effects of varying levels of treatment integrity on child compliance during treatment with a three-step prompting procedure. *Journal of Applied Behavior Analysis, 39*, 369–373. doi: [10.1901/jaba.2006.144-05](https://doi.org/10.1901/jaba.2006.144-05).
- Williams, K. T. (1997). *Expressive vocabulary test*. Circle Pines, MN: American Guidance service.
- Wolery, M., & Gast, D. L. (1984). Effective and efficient procedures for the transfer of stimulus control. *Topics in Early Childhood Special Education, 4*, 52–77. doi: [10.1177/027112148400400305](https://doi.org/10.1177/027112148400400305).
- Wolery, M., & Shuster, J. W. (1997). Instructional methods with students who have significant disabilities. *The Journal of Special Education, 31*, 61–79. doi: [10.1177/002246699703100106](https://doi.org/10.1177/002246699703100106).