Effects of Instruction, Goals, and Reinforcement on Academic Behavior: Assessing Skill versus Reinforcement Deficits

Claire St. Peter Pipkin, Shari M. Winters & James W. Diller

Abstract

We attempted to determine the effects of instructions, goal setting, and reinforcement, in isolation and combination, on the letter naming proficiency of 2 underperforming Kindergarten students. During a no-intervention baseline, both students’ accuracy was low or declining. Instructions alone produced increases over baseline responding, but the effects were not maintained; goal setting and reinforcement alone did not produce any initial effects. When reinforcement was combined with instructions and goal setting, the performance of both participants improved in all conditions, with greatest initial improvements in the instruction plus reinforcement condition. Data from this phase showed some evidence of carryover between conditions. These findings suggest that single interventions were not sufficient for performance improvement, but that combined interventions were effective.

Keywords: academic performance, goal setting, instruction, Kindergarten students, letter naming, literacy, motivational deficits, skill deficits, reinforcement

There are several possible causes when students underperform or otherwise fail to meet academic standards. Gilbert (1978) proposed six variables that may be involved in underperformance: information, instruments, incentives, knowledge, capacity, and motives. Of these domains, information, incentives, knowledge, and motivation may be of particular interest in educational settings. Informational deficits (and deficits in the related knowledge domain) involve individuals not knowing how to perform the skill, not knowing what performance standards are in place, or not receiving feedback about their behavior relative to the expectations. According to Gilbert, failures of incentives involve weak or poorly scheduled reinforcers for performing well or competing reinforcers for poor performance. When students underperform, the question of whether the deficits are due to a lack of instruction or to mismanaged contingencies is important to consider; does the performer know how to do what they are being asked to do? Is there sufficient reinforcement available for performing the expected task? When these questions are answered correctly, an intervention aimed at improving performance may be developed.

Several studies have examined assessments and interventions based on differences between skill and performance deficits (e.g., Daly, Martens, Hamler, Dool, & Eckert, 1999; Duhon et al., 2004; Eckert et al., 2000; Lerman, Vondran, Addison, & Kuhn, 2004; Noell et al., 1998). The efficacy of these interventions has been demonstrated across different populations (including children with autism, developmental disabilities, and typical development), responses (including math and reading at a variety of levels), and settings (including homes, schools, and clinics). Interventions typically involve some combination of instructional strategies, goal setting, performance feedback, and reinforcement, in isolation or combination. For example, Duhon et al. (2004) examined the utility of a brief assessment (similar to the one described by Northup et al., 1991) for predicting performance in extended interventions targeting math and writing skills. Brief assessments were conducted to establish baseline levels of performance. Instructions were provided but no explicit reinforcement contingency was in place. Following the assessment, the experimenters conducted out-of-class sessions to examine effects of goal setting and rewards. Two participants performed more accurately during the goal/reward sessions than during the initial assessment, suggesting a performance deficit, while the other 2 participants showed no change between instructional and goal/reward sessions, suggesting a skill deficit. During the extended intervention component of the study, the experimenters employed an alternating-treatment design (e.g., Barlow & Hersen, 1984) to evaluate the relative effectiveness of performance- or skill-based treatments.
The performance-based treatment involved stating a goal and allowing access to a small reward if the goal was met. The skill-based treatment involved pre-session practice or the use of an instructional aid. Students with hypothesized performance deficits improved in the performance-based treatment and students with hypothesized skills deficits improved in the skill-based treatment.

Like Duhon et al. (2004), most research assessing reinforcement- or skill-deficits has used combinations of treatment components. For example, Duhon et al. incorporated both goals and reinforcement in their performance-based intervention. Eckert, Adroin, Daisey, and Scarola (2000) assessed various combinations of six components commonly used in performance- or skill-based interventions for reading: previewing the reading passage, practicing the passage, performance feedback (on the amount of time only or amount of time plus number of errors), goal setting (with goals set by the student or the experimenter), performance charting, and reinforcement. Reading fluency was improved in conditions combining skill-based interventions (preview plus practice) with performance-based interventions (feedback, goal setting, charting, or reinforcement), as compared to baseline. However, results were largely undifferentiated between the various combined skill- and performance-based interventions. Additionally, each condition included at least two components, precluding determination of the effect of any individual component on responding.

Prior literature has suggested that both skill and performance deficits may contribute to poor academic outcomes, and that interventions can be developed to address these deficits (e.g., Bonfiglio, Daly, Marten, Lin, & Corsaut, 2004; Eckert et al., 2000; Eckert et al., 2002; Lerman et al., 2004; Noell et al., 1998). Authors have acknowledged that the interactions between the antecedent-based and consequence-based interventions may be difficult to separate (cf., Bonfiglio et al., 2004). Unfortunately, most research exploring performance or skills deficits combined several individual components into a larger treatment package (e.g., Eckert et al., 2000). In particular, goal setting was frequently combined with reinforcement or performance feedback, making the isolated effects of these components difficult to discern (e.g., Bonfiglio et al., 2004; Duhon et al., 2004; Eckert et al., 2000; 2002). Goal setting in isolation may be an easy and inexpensive method of improving students’ academic behavior. However, if goals alone are not sufficient to promote appropriate behavior, additional intervention components may be required.

The present study examined effects of instruction, goal setting, and reinforcement, in isolation and combination, on letter naming with typically developing Kindergarten students. Two interventions were used to determine whether participant noncompliance was due to skill deficits, performance deficits, or some combination of the two. During the first intervention, three conditions were compared in an alternating-treatments design: (1) pre-session instruction was provided but no consequences were programmed, (2) goals were stated but no consequences were programmed for meeting those goals, or (3) no pre-session antecedents (i.e., goals or instructions) were used but access to reinforcers was provided contingent on improved performance. The purpose of this comparison was to determine the effects of instruction, goals, and reinforcement in isolation. During the second intervention, access to small edibles was included in all conditions to determine if combined skill and performance deficits contributed to noncompliance with academic tasks.

Method

Participants, Setting, and Materials
Two 5-year-old children were referred by teachers and caregivers for skill deficits in alphabet letter recognition and disruptive behavior during class instruction. Jake and Theresa were both Kindergarten students in regular-education classrooms. Both were of at least average intelligence, but their disruptive behavior in the classroom resulted in below-average grades and put them at risk for retention in Kindergarten for the next academic year.

Sessions were conducted in small testing rooms located near the participants’ classrooms. Testing rooms contained tables, child- and adult-sized chairs, storage cabinets, and a variety of educational items, such as books and games. Small (1 in. by 1 in.) letter tiles were used throughout the experiment.

Response Measurement and Reliability

Observers were undergraduate and graduate students, who were trained using a real-time data collection program (InstantData v1.1 for PC) on laptop computers. All observers had previously attained interobserver agreement scores (calculated as described below) of at least 90% for three consecutive sessions. Observers recorded the duration of each session, the number of letters identified correctly, the number of letters identified incorrectly, and the number of times the participant said “I don’t know.” Correct responding was operationally defined as naming the letter that was presented. Incorrect responding was operationally defined as naming a different letter than the one presented. Percentage of correct responses was calculated by dividing the number of correct responses by the sum of incorrect and “I don’t know” responses, and multiplying by 100.

For Jake, data were collected using a real-time data collection program (InstantData v1.1 for PC) installed on laptop computers. Because computers were not available for sessions with Theresa, data were collected by tallying the number of correct, incorrect, and “I don’t know” responses using a paper-and-pencil data sheet.

Two observers independently observed and scored responses for 25% of sessions across both participants. Interobserver agreement (IOA) was calculated by dividing the lesser number of responses scored by an observer by the greater number (scored by the other observer) and multiplying the quotient by 100. Interobserver agreement for correct letter identification averaged 94.2% (range 81.8% to 100%), incorrect letter identification averaged 93.2% (range 66.7% to 100%), and “I don’t know” responses averaged 96.0% (range 50.0% to 100%).

Procedure

Experimental Design. This experiment consisted of a nonconcurrent multiple-baseline-across-participants design with an embedded reversal for Jake. Multiple sessions, each 2 min in duration, were conducted two to three days per week. Prior to the start of each session, the child was led into the room and asked to choose where he or she would like to sit. An experimenter and an observer sat at the same table, approximately 0.5 m away from the child. All experimenters were doctoral students. Sessions were conducted until participants attained 80% accuracy on at least one letter set, with no evidence of a downward trend.

Pre-experimental Letter-Naming Assessment. We first conducted a pre-experimental assessment to determine which letters the child was able to correctly identify. Participants were told that if they did not know the name of a letter, they could say “I don’t know.” Letter tiles were presented in random order. Upon presentation of each letter tile, the experimenter prompted the child to identify the name of the letter by asking “what letter?” If the child did not identify the letter within 10 s, the experimenter prompted the child again. The experimenter continued to prompt every 10 s until the child responded. When the child responded (correctly or incorrectly), the letter was removed, and the experimenter immediately presented
another letter. Each session lasted until all 26 letters were presented 10 times. Following the session, the experimenter provided the child with noncontingent access to leisure items for approximately 2 min before starting the next session. Three sessions were conducted for each participant during the pre-experimental assessment. The letters that were incorrectly identified on each presentation for all three sessions were randomly assigned to one of three sets that would be used in subsequent analyses. Letter sets contained four and six letters for Jake and Theresa, respectively.

Baseline. Baseline sessions were conducted for each participant to ensure that each of the sets contained letters that were all unknown to the participants. The child was asked to name as many letters as possible in 2-min sessions by using the same verbal prompt as in the pre-experimental assessment. As in the pre-experimental assessment, the child was re-prompted every 10 s if he or she not respond within that period. If the child responded (correctly or incorrectly), the letter was removed, and the experimenter immediately presented another letter. The letter tiles were presented randomly without replacement and were cycled through until 2 min elapsed. There were no programmed consequences for responses, either correct or incorrect, during the baseline phase.

Intervention 1. During Intervention 1, antecedents and consequences differed across the three sets of letters. In one set (hereafter, the “instruction set”), the experimenter provided pre-session instruction on the names of the letters in that set. If underperformance was due to skill deficits alone, accuracy should improve in this condition. Instruction consisted of the experimenter showing the child a letter and saying the name of the letter. The experimenter then prompted the child to repeat the letter name. If the child did not repeat the name of the letter within 10 s, the experimenter repeated the name of the letter and prompted the child to repeat the letter name. Pre-session instruction continued until all letters in the set had been shown four times. An instructional period was conducted immediately before each session with the instruction set.

The experimental session began immediately after the instructional period. The experimenter presented a letter and prompted the child to name the letter. A prompt occurred every 10 s until the child responded. Following a response, either correct or incorrect, the letter was removed, and a new letter was presented. The letter tiles were presented without replacement and were cycled through until 2 min elapsed. No pre-session goals were stated, and no reinforcers were available.

For another set of letters (hereafter, the “goal set”), the experimenter provided a goal before the start of the session by stating “Try to get [goal value] right.” The goal was determined by adding one letter to the number of letters identified correctly in the previous session with the goal set. For example, if the child identified 17 letters correctly in the previous session with the goal set, the goal would be to identify 18 letters correctly. The purpose of this manipulation was to determine the effect of goals without instruction or reinforcement on letter naming; if underperformance was due to performance deficits alone, accuracy might improve in this condition. The experimental session began immediately after the goal was stated. The experimenter presented a letter and prompted the child to name the letter. A prompt occurred every 10 s until the child responded. Following a response, either correct or incorrect, the letter was removed, and a new letter was presented. The letter tiles were presented randomly without replacement and were cycled through until 2 min elapsed. No instructions, feedback, or reinforcement were available.
The final set of letters (hereafter, the “reinforcement set”) was not preceded with any instruction or goal. However, if the child reached the reinforcement criterion described below, he or she was provided with access to edible items without explanation of how those items were earned. The reinforcement criterion was determined by adding one letter to the number of letters identified correctly in the previous session with the reinforcement set (the same calculation used to determine the goal, only based on performance in a different set). If the child reached the reinforcement criterion, he or she was provided with five small edible items that had been identified as preferred through both parental and child nomination. Reinforcers for attaining the criteria for Jake and Theresa were Superman® gummy fruit snack and Goldfish® crackers, respectively. The purpose of this manipulation was to determine the effects of reinforcement alone on letter naming; if underperformance was due to reinforcement deficits alone, accuracy should improve in this condition. A prompt occurred every 10 s until the child responded. Following a response, either correct or incorrect, the letter was removed, and a new letter was presented.

The letter tiles were presented randomly without replacement and were cycled through until 2 min elapsed.

*Intervention 2.* During Intervention 2, the antecedents for each of the sets were identical to those described in Intervention 1. However, the consequence for each letter set was identical to that of the reinforcement set described in Intervention 1. Therefore, if the child correctly identified in one session the same amount as the previous session with that set plus one, the experimenter provided the child with five edible items. The antecedents and consequences associated with the reinforcement set were identical to Intervention 1; no antecedents were used and reinforcers were available for improved performance.

*Reinforcement Only.* Jake was exposed to a final phase, to determine the effect of reinforcement only on letter naming with the sets previously associated with instructions, goals, and reinforcement. During this phase, there were no programmed antecedents (instruction or goals), but reinforcers were available for all sets contingent on improved performance (as in Intervention 2). During all sessions, a prompt occurred every 10 s until the child responded. Following a response, either correct or incorrect, the letter was removed, and a new letter was presented. The letter tiles were presented randomly without replacement and were cycled through until 2 min elapsed. Attainment of the reinforcement criterion (as previously described) resulted in the presentation of edible items to the child.

**Results**

Results from the pre-experimental assessment (not shown in Figure) identified 12 letters incorrectly named by Jake and 19 letters incorrectly named by Theresa. To create equal numbers of letters across the three sets, one eligible letter was excluded from Theresa’s analysis. Thus, each letter set for Jake and Theresa contained four and six letters, respectively, depicted in both upper and lower case. For Jake, the instruction set contained the letters E, G, U, and Y; the goal set contained H, P, Q, and W; the reinforcement set contained F, M, T, and R. For Theresa, the instruction set contained the letters B, E, H, N, U and X. The goal set contained A, D, G, R, L, and W, and the reinforcement set contained C, F, I, Q, V, and Y. These letter sets remained constant through all baseline and intervention phases.
Results for the baseline and intervention phases for Jake and Theresa are shown in Figure 1.

Figure 1. Baseline and intervention results for Jake (upper panel) and Theresa (lower panel). Triangles denote responding in the instruction set. Squares denote responding in the goal set. Circles denote responding in the reinforcement set. Shaded symbols show sessions in which reinforcement contingencies were added to the antecedent manipulations.

The accuracy of responding, shown as percentage of letters named correctly, is along the y-axis, with consecutive sessions along the x-axis. Each data path shows results for a different set of letters: the instruction set by triangles, the goal set by squares, and the reinforcement set by circles.
The results for Jake are shown in the upper panel. During baseline, Jake correctly responded to more letters in the instruction set than any other set. This is surprising given that Jake was not able to correctly identify any of the letters (in any set) during the pre-experimental assessment and that we provided no instruction on any of the letters during the baseline phase. Nonetheless, Jake’s performance on all sets was stable or decreasing by the end of the phase.

Intervention 1 involved manipulations of antecedents for the instruction and goal sets and a manipulation of consequences for the reinforcement set. Jake’s performance initially improved in the instruction set, but declined over the course of the phase. Jake’s performance in the goal and reinforcement sets did not improve, despite contacting the reinforcer following session 12. These results suggest that, for Jake, instruction may be necessary but insufficient for performance increases.

Intervention 2 involved manipulations of antecedents and consequences for the instruction and goal sets (opportunity to earn a reinforcer was added to the previously used antecedents). To highlight the difference in consequences, we shaded the triangles (instruction set) and squares (goal set). The antecedents and consequences in the reinforcement set remained identical to Intervention 1. Jake’s performance improved on all three sets during Intervention 2. The greatest immediate gain occurred in the set in which Jake received instruction before the session and a reinforcer for meeting the performance criterion. However, immediate gains were also observed in the goal set and the reinforcement set. By the middle of Intervention 2, Jake was performing with high accuracy on all three sets, with no differentiation in responding between the sets.

Because of this finding, we reversed to baseline to determine the role of reinforcement in Jake’s performance. Jake’s accuracy continued to improve on the reinforcement set, but declined somewhat on the other two sets. Finally, we instituted a reinforcement-only phase, in which no antecedent interventions were used. During the reinforcement-only phase, Jake’s performance improved from the previous baseline phase for all three sets. However, improvements were slower in the set previously associated with goals than in the other two sets. This difference seemed to be related to the particular letters that were assigned to that set. Specifically, the goal set contained the letters P and Q. Jake consistently answered “P” when shown a lower case Q. Therefore, we conducted a brief instructional period immediately before session 93. During the instruction period, the experimenter reviewed upper and lower case P and Q with Jake, using the strategy previously used with the instruction set, and then asked Jake to independently name the letters. Feedback was provided for correct and incorrect naming during the instructional period. The instructional period continued until Jake was able to accurately name both upper and lower case letters correctly two consecutive times. Following this instructional period, Jake’s accuracy on the set previously associated with goals increased to above 80% correct.

Theresa’s results are shown in the lower panel of Figure 1. During baseline, Theresa correctly named a letter on only one opportunity. There were no clear differences between the three sets. During Intervention 1, Theresa’s performance on the instruction set increased initially, but decreased by the end of the phase. This result suggests that instruction alone was insufficient to promote and maintain accurate responding, but may have been a necessary component of an intervention, as it was for Jake. Theresa’s performance did not improve in the goal and reinforcement sets during Intervention 1. However, Theresa never contacted the reinforcement contingency during this phase, so the effects of reinforcement alone cannot be determined from these data. During Intervention 2, Theresa’s performance in the instruction set exceeded her performance in the other two sets. However, accuracy increased above baseline and Intervention 1 levels for all sets. Unlike Jake’s performance, Theresa’s letter naming in the goal set and reinforcement set declined through the course of the phase.
Discussion

We evaluated the effects of instructions, goal setting, and reinforcement, in isolation and combination, on the letter naming performance of 2 typically developing Kindergarten students. Instruction, goal setting, and reinforcement were not effective in isolation, suggesting that the student’s poor classroom performance was due to both instructional and reinforcement deficits. When reinforcement was combined with instruction and with goal setting, performance in all conditions improved. For Jake, exposure to the combined interventions seemed initially necessary, but reinforcement alone was sufficient by the end of the experiment. For Theresa, a combination of instruction and reinforcement was necessary to maintain performance.

Both participants demonstrated improved accuracy on all sets when a reinforcement contingency was added to instructions and goal setting. In other words, performance improved even on the letter set that remained unchanged from the previous phase (the reinforcement set). This suggests that our initial conclusion from Intervention 1—that instruction was necessary but not sufficient to maintain performance—was incorrect. Had instruction been necessary but not sufficient, the students’ performances would have increased on the instruction set, but not the goal or reinforcement sets, during Intervention 2. The increase in accuracy on the goal and reinforcement sets suggests that the students “knew” more than they were demonstrating during the previous phase. This change in behavior may be due to changes in discriminative stimuli, in that the edible reinforcers were placed on the table within the participant’s line of sight, during Intervention 2. Perhaps the presence of the edibles signaled the availability of reinforcers across conditions and was sufficient to increase responding. It is also possible that some other feature of Intervention 2, such as an increased overall reinforcement rate or generalization between conditions, led to improved performance.

Theresa’s performance in the goal and reinforcement sets did not maintain during Intervention 2. Her performance in both sets declined steadily after initial improvements. It is possible that the edible item (Goldfish® crackers) lost its reinforcing efficacy over time. Although no formal reinforcer assessment was conducted, Theresa frequently requested Goldfish® after the sessions during Interventions 1 and 2. However, a formal reinforcer assessment may have clarified these results. It is also possible that Theresa forgot some of the letters across time (as it seems that Jake “forgot” the difference between lower case P and Q between Intervention 2 and the reinforcement-only phases). This “forgetting” could have been attenuated in the instruction set, in which we told Theresa the names of the letters before the start of each session.

The persistence of accurate responding during the reversal to baseline for Jake warrants further investigation. Jake’s behavior in the set previously associated with reinforcement showed more resistance to extinction during the reversal to baseline than did behavior in the other two sets. It is possible that the change from Intervention 2 to baseline was less discriminable for the reinforcement set than for the other two sets. For both the instruction and goal sets, the return to baseline involved two changes: the removal of the antecedents (instruction or goals, respectively) and the removal of the reinforcement contingencies. The change in antecedents may have made the transition to baseline more discriminable than with the reinforcement set, which involved the removal of the reinforcement contingency, but no change in the antecedent verbal behavior of the experimenter. It is also possible that the distribution of letters across the sets influenced this outcome. The reinforcement set contained letters appearing more commonly in English words; the average frequency of letters in the reinforcement set was 256.8 per 1000, in contrast to 220.3 per 1000 for the instruction set and 99.8 per 1000 for the goal set (calculated using data from Pratt, 1996). Given that declines in Jake’s performance across sets matched the average frequency of the letters in that set, it seems plausible that the frequency with which Jake was likely to encounter these letters outside of the experiment may have influenced his performance during baseline conditions.
It is also possible that some of the initial improvements in performance were due to events occurring outside the experimental situation. For example, both participants were exposed to a regular-education Kindergarten curriculum outside of the experimental sessions. Therefore, although we did not provide instruction on the letters in the goal and reinforcement sets, both participants were undoubtedly exposed to some instruction on these letters in their classrooms. The influence of this extra-experimental history cannot be eliminated as a contributing factor to the results of this study. We chose to target letter naming in this experiment because of its social significance to the children and their parents; future research could use arbitrary stimuli, for which participants would receive no extra-experimental exposure.

For Jake, interaction between the sets during Intervention 2 seems likely. Jake’s performance in the goal and reinforcement sets did not improve until he received a reinforcer in the instruction set. In addition, his performance on the goal and reinforcer sets improved dramatically between sessions 46 and 48. These sessions were conducted on the same day; there is little chance that this improvement was due to events occurring outside the experiment (such as extra instruction at home or in the classroom, or a change in the reinforcing efficacy of the Gummies®). It is difficult to determine what factors led to the dramatic improvement in performance over the course of the day’s sessions. Recall that Jake was never given instructions on how to earn reinforcers; it is possible that Jake “figured out” (developed some self-generated rules) that reinforcers were reliably available for improved performance. It is also possible that Jake’s behavior in the instruction set generalized across the other two sets. Future research could examine the effects of instructions, goals, and reinforcement using a reversal design to reduce the possibility of carryover or interaction between the conditions.

One limitation of this study is that Theresa was not exposed to a reversal to baseline. Jake’s behavior during the reversal to baseline may be atypical for performance under these conditions, and it would have been helpful to observe changes in Theresa’s performance. However, Theresa’s data showed clear differentiation between the sets, and she met the study termination criterion, during Intervention 2. It also would have been interesting to determine if Theresa’s behavior in the instruction set would maintain during a reinforcement-only phase. If this was the case, it would suggest that continuing instruction was not necessary to maintain performance.

A second limitation of the study is the experimental design. We used a nonconcurrent multiple-baseline-across-participants design because of the timing of the referrals. When we began collecting Theresa’s data, there were no other Kindergarten students available who were seemingly capable, but underperforming. Theresa’s teacher requested that we begin working with her immediately because of an escalation of her noncompliance in the classroom. However, use of a concurrent multiple-baseline-across-participants design would have permitted stronger demonstration of experimental control. Nonetheless, we were able to demonstrate differentiation in responding between the experimental conditions, with results that replicated (at least to some extent) across participants.

The results of this study suggest that a single intervention strategy, including goal setting alone, may be ineffective at improving the academic performance of Kindergarten students. This study replicates prior work suggesting that multiple-component interventions may be necessary to improve performance (cf. Eckert et al., 2000), and extends the current literature by demonstrating that goal setting was not effective in isolation. We hope that these results underscore the importance of implementing effective instructional strategies combined with reinforcement contingencies to support academic behavior. We also hope that this study will provide a foundation for future research examining the interaction between rapidly alternating instructional conditions (such as instruction plus reinforcement and reinforcement only). If future research determines that these interactions improve performance, teachers could use that information to determine effective instructional designs. For example, a teacher could intersperse periods of instruction plus reinforcement with periods of reinforcement alone. This strategy may be particularly useful when a teacher can only provide instruction to select students (i.e., during small-group instruction).
References


Author Contact Information:

Peter Pipkin
Department of Psychology
West Virginia University,
Box 6040, Morgantown, WV 26506