

5-2020

## Test-Driving Interventions for Teachers to Improve Treatment Integrity

Viktoría Papp

Follow this and additional works at: [https://aquila.usm.edu/honors\\_theses](https://aquila.usm.edu/honors_theses)



Part of the [School Psychology Commons](#)

---

### Recommended Citation

Papp, Viktoria, "Test-Driving Interventions for Teachers to Improve Treatment Integrity" (2020). *Honors Theses*. 732.

[https://aquila.usm.edu/honors\\_theses/732](https://aquila.usm.edu/honors_theses/732)

This Honors College Thesis is brought to you for free and open access by the Honors College at The Aquila Digital Community. It has been accepted for inclusion in Honors Theses by an authorized administrator of The Aquila Digital Community. For more information, please contact [Joshua.Cromwell@usm.edu](mailto:Joshua.Cromwell@usm.edu), [Jennie.Vance@usm.edu](mailto:Jennie.Vance@usm.edu).

The University of Southern Mississippi

Test-Driving Interventions for Teachers to Improve Treatment Integrity

by

Viktoría Papp

A Thesis  
Submitted to the Honors College of  
The University of Southern Mississippi  
in Partial Fulfillment  
of Honors Requirements

May 2020



Approved by:

A handwritten signature in black ink, appearing to read "Joe", with a small dot above the final letter.

---

D. Joe Olmi, Ph.D., Thesis Adviser  
Professor of Psychology

---

Sara Jordan, Ph.D., Director  
School of Psychology

---

Ellen Weinauer, Ph.D., Dean  
Honors College

## **Abstract**

Teachers often have to implement interventions in the classroom to address challenging student behaviors. Such issues are often addressed during behavioral consultation which involves working with a specialized consultant, such as a school psychologist, to determine the most robust strategy that will predictably have the greatest impact in addressing the problematic behavior. Equally important in this intervention process is the application of any intervention technique or strategy in the manner in which it has been designed and investigated to work based on supporting research.

There is an increasing need for consultants to consider utilizing strategies that promote the adherence to treatment procedures. Treatment integrity is the adherence to accurate intervention implementation. TI is gaining recognition with regards to its influence on intervention effectiveness and student outcomes. These methods can be consequent or antecedent methods, and the teacher plays a critical role in the selection of either. Consequent techniques for TI improvement are only utilized after the teacher demonstrated poor adherence to implementation. In cases like this, intervention effectiveness and outcomes could be compromised before the teacher receives feedback and training to correct inaccuracies in implementation, and the intervention can occur as intended. Therefore, it is important to explore antecedent options. In contrast to consequent methods, antecedent techniques for TI improvement aim to forego insufficient implementation, and are used in order to prevent poor TI levels from occurring. Thus, antecedent methods can produce more effective and efficient intervention outcomes. Test-driving is an antecedent method that provides teachers the initial opportunity to assess the acceptability of different treatment procedures, and its

goal is to promote high levels of treatment integrity from the beginning of the implementation. The purpose of this study is to evaluate the effectiveness of test-driving interventions, and the role it plays in intervention integrity.

Keywords: treatment integrity, test-driving, antecedent methods, intervention implementation, behavioral consultation, school psychology

## **Acknowledgements**

I would like to thank Stefanie Schrieber, graduate student in the School Psychology Program at The University of Southern Mississippi, for her assistance and for her allowing me to participate in this project. My experiences have been invaluable. My exposure to information pertaining to single-case design, intervention design within school settings, issues associated with treatment integrity, and school-based consultation will help prepare me for my graduate program in Marriage and Family Therapy, which I will begin this upcoming fall.

First, I would like to thank Dr. D. Joe Olmi for guiding me through this process. It was a pleasure working together on my thesis project, and I greatly appreciate his help and support. I am also thankful to Dr. Lauren McKinley for getting me started on this journey and for assisting me while I was only beginning to explore my interests and possible research topics. I would like to thank Stefanie Schrieber for letting me join her dissertation project and for allowing me to gain real research experience. Last but not least, I would like to say thank you to Dr. Jonathan Barron for encouraging me to apply to the Honors College; it has been an adventure.





# Table of Contents

List of Figures .....	x
List of Abbreviations .....	xi
Chapter 1: Introduction .....	1
Treatment Integrity .....	2
Methods for Improving TI .....	3
Test-Driving Interventions.....	4
Purpose.....	5
Chapter 2: Methods.....	7
Setting and Participants.....	7
Design and Procedures.....	8
Dependent Measures .....	12
Observer Training and Interobserver Agreement (IOA) .....	13
Chapter 3: Results .....	14
Data Collection and Analysis.....	14
Results to Date and Anticipated Results.....	16
Chapter 4: Discussion .....	19
Implications.....	19
Limitations .....	21
References.....	24
Appendices.....	27
Appendix A: IRB Approval .....	27
Appendix B: Treatment Integrity Forms .....	28

Appendix C: Observation Form.....	29
Appendix D: Procedural Integrity Checklist.....	30
Appendix E: Treatment Overviews .....	31
Appendix F: URP-IR (Adapted).....	33
Appendix G: URP-A .....	34
Appendix H: Operational Definitions.....	36

## List of Figures

Figure 1 .....	14
Figure 2 .....	15
Figure 3 .....	15
Figure 4 .....	16

## **List of Abbreviations**

AEB	Academically Engaged Behavior
BC	Behavioral Consultation
BST	Behavior Skills Training
CICO	Check-In/Check-Out
DB	Disruptive Behavior
NCR	Non-Contingent Reinforcement
POT	Passive Off-Task
SDO	Systematic Direct Observation
TI	Treatment Integrity
TIPP	Treatment Integrity Planning Protocol
URP-A	Usage Rating Profile – Assessment
URP-IR	Usage Rating Profile – Intervention, Revised
WWC	What Works Clearinghouse

## **Chapter 1: Introduction**

In the field of psychology, professionals such as school psychologists often assist teachers with designing and implementing methods or strategies to improve student behavior using various interventions in the classroom. The interpretation of intervention effectiveness is beginning to change as the body of research increases regarding other dimensions of treatment implementation and related concerns. In most instances, teachers only receive indirect instructions regarding implementation of classroom-based interventions (Fallon et al., 2018) and have little say-so as to which interventions are ultimately selected for implementation, which may lead to lower levels of treatment integrity or implementation of the intervention as designed.

Treatment integrity (TI) is the extent to which an intervention is applied as intended or as designed (Gresham, 1989). Intervention effectiveness is highly dependent on treatment integrity (Noell, Gresham, & Gansle, 2002); therefore the level of adherence to the components of the intervention is important information to be considered while making decisions and adjustments to interventions based on effectiveness. According to Fallon and colleagues (2018), it is crucial for professionals to receive and evaluate accurate reports of treatment integrity in order to make appropriate treatment decisions. TI data are required to analyze factors of intervention delivery, such as the degree of application of the independent variable and the relationships between the interventions and outcomes (Fallon et al., 2018).

There is an increasing need for focusing on TI and how to improve it. In school settings, it is critical that teachers receive training that promotes adherence to interventions as designed in order to ensure successful implementation (Sanetti &

Kratochwill, 2009). This study will explore test-driving interventions as a means of determining intervention preference, which then may have an effect on TI.

### **Treatment Integrity**

As mentioned above, TI is the extent that interventions are implemented in accordance with established implementation guidelines (Gresham, 1989). One way to assess TI is to create a checklist of the components of the intervention and calculate the percentage of adherence to those components by dividing the number of correctly implemented steps by the number of total steps involved in implementing the intervention. Systematic direct observation (SDO) is another method used to assess TI by having an observer present with the teacher implementing the treatment and having that observer record the implemented steps on the checklist (Lane et al., 2004). After that, the observer converts the results to a percentage and determines the level of adherence to the intervention.

Other methods to assess TI are self-report and permanent products. Teacher self-report is the most commonly used method in school settings. Since it does not require a professional to be present in the classroom during implementation, it is less time and resource intensive than SDO (Fallon et al., 2018). However, it is often less accurate and reliable and often results in an overestimate of TI. Fallon and colleagues (2018) offered three possible explanations why teachers tend to overestimate their adherence to the implementation procedures. One of these was that self-reporting is inaccurate due to the need to remember one's activities, which could be influenced by memory limitations. Another possibility was that self-reporting practices might be impacted by a social desirability response bias, which the researchers explained as the tendency to make

oneself appear in the most favorable way. The final explanation was the possibility that teachers lack the skills and understanding of the intervention and reporting processes, which could lead to inaccurate self-report ratings. Both self-report and permanent products were found to be less reliable methods of TI measurement than SDO (Gresham, Dart & Collins, 2017). Since TI is the dependent variable in this study and since SDO is the gold standard of behavioral assessment, SDO was used.

### **Methods for Improving TI**

Strategies that are used to enhance TI are usually consequent methods that focus on improving TI once an intervention has been implemented incorrectly, meaning that consequent strategies respond to low TI once it has been demonstrated and identified as a problem. Then, the teacher receives feedback and additional training accordingly.

Implementation planning, treatment integrity planning protocols (TIPP), performance feedback, self-monitoring, participant modeling and role-play, and motivational interviewing are all methods that are utilized by a consultant to correct poor TI (Noell, Witt, Gilbertson, Ranier, & Freeland, 1997; Sanetti & Collier-Meek, 2015; Sanetti, Collier-Meek, Long, Byron, & Kratochwill, 2015; Sanetti & Kratochwill, 2009).

In contrast, the goal of antecedent methods is to ensure that interventions are implemented with high levels of TI from the beginning, without waiting for treatment implementation errors that require correction. Antecedent strategies are used prior to intervention implementation to prevent low levels of TI. However, antecedent methods are much less commonly utilized by consultants. The two methods that fall in this category are direct training and test-driving interventions. For the purposes of this study,

the effects of test-driving an intervention as an antecedent method for promoting high TI levels will be assessed and evaluated.

### **Test-Driving Interventions**

There is limited, if any, substantive research investigating test-driving an intervention as a means of enhancing TI as it has been examined in only one study thus far (Dart et al., 2012). The purpose of this current study is to replicate and further examine the effects of test-driving interventions as a means of increasing and maintaining high levels of TI. Test-driving allows teachers to test, rate, and then choose a preferred intervention prior to the final selection and implementation of that particular intervention. Having the opportunity to test multiple interventions allows one to obtain a clear picture of what each intervention procedure entails and gives one an opportunity to rate the acceptability of those various interventions based on real direct experiences with those procedures.

Dart and colleagues (2012) provided teachers the opportunity to test-drive four interventions: self-monitoring, modified Check-in/Check-out, response cost, and behavior specific praise. After the test-driving phase teachers ranked the interventions based on acceptability from most to least favorite. The authors found that teachers demonstrated higher TI levels when asked to re-implement their highest ranked intervention. TI levels significantly increased during the implementation of the preferred intervention compared to the other interventions. Student outcomes also improved during the preferred interventions.



## **Purpose**

As mentioned above, TI is an essential component in intervention effectiveness and should be treated as such when evaluating treatment results. However, it is largely underrepresented in the literature. Intervention effectiveness is often defined by student outcomes, while TI and the adherence to implementation plans should also be considered in determining intervention effectiveness.

As noted previously, consequent methods to promote TI are more commonly used than antecedent methods. The limitation of consequent methods is that they allow teachers to demonstrate low levels of TI. The goal of antecedent methods is to achieve high TI levels before poor implementation occurs. The only two antecedent methods are direct training and test-driving. The effects of test-driving an intervention by the teacher and assessing its effect on TI requires further examination. So far, test-driving is the only antecedent strategy that involves direct teacher experience with implementing an intervention and thus helps teachers evaluate their experiences with that intervention before selecting or choosing it as the behavior change procedure. The purpose of the study is to further evaluate the effects of test-driving interventions as related to TI.

### **Research Questions.**

1. Will teachers who previously implemented interventions with 50% or lower treatment integrity show improvements in treatment integrity when utilizing teacher-choice interventions based on test-driving?
2. Will there be collateral improvements in student academically engaged behavior when the teacher is utilizing teacher-choice interventions based on test-driving?

3. Will teachers' ratings on the URP-IR and URP-A be highest for teacher-choice interventions with respect to dimensions of social validity, treatment acceptability, feasibility, and understanding?

## Chapter 2: Methods

### Setting and Participants

The study was conducted in the Southeastern United States, in a rural high school. Student enrollment of the school is 569 students, of which 47.8% are female and 52.2% are male. Twenty-six percent of the school's students are African American, 68.72% are white, and 2.99% are two or more races. Approximately 69.6% of the students are part of the free/reduced lunch program. After teacher referrals were received by the school's behavior consultants regarding student behavior concerns, four teacher-student dyads were selected. The Institutional Review Board (IRB) granted approval for the conduct of the study (see Appendix A). The school administration, participating teachers, and students' parents/guardians all consented to the project. The primary researcher interviewed the participating teachers, identified the students' problem behaviors, and collected basic demographic information relative to the teachers and students. For the purpose of the study, low levels of TI were defined as 50% or less adherence to the intervention components. Teachers who demonstrated low TI levels based on this definition consented to participate in the study.

Teacher A was a white female with fourteen years of teaching experience. She taught the third period biology class with ten male students and nine female students. Student A in Teacher A's classroom was a black male in the ninth grade who was referred to behavior consultants for behavior problems such as talking when not permitted and sleeping in class. Teacher B was a white female with two years of teaching experience. She taught the first period English class of 12 male and 12 female students. Student B in Teacher B's class was a black male in the ninth grade referred for talking

when not permitted and not completing schoolwork. Teacher C was a white female with nine years of teaching experience. She taught the second period English class with six male and eleven female students. Student C in Teacher C's class was a white female in the ninth grade, referred for talking when not permitted, and different disruptive out of seat behaviors such as dancing and walking around the classroom. Teacher D was a white female with six years of teaching experience. She taught second period algebra in a class of eight male and two female students. Student D in Teacher D's class was a white male in the ninth grade, referred for sleeping and talking.

### **Design and Procedures**

A multiple baseline across four teacher-student dyads design was used during the study. To eliminate threats to interval validity, teacher-student dyad two was yoked to dyad one, dyad three was yoked to dyad two, and dyad four was yoked to dyad three, therefore control participants were utilized. Data were staggered for a minimum of two sessions across participants. All phases except for the test-drive phase met What Works Clearinghouse (WWC) standards for multiple baseline design including: (a) a minimum of five data points per phase, (b) a minimum of 80% IOA for at least 20% of observations across participants and phases, (c) a minimum of three replications, and (d) systematic manipulation of the independent variable (Kratochwill et al., 2010). The test-drive phase only included two to three data points instead of five due to the nature of the intervention.

TI and student behavior data were collected during 20-minute sessions using SDO. Within five minutes of entering the room as the implementation period started, observers began monitoring and recording data while standing in the back of the classroom. Observers used different checklists for each intervention to monitor and assess

TI (see Appendix B). Student behavior (AEB, DB, and POT) were measured using 10-second momentary time sampling (see Appendix C). At the beginning of each interval, the behavior of the target student was recorded on the Observation Form. Examples of academically engaged behavior (AEB) included writing, reading the assigned material, raising hand, talking to the teacher or peers about the assigned material when permitted, and listening to a lecture or teacher instruction. Disruptive behaviors (DB) included making audible sounds, talking to peers or teacher when not permitted, talking about topics unrelated to the assigned task, out-of-seat behavior, manipulating objects, drawing or writing unrelated to the assigned task. Passive off-task (POT) behaviors included looking around the room or out the window, sleeping, listening to peers talk about issues unrelated to the assigned task, and sitting quietly not engaging in any assigned activity. After the observation sessions, observers calculated the percentage of the target student's AEB, DB, and POT behaviors, as well as the percentage of TI.

**Interventions.** Four interventions were used in the study: self-monitoring, modified Check-in/Check-out (CICO), response cost, and non-contingent reinforcement (NCR). Self-monitoring required both the student and the teacher to record student behavior in four-minute fixed intervals that were signaled by a MotivAider®. After the student received brief instructions for the intervention, they started their timers. The MotivAider® vibrated in set intervals, and the student and the teacher marked on or off-task behavior on the self-monitoring form. At the end of the session, an honesty check was conducted, and their ratings were compared. Reinforcement was given to the student if at least 80% accuracy was demonstrated on the self-monitoring form.

Modified Check-in/Check-out (CICO) involved the teacher checking in with the student at the beginning of the observation session. During this time, the teacher provided encouragement, three behavior expectations, and a behavior goal. The teacher conducted a direct behavior rating at the end of the session (DBR; Gresham, 2010; Kilgus, Riley-Tillman, Chafouleas, Christ & Welsh, 2014), and checked out with the student. Check-out included providing encouragement and corrective feedback. Additionally, the student received reinforcement if at least 80% of the points was achieved on the DBR for on-task behavior.

Response cost required the teacher to record disruptive behavior by marking an X in one of the five boxes on the form. The student could earn reinforcement if at least 80% of the boxes were empty at the end of the session. Non-Contingent Reinforcement (NCR) involved the teacher providing behavior specific praise to the target student every four minutes, while ignoring minor disruptive behaviors. No additional reinforcement was provided in this procedure.

**Procedures.** There were four different phases in the study: treatment-as-usual, test-drive, preferred intervention, and yoked intervention. In the treatment-as-usual phase, the behavioral consultant selected and assigned the most appropriate behavioral intervention from the list above, based on the problems and target behaviors identified following referrals. Teachers were trained on the procedures using behavioral skills training (BST), in which the teacher was provided information on the intervention, the consultant modeled the intervention, and then the teacher rehearsed the procedures and received corrective feedback and praise until demonstrating 100% accuracy in implementation. The teachers implemented their intervention for at least five sessions,

during which TI and student behavior data were collected utilizing SDO. Teachers demonstrating 50% or lower TI levels across the phase were included in the study.

During the test-drive phase teachers had to demonstrate 100% TI for each intervention. Before test-driving, teachers received BST on intervention procedures. TI data were collected using procedural integrity checklists (see Appendix D). After demonstrating 100% adherence to procedures on the first implementation of each intervention, teachers accessed a small reward previously determined by preference assessment, and rated interventions on the URP-IR. Then, they ranked interventions from most to least preferred. In the preferred intervention phase, teachers implemented their most preferred intervention for a minimum of five sessions, without receiving feedback on their implementation. During the yoked intervention phase, teacher-student dyads implemented the other teachers' preferred interventions for a minimum of five data points.

**Materials.** Four Treatment Overviews were used to introduce the intervention procedures (see Appendix E). Treatment Overviews included brief instructions for all of the interventions (i.e. self-monitoring, response cost, modified Check-in/Check-out, and non-contingent reinforcement). Procedural Integrity Checklists were utilized for each intervention to record the steps implemented in the test-drive sessions (see Appendix D). The checklists consisted of step-by-step breakdowns for all interventions. Similarly, Treatment Integrity Forms included intervention steps and were used to monitor and record steps correctly implemented by the teachers (see Appendix B). The Observation Form included all of the intervals during a 20-minute observation session and was used to record students' academically engaged behaviors, disruptive behaviors, and passive off-

task behaviors at the beginning of each 10-second interval (see Appendix C). The Usage Rating Profile – Intervention, Revised (URP-IR; Briesch, Chafouleas, Neugebauer, & Riley-Tillman, 2013) was used to measure the interventions’ social validity. Teachers completed items related to Acceptability, Understanding, and Feasibility on the URP-IR rating scale, marking items between 1 and 6 from strongly disagree to strongly agree (see Appendix F). Additionally, the Usage Rating Profile – Assessment rating scale was used to allow teachers to evaluate the acceptability of the interventions (URP-A; Chafouleas, Miller, Briesch, Neugebauer, & Riley-Tillman, 2012). Teachers completed all items on this rating scale without exclusions, and rated items on a Likert scale between 1 and 6 from strongly disagree to strongly agree (see Appendix G).

Student rewards were determined by preference assessments before the implementation of interventions began in the treatment-as-usual phase, and these items included candy, snacks, t-shirts, bonus points, homework passes, and free time. Additionally, MotivAider® devices were set to four-minute intervals during self-monitoring procedures, and were used to signal fixed intervals by vibration. The buzzing was a cue for the student and the teacher to record the student’s momentary on-task or off-task behavior.

### **Dependent Measures**

The primary dependent variable was TI, which is the degree of adherence to intervention components. To measure TI, checklists were used and implementation was evaluated based on the treatment steps implemented as designed. As mentioned above, TI was determined by calculating the percentage of procedure components implemented correctly. Secondary dependent variables in the study are students’ academically engaged



behavior (AEB), disruptive behavior (DB), and passive-off task behavior (POT), as defined above (see Appendix H).

### **Observer Training and Interobserver Agreement (IOA)**

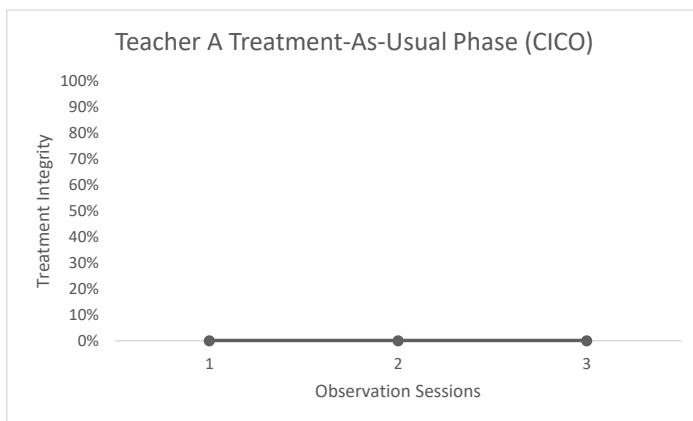
The training for secondary observers required the demonstration of at least 90% interobserver agreement (IOA) on a 10-minute video sample used for that purpose. Secondary observers were trained on dependent and independent variables and data collection procedures. Operational definitions of student behavior were provided (see Appendix H). IOA data were collected at least 20% of observations for all dyads in all phases, and was calculated using component by component (TI) and interval by interval (student behavior) methods. Both are calculated by dividing the number of components or intervals agreed upon by the number of total possible components or intervals, and converting the value to a percentage. Throughout the observations, secondary observers had to maintain a minimum of 90% agreement in adherence to graduate program requirements, or they would be retrained using the same methods until requirements were met before conducting other observations.

## Chapter 3: Results

### Data Collection and Analysis

Treatment integrity data were collected for four different teacher-student dyads and four different interventions in the treatment-as-usual phase. Data collection was abruptly interrupted by the emergence of the global pandemic COVID-19. Due to the situation, no further data are available at this point.

Teacher A implemented modified Check-in/Check-out in the treatment-as-usual phase. Throughout all of the three sessions observed, she maintained 0% treatment integrity (see Figure 1) indicating she failed to implement the treatment with any degree of integrity.



*Figure 1. Teacher A's treatment integrity data of implementing modified Check-in/Check-out in the treatment-as-usual phase.*

Teacher B implemented self-monitoring in the treatment-as-usual phase, she was observed for five sessions. She began with higher TI levels of about 75%; however, later her TI decreased significantly and stabilized around 25% (see Figure 2) suggesting that her TI levels fell to unacceptable levels following traditional consultation efforts.

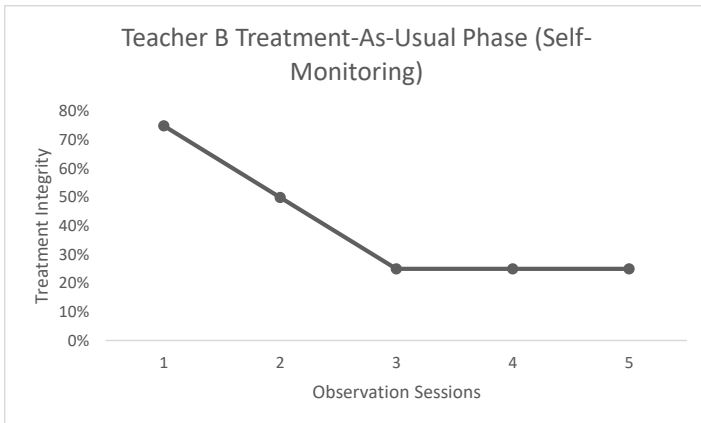


Figure 2. Teacher B's treatment integrity data of implementing self-monitoring in the treatment-as-usual phase.

Teacher C implemented behavior specific praise in the treatment-as-usual-phase. TI data were recorded for six sessions and were varied between 40% and 80% (see Figure 3). TI levels were variable.

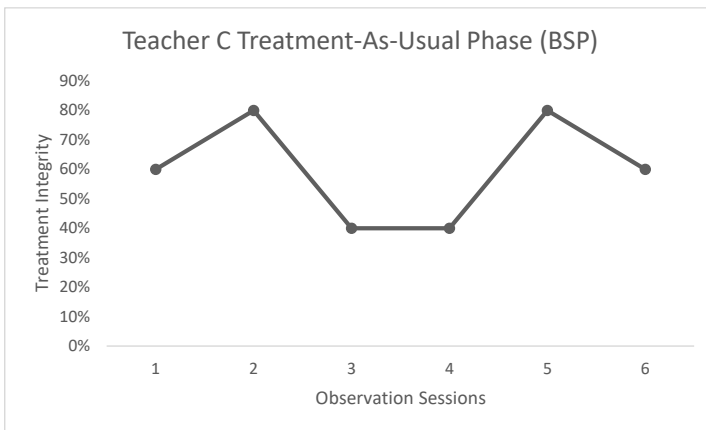
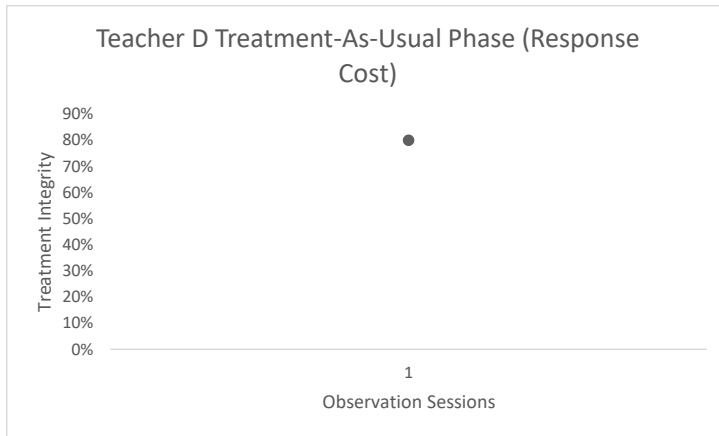


Figure 3. Teacher C's treatment integrity data of implementing behavior specific praise in the treatment-as-usual phase.

Teacher D implemented response cost in the treatment-as-usual phase, and only one session was observed. TI was 80% (see Figure 4). Summary statements cannot be made given the limited data.



*Figure 4. Teacher D's treatment integrity data of implementing response cost in the treatment-as-usual phase.*

## **Results to Date and Anticipated Results**

**Research Question 1:** Will teachers who previously implemented interventions with 50% or lower treatment integrity show improvements in treatment integrity when utilizing teacher-choice interventions based on test-driving?

After the treatment-as-usual phase, Teacher A will begin implementing all of the interventions in the test-drive phase. Since teacher-student pairs are yoked to each other in order to control for threats to internal validity, Teachers B, C, and D will begin entering the other phases of the study in a staggered fashion relative to each other. Thus, only when Teacher A completes the test-drive phase will Teacher B begin implementing Teacher A's preferred intervention in the yoked intervention phase, simultaneously with Teacher A beginning to implement her own preferred intervention in the teacher-preferred phase. Then, Teacher B will complete the test-drive phase and implement her preferred intervention after, along with Teacher C also implementing Teacher B's preferred intervention. The cycle goes until each teacher implements another teacher's preferred intervention in the yoked intervention phase, test-drives interventions, and completes their own teacher-preferred intervention phase. This is to ensure that increases

in TI levels in the teacher-preferred phase are due to the benefits of test-driving and selecting the teacher choice intervention, and not just some factor about the intervention itself.

It is expected, that after teachers test-drive all interventions with 100% treatment integrity and select their teacher choice intervention based on test-driving the intervention, TI levels will significantly increase in the teacher-preferred phase of each teacher. However, teachers implementing another teacher's most preferred intervention in the yoked intervention phase will continue to implement the interventions demonstrating poor TI levels. Therefore, the test-driving procedures and preference rating will positively impact TI levels.

**Research Question 2:** Will there be collateral improvements in student academically engaged behavior when the teacher is utilizing teacher-choice interventions based on test-driving?

As a result of better, more accurate intervention implementation, student behavior is also expected to improve in the teacher-preferred or teacher choice phase.

Academically engaged behaviors will increase, and disruptive and passive-off task behaviors will decrease compared to those in other phases where teachers implemented the interventions with lower levels of TI.

**Research Question 3:** Will teachers' ratings on the URP-IR and URP-A be highest for teacher-choice interventions with respect to dimensions of social validity, treatment acceptability, feasibility, and understanding?

Additionally, teachers will have the highest ratings on URP-IR and URP-A of social validity, treatment acceptability, feasibility, and understanding for their most preferred intervention also implemented in the teacher-preferred phase following test-driving interventions.

## Chapter 4: Discussion

### Implications

The purpose of this study was to assess the impact of test-driving interventions on teachers' TI levels. Specifically, to what degree improvements in TI can be attributed to teacher choice or preferred interventions based on test-driving those interventions for selection purposes. It is expected that test-driving interventions will produce similar effects on treatment integrity as it did in the study conducted by Dart and colleagues (2012). Based on the findings of that study, it is expected that the levels of TI will improve once teachers are given the chance to implement their preferred or teacher choice interventions based on test-driving that intervention.

With respect to the data collected across the four participating teachers in this study, there was variable TI across all irrespective of intervention, but none were as low as Teacher A who had 0% TI. While one might expect variation in TI, one would not expect zero levels. One might hypothesize about why such was the case. It could be that there was no buy-in by her to the intervention process, a lack of feedback from the consultant during intervention implementation, a poor relationship between the teacher and the consultant, a lack of administrative oversight of the intervention process, or a poor relationship with the target student. It remains that additional work by the consultant may be necessary to have Teacher A totally buy in to the project.

A second objective of the study is to determine the collateral effect of teachers' treatment integrity on student academically engaged behavior, disruptive behavior, and passive off-task behavior. Based on the previously cited Dart et al. (2012) study, it is expected that as TI levels improve in the teacher-preferred phases, students' academically

engaged behavior will also improve and disruptive behavior along with passive off-task behavior will decrease as a result of better treatment implementation. Therefore, higher TI levels should produce better treatment outcomes with respect to the degree of improvements in student behavior.

Lastly, the study was designed to determine whether or not teachers' ratings of the interventions on the URP-IR and URP-A would be in accordance with their preferred interventions. It is expected that teachers will have the highest ratings of social validity, treatment acceptability, feasibility, and understanding for their teacher-choice interventions, a finding consistent with the results of the Dart et al. (2012) study.

Teacher A demonstrated 0% treatment integrity throughout all of her sessions in the treatment-as-usual phase. This indicates that utilizing traditional behavioral consultation methods in this case is not enough to ensure accurate and successful implementation of this specific intervention in her classroom. Based on her TI data, Teacher A had difficulties adhering to intervention steps following BST alone. Since the intervention was not implemented correctly, it is likely that there would not be a collateral positive effect on student behaviors using traditional behavioral consultation. Low levels of TI are generally expected in the treatment-as-usual phase; however, 0% TI throughout all observations sessions is exceptionally low. It is likely that Teacher A will benefit from test-driving interventions as an antecedent method to promote and maintain high levels of TI before the implementation needs correction or she may need additional attention by the consultant with respect to other implementation issues. It is expected that Teacher A's TI levels will significantly increase in the teacher-preferred phase, and



student academically engaged behavior will also improve while disruptive behavior and passive off-task behavior will decrease.

Teacher B implemented her intervention with relatively high TI levels. However, TI decreased over multiple observations and stabilized at 25% TI. It is anticipated that test-driving interventions and selecting a teacher-choice intervention will result in higher levels of TI for Teacher B with high levels maintained over time. It is expected that if Teacher B demonstrates high levels of TI and maintains them in the teacher-preferred phase, then improvements in student behavior will also be evident.

Teacher C's implementation in the treatment-as-usual phase was varied. She demonstrated higher TI levels during some observation sessions and low levels in others. This is also a usual pattern following traditional BC. She is also expected to benefit from test-driving interventions and will likely produce higher and more stable levels of TI in the teacher-preferred phase. Student outcomes are also expected to improve in the teacher-preferred phase as the implementation improves.

Because of the abrupt interruption of the study caused by the pandemic, Teacher D was only observed for one session. She demonstrated a high TI level of 80%; however, it is unknown how that would have changed in future sessions. Therefore, conclusions cannot be drawn given the limited data available. Teachers were similar in that none of them engaged in high stable TI levels. The degree of non-adherence to intervention was slightly different among them, but overall all of the participating teachers demonstrated poor levels of TI. Teacher A's TI was exceptionally low relative to the other participating teachers. A few explanations for this occurrence include reasons such as a poor understanding of the intervention procedures and goals, no buy-in to the intervention

process, or experiencing the implementation as an inconvenience during class. It could also be due to the lack of feedback from the behavior consultants, lack of administrative oversight and supervision of intervention implementation, or poor relationship with either the behavior consultant or the target student. As noted previously, Teacher D's TI was exceptionally high; however her TI level was determined based on a single observation session of intervention implementation.

### **Limitations**

Due to the global pandemic COVID-19, data collection was discontinued. Data collection in the subsequent conditions was not possible; therefore arriving at answers to the research questions is not possible at this time.

It would certainly be crucial to complete this study and conduct other similar studies related to treatment integrity in the future. Treatment integrity is crucially important in successfully implementing interventions; however, methods for improving TI levels are still unexplored. Students at risk for academic issues and behavior problems cannot receive effective help if the treatments are not correctly implemented causing student outcomes to be severely compromised. Teacher choice could be a simple yet powerful tool in promoting high TI levels. Therefore, there is an increasing need for related studies to further examine the effects of teacher selection of interventions and the impacts it has on adherence to treatment components. Test-driving interventions could be further expanded and more deeply examined, for example by including more types of interventions and involving teachers with even more diverse levels of teaching experience, age, subjects taught, grades and ages of students taught.

In conclusion, although there are limited data available here, it is important to continue this line of research. As the Dart (2012) and other studies suggested traditional consultation strategies likely result in low TI. Test-driving interventions may very well result in higher levels of TI, therefore significantly impacting student response to intervention leading to improved student performance behaviorally and academically.

## References

- Briesch, A. M., Chafouleas, S. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2013). Assessing influences on intervention use: Revision of the Usage Rating Profile-Intervention. *Journal of School Psychology, 51*, 81–96.
- Chafouleas, S. M., Miller, F. G., Briesch, A. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2012). Usage Rating Profile – Assessment. Storrs, CT: University of Connecticut.
- Dart, E. H., Cook, C. R., Collins, T. A., Gresham, F. M., & Chenier, J. S. (2012). Test driving interventions to increase treatment integrity and student outcomes. *School Psychology Review, 41*, 467–481.
- Fallon, L. M., Kurtz, K. D., & Mueller, M. R. (2018). Direct training to improve educators' treatment integrity: A systematic review of single-case design studies. *School Psychology Quarterly, 33*(2), 169–181.
- Fallon, L. M., Sanetti, L. M. H., Chafouleas, S. M., Faggella-Luby, M. N., & Briesch, A. M. (2018). Direct training to increase agreement between teachers' and observers' treatment integrity ratings. *Assessment for Effective Intervention, 48*, 196–211.
- Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review, 18*, 37–50.
- Gresham, F. M. (2010). Data-based decision making for students' social behavior. *Journal of Evidence-Based Practices for Schools, 11*, 149-168.
- Gresham, F. M., Dart, E. H., & Collins, T. A. (2017). Generalizability of multiple measures of treatment integrity: Comparisons among direct observation, permanent products, and self-report. *School Psychology Review, 46*, 108-121.

- Kilgus, S. P., Riley-Tillman, T. C., Chafouleas, S. M., Christ, T. J., & Welsh, M. E. (2014). Direct behavior rating as a school-based behavior universal screener: Replication across sites. *Journal of School Psychology, 52*, 63-82.
- Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved from What Works Clearinghouse web site: [http://ies.ed.gov/ncee/wwc/pdf/wwc\\_scd.pdf](http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf).
- Lane, K. L., Bocian, K. M., MacMillan, D. L., & Gresham, F. M. (2004). Treatment integrity: An essential but often forgotten component of school-based interventions. *Preventing School Failure, 48*, 36–43.
- Noell, G. H., Witt, J. C., Gilbertson, D. N., Ranier, D. D., & Freeland, J. T. (1997). Increasing teacher intervention implementation in general education settings through consultation and performance feedback. *School Psychology Quarterly, 12*, 77-78.
- Noell, G.H., Gresham, F.M., Gansle, K.A (2002). Does treatment integrity matter? A preliminary investigation of instructional implementation and mathematics performance. *Journal of Behavioral Education, 11*, 51–67.
- Sanetti, L. M. H., & Collier-Meek, S. M. (2015). Data-driven delivery of implementation supports in a multi-tiered framework: A pilot study. *Psychology in the Schools, 52*, 815–828.
- Sanetti, L. M. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to a special series. *School Psychology Review, 38*, 445–459.

Sanetti, L. M. H., Collier-Meek, M. A., Long, A. C. J., Byron, J. R., & Kratochwill, T. R.

(2015). Increasing teacher treatment integrity of behavior support plans through consultation and implementation planning. *Journal of School Psychology, 53*, 209–229.

Schrieber, S. (2020). *Test-Driving Interventions for Teachers: Incorporating Treatment Acceptability to Improve Treatment Integrity* [Unpublished doctoral dissertation]. The University of Southern Mississippi.

# Appendices

## APPENDIX A

### IRB Approval

Date: 3-10-2020

IRB #: IRB-19-373

Title: Test-Driving Interventions for Teachers: Incorporating Treatment Acceptability to Improve Treatment Integrity

Creation Date: 7-31-2019

End Date:

Status: **Approved**

Principal Investigator: Stefanie Schrieber

Review Board: Sacco (Exempt/Expedited Board)

Sponsor: Society for the Study of School Psycholo

#### Study History

Submission Type	Initial	Review Type	Expedited	Decision	<b>Approved</b>
Submission Type	Modification	Review Type	Expedited	Decision	<b>Approved</b>

#### Key Study Contacts

Member	Brad Dufrene	Role	Co-Principal Investigator	Contact	brad.dufrene@usm.edu
Member	Stefanie Schrieber	Role	Primary Contact	Contact	stefanie.schrieber@usm.edu
Member	Mary Ware	Role	Co-Principal Investigator	Contact	mary.ware@usm.edu
Member	Taylor Ben	Role	Co-Principal Investigator	Contact	taylor.ben@usm.edu
Member	Kenneth Ross	Role	Co-Principal Investigator	Contact	kenneth.j.ross@usm.edu
Member	Chelsea Johnson	Role	Co-Principal Investigator	Contact	chelsea.n.johnson@usm.edu
Member	Stefanie Schrieber	Role	Principal Investigator	Contact	stefanie.schrieber@usm.edu

APPENDIX B  
Treatment Integrity Forms

**Self-Monitoring**

Teacher:

Phase:

Date:

	YES	NO
1. Provided student with necessary materials (timer, monitoring form, instructions)		
2. Recorded student behavior every 5 minutes using teacher form		
3. Completed honestly check with target student to ensure accuracy		
4. Provided earned rewards or corrective feedback		

**Response Cost**

Teacher:

Phase:

Date:

	YES	NO
1. Provided student with necessary materials (response-cost monitoring form)		
2. Marks an X when student engages in target behavior		
3. Deliver reward contingent on previously agreed upon criteria		

**Modified Check-in/Check-out**

Teacher:

Phase:

Date:

	YES	NO
1. Meet and review behavior expectations		
2. Remained vigilant of student's behavior throughout session		
3. Completed DBR at the end of session		
4. Reviewed DBR performance with student and provided reward if met		

**Non-Contingent Reinforcement**

Teacher:

Phase:

Date:

	YES	NO
1. Deliver BSP		
2. Deliver BSP		
3. Deliver BSP		
4. Deliver BSP		
2. Ignore all instances of minor disruptive behavior		



APPENDIX C  
Observation Form

Date: \_\_\_\_\_ Classroom: \_\_\_\_\_ Observer: \_\_\_\_\_ Phase: \_\_\_\_\_

Interval	1.1	1.2	1.3	1.4	1.5	1.6	2.1	2.2	2.3	2.4	2.5	2.6
AEB												
DB												
Interval	3.1	3.2	3.3	3.4	3.5	3.6	4.1	4.2	4.3	4.4	4.5	4.6
AEB												
DB												
Interval	5.1	5.2	5.3	5.4	5.5	5.6	6.1	6.2	6.3	6.4	6.5	6.6
AEB												
DB												
Interval	7.1	7.2	7.3	7.4	7.5	7.6	8.1	8.2	8.3	8.4	8.5	8.6
AEB												
DB												
Interval	9.1	9.2	9.3	9.4	9.5	9.6	10.1	10.2	10.3	10.4	10.5	10.6
AEB												
DB												
Interval	11.1	11.2	11.3	11.4	11.5	11.6	12.1	12.2	12.3	12.4	12.5	12.6
AEB												
DB												
Interval	13.1	13.2	13.3	13.4	13.5	13.6	14.1	14.2	14.3	14.4	14.5	14.6
AEB												
DB												
Interval	15.1	15.2	15.3	15.4	15.5	15.6	16.1	16.2	16.3	16.4	16.5	16.6
AEB												
DB												
Interval	17.1	17.2	17.3	17.4	17.5	17.6	18.1	18.2	18.3	18.4	18.5	18.6
AEB												
DB												
Interval	19.1	19.2	19.3	19.4	19.5	19.6	20.1	20.2	20.3	20.4	20.5	20.6
AEB												
DB												

Dependent Variable	Percentage of Intervals	IOA: Yes / No
<b>AEB:</b>	_____ / 120 = _____ %	_____ / 120 = _____ %
<b>DB:</b>	_____ / 120 = _____ %	_____ / 120 = _____ %
<b>POT:</b>	_____ / 120 = _____ %	_____ / 120 = _____ %

APPENDIX D  
Procedural Integrity Checklist

**Teacher:**

**Observer:**

<b>Response Cost</b>	<b>Date implemented:</b>	<b>YES</b>	<b>NO</b>	<b>N/A</b>
1. Provided student with necessary materials (response-cost monitoring form)				
2. Marks an X when student engages in target behavior				
3. Deliver reward contingent on previously agreed upon criteria				

<b>Self-Monitoring</b>	<b>Date implemented:</b>	<b>YES</b>	<b>NO</b>	<b>N/A</b>
1. Provided student with necessary materials (timer, monitoring form, instructions)				
2. Recorded student behavior every 5 minutes using teacher form				
3. Completed honesty check with target student to ensure accuracy				
4. Provided earned rewards or corrective feedback				

<b>Modified Check-in/Check Out</b>	<b>Date implemented:</b>	<b>YES</b>	<b>NO</b>	<b>N/A</b>
1. Meet and review behavior expectations				
2. Remained vigilant of student's behavior throughout session				
3. Completed DBR at the end of session				
4. Reviewed DBR performance with student and provided reward if met				

<b>Non-Contingent Reinforcement</b>	<b>Date implemented:</b>	<b>YES</b>	<b>NO</b>	<b>N/A</b>
1. Deliver BSP				
2. Deliver BSP				
3. Deliver BSP				
4. Deliver BSP				
5. Ignore all instances of minor disruptive behavior				

Total number of steps implemented correctly

Total intervention steps

Percentage of steps implemented correctly


APPENDIX E  
Treatment Overviews

**Self-Monitoring**

Provide the student with a self-monitoring form and MotivAider set to four-minute fixed intervals. Tell the student that each time the MotivAider buzzes, they will determine whether they were on or off-task. If the student is on-task, they should place a check mark in the box. If the student is off-task, they should place an “X” in the box.

The teacher will also monitor the student’s behavior using a MotivAider set to the same interval, and a behavior monitoring form. The teacher will also place a check or an X in the box each time the MotivAider buzzes.

*Honesty Check:* At the end of the 20-minute period, the teacher will conduct an honesty check with the student, in which the teacher will determine whether the student’s monitoring form is accurate. If the student’s monitoring form matches 100% with the teacher’s, and they have four or more check marks, they will be provided with a reward.

Student:	Date:	Teacher:		
Place a <b>check mark</b> in the box if you were <b>on-task</b> when the MotivAider Buzzed Place a <b>"X"</b> in the box if you were <b>off-task</b> when the MotivAider Buzzed				
1	2	3	4	5
% Agreement with teacher:				

**Response Cost**

Provide the student with a response-cost form. Tell the student that each time they engage in the disruptive behavior, the teacher will place an “X” in one of the boxes on the form.

During the 20-minute session, mark an “X” in the box immediately after you observe the student engage in the disruptive behavior. If the student has received one or fewer X’s, at the end of the 20-minute session, then provide the student with a reward.

Student:	Date:	Teacher:		
The teacher will place an <b>"X"</b> in the box if you were being disruptive to the class				

### **Modified Check-in/Check-Out**

Meet with the student to review behavior expectations and let them know they must remain “on-task” in the classroom for the majority of the session to earn a reward.

Remain aware of the student’s behavior during the next 20 minutes. At the end of the 20-minute session, rate the student’s on-task behavior on a scale of 0-100% using the Direct Behavior Rating (DBR) form.

After completing the DBR, review the rating with the student, and either provide them with praise and a reward for meeting their goal, or corrective feedback and encouragement to try again later.

Student:	Date:	Teacher:			
Rate how often the student was on-task and circle the appropriate percentage					
0%	20%	40%	60%	80%	100%

### **Non-Contingent Reinforcement**

Provide a statement of behavior specific praise (BSP) to the target student at least once every four minutes (or five times in the 20-minute period). Statements of BSP must include a general praise statement such as “good job” or “awesome job” paired with a behavior “being on-task,” or “completing your work. Examples of BSP include “Great job working today,” and “Thank you for staying on-task.” All instances of the student engaging in minor disruptive behavior should be ignored.

APPENDIX F  
URP-IR (Adapted)

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree
1 This intervention is an effective choice for addressing a variety of problems	1	2	3	4	5
3 I would be able to allocate my time to implement this intervention	1	2	3	4	5
4 I understand how to use this intervention	1	2	3	4	5
6 I am knowledgeable about the intervention procedures	1	2	3	4	5
7 The intervention is a fair way to handle the child's problem behavior	1	2	3	4	5
8 The total time required to implement the intervention procedures would be manageable	1	2	3	4	5
9 I would not be interested in implementing this intervention	1	2	3	4	5
11 I would have positive attitudes about implementing this intervention	1	2	3	4	5
12 This intervention is a good way to handle the child's behavior problem	1	2	3	4	5
13 Preparation of materials needed for this intervention would be minimal	1	2	3	4	5
17 material resources needed for this intervention are reasonable	1	2	3	4	5
18 I would implement this intervention with a good deal of enthusiasm	1	2	3	4	5
19 This intervention is too complex to carry out accurately	1	2	3	4	5
21 This intervention would not be disruptive to other students	1	2	3	4	5
22 I would be committed to carrying out this intervention	1	2	3	4	5
23 The intervention procedures easily fit in with my current practices	1	2	3	4	5
27 The amount of time required for record keeping would be reasonable	1	2	3	4	5

APPENDIX G  
URP-A



# URP-Assessment

Directions: Consider the described assessment when answering each of the following statements. Circle the number that best reflects your agreement with the statement, using the scale provided below.

	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
1. This assessment is an effective choice for understanding a variety of problems.	1	2	3	4	5	6
2. I would need additional resources to carry out this assessment.	1	2	3	4	5	6
3. I would be able to allocate my time to implement this assessment.	1	2	3	4	5	6
4. I understand how to use this assessment.	1	2	3	4	5	6
5. A positive home-school relationship is needed to use this assessment.	1	2	3	4	5	6
6. I am knowledgeable about the assessment procedures.	1	2	3	4	5	6
7. The assessment is a fair way to evaluate the child's behavior problem.	1	2	3	4	5	6
8. The total time required to implement the assessment procedures would be manageable.	1	2	3	4	5	6
9. I would not be interested in implementing this assessment.	1	2	3	4	5	6
10. My administrator would be supportive of my use of this assessment.	1	2	3	4	5	6
11. I would have positive attitudes about implementing this assessment.	1	2	3	4	5	6
12. This is a good way to assess the child's behavior problem.	1	2	3	4	5	6
13. Preparation of materials needed for this assessment would be minimal.	1	2	3	4	5	6
14. Use of this assessment would be consistent with the mission of my school.	1	2	3	4	5	6

URP-A was created by Sandra M. Chafouleas, Faith G. Miller, Amy M. Briesch, Sabina Rak Neugebauer, & T. Chris Riley-Tillman. Copyright © 2012 by the University of Connecticut. All rights reserved. Permission granted to photocopy for personal and educational use as long as the names of the creators and the full copyright notice are included in all copies.



	Strongly Disagree	Disagree	Slightly Disagree	Slightly Agree	Agree	Strongly Agree
15. Parental collaboration is required in order to use this assessment.	1	2	3	4	5	6
16. Material resources needed for this assessment are reasonable.	1	2	3	4	5	6
17. I would implement this assessment with a good deal of enthusiasm.	1	2	3	4	5	6
18. This assessment is too complex to carry out accurately.	1	2	3	4	5	6
19. These assessment procedures are consistent with the way things are done in my system.	1	2	3	4	5	6
20. Use of this assessment would not be disruptive to students.	1	2	3	4	5	6
21. I would be committed to carrying out this assessment.	1	2	3	4	5	6
22. The assessment procedures easily fit in with my current practices.	1	2	3	4	5	6
23. I would need consultative support to implement this assessment.	1	2	3	4	5	6
24. I understand the procedures of this assessment.	1	2	3	4	5	6
25. My work environment is conducive to implementation of an assessment like this one.	1	2	3	4	5	6
26. The amount of time required for record keeping would be reasonable.	1	2	3	4	5	6
27. Regular home-school communication is needed to implement these assessment procedures.	1	2	3	4	5	6
28. I would require additional professional development in order to implement this assessment.	1	2	3	4	5	6

## APPENDIX H Operational Definitions

AEB – student actively or passively attending to assigned work – examples

- Writing.
- Reading aloud.
- Raising a hand.
- Talking to the teacher about the assigned material.
- Talking to a peer about the assigned material.
- Looking up a word in a dictionary.
  
- Listening to a lecture.
- Looking at an academic worksheet.
- Reading assigned material silently.
- Looking at the blackboard during teacher instruction.
- Listening to a peer respond to a question.

DB – motor or verbal activity that is not directly associated with assigned academic task – examples

- Making any audible sound, such as whistling, humming, or forced burping.
- Talking to another student about issues unrelated to an assigned academic task.
- Talking to another student about an assigned academic task when such talk is prohibited by the teacher.
- Making unauthorized comments or remarks.
- Calling out answers to academic problems when the teacher has not specifically asked for an answer or permitted such behavior.
  
- Engaging in any out-of-seat behavior (defined as buttocks not in contact with the seat).
- Aimlessly flipping the pages of a book.
- Manipulating objects not related to the academic task (e.g., playing with a paper clip, throwing paper, twirling a pencil, folding paper).
- Physically touching another student when not related to an academic task.
- Bending or reaching, such as picking up a pencil off the floor.
- Drawing or writing that is not related to an assigned academic activity.
- Turning around in one's seat, orienting away from the classroom instruction.
- Fidgeting in one's seat (i.e., engaging in repetitive motor movements for at least 3 consecutive seconds; student must be off-task for this category to be scored).

POT – when student is passively not attending to assigned academic activity – examples

- Sitting quietly in an unassigned activity.
- Looking around the room.
- Staring out the window.
- Passively listening to other students talk about issues unrelated to the assigned academic activity.
  
- Sleeping