

Project VIABLE:

Evaluating the Effectiveness of a Direct Behavior Rating Training Module

Sayward E. Harrison, Rose Jaffery, & Austin Johnson

Contributors: T. Chris Riley-Tillman, Sandra M. Chafouleas, Theodore J. Christ

ABSTRACT

Direct Behavior Rating (DBR) is a flexible, repeatable, and defensible behavioral assessment method that involves making a brief rating of student behavior immediately following an activity. Results from prior research have guided the development of an **online training module** which utilizes modeling, practice, and feedback to train raters in the use of DBR. Attendees will become familiar with the module and will learn results from a study which demonstrated the module's effectiveness for improving rater accuracy.

Project VIABLE

This study represents one of the investigations initiated under Project VIABLE. Through Project VIABLE, empirical attention is being directed toward the development and evaluation of formative measures of social behavior involving a direct behavior rating (DBR). The goal of Project VIABLE is to examine DBR through 3 phases of investigation including 1) foundations of measurement, 2) decision making and validity, and 3) feasibility.

PURPOSE OF THE PRESENT STUDY

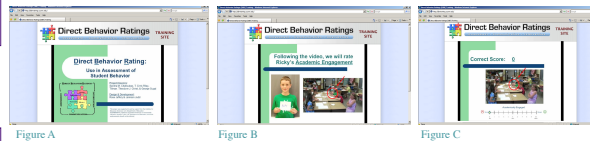
Previous research has demonstrated the efficiency and technical adequacy of DBR as a method of assessing student behavior (Chafouleas, Riley-Tillman, & Christ, 2009). Prior research has also explored the various levels of training intensity and the components needed to improve rater accuracy using DBR (Chafouleas et al., 2009; LeBel, Kilgus, Briesch, & Chafouleas, 2010; Schlientz, Riley-Tillman, Briesch, Walcott, & Chafouleas, 2009). Results from these studies helped to guide the development of an online training module through which potential raters can receive training on how to use DBR to assess student behavior.

The purpose of this study was to examine the effectiveness of this online training module in improving rater accuracy using DBR. It was hypothesized that participants who completed the online training module would demonstrate greater accuracy in rating student behavior than participants who learned about DBR by reading a behavioral assessment text.



DBR TRAINING MODULE

An internet-based interactive module was created to provide training on DBR (Figure A; <http://dbrrtraining.uconn.edu/>). The training module provides participants with an **overview of DBR** and a **demonstration** section that models how to make ratings of student behavior using a virtual DBR form. The module then provides trainees with a **practice and feedback** section in which participants watch 30-60 second clips of students participating in classroom activities and practice rating three specified target behaviors: *Academically Engaged*, *Disruptive*, and *Respectful* (Figure B). Participants receive corrective feedback regarding the accuracy of their ratings in the form of a voiceover within the module (Figure C). The module's practice and feedback section is also tailored to the performance of the trainee. Thus, if a trainee rates *Academic Engagement* inaccurately compared to the actual rating (pre-determined via systematic direct observation procedures), then the trainee again views and rates a clip displaying a similar level of *Academic Engagement* in increase his/her accuracy. The training module takes approximately 35 minutes to complete, and once trainees have progressed through the module, they are provided with a certificate of completion.



PROCEDURE

Ninety students from a large university in the Southeast blindly self-selected into one of six study sessions, which had been randomly assigned beforehand to either the experimental or control condition. All participants initially completed a DBR Experience Questionnaire to assess previous experience with DBR. Participants then proceeded to take a pre-test in which they watched nine one-minute video clips on a large screen. For each clip, participants rated one of three target behaviors (i.e., *Academically Engaged*, *Disruptive*, *Respectful*) for a target student. For the three control sessions, participants (n=47) next read two chapters from the book *School-based Behavioral Assessment* (Chafouleas, Riley-Tillman, & Sugai, 2007), which provided participants with an introduction to behavioral assessment within the schools and specific instruction on the use of DBR. In contrast, participants in the three experimental sessions (n=43) completed the internet-based DBR training module on individual computers in the laboratory. After 35 minutes had passed in either condition, the researcher instructed participants to again direct their attention to the large screen at the front of the laboratory to complete the post-test. Nine more one-minute clips were shown. After each clip, participants again rated a student in the clip on one of the three target behaviors.

RESULTS AND SUMMARY

The outcome variable of interest was the degree of accuracy with which participants in both conditions rated the target students' behaviors in each clip. Difference scores were computed by subtracting participants' mean DBR data from "expert" DBR (DBR_{exp}) data as well as from systematic direct observation (SDO) data for each clip. Rater error was evaluated by taking the absolute value of these difference scores and comparing them statistically. Overall MANOVA and post-hoc independent samples *t*-tests were conducted to assess for significant differences between the experimental and control groups' ratings.

Post-test. For post-test rater error scores, overall results of repeated measures MANOVA indicated that raters in the experimental group were significantly more accurate than the comparison group when using DBR_{exp} as the criterion ($F[9, 80] = 5.509, p < .001, \text{partial } \eta^2 = .383$) or SDO ($F[9, 80] = 6.281, p < .001, \text{partial } \eta^2 = .414$). *T*-tests comparing the control and experimental groups' post-test accuracy for each clip, using DBR_{exp} as the criterion, indicated ratings in the experimental group were statistically closer to the criterion score for the High Respect (RS) clip at the $< .001$ significance level (Table 1). The control and experimental groups' accuracy scores for the other post-test clips were not significantly different. The control and experimental groups' post-test accuracy was also evaluated using SDO as the criterion. All ratings in the experimental group were closer to the SDO criterion score. Participant ratings for the Med and High RS clips were statistically closer to the SDO criterion score at the $< .0056$ significance level. Overall, the standard deviation of the mean accuracy scores was smaller for the experimental group, indicating less rater variability for participants who were trained using the module. **Summary.** As expected, participants who utilized the online interactive DBR training module evidenced greater accuracy and displayed reduced variability in their ratings of student behavior, suggesting that the completion of an online training module may be a more effective method of training raters in the use of DBR than reading a traditional behavioral assessment text.

Table 1. Post-Test Descriptive Statistics and Independent Samples *T*-tests with Bonferroni Correction

| Behavior | Level | Group | DBR _{exp} - DBR _{ctrl} | | | | DBR _{exp} - SDO | | | |
|---------------------------|-------|--------------|--|------|--------|-------|--------------------------|------|--------|-------|
| | | | M | SD | t | p | M | SD | t | p |
| Academically Engaged (AE) | Low | Comparison | 2.45 | 2.12 | | | 2.13 | 1.85 | | |
| | | Experimental | 1.53 | 1.76 | -2.204 | .030 | 1.59 | 1.34 | -1.629 | .107 |
| | Med | Comparison | 2.62 | 1.81 | | | 2.40 | 1.43 | | |
| | | Experimental | 2.74 | 1.71 | 0.342 | .733 | 2.32 | 1.40 | -0.266 | .791 |
| | High | Comparison | 0.83 | 1.49 | | | 0.89 | 1.35 | | |
| | | Experimental | 0.16 | 0.53 | -2.869 | .006 | 0.32 | 0.43 | -2.786 | .007 |
| Respectful (RS) | Low | Comparison | 1.28 | 1.25 | | | 1.59 | 0.81 | | |
| | | Experimental | 1.07 | 1.64 | 1.301 | .197 | 1.46 | 1.35 | -0.535 | .595 |
| | Med | Comparison | 1.91 | 1.18 | | | 3.82 | 1.43 | | |
| | | Experimental | 2.99 | 1.80 | 0.561 | .576 | 2.92 | 1.56 | -2.860 | .005* |
| | High | Comparison | 2.79 | 2.40 | | | 2.79 | 2.40 | | |
| | | Experimental | 0.60 | 1.47 | -5.262 | .000* | 0.60 | 1.47 | -5.262 | .000* |
| Disruptive (DB) | Low | Comparison | 1.77 | 1.95 | | | 1.87 | 2.02 | | |
| | | Experimental | 1.58 | 1.45 | -0.513 | .610 | 1.55 | 1.47 | -0.884 | .379 |
| | Med | Comparison | 1.91 | 1.35 | | | 2.46 | 1.45 | | |
| | | Experimental | 1.53 | 1.44 | -1.294 | .199 | 1.70 | 1.41 | -2.59 | .011 |
| | High | Comparison | 1.04 | 1.38 | | | 1.20 | 1.23 | | |
| | | Experimental | 1.16 | 1.56 | 0.388 | .699 | 1.20 | 1.34 | -0.016 | .987 |

* $p < .00556$