Evaluating the Function of Problem Behaviors using Direct Behavior Ratings

Crystal N. Taylor & Stephen P. Kilgus
University of Missouri

Jennifer S. Kazmerski
East Carolina University

Nathaniel P. von der Embse
Temple University

NASP 2015, Orlando, FL
Implementation of Tier 2 Interventions

- Behavior multi-tiered systems of support
  - Tier 2 → prevention of further symptom development

- **Standard protocol** (Yong and Cheney, 2013)
  - Single intervention implemented in standard fashion

- **Flexible protocol** (Hawken, Adolphson, MacLeod, and Schumann, 2009)
  - Problem-solving model
Standard Protocol

• Single Tier 2 intervention
• Implemented in a common way for all students assigned to Tier 2
• Decisions:
  – Highly responsive → return to Tier 1
  – Somewhat responsive → stay at Tier 2
  – Unresponsive → go to Tier 3
• Ex. Check In/Check Out (CICO)
  – Morning check in with coordinator
  – Ongoing performance feedback from teacher throughout the day
  – Afternoon check out with coordinator

(Yong and Cheney, 2013)
Flexible Protocol

• Problem solving approach
  – Menu of intervention strategies; modifiable interventions
  – Use data to inform selection/modification
• Intervention procedures depend upon function of behavior
• Identify the function of behavior through brief FBA instruments
  – Function = purpose the behavior serves
  – Assumption = intervention will be more effective if it matches a student’s function

(Hawken, Adolphson, MacLeod, and Schumann, 2009)
Standard vs. Flexible

- **Ex. Check In/ Check Out (CICO)**
- McIntosh, Campbell, Carter, & Dickey, 2009
  - Behavior maintained by Adult Attention
  - Behavior maintained by Escape/Avoidance
- Modifications have proven effective (Campbell & Anderson, 2008; Kilgus, Fallon, & Feinberg, 2015; Turtura, Anderson, & Boyd, 2014)
Take Home Message

• Limitations associated with standard protocol approach
  – SP intervention is unlikely to be effective for a subgroup of students

• SP protocol implications
  – Student who could respond to Tier 2 interventions might be moved on to Tier 3

• Need for flexibility in intervention selection/ modification
  – Need to collect functional behavior assessment (FBA) data
Functional Behavioral Assessment Tools

• What is required of a FBA tool at Tier 2?
  – Efficiency
  – Accurate portrayal of the function of behavior
  – Direct
  – Easy integration into other forms
Functional Behavioral Assessment Tools

• Functional Assessment Checklist for Teachers and Staff (FACTS) (March et al., 2000)
  – Rating Scale + Semi-Structured Interview
  – Indirect methodology
  – Limited evidence (McIntosh et al., 2008; Zaja, Moore, van Ingen, & Rojahn, 2011)
Functional Behavioral Assessment Tools

- Functional Analysis Screening Tool (FAST) (Iwata & DeLeon, 1995)
  - Rating scale
  - Iwata, DeLeon & Roscoe (2013)
    - Miscalculated function 1/3 cases
    - Indirect methodology
Systematic Direct Observations

• Example: ABC recording, time sampling procedures, scatterplot

• Good \(\rightarrow\) highly direct, low inference
  – Collected at time and place in which behavior is exhibited

• Bad \(\rightarrow\) takes a large amount of time and effort
  – Limited applicability at Tier 2 due to inefficiency
Direct Behavior Ratings

• Direct method for monitoring social behavior
  – SDO + Behavior Rating Scale
  – Direct $\rightarrow$ short latency, low inference
  – Brief teacher ratings
  – Progress monitoring
  – Corresponds to operationally defined behaviors
  – Minimum training
  – DBR-SIS* and DBR-MIS

(Chafouleas, Riley-Tillman, and McDougal 2002, Chafouleas, Riley-Tillman, & Christ, 2009)
Direct Behavior Ratings

DBR-SIS

- Psychometric defensibility in assessing social behavior
  - Sensitivity to change, validity, and reliability
- Could potentially collect data regarding consequences at the same time as behaviors
  - Use in FBA?

(Chafouleas, Riley-Tillman, and McDougal 2002, Chafouleas, Riley-Tillman, & Christ, 2009)
DBR-SIS in FBA: Interpretation & Use

- **Interpretation** → akin to *conditional probabilities*
  - The percentage of problem behavior instances followed by each consequence

- **Use** → collect at same time as baseline progress monitoring (e.g., re: disruptive behavior) to inform subsequent intervention decisions
Purpose

DBR-SIS utility in FBA?

• Can the DBR-SIS generate accurate ratings of behavioral consequences?
• What level of training is needed for accurate DBR-SIS ratings?
• Can users collect both ratings of behavior and consequences and still remain accurate?
Experiment 1: Method

- Participants
  - 178 undergraduates
- Randomly assigned
  - Training with feedback
  - Training no feedback
  - Pretest-Posttest only
  - Posttest only
Experiment 1: Method

• Materials
  – Book Chapter
  – Video Clips
  – DBR-SIS
Experiment 1: Method

Directions: Place a mark along the line that best reflects the percentage of total time the student exhibited disruptive behavior.

Disruptive behavior is a student action that interrupts regular school or classroom activity. For example, out of seat, fidgeting, playing with objects, acting aggressively, talking/yelling about things that are unrelated to classroom instruction.

Disruptive Behavior:

% of Total Time

0% 1 2 3 4 5 6 7 8 9 10
Never Sometimes Always

Directions: Place a mark along the line that best reflects the percentage of disruptive behaviors that were followed by each consequence.

Adult Attention: Positive, negative, or neutral adult reaction that can be either verbal or nonverbal. Examples: reprimand, redirection to work, praise, discussion, high-fives, or shushing.

Peer Attention: Positive, negative, or neutral peer reaction that can be either verbal or nonverbal. Examples: talking, laughing, arguing, high-fives, hitting, kicking, or yelling.

Escape/Avoidance: Removal of task, activity, or performance expectations. Examples: removal of academic materials, allowance to delay task completion, permission to leave room, or elimination of task demands.

Access to Tangibles or Activities: Acquisition of items or activities. Examples: toys, food, prizes, games, preferred tasks, sleep, technology, or homework pass.

Adult Attention:

% of Total Time

0% 1 2 3 4 5 6 7 8 9 10
Never Sometimes Always
Experiment 1: Method
## Experiment 1: Method

True score and inter-observer agreement for contrived videos in experiments 1 and 2

<table>
<thead>
<tr>
<th>True Score Rating</th>
<th>Study 1</th>
<th>Study 2</th>
<th>IOA Kappa *</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DB</td>
<td>AA</td>
<td>PA</td>
</tr>
<tr>
<td>Clip 1</td>
<td>7</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Clip 2</td>
<td>2</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Clip 3</td>
<td>1</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Clip 4</td>
<td>5</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Clip 5</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Clip 6</td>
<td>1</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Clip 7</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Clip 8</td>
<td>2</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

Note: Disruptive Behavior (DB), Adult Attention (AA), Peer Attention (PA), Escape/Avoidance (EA), and Access to Tangibles/Activities (TA).

*Kappa scores reflect disruptive behavior agreement only.*
Experiment 1: Method

• Procedure
  – 40-45 minute presentation including pretest, post test, and practice videos
  – Training with feedback
  – Training no feedback
  – Pretest-Posttest only
  – Posttest only
## Experiment 1: Results

<table>
<thead>
<tr>
<th>Function</th>
<th>Kruskal-Wallis ANOVA</th>
<th>Repeated Measures MANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>Wilks’ Lambda F (Time*Group)</td>
</tr>
<tr>
<td>Adult Attention</td>
<td><strong>97.98</strong></td>
<td><strong>56.59</strong></td>
</tr>
<tr>
<td>Peer Attention</td>
<td><strong>84.30</strong></td>
<td><strong>53.80</strong></td>
</tr>
<tr>
<td>Escape</td>
<td><strong>92.45</strong></td>
<td><strong>67.63</strong></td>
</tr>
<tr>
<td>Access to Items</td>
<td><strong>72.90</strong></td>
<td><strong>40.31</strong></td>
</tr>
<tr>
<td>Disruptive</td>
<td><strong>74.27</strong></td>
<td><strong>48.30</strong></td>
</tr>
</tbody>
</table>

Note: Dependent variables correspond to corrected (absolute) accuracy scores

- **Mann Whitney U** → Statistically significant difference ($p < .001$) between **Training with Performance Feedback** and all other groups across all functional targets.
## Experiment 1: Results

<table>
<thead>
<tr>
<th>Comparison Score</th>
<th>Group</th>
<th>Adult Attention</th>
<th>Peer Attention</th>
<th>Escape/Avoidance</th>
<th>Access to Tangibles/Activities</th>
<th>Disruptive Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1</td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Pretest Uncorrected</td>
<td>PO*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PP</td>
<td>-2.30</td>
<td>2.02</td>
<td>-1.86</td>
<td>2.13</td>
<td>-3.26</td>
<td>2.05</td>
</tr>
<tr>
<td>T</td>
<td>-2.03</td>
<td>1.89</td>
<td>-2.06</td>
<td>1.83</td>
<td>-2.71</td>
<td>1.41</td>
</tr>
<tr>
<td>TF</td>
<td>-2.29</td>
<td>2.12</td>
<td>-2.13</td>
<td>1.93</td>
<td>-2.91</td>
<td>1.59</td>
</tr>
<tr>
<td>Posttest Uncorrected</td>
<td>PO</td>
<td>-0.61</td>
<td>1.78</td>
<td>-3.61</td>
<td>1.89</td>
<td>-3.87</td>
</tr>
<tr>
<td>PP</td>
<td>-0.66</td>
<td>1.72</td>
<td>-4.07</td>
<td>1.88</td>
<td>-4.73</td>
<td>2.28</td>
</tr>
<tr>
<td>T</td>
<td>-0.09</td>
<td>1.61</td>
<td>-3.07</td>
<td>1.89</td>
<td>-3.35</td>
<td>2.15</td>
</tr>
<tr>
<td>TF</td>
<td>3.14</td>
<td>1.09</td>
<td>0.02</td>
<td>1.40</td>
<td>-0.38</td>
<td>0.91</td>
</tr>
<tr>
<td>Pretest Corrected</td>
<td>PO</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PP</td>
<td>2.70</td>
<td>1.57</td>
<td>4.73</td>
<td>1.21</td>
<td>3.26</td>
<td>2.05</td>
</tr>
<tr>
<td>T</td>
<td>2.34</td>
<td>1.67</td>
<td>5.10</td>
<td>1.29</td>
<td>2.71</td>
<td>1.41</td>
</tr>
<tr>
<td>TF</td>
<td>2.87</td>
<td>1.77</td>
<td>4.83</td>
<td>1.53</td>
<td>2.91</td>
<td>1.59</td>
</tr>
<tr>
<td>Posttest Corrected</td>
<td>PO</td>
<td>2.56</td>
<td>0.96</td>
<td>3.78</td>
<td>1.68</td>
<td>3.87</td>
</tr>
<tr>
<td>PP</td>
<td>2.45</td>
<td>0.88</td>
<td>4.20</td>
<td>1.63</td>
<td>4.73</td>
<td>2.28</td>
</tr>
<tr>
<td>T</td>
<td>2.63</td>
<td>1.03</td>
<td>3.30</td>
<td>1.63</td>
<td>3.35</td>
<td>2.15</td>
</tr>
<tr>
<td>TF</td>
<td>3.27</td>
<td>0.91</td>
<td>1.43</td>
<td>0.79</td>
<td>0.46</td>
<td>0.88</td>
</tr>
</tbody>
</table>
Experiment 1

- Training with feedback provided the most accurate ratings
  - Within 10% of SDO true scores
  - However, adult attention was less accurate
    - Training Modification
Motivation for a second experiment

• % of target student disruptions met with each consequence was similar among practice clips
  – Bias in posttest?
• Increased focus on FBA in general
  – More focus on DBR-SIS in particular
• Similar posttest clips
  – Inadequate sampling of performance
Experiment 2: Method

• Participants
  – 213 undergraduates

• Randomly assigned
  – Training with feedback
  – Training no feedback
  – Pretest-Posttest only
  – Posttest only
Experiment 2: Method

• Changes to PowerPoint
  – Less FBA
  – More detailed examples of rating
  – Clip order was modified

• Changes to Videos
  – 2 Videos added
  – Specific Script
## Experiment 1: Results

<table>
<thead>
<tr>
<th>Function</th>
<th>Kruskal-Wallis ANOVA</th>
<th>Repeated Measures MANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>Wilks’ Lambda F (Time*Group)</td>
</tr>
<tr>
<td>Adult Attention</td>
<td><strong>43.22</strong></td>
<td>*3.12</td>
</tr>
<tr>
<td>Peer Attention</td>
<td><strong>20.12</strong></td>
<td>**13.07</td>
</tr>
<tr>
<td>Escape</td>
<td><strong>86.45</strong></td>
<td>**10.43</td>
</tr>
<tr>
<td>Access to Items</td>
<td><strong>27.56</strong></td>
<td>*6.50</td>
</tr>
<tr>
<td>Disruptive</td>
<td><strong>29.49</strong></td>
<td>*3.09</td>
</tr>
</tbody>
</table>

Note: Dependent variables correspond to corrected (absolute) accuracy scores

- **Mann Whitney U** → Statistically significant difference ($p < .001$) between **Training with Feedback** and all other groups across all functional targets
## Experiment 2: Results

<table>
<thead>
<tr>
<th>Comparison Score</th>
<th>Group</th>
<th>Adult Attention</th>
<th>Peer Attention</th>
<th>Escape/Avoidance</th>
<th>Access to Tangibles/Activities</th>
<th>Disruptive Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Pretest Uncorrected</td>
<td>PO*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>PP</td>
<td>-4.90</td>
<td>1.87</td>
<td>-2.21</td>
<td>1.70</td>
<td>-3.87</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>-5.18</td>
<td>1.56</td>
<td>-2.19</td>
<td>1.32</td>
<td>-3.53</td>
</tr>
<tr>
<td></td>
<td>TF</td>
<td>-4.88</td>
<td>1.92</td>
<td>-2.05</td>
<td>1.85</td>
<td>-2.97</td>
</tr>
<tr>
<td>Posttest Uncorrected</td>
<td>PO</td>
<td>-2.50</td>
<td>1.90</td>
<td>-0.25</td>
<td>1.85</td>
<td>-3.87</td>
</tr>
<tr>
<td></td>
<td>PP</td>
<td>-3.05</td>
<td>1.96</td>
<td>-0.78</td>
<td>1.79</td>
<td>-3.65</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>-2.62</td>
<td>2.00</td>
<td>-0.14</td>
<td>2.14</td>
<td>-2.31</td>
</tr>
<tr>
<td></td>
<td>TF</td>
<td>-0.86</td>
<td>1.72</td>
<td>0.92</td>
<td>2.11</td>
<td>-0.82</td>
</tr>
<tr>
<td>Pretest Corrected</td>
<td>PO</td>
<td>-</td>
<td>-</td>
<td>0.92</td>
<td>2.11</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>PP</td>
<td>5.15</td>
<td>1.49</td>
<td>2.65</td>
<td>1.10</td>
<td>4.10</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>5.20</td>
<td>1.52</td>
<td>2.56</td>
<td>1.07</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td>TF</td>
<td>4.89</td>
<td>1.92</td>
<td>2.92</td>
<td>1.15</td>
<td>3.34</td>
</tr>
<tr>
<td>Posttest Corrected</td>
<td>PO</td>
<td>3.07</td>
<td>1.50</td>
<td>3.58</td>
<td>1.46</td>
<td>3.87</td>
</tr>
<tr>
<td></td>
<td>PP</td>
<td>3.30</td>
<td>1.83</td>
<td>3.90</td>
<td>1.23</td>
<td>3.65</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>2.77</td>
<td>1.82</td>
<td>3.72</td>
<td>1.33</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td>TF</td>
<td>1.94</td>
<td>1.23</td>
<td>2.78</td>
<td>1.12</td>
<td>0.88</td>
</tr>
</tbody>
</table>
Percentage of Disruptions met with each Consequence

<table>
<thead>
<tr>
<th>Consequence</th>
<th>SDO</th>
<th>DBR-Fx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult Attention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peer Attention</td>
<td>70</td>
<td>40</td>
</tr>
<tr>
<td>Escape</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access to Item</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disruptive Behavior</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Posttest Clip 1
Posttest Clip 2

Percentage of Disruptions met with each Consequence

Consequence

- Adult Attention
- Peer Attention
- Escape
- Access to Item
- Disruptive Behavior

Consequence

PosVest
Clip 2
SDO
DBR-Fx
Discussion

1. Can the DBR-SIS generate accurate ratings of behavioral consequence?
   1. Yes – within 10-20% of SDO data
2. What level of training is needed for accurate DBR-SIS ratings?
   1. Training with practice and feedback
3. Can users collect both ratings of behavior and consequences and still remain accurate?
   1. Yes – ratings of behavior and consequences both fell within 10-20% of SDO data
   2. Behavior accuracy similar to that found in previous training DBR-SIS studies (e.g., Chafouleas et al., 2012)
Discussion

• Accurate functional assessment instrument within Tier 2 (with teacher training w/ feedback)
  – Collect DBR-SIS disruptive behavior + behavioral consequences
  – Use data to plan function-based interventions
  – Continue to progress monitor with DBR-SIS
Limitations

- Participant population
- Observation period not analogous to traditional DBR-SIS periods.
- Higher levels of adult and peer attention
- Utilization of student actors, not a typical classroom setting
Future Research

• DBR-SIS ratings compared to a comprehensive FBA
• Treatment Utility
Thank you!