

## **The Effects of a Treatment Package in Establishing Independent Academic Work Skills in Children with Autism**

Lillian V. Pelios

Bancroft Center for Autism Research and  
Educational Services

Gregory S. MacDuff

Princeton Child Development Institute

Saul Axelrod

Temple University

### **Abstract**

The purpose of this study was to evaluate the effectiveness of a treatment package in producing independent work by three children with autism with minimal supervision by an adult. The package included: (a) delayed reinforcement for on-task and on-schedule responding, (b) fading of instructional prompts and of the instructor's presence, (c) unpredictable supervision, (d) response cost for off-task responding. It was found that the treatment package resulted in increased levels of on-task and on-schedule responding during treatment for all three children with a supervising adult only occasionally present. Two children required minimal adult supervision in maintenance. Generalization probes showed that the behavior of all three children transferred across both novel material and a novel setting in the absence of adult supervision.

Key words: Delayed reinforcement, fading supervision, response cost, autism, independent skills, intervention package, unpredictable supervision.

★ ★ ★

Numerous publications describe the variety of procedures that have been developed over decades to ameliorate the behavioral deficits and/or excesses that are typically displayed by persons with autism. The most effec-

---

Address Correspondence to: Lillian Pelios, Bancroft CARES, Bancroft Neurohealth, 201 S. Kings Highway, Cherry Hill, NJ 08304, Phone: (856) 216-8090, Fax: (856) 216-9240, E-mail: lpelios@bnh.org

Authors' Note: We extend our appreciation to Pat J. Krantz and Lynn E. McClannahan who facilitated the completion of this study. We also thank Meg MacDuff, who served as the teacher. Her enthusiasm for research and her excellent skills as a teacher made it so much easier to complete this project.

tive studies have taught precisely targeted behaviors in structured educational settings (DeMyer, Hingtgen, & Jackson, 1981; Dunlap, Koegel, & O'Neil, 1985; Koegel, Rincover, & Egel, 1982). Despite these successes, learned behavior has not generalized to novel settings and/or persisted once the treatment contingencies have been withdrawn (Dunlap & Johnson, 1985; Dunlap, Koegel, & Johnson, 1987; Stahmer & Shreibman, 1992). Additionally, persons with autism frequently do not independently initiate tasks or remain appropriately engaged with materials even though they may have displayed mastery of the specific tasks in structured settings (Dunlap & Plenis, 1988; Fowler, 1988; MacDuff, Krantz, & McClannahan, 1993; Sailor, Goetz, Anderson, Hunt, & Gee, 1988). The presence of a supervising adult who will prompt engagement continues to be a requirement for many persons with autism to function in their environments (Dunlap & Johnston, 1985; Dunlap, Koegel, & Johnson, 1987; Stahmer & Shreibman, 1992). Furthermore, people with autism may also have difficulty acquiring lengthy response sequences. Thus, even though a child with autism may have learned to complete a morning routine (e.g., brush teeth, get dressed, and go to breakfast), she may need prompts to initiate each activity or each part of an activity (MacDuff, Krantz, & McClannahan, 1993). Again, successful completion of the sequence requires monitoring from a supervising adult.

Integration of persons with autism and developmental disabilities into mainstream society continues to be a dominant priority in the field of education. Simply teaching children with autism a variety of skills is not enough to prepare them for a functional life in society. For instance, it is important that persons with autism learn to display already acquired skills in the absence of a supervising adult. Krantz and McClannahan (1999) suggest that in order to program for integration the following six response classes should be addressed: (a) following adults' instructions, (b) exhibiting generative language, (c) emitting low rates of inappropriate behavior, (d) sustaining engagement, (e) responding to temporally delayed consequences and (f) generalizing skills across settings. If children with autism are to share the same experiences as their nondisabled peers, it is imperative to identify procedures to successfully teach these skills to them.

As mentioned earlier, typically, the removal of close supervision, whether in group or in individualized settings, leads to the recurrence of stereotypic and off-task behaviors and a decline in appropriate, productive responding (Dunlap & Johnson, 1985; Marholin & Steinman, 1977). One possible explanation for the failure of appropriate responding to generalize has been related to the strict stimulus control that is typically established by structured training conditions (Billingsley & Romer, 1983; Snell, 1983). According to this hypothesis, the stimuli under which behavioral training is implemented acquire discriminative characteristics and come to occasion appropriate behavior because such stimuli may be associated

with reinforcement of targeted responses (Koegel & Rincover, 1977; Pierce & Schreibman, 1994). In contrast, the absence of such stimuli may serve to signal the lack of controlling contingencies, thereby, resulting in behavioral variation rather than reliable emission of the target response (Corte, Wolf, & Locke, 1971; Marholin & Steinman, 1977; Risley, 1968). Thus, the departure of an instructor may result in increased off-task behavior in a child with autism by virtue of the predictable absence of contingencies for that period of time. To address these stimulus control considerations, some researchers have begun to investigate how the use of indistinguishable or unpredictable contingencies as well as delayed contingencies can facilitate generalization and maintenance of desirable behavior (Dunlap & Johnson, 1985; Dunlap, Koegel, & Johnson, 1987; Schwarz & Hawkins, 1970; Stokes & Baer, 1977). For instance, Dunlap and Johnson (1985) demonstrated that when an unpredictable schedule of supervision was used with three children with autism, on-task levels were much higher during the therapist's absence than when a predictable supervision-schedule was used. Moreover, Dunlap, Koegel, and Johnson (1987) demonstrated that the therapist could be removed from the treatment environment of three children with autism, although other adults remained present. Also, the authors showed that appropriate behavior could be successfully maintained in community settings, again in the presence of adults other than the experimenters, with infrequent and delayed contingencies. In another study, MacDuff, Krantz, and McClannahan (1993) taught children with autism to complete the steps in a complex task and to change tasks independently, by using photographic cues (i.e., albums depicting after-school activities) in combination with graduated guidance to increase on-task and on-schedule responding. With this two-component package, the authors taught four youths with autism to display sustained engagement and lengthy response chains, to independently change activities, and to change activities in different group home settings in the absence of immediate supervision and prompts. However, as with the Dunlap et al. study, in the MacDuff et al. study adults other than the experimenter were present during intervention and generalization probes.

Several researchers have demonstrated the effectiveness of pictorial and/or textual cues, used alone or in combination with a variety of prompting/fading procedures and/or self-management procedures, in teaching persons with disabilities to initiate tasks. Thus far, however, research with pictorial/textual cues has focused primarily on adolescents or adults rather than younger children (Frank, Wacker, Berg, & McMahon, 1985; Johnson & Cuvo, 1981; Krantz, MacDuff, & McClannahan, 1993; MacDuff, Krantz, & McClannahan, 1993; Nietupski, Clancy, & Christiansen, 1984; Sowers, Verdi, Bourbeau, & Sheehan, 1985; Thinesen & Bryan, 1981; Wacker & Berg, 1983, 1984). Furthermore, intervention packages using pictorial/textual cues typically have not attempted to investigate on-task and on-schedule

responding in the absence of adult supervision. The purpose of this study was to investigate the effectiveness of a treatment package in producing independent work by three children with autism in the absence of a supervising adult. The package included: (a) delayed reinforcement for on-task and on-schedule responding, (b) fading of instructional prompts and of the instructor's presence, (c) unpredictable supervision, and (d) contingent response cost for off-task responding.

## Method

### *Participants*

The participants of this study were Zeb and David, who were 9, and Carl, who was 5 years old. All three children met the DSM IV (APA, 1994) criteria for the diagnosis of autism. The children's scores on the Peabody Picture Vocabulary Test ranged from 2.5 to 4.5 years ( $M = 3$ ). Age equivalent scores on the Vineland Adaptive Behavior Scale were 4.11 and 1.4 for David and Zeb respectively. No scores were obtained for Carl. Informed parental consent to participate in this study was obtained for each participant.

Zeb and David had been using activity schedules for up to seven years; Carl had been using activity schedules for approximately one year. All three children were proficient schedule followers with adult supervision. All three children had severe language deficits; they displayed high rates of echolalia, vocal noise, noncontextual speech, and limited spontaneous language. All three children remained on-task and on-schedule, with irregular performance, in the presence of supervising adults, but engaged in stereotypy (e.g., verbal and motoric) in the absence of adults. All participants had acquired picture object correspondence prior to the study, had past experience with monetary reinforcement systems, such as token boards, and were accustomed to tokens embedded in their schedules.

### *Settings and Tasks*

Sessions took place in a 3 x 4 m classroom with no windows excluding the one-way mirror that was built in the door. The room was furnished with a desk, a chair, and a long table. The activities were arranged in folders on the long table in the order indicated by the schedule. No other children were present in the classroom. Schedules consisted of a 7 x 9 inch black binder that held 7 to 16 pages. A textual cue indicating which activity was next (e.g., "Math," "Spelling," "Coloring") was displayed on every other page while a penny was embedded on the remaining pages. The activities selected had been included in the participants' past activity schedules and had been mastered by the participants, sometimes as long as a

year ago. These activities were currently used in maintenance schedules that the children used at home or at school when they had completed their current schedules. Some examples of activities were: completing math worksheets, completing a puzzle, coloring a picture, and listening to music while looking at a photo album.

### *Dependent Variables*

A 15-s partial interval recording procedure was used to successively record all the dependent variables. The observers checked a box to indicate the occurrence or non-occurrence of each variable.

*On/Off-task responding.* On-task was recorded if the participant was visually attending to play or work materials, looking at his activity schedule, and manipulating play or work materials appropriately (i.e., as they were designed to be used), or moving from one scheduled activity to another. Off-task was scored if the participant used materials in a manner other than that for which they were designed, manipulated but did not visually attend to the materials, engaged in inappropriate behavior (e.g. aggression, tantrums, stereotypies), or did not engage in activities or use materials.

*On/Off schedule responding.* On-schedule was recorded if the participant was engaged in the activity depicted on the page to which his activity schedule was open. Off-schedule was scored if the participant was engaged in an activity other than what was depicted on his activity schedule.

### *Independent Variables*

As with the dependent variables, a 15 s partial interval recording procedure was used to simultaneously record the independent variables. Independent variables consisted of a package including: (a) manual prompts (verbal and gestural prompts were also recorded for purposes of treatment integrity), (b) presence of the instructor, (c) contingent/delayed reinforcement, and (d) response cost. The observers checked a box to indicate the occurrence or non-occurrence of each variable.

*Manual prompts.* All physical contact between the teacher and the participants, designed to assist the participants in the completion of a task, was scored as a manual prompt. Examples of such prompts are physically guiding the participant to pick up materials or to turn the page via hand-over-hand prompts and light touches such as those that may occur during fading graduated guidance.

*Verbal prompts.* Verbal prompts included instructions such as "Do your work," "Work nicely," "Look at your work" and any other verbal statement directing the participants to be engaged.

*Gestural prompts.* Gestural prompts included all pointing, motioning,

or nodding toward children or materials, as well as pointing to specific toys or photographs that represent the next task in a sequence. Thus, non-specific and specific gestures (e.g., pointing to a child or pointing to the last piece to be placed in the puzzle) were scored as gestural prompts.

*Presence of the instructor.* The instructor was scored as present in all those instances in which she was in the room. She was also scored as present when she entered the room for 3 s on a random schedule to observe the participants, as was the case during the later steps in treatment and during maintenance. Once the instructor's presence was faded to outside the room, the time-delay schedule was implemented initially by entries that were predictable (e.g., every 1 min, 2 min, 3 min) and subsequently by entries that were unpredictable (e.g., entering randomly every 2-5 and every 3-5 min).

*Contingent/delayed consequences.* Self-administration of pennies by the participants was scored as a contingent reward. The exchange of pennies for preferred items or activities at the end of the session was scored as delayed reward.

*Response-cost.* The instructor implemented response cost via removal of all acquired pennies if the participants were off-task for 3 s or more and manually prompted the participants to start the schedule over.

*Sessions.* Sessions were 15 min long in duration. Data were collected on both dependent and independent variables for the duration of the session. No more than two sessions per child were run on any given day. On those occasions where two sessions were conducted a 30 min break between sessions was reliably observed. This study began in January and ran until the end of the academic year that is until August. The experimenters did not systematically manipulate the number of sessions conducted per day. Rather, this was a function of student absenteeism and/or time restrictions for the completion of the project.

#### *Interobserver Agreement*

Interobserver agreement data were collected in vivo through a two-way mirror built into the door of the classroom. The two observers who stood outside the door simultaneously but independently scored data cued by a tape-recorded signal. Interobserver agreement data were obtained for the dependent variables on 30 % or more of sessions across all conditions. An agreement was scored when both observers recorded on-task and on-schedule as occurring at the end of an interval or as not occurring. Interval-by-interval percentage interobserver agreement was calculated by dividing the number of agreements by the number of agreements plus disagreements and multiplying by 100 % (Cooper, Heron, & Heward, 1987). Mean interobserver agreement for all dependent variables was above 98 % (range = 98-100 %).

### *Treatment Integrity- Independent Variables*

Treatment integrity consisted of operationally defining the independent variables and of direct observation of those variables throughout the study to ensure that the procedure was implemented in the manner in which it was intended. Levels of treatment integrity were calculated in the same manner as discussed in Interobserver Agreement. Mean interobserver agreement for all six independent variables was scored at 98 % or higher (range = 98-100 %).

## **Procedures**

### *Experimental Design*

A multiple probe design (Horner & Baer, 1978) across participants was used to assess the effects of the treatment package across all conditions (e.g., baseline, treatment, resequencing, maintenance, and generalization).

### *Experimental Conditions*

All experimental conditions occurred in the same classroom excluding the generalization probes across rooms. All sessions in each condition lasted 15 min.

*Baseline.* During this condition, the teacher escorted the participant to a classroom in which the activity schedule and the required materials to complete the tasks were located. The teacher delivered the instruction "Do your schedule" and left the room. The teacher returned to the classroom after 15 min had elapsed and announced: "Doing your schedule is over."

*Treatment.* The treatment package consisted of several steps.

(a) *Manual prompts:* the teacher manually prompted high rates of on-task and on-schedule behavior via graduated guidance. Graduated guidance required that manual prompts be faded via spatial fading and shadowing. According to this procedure, the teacher initially used hand-over-hand prompts to direct the participant. As the participant became increasingly more independent with the task, the teacher gradually changed the location of manual prompts from the hands, to the wrists, the elbows, and finally the shoulders. Once the participant demonstrated independent responding with the task at this level of prompting, the teacher gradually began to fade proximity until she was outside the room;

(b) *Instructor's presence:* when 80% of the intervals or better had been scored for on-task and on-schedule for two consecutive sessions and response-cost had not been implemented, the teacher's presence was faded to the next step. Fading occurred via shadowing the participant to gradu-

ally increasing physical distance until the instructor was outside the room. Steps in fading the instructor's proximity were the following: (step 0) shadowing the student at a distance less than 5 inches, (step 1) shadowing the student at 6 inches, (step 2) shadowing the student at 12 inches, (step 3) shadowing the student at 18 inches, (step 4) shadowing the student at 24 inches or standing by the door, and (step 5) standing by the door with the door half open. Once outside the room, the teacher reentered the classroom for approximately 2 s according to a time-delay schedule, observed the child and exited without saying anything to the participant. Steps in the time-delay schedule were the following: (step 6) entering the room every 1 min, (step 7) entering every 2 min, (step 8) entering every 3 min, (step 9) entering randomly every 2-5 minutes, (step 10) entering randomly every 3-5 minutes. The criterion for moving to the next level of fading was 80% of the intervals scored for on-task and on-schedule in the absence of response-cost for two consecutive sessions;

(c) *Contingent/delayed reinforcement*: immediate rewards consisted of pennies embedded on separate pages between activities in the activity schedule. During intervention if the child administered a penny without first completing the work the instructor would manually prompt the child to return the penny and follow the correct sequence. At the end of the session, pennies were exchanged for preferred snacks or activities;

(d) *Response cost*: when off-task behavior and/or stereotypy occurred for a minimum of 3 s, the teacher immediately implemented response-cost by removing all acquired coins from the token board and manually guiding the participants to start the schedule from the beginning. When off-task responding and/or stereotypy occurred during those sessions in which the teacher's presence had been faded from the classroom, the teacher would reenter the room, implement the response cost procedure, manually prompt the participant to start the schedule over and leave the room.

*Resequencing*. As each participant demonstrated stable levels (i.e., 80% or more of the intervals scored) of on-task and on-schedule responding, the activities in the schedule, excluding the first page were resequenced according to a table of random numbers. Prompting was delivered according to the specifications in the treatment section. The purpose of this condition was to evaluate whether the participants had acquired a generalized repertoire of schedule following or whether they were simply following a memorized sequence of activities.

*Maintenance*. In maintenance the teacher checked on the student once per session, at the first to 14<sup>th</sup> minutes, assigned randomly. The teacher was present in the room for no longer than 2 s.

*Generalization*. Generalization was assessed across (a) novel stimuli that were cued by the activity schedule and (b) a novel setting. Probes were conducted across novel materials and settings pre- and post-treatment as well as during maintenance for two of the three participants. During probes



the teacher did not enter the room after the initial instruction and no prompts were delivered. If the student completed the activity schedule without engaging in off-task responding for 3 s or more, he was allowed to choose a reward.

## Results

### *Dependent Variables*

Figure 1 displays levels of on-task and on-schedule performance for the three participants across the baseline, treatment, resequencing, and maintenance conditions. During baseline Zeb was almost never scored on-task or on-schedule. His mean performance was 0.25% for both on-task and on-schedule (range = 0-2% for both).

David and Carl displayed considerable variability in performance during baseline measures for both on-task and on-schedule responding across sessions. Mean performance for David during baseline was 63% on-task (range = 0-98%) and 55% on-schedule (range = 0-98%). Mean performance for Carl during baseline was 9% on-task (range = 0-33%) and 3% on-schedule (range = 0-17%).

During treatment, mean on-task performance for all three children ranged between 95% and 99%. Specifically, Zeb's mean on-task and on-schedule performance was 99% (on-task and on-schedule range = 85-100%). Mean on-task performance for David during treatment was 98% (range = 92-100%) of the intervals scored while his mean on-schedule performance was 97% (range = 92-100%). Finally, Carl was scored on-task for 97% of the intervals (range = 87-100%) during treatment while on-schedule was scored for 96% of the intervals (range = 80-100%).

During resequencing of photographs and the maintenance condition, all three participants continued to display high and stable rates of on-task and on-schedule performances. As shown in Figure 1, Zeb's mean performance for on-task during resequencing was 98% (range = 97-100%) while his mean on-schedule performance was scored at 96% (range = 87-100%). David's mean performance during resequencing was 97% for on-task responding and 96% for on-schedule responding (range = 90-100% for both on-task and on-schedule). Carl's mean on-task performance during resequencing was 97% for both on-task and on-schedule (on-task range = 87-100%; on-schedule range = 80-100%).

Levels of on-task and on-schedule responding during maintenance can be seen in Figure 1. Mean performance for Zeb was scored at 99% for both on-task and on-schedule responding (range = 97-100%). During maintenance, David's performance for both on-task and on-schedule was scored at 98% (range = 98-100% for both on-task and on-schedule). No maintenance data are available for Carl because the academic year ended.

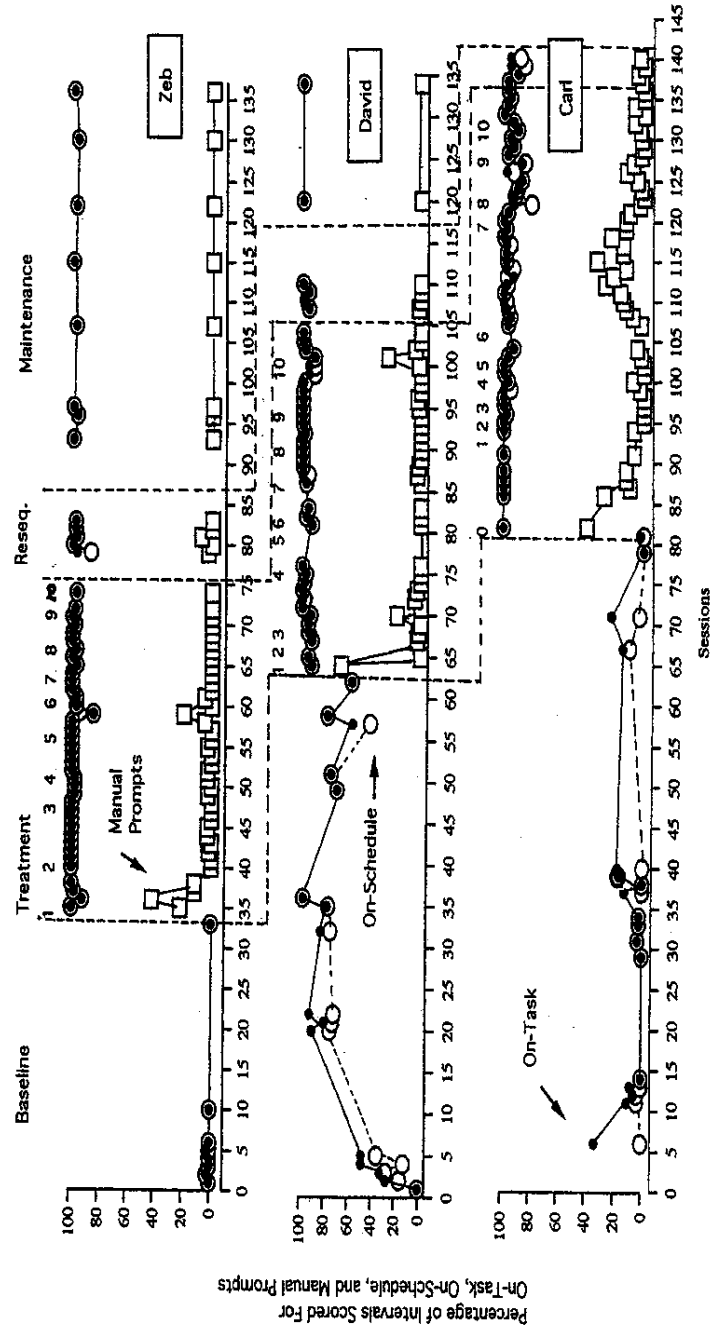


Figure 1. Displays the percentage of intervals scored for on-task and on-schedule responding and manual prompts. The numbers above the data points in the treatment condition illustrate the steps required for fading the instructor's presence. For instance the number "1" indicates that the instructor was shadowing the participant at 6 inches while the number "2" indicates that the instructor was shadowing at 12 inches and so on.

Figure 2 displays levels of on-task and on-schedule responding during generalization probes conducted with the same material but in a different room across baseline, treatment and maintenance conditions for all three children. In baseline Zeb was never on-task or on-schedule. In treatment he was scored on-task and on-schedule during 99% of the intervals scored (range = 97-100% for both on-task and on-schedule). In maintenance Zeb was scored on-task for 92% of the intervals and on-schedule for 91% of the intervals scored (range = 82-98% for on-task; 78-98% for on-schedule).

During generalization probes across a novel setting, David's mean on-task and on-schedule performances in baseline were 73% and 41% respectively (range = 57-92% for on-task and 21-77% for on-schedule). Generalization probes conducted while David was still in treatment show that his mean score for on-task was 85% of the intervals scored (range = 70-100%) and for on-schedule 83% of the intervals scored (range = 63-100%). Generalization probes across a novel setting showed that David's mean performance for both on-task and on-schedule in maintenance was 99% (range = 98-100% for both).

Carl's mean on-task and on-schedule performances in baseline, during generalization probes across a novel setting were scored at 0.8% and 0.3% respectively (on-task range = 0-3% and on-schedule range = 0-2%). Generalization probes across a novel setting show that Carl's mean score for on-task and on-schedule was 47% of the intervals scored while in treatment (range = 0-95%). When looking at individual data points, however, it can be noted that the first two data points in treatment were at 0% for both on-task and on-schedule responding while the last two data points were at 95 and 93% for on-task and at 72% and 92% for on-schedule responding. It is clear that responding for both on-task and on-schedule increased in the later stages of treatment. No maintenance data are available for Carl.

Figure 3 displays levels of on-task and on-schedule responding during generalization probes across novel material but in the original instructional setting. During these probes, Zeb was never scored on-task or on-schedule while in baseline. During generalization probes across novel material, Zeb's mean score for on-task and on-schedule was 100% of the intervals scored while in treatment. Generalization probes across novel material conducted during maintenance show that his mean score for both on-task and on-schedule was 93% of the intervals (range = 85-100% for both).

David's mean on-task performance in baseline during generalization probes across novel material was 82% (range = 70-93%), while his mean on-schedule performance was scored at 56% (range = 32-80%). Generalization probes across novel material conducted in treatment indicated that David's mean score for on-task was 95 % of the intervals scored (range = 87-100%), while his mean on-schedule score was 94% (range = 83-100%). Similarly, generalization probes across novel material during maintenance show that David's mean score for on-task was 98% and for on-schedule

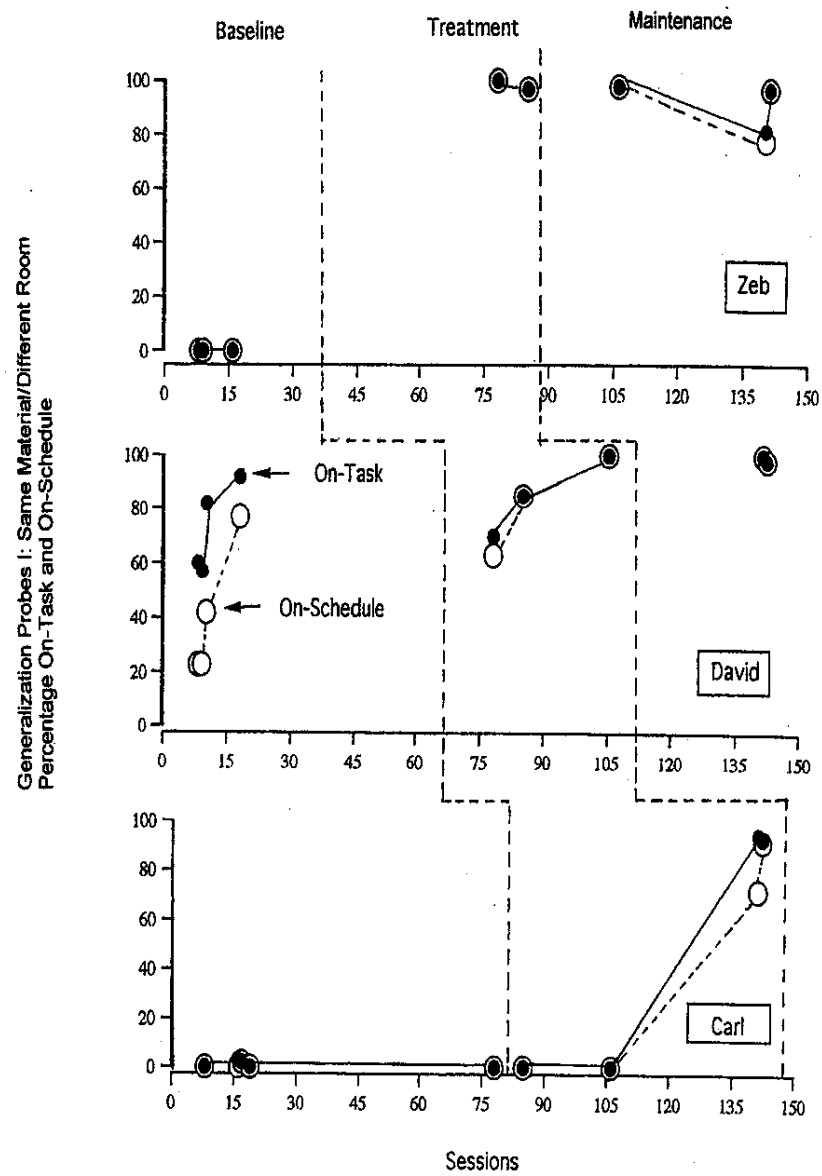


Figure 2. Generalization probes across same material but different room. Percentage of intervals scored for on-task and on-schedule responding.

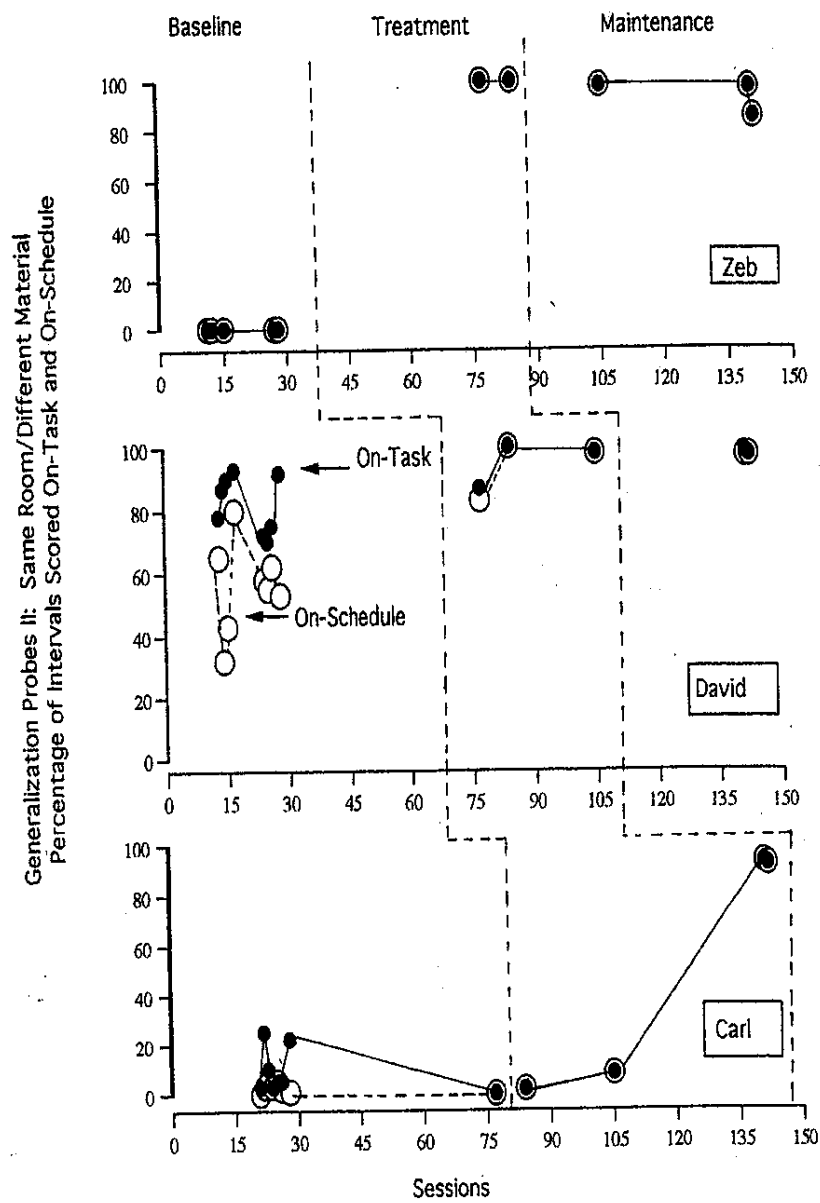


Figure 3. Generalization probes across same room but novel material. Percentage of intervals scored for on-task and on-schedule responding.

was 97% of the intervals scored during maintenance (on-task range = 97-98%).

During generalization probes of novel material, Carl's mean performance for on-task in baseline was scored at 9% (range = 0-25%) while his mean on-schedule performance was scored at 3% (range = 1-5%). Generalization probes for novel material conducted during treatment indicate that Carl's mean scores for both on-task and on-schedule was 49% of the intervals scored (on-task and on-schedule range = 8 - 93%). Upon examination of individual data points collected during treatment in the generalization across novel material condition, it can be noticed that his scores were higher in the later stages of treatment. For instance, the first two data points for both on-task and on-schedule were scored at 2 and 8% while the last two data points in treatment for on-task and on-schedule were scored at 93 and 92%. The academic year ended before maintenance data for Carl could be collected.

During baseline manual, verbal, and gestural prompts, the instructor's presence, use of reinforcement, and response cost was never scored for any of the three participants. Percentages of the instructor's presence, use of reinforcement, and response cost across all conditions can be seen in Table 1. Manual prompts can be seen on Figure 1 across the remaining conditions.

Table 1.

*Mean percentages of intervals scored for the instructor's presence, delivery of reinforcers, and implementation of response cost (RC) across baseline, treatment, resequencing, and maintenance conditions.*

Conditions	Participants		
Baseline	Zeb	David	Carl
Instructor Present	0 %	0 %	0 %
Reinforcers	0 %	0 %	0 %
RC	0 %	0 %	0 %
Treatment	Zeb	David	Carl
Instructor Present	67 %	56 %	44 %
Reinforcers	8 %	10 %	6 %
RC	1 %	0.65 %	2.5 %
Resequencing	Zeb	David	Carl
Instructor Present	7 %	14 %	9 %
Reinforcers	10 %	10 %	7 %
RC	0.6 %	7 %	0.5 %
Maintenance	Zeb	David	Carl
Instructor Present	2 %	2 %	-
Reinforcers	10 %	11 %	-
RC	0 %	0 %	-

### Discussion

This study demonstrated that the intervention package resulted in higher levels of on-task and on-schedule responding for all three participants. These high levels of responding were accomplished with a supervising adult present only for brief and intermittent periods during the later sessions in treatment and during the maintenance phase. Generalization probes showed that the behavioral gains of all three children transferred across both novel material and a novel setting in the complete absence of a supervising adult. Furthermore, both on-task and on-schedule responding maintained at high rates, that is above 80 %, with minimal adult supervision over the course of seven weeks for Zeb and four weeks for Daniel as indicated by the maintenance data. It is important to note that neither the instructor nor any other adult was present during the generalization probes, while during maintenance the teacher was present no more than 2 s for the duration of the session. This point distinguishes the present study from other studies in which an adult was always present in the room with the children.

Overall response gains, as indicated by comparing baseline to treatment and maintenance measures, were substantial. With each application of teaching the schedule, on-task and on-schedule performance increased for all children while the presence of the instructor and the frequency of manual prompts and response-cost were gradually decreased. Individual data points for Carl indicated that his performance increased considerably in the later stages of treatment across both generalization conditions. Therefore, although Carl did not reach criterion on all his sessions, a considerable increase in his on-task and on-schedule responding was noted.

An interesting observation was made regarding the number of sessions required across children to fade the instructor's proximity and to complete the time-delay schedule. Specifically, when comparing the number of sessions required to fade the proximity of the instructor to the number of sessions required to complete the time-delay schedule, it was observed that Carl required far more sessions to meet criterion during the delay schedule. Carl's performance during the generalization probes further supports this point since his performance did not meet criterion until the later sessions. It is expected, however, that responding across a novel setting and novel material would be delayed until responding in the treatment condition was reliable. In other words, Carl's initial lack of generalized responding across materials and location may have been due to the fact that generalization probes for him were conducted much closer to the initiation of treatment as compared to David and Zeb. Specifically, Carl's first generalization probe across both materials and location was conducted upon completion of his third session in treatment. Zeb's first generalization probe, on the other hand, was conducted at the end of treatment and David's first

generalization probe was conducted upon completion of session 11 in treatment. Providing David and Zeb with more treatment sessions prior to probing for generalization may have been responsible for their performance meeting criterion within fewer sessions both with the instructor's proximity faded to the door, as well as with the time-delay schedule. Additionally, the noted disparity between Carl and the other two children may be compounded by the fact that Carl was the youngest of the three and had the least experience in following activity schedules. Furthermore, it is likely that David and Zeb, due to their schooling, have had more incidental experiences with unpredictable supervision and response cost in the classrooms and their homes. Thus, prior experience with natural contingencies for the two older children, although not systematic enough to produce reliable changes in responding, may have played a role in facilitating higher and reliable rates of responding more quickly once implemented systematically.

Interestingly, during baseline Carl, who had only one year of experience in schedule following and Zeb, who had seven years of experience, spent almost the entire 15-min period engaging in vocal and motor stereotypy. Additionally David, who had seven years of experience in following activity schedules, during baseline only occasionally engaged in schedule following. Furthermore, his performance was so variable that ultimately his responding was unreliable (e.g., 10 of the 16 points collected in baseline were scored below criterion for both on-task and on-schedule responding). Interestingly, David's on-task responding was higher than his on-schedule responding. During observations, it was noted that David would select activities and sometimes complete them out of sequence, rather than follow the sequence indicated by the schedule. The reduction of variability in his responding once the treatment was implemented indicates that the procedure was effective in establishing a reliable schedule following repertoire. Although, David's preference for certain activities is not a problem per se, eventually, his ability to participate in society as a productive member will in part depend on his ability to follow a work schedule.

During resequencing, two participants erred on the first activity in the new sequence. While collecting data, the observers noted that the participants attempted to complete the sequence in the original schedule, indicating that the children may have been following a familiar routine rather than responding to the textual stimuli in the schedule. One prompt per participant was sufficient to establish the new sequence. When the schedule was resequenced to its original order during the maintenance condition and during the generalization probes, the participants did not require any additional prompts to follow the original sequence. This indicates that the textual cues were established as discriminative stimuli and that the children were not merely engaging in familiar routines. Learning to respond to textual cues in a discriminated fashion can be a very important



skill that presents both immediate and future gains. For younger children it may facilitate generalization to novel schedules and settings whereas for youths and adults with autism it may increase engagement in vocational tasks and facilitate the acquisition and maintenance of complex sequences.

Because the intervention used in this study was a treatment package consisting of a number of different variables that were implemented simultaneously, it is impossible to identify which specific variables were responsible for the behavioral gains. Nevertheless, this study clearly demonstrated that children with autism could maintain high and constant levels of both on-task and on-schedule responding cued via an activity schedule. Furthermore, these levels of responding were achieved with minimal adult supervision during the intervention and maintenance phases and with no supervision during the generalization probes. Previous studies investigating the impact of activity schedules, delay schedules, prompt fading, and unpredictable supervision as independent variables did not attempt to assess responding in the absence of all adults (Dunlap & Johnson, 1985; Dunlap, Koegel, & Johnson, 1987; Krantz, MacDuff, & McClannahan, 1993; MacDuff, Krantz, & McClannahan, 1993; Pierce & Schreibman, 1994). In these studies, adults other than the experimenter were present during the implementation of the intervention and during the probes. As mentioned earlier this could be a confounding variable since "adults" as a class could have acquired discriminative qualities for access to reinforcement. In the present study we conducted all sessions in an isolated classroom in the absence of other children and adults unless otherwise specified by the intervention. We established increased levels of on-task and on-schedule responding with adults only occasionally present. For instance in the maintenance condition, the adult was present only once for as little as 2 sec throughout the 15 min session. Similar findings were demonstrated by the generalization probes with no adult supervision.

Several additional components might have contributed to the success of the intervention package. For instance, the systematic fading of instructional prompts and fading the instructor's presence may have resulted in transfer of stimulus control to the schedule and the materials. Additionally, establishing unpredictable supervision may have contributed to the reduction of off-task responding and stereotypy during these conditions as well as during the generalization probes. The low levels of off-task responding and stereotypy indicate that although response cost had never been implemented during the generalization probes, the possibility for the instructor to enter the room and implement the procedure at any time existed.

Clinical data that were available from the children's center based program indicated that in the past reinforcement contingencies alone were not sufficient to eliminate off-task responding and stereotypy in the absence of an instructor. Therefore, we programmed the implementation of

response cost for these responses. Future research might investigate whether reinforcing desirable behavior and concurrently punishing undesirable behavior are both necessary to produce desirable and lasting behavioral change in the absence of continuous adult supervision while simultaneously eliminating off-task responding and stereotypy. Another point that requires further investigation involves the effectiveness of delaying consequences, specifically response cost, in producing lasting behavioral change. For example, although we arranged for reinforcement to be contingent via self-delivery of tokens, access to the reinforcer was delayed until the end of the 15-min session. Response-cost, however, was implemented contingent on each and every occurrence of off-task responding. It would be important to assess the effectiveness of the intervention package when response-cost was faded to the end of the session. Successfully fading response cost would accomplish two things. First, it would add testament to the value of the intervention package because it would signify that behavior is under the control of more naturalistic contingencies. Second, it would increase the practical value of the intervention package because the need for continuous observation to determine whether response cost should be contingently implemented or not would be eliminated.

Bringing behavior under the control of naturalistic contingencies is important in ensuring children's inclusion in typical settings. This is especially true since the inappropriate behavior of children with autism is not always observed in the process of occurring and delayed punishment is often not effective in reducing such behavior. One way to assess the effectiveness of noncontingent and delayed implementation of response cost would be to collect data on on-task/schedule and off-task/schedule responding during an extended maintenance phase. Such a phase would be engineered to simulate the natural environment in which children are observed engaging in inappropriate behavior that would be consequated occasionally at a later point in time (delayed punishment). It would be interesting to compare rates of responding over time in this condition to rates of responding maintained by continuous and immediate aversive consequences. Finally, another important point to be examined is whether high rates of on-task and on-schedule responding would be maintained in a classroom type setting with the use of this intervention package.

Dunlap, Koegel, Johnson, and O'Neil (1987) in their study successfully faded both reinforcing and punishing consequences that were initially delivered contingently (e.g., praise and reprimands) to the end of the session. The authors concluded that systematic fading of reprimands was necessary to increase the durability of appropriate responding. It should be noted, however, that the presence of various adults other than the experimenter in all conditions throughout this study might have confounded the outcome. In contrast to the Dunlap et al. study, the present study seems to

provide some evidence that durable behavior can be produced without delaying the implementation of response cost. For instance, in the present study all three children showed increased levels of on-task and on-schedule responding during both generalization probes across novel material and novel settings. In fact, response-cost was never delivered during generalization probes. Additionally, two of the children showed increased responding during maintenance, without the need for fading response-cost. Similarly, during maintenance response-cost was never implemented on Zeb's responding and it was implemented rarely (0.7 % of the intervals scored) for David's off-task responding. A more prolonged maintenance phase in the present study, however, across all three children would have been helpful in further supporting this argument.

Many researchers have criticized the employment of packages arguing that use of multiple elements as an independent variable makes it impossible to specify relationships between dependent and independent variables (Michael, 1980). At times, however, a scientist/practitioner may choose to use a package, when it is expected that it is more likely to produce a desirable outcome for the participant rather than an intervention that is more analytic (Azrin, 1977). The argument in favor of packages is based on the philosophy that focus should be directed to outcome. Thus, the development of methods that maximize speed of acquisition and increase the number of persons who may benefit represents a worthy endeavor. This argument has direct implications regarding the present study.

The importance of identifying effective means of teaching persons with autism to work independently is indisputable. Yet only a limited number of studies was identified that investigate the effects of components similar to the ones used in the current intervention. Furthermore, most of these studies were conducted with adults. Finally, in all of the studies identified, a supervising adult was always present in the room with the participants. This could have been a confounding variable with respect to conclusions regarding performance in the absence of a supervising adult. The present study has provided some evidence that young children with autism can produce high rates of on-task and on-schedule responding in the absence of a supervising adult and maintain such rates over time with minimal adult supervision. It remains the burden of experimenters/practitioners to extend these findings to additional children and variable settings as well as to investigate whether all the components used in the present study are necessary to replicate these outcomes.

### References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Azrin, N. H. (1977). A strategy for applied research: Learning based but outcome oriented. *American Psychologist*, 32, 140-149.

- Billingsley, F. F., & Romer, L. T. (1983). Response prompting and the transfer of stimulus control: Methods, research, and a conceptual framework. *Journal of the Association for the Severely Handicapped*, 8, 3-12.
- Corte, H. E., Wolf, M. M., & Locke, B. J. (1971). A comparison of procedures for eliminating self-injurious behavior of retarded adolescents. *Journal of Applied Behavior Analysis*, 4, 201-214.
- DeMyer, M. K., Hingtgen, J. N., & Jackson, R. K. (1981). Infantile autism reviewed: A decade of research. *Schizophrenia Bulletin*, 7, 388-451.
- Dunlap, G., & Johnson, J. (1985). Increasing the independent responding of autistic children with unpredictable supervision. *Journal of Applied Behavior Analysis*, 18, 227-236.
- Dunlap, G., Koegel, R. L., & Johnson, J. (1987). Maintaining performance of autistic children in community settings with delayed contingencies. *Journal of Applied Behavior Analysis*, 20, 185-191.
- Dunlap, G., Koegel, R. L., & O'Neil, R. E. (1985). Pervasive developmental disorders. In P. H. Bornstein & A. E. Kazdin (Eds.), *Handbook of clinical behavior therapy with children*, (pp. 499-540). New York: Dorsey Press.
- Dunlap, G., & Plenis, A. J. (1988). Generalization and maintenance of unsupervised responding via remote contingencies. In R. H. Horner, G. Dunlap, & R. L. Koegel (Eds.), *Generalization and maintenance: Life-style changes in applied settings* (pp. 121-142). Baltimore: Paul H. Brookes.
- Fowler, S. A. (1988). The effects of peer-mediated interventions on establishing, maintaining, and generalizing children's behavior changes. In R. H. Horner, G. Dunlap, & R. L. Koegel (Eds.), *Generalization and maintenance: Life style changes in applied settings* (pp. 143-170). Baltimore: Paul H. Brookes.
- Frank, A., Wacker, D., Berg, W., & McMahon, C. (1985). Teaching selected microcomputer skills to retarded students via picture prompts. *Journal of Applied Behavior Analysis*, 18, 179-185.
- Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis*, 11, 189-196.
- Johnson, B., & Cuvo, A. (1981). Teaching mentally retarded adults to cook. *Behavior Modification*, 5, 187-202.
- Koegel, R. L., & Rincover, A. (1977). Research on the difference between generalization and maintenance in extra-therapy settings. *Journal of Applied Behavior Analysis*, 10, 1-16.
- Koegel, R. L., Rincover, A., & Egel, A. L. (1982). *Educating and understanding autistic children*. San Diego: College-Hill Press.
- Krantz, P. J., MacDuff, M. T., & McClannahan, L. E. (1993). Programming participation in family activities for children with autism: Parent's use of photographic activity schedules. *Journal of Applied Behavior Analysis*, 26, 137-138.
- Krantz, P. J., & McClannahan, L. E. (1999). Strategies for integration: Building repertoires that support transitions to public schools. In P. M. Ghezzi, W. L. Williams, & J. E. Carr (Eds.), *Autism: Behavior Analytic Perspectives* (pp. 221-223). Reno, NV: Context Press.
- MacDuff, G. S., Krantz, P. J., & McClannahan, L. E. (1993). Teaching children with autism to use photographic activity schedules: Maintenance and generalization of complex response chains. *Journal of Applied Behavior Analysis*, 26, 89-97.
- Marholin, D., & Steinman, L. (1977). Stimulus control in the classroom as a function of the behavior reinforcers. *Journal of Applied Behavior Analysis*, 10, 456-478.
- Michael, J. (1980). Flight from behavior analysis. Presidential address ABA 1980. *The Behavior Analyst*, 3, 1-22.
- Nietupski, J., Clancy, P., & Christiansen, C. (1984). Acquisition, maintenance, and generalization of vending machine purchasing skills by moderately handicapped students. *Education and Training of the Mentally Retarded*, 19, 91-96.
- Pierce, K., & Schreibman, L. (1994). Teaching daily living skills to children with autism in unsupervised settings through pictorial self-management. *Journal of Applied Behavior Analysis*, 27, 471-481.

- Risley, T. R. (1968). The effects and side effects of punishing the autistic behaviors of a deviant child. *Journal of Applied Behavior Analysis*, 1, 21-34.
- Sailor, W., Goetz, L., Anderson, J., Hunt, P., & Gee, K. (1988). Research on community intensive instruction as a model for building functional, generalized skills. In R. H. Horner, G. Dunlap, & R. L. Koegel (Eds.), *Generalization and maintenance: Life-style changes in applied settings* (pp. 113-145). Baltimore: Paul H. Brookes.
- Schwarz, M. L., & Hawkins, R. P. (1970). Application of delayed reinforcement procedures to the behavior of an elementary school child. *Journal of Applied Behavior Analysis*, 3, 85-96.
- Snell, M. E. (1983). Implementing and monitoring the IEP: Intervention strategies. In M. E. Snell (Ed.), *Systematic instruction of the moderately and severely handicapped* (pp. 119-145). Columbus, OH: Merrill.
- Sowers, J., Verdi, M., Bourbeau, P., & Sheehan, M. (1985). Teaching job independence and flexibility to mentally retarded students through the use of a self-control package. *Journal of Applied Behavior Analysis* 18, 81-85.
- Stahmer, A. C., & Schreibman, L. (1992). Teaching Children with autism appropriate play in unsupervised environments using a self-management package. *Journal of Applied Behavior Analysis*, 25, 447-459.
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, 10, 349-367.
- Thinesen, P., & Bryan, A. (1981). The use of sequential picture cues in the initiation and maintenance of grooming behaviors with mentally retarded adults. *Mental Retardation*, 19, 247-250.
- Wacker, D. P., & Berg, W. K. (1983). Effects of picture prompts on the acquisition of complex vocational tasks by mentally retarded adolescents. *Journal of Applied Behavior Analysis*, 16, 417-433.
- Wacker, D. P., & Berg, W. K. (1984). Training adolescents with severe handicaps to set up job tasks independently using picture prompts. *Analysis and Intervention in Developmental Disabilities*, 4, 353-365.