Project VIABLE: Overview of Work Related to Development and Evaluation of Direct Behavior Rating Single Item Scales (DBR-SIS)

Sandra M. Chafouleas[1], T. Chris Riley-Tillman[2], Theodore J. Christ[3], & George Sugai[1]

University of Connecticut[1]   East Carolina University[2]   University of Minnesota[3]

University of Connecticut
East Carolina University
UNIVERSITY OF MINNESOTA

PROJECT GOAL:
Develop and Evaluate Direct Behavior Rating (DBR)

Phases I & II: Develop instrumentation and procedures; evaluate defensibility of DBR in decision-making
• Large datasets; repeated observations of student behavior
• Understanding critical factors (e.g. scale format, behavior targets, training requirements)
• Pilot testing various aspects with classroom teachers

Phase III: Evaluate feasibility and utility of DBR in school settings at small scale
• Packaging what we have learned to train users
• Establish groups of teachers/schools willing to participate in DBR training and use
• Evaluate data/feedback

Defining Direct Behavior Rating

Direct – Rating occurs in close proximity to the time and place of the observation. Thus, the rater must observe the target for a "sufficient" portion of the observation period.
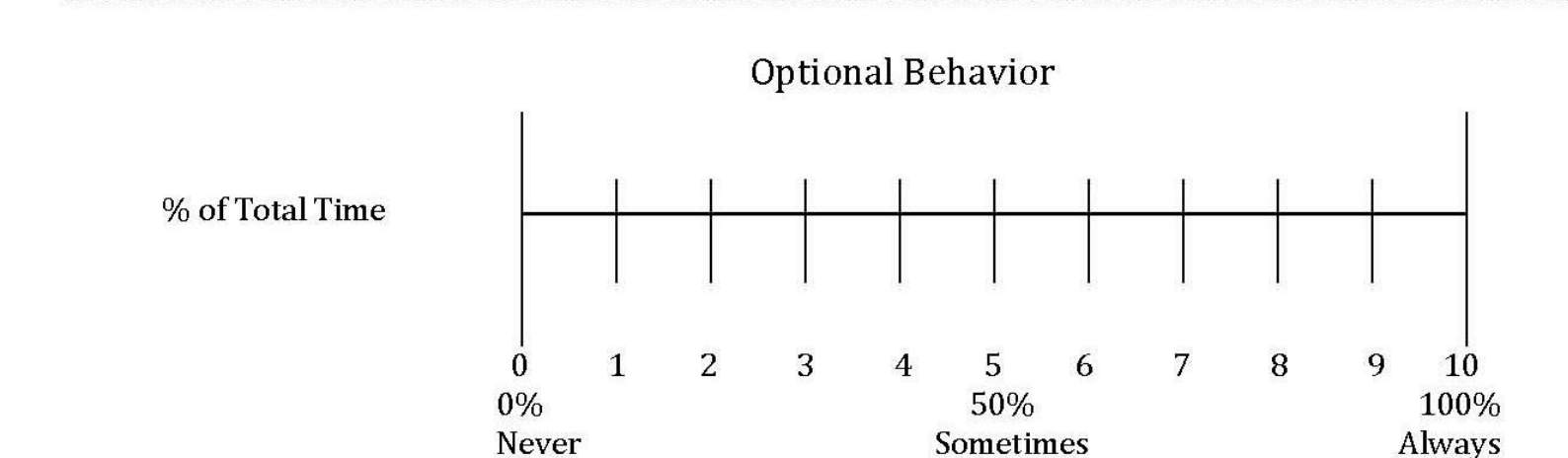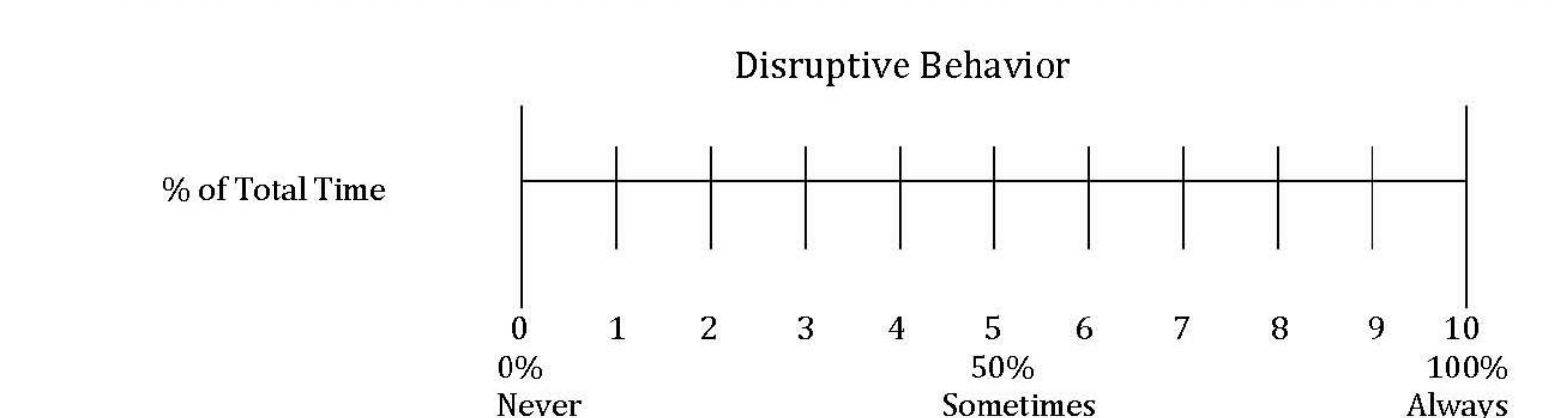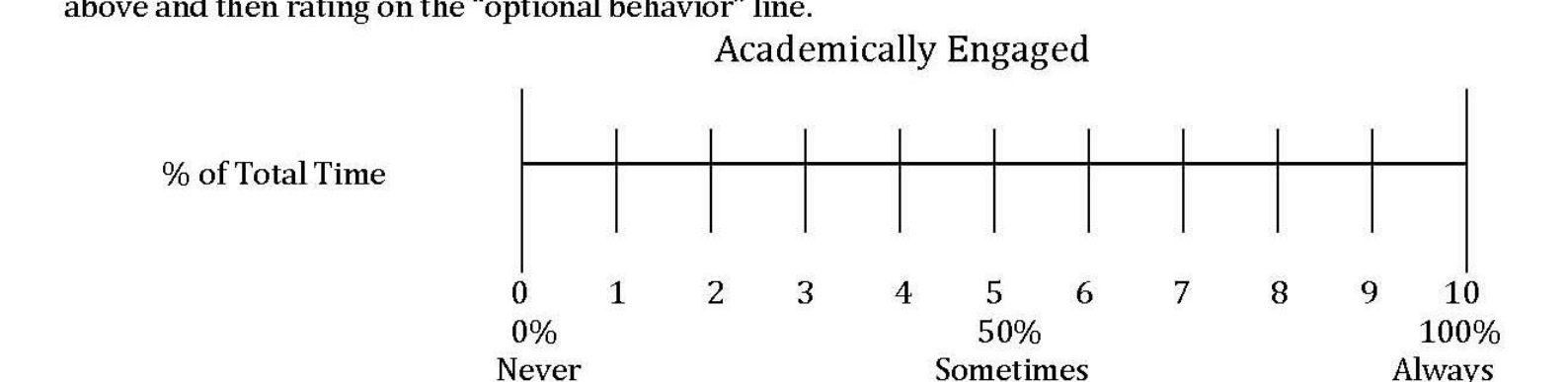*Behavior* – The target of rating must be well-defined and accessible for observation.
Rating – The rating component quantifies rater perception of the target behavior.
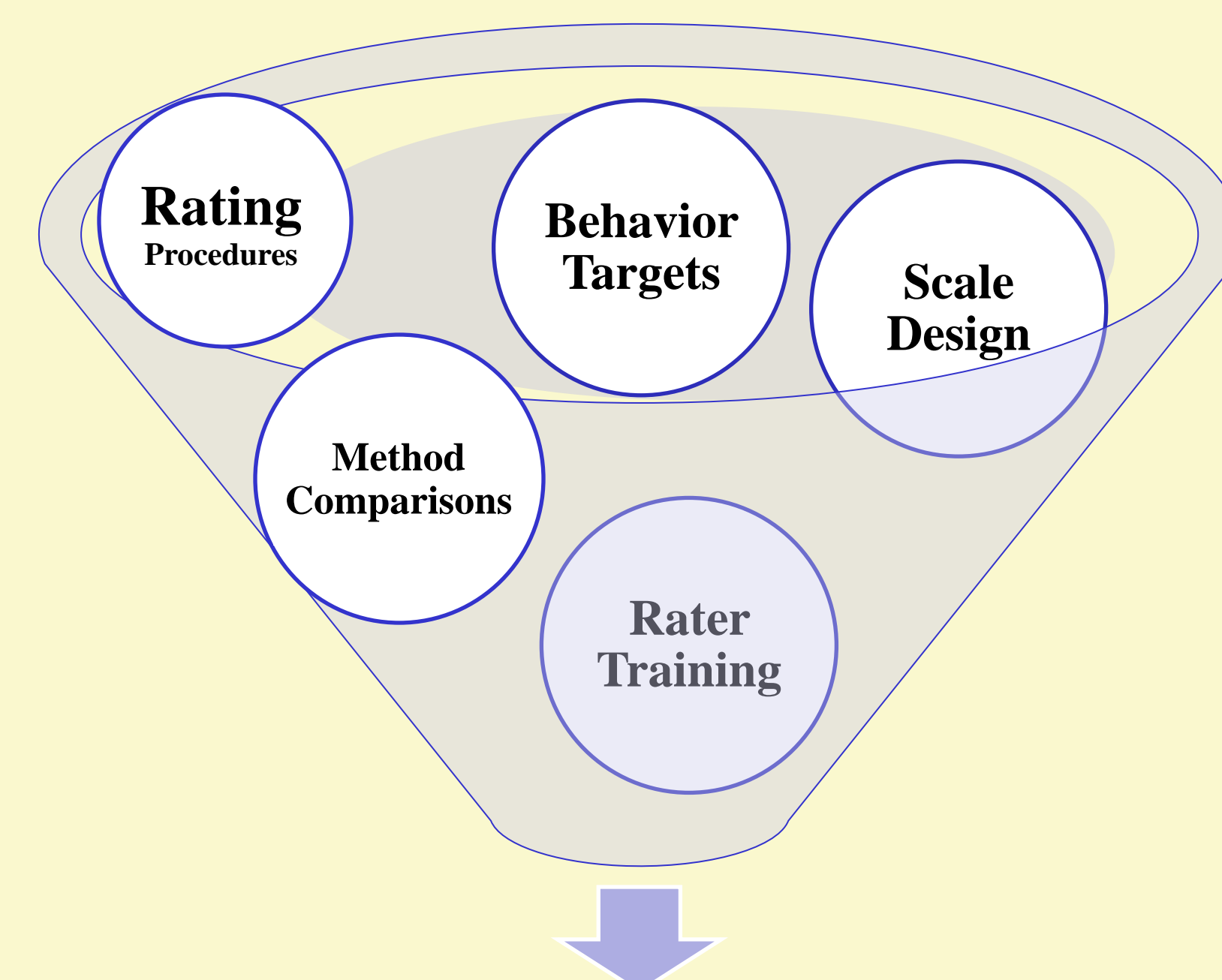
Example DBR-SIS Form

PROCESS OVERVIEW

Development through Multiple "Silos"…
then Initial Evaluation on Small Scale

Rating Procedures
Behavior Targets
Scale Design
Method Comparisons
Rater Training

DBR is a Defensible and Usable Method
in School-based Assessment

Comparisons across Methods and Raters

A Few Examples:

Briesch, Chafouleas, & Riley-Tillman (in press)
Sample: 2 teachers in a full day inclusive K classroom, 14 students
Measures: researcher-completed SDO, teacher-completed DBR-SIS of Academic Engagement
Analyses: Generalizability Theory
Conclusion: Both methods were equally sensitive to intra-individual differences in academic engagement however, differences were noted with regard to the influences of both rater and time. SDO rating variance was explained by changes in student behavior across days and rating occasions, whereas rater-related effects accounted for the greatest proportion of DBR variance.

Chafouleas, Briesch, Riley-Tillman, Christ, Black, & Kilgus (2010)
Sample: 2 teachers and 2 research assistants – 7 middle school students in the same Language Arts classroom
Measures: researcher-completed and teacher-completed DBR-SIS for Academic Engagement and Disruptive Behavior over 6 days (3x/period)
Analyses : Multiple imputation to handle substantial missing data, Generalizability theory
Conclusion: Degree of reliability-like estimates can differ substantially depending on individual rater. In the absence of estimates of rater reliability and firm recommendations regarding rater training, ratings obtained from DBR-SIS, and subsequent analyses, be conducted within rater.

Chafouleas, Hagermoser-Sanetti, Kilgus, & Maggin (in prep)
Sample: 20 teacher-student dyads in elementary grades
Design and Intervention: A-B intervention involving behavioral consultation and DRC-based intervention. Five options for "change metrics" were calculated.
Measures: researcher-completed SDO, teacher-completed DBR-SIS
Conclusion: Change (in expected directions) in student behavior across phases and sources. High correspondence between DBR-SIS and BOSS absolute change metrics suggests that students were ranked similarly across the two measures with regard to intervention responsiveness. Provides preliminary support for the use of DBR-SIS to differentiate between those who have or have not responded to intervention.
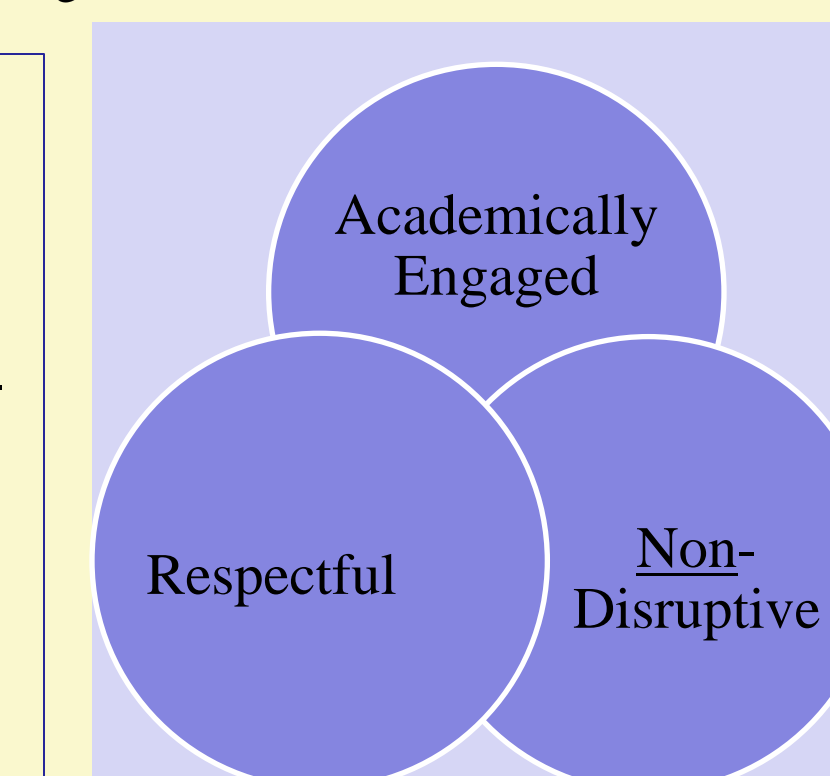
Behavior Targets

Considerations:
Molar v. molecular wording? - E.g. Disruptive – Out of Seat
Negative v. positive wording? - E.g. Disrespectful – Respectful
General Outcome v. Individualized Targets - Applicable to all – Relevant to Some/Few

Riley-Tillman, Chafouleas, Christ, Briesch, & LeBel (2009)
The first attempt… DBR vs. SDO comparisons of 3 target constructs and 2 wording. Wording and specificity of target construct can impact rater accuracy. Molar wording resulted in stronger correspondence and positive phrasing was stronger for academic engagement yet unclear for disruptive. "Compliance" definition needed revision.

Christ, Riley-Tillman, Chafouleas, & Jaffery (in review)
Adding on… analyses to separate rater bias and error, and influence of base rates. High correspondence between DBR and SDO for Academic Engagement and Disruptive Behavior, but results for molecular behaviors were weak. Substantial rater bias was present (underestimate desirable and vice versa).

Chafouleas, Riley-Tillman, Jaffery, Sen, Music, & Christ (2010)
And adding further… only molar behaviors of academic engagement, disruptive, and respectful. Comparisons with SDO and DBR-Expert and controlled the clips (base rates). DBR-Expert resulted in closer correspondence than SDO, Stronger evidence for Academic Engagement and Disruptive than Respectful, Medium levels of behavior harder to rate than low and high

"The BIG 3" General Outcomes to Evaluate in DBR-SIS…
yet still possible to flexibly select targets

Academically Engaged
Respectful
Non-Disruptive

Rater Training

Considerations:
What level of accuracy might be expected in the absence of training?
Are some behaviors more difficult to rate accurately?
What improvement might be expected given training involving…
Information about DBR, Information about Rater Bias, Modeling, Demonstration, Performance feedback?

Schlientz, Riley-Tillman, Briesch, Walcott, & Chafouleas (2008)
A single training session involving practice and feedback resulted in greater accuracy compared to a brief familiarization session

Harrison & Riley-Tillman (2010)
Adding on… initial comparison of behaviors and base rates
Training with **practice and feedback** resulted in improved accuracy for rating disruptive behavior, and higher for disruptive and compliance when base rates low or high.

Chafouleas, Kilgus, Riley-Tillman, & Jaffery (2010)
Adding on… impact of Frame of Reference and Rater Error Training added … control of base rates of behavior and varied "exposure" to performance feedback
"**Exposure**" mattered for some clips… thus, "Standard Training" should suffice as long as sufficient opportunities for practice and feedback are provided.

An Efficient On-Line Training with 3 Modules:
(a) overview, (b) modeling, & (c) practice/feedback

Direct Behavior Ratings
TRAINING SITE

Scale Design

Considerations:
Number of Gradients?
Anchors?
Qualitative Descriptors?
Visual Cue?

Christ & Boice (2009); Christ, Riley-Tillman, & Chafouleas (2009)
Scales should be comprised of at least 6 gradients yet 10 appears optimal to facilitate ease of scale interpretation and utility within visual analysis of formative data.
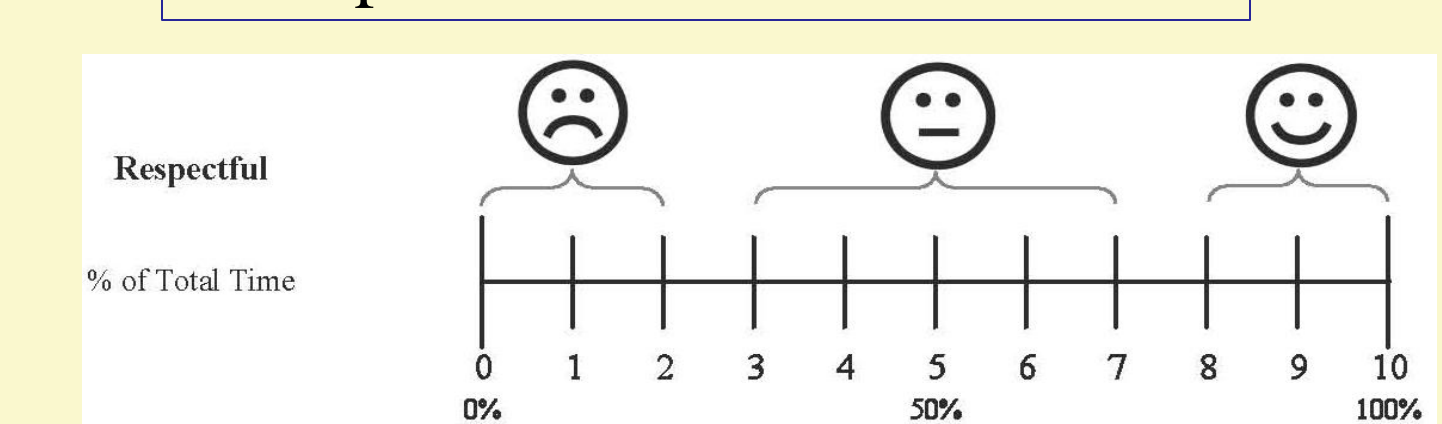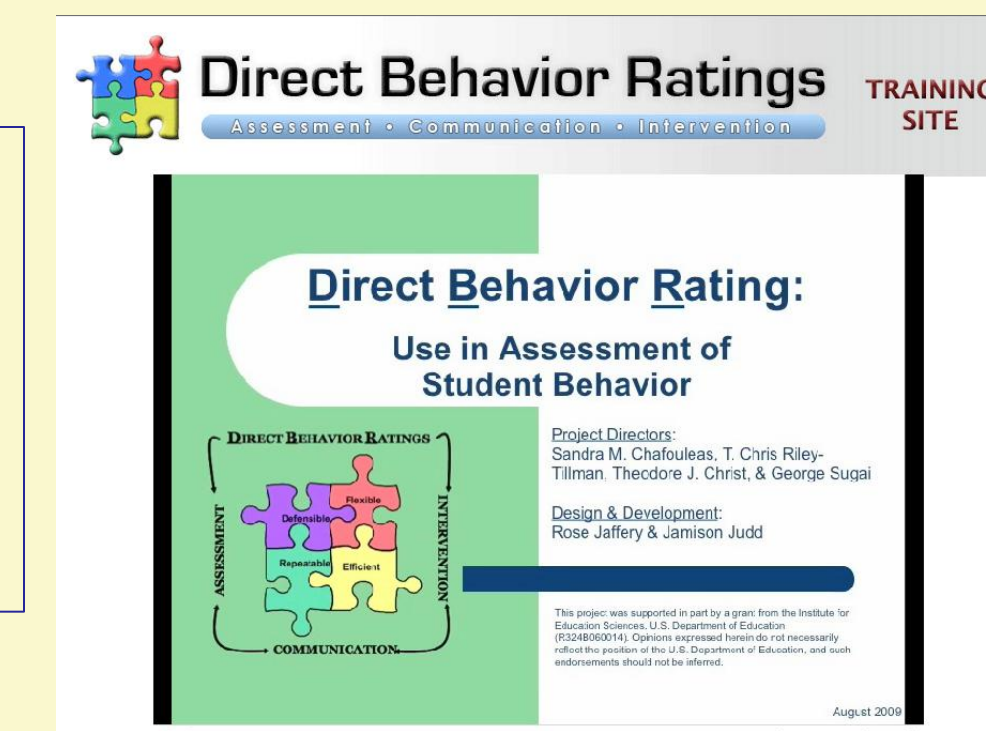
Briesch, Kilgus, Chafouleas Riley-Tillman, & Christ (2010); Christ & Boice (2009)
Scales can use a variety of physical options. A line can be used to provide a visual cue toward rating, although the total length of the line does not impact reliability or accuracy.

Riley-Tillman, Christ, Chafouleas, Boice, & Briesch (2009); Riley-Tillman, Chafouleas, & Music (2009)
Scales may vary with regard to WHAT is rated (duration, proportion), and no strong preferred design has emerged among teachers

Example Format for DBR-SIS

Respectful
% of Total Time
0   1 2 3 4 5 6 7 8 9 10
0%   50%   100%

SUMMARY POINTS:

SUMMARY
• DBR is a viable new method for behavior assessment with utility in problem-solving models (emphasis on screening and progress monitoring)
• Work to date has focused on development related to DBR-SIS, along with initial evaluations at smaller scale
• DBR-SIS offers a format with defining characteristics that include defensible, flexible, efficient, and repeatable
• DBR-SIS may be useful in assessment across all "tiers" in multi-tiered problem-solving frameworks – the general outcomes are applicable to all and additional targets may be selected as needed for idiographic assessment

NEXT STEPS
• Understand Variability Across Time and Grade
• Understand Risk "Cut-points"
• Enhanced Efficiency of Repeated Measurement - Web-Based Application
• Full Evaluation of Data Use and Interpretation – Schools, Teachers, Students at Scale

DIRECT BEHAVIOR RATINGS
ASSESSMENT   INTERVENTION
Flexible   Defensible   Repeatable   Efficient
COMMUNICATION
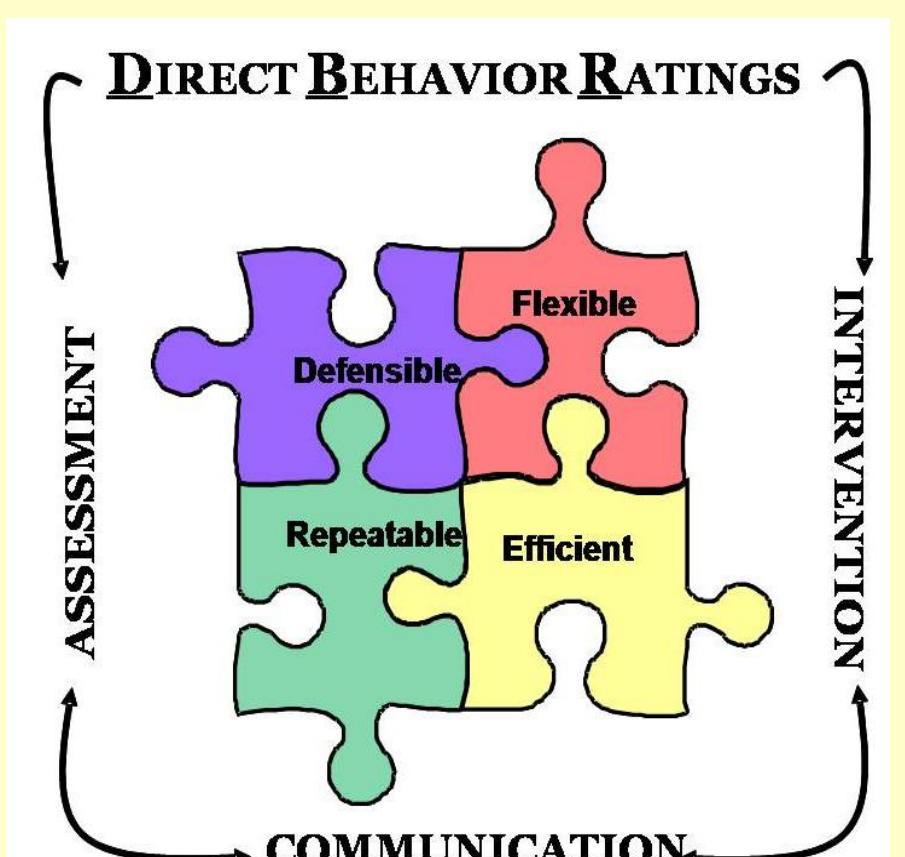
For additional information, please visit www.directbehaviorrating.org
Email correspondence regarding the project should be directed to Sandra Chafouleas at sandra.chafouleas@uconn.edu.