Advances in Wraparound fidelity monitoring: Pulling it all together

Jennifer Schurer Coldiron, MSW, PhD
Eric J. Bruns, PhD
April Sather, MPH
Alyssa Hook, BS

Tuesday, March 24, 2015 • 10:30-11:30am
Agenda for Today’s Symposium

• WFAS Overview
• Reviving the DRM
• Refining the TOM
• WrapSTAR
The Wraparound Fidelity Assessment System (WFAS)

- A multi-method approach to assessing the quality and context of individualized care planning and management for children and youth with complex needs and their families

- Interview: Wraparound Fidelity Index, v. 4
- Survey: short form, WFI-EZ
- Team Observation Measure
  - Version 2.0 currently being piloted
- Document Review Measure
  - Version 2.0 being developed and piloted
- Community Supports for Wraparound Inventory

www.wrapinfo.org
The original suite of 4 tools were developed in 2007 with NIH funding

- National Wraparound Initiative experts, with funding from the NIH, developed four prototype instruments
  - Constructed initial indicator pools and revised using a Delphi process
  - Iterative process of solicitation and receipt of feedback from approximately 15 individuals spanning roles such as national and local Wraparound trainers, researchers, and implementation leaders

- Intended primarily for use by program evaluators, local quality assurance staff, and researchers
Once WFAS was developed, it was pilot tested with NIH (STTR) funding

- User testing (NWI Experts) and pilot communities
  - Focus Groups
  - Items flagged/revised
- Larger sample of sites piloted again
  - 15 sites tested the WFAS tools
  - Psychometric data was gathered (presented later)
    - Feasibility
    - Acceptability
    - Reliability
    - Variance
WFAS Tools are now being used around the country

Wraparound Fidelity Tools Used in the U.S.

Legend:  
- WFAS tool(s) used statewide
- WFAS tool(s) used by one or more local sites
Agenda for Today’s Symposium

- WFAS Overview
- Reviving the DRM
- Refining the TOM
- WrapSTAR
Reviving the Wraparound Document Review Measure (DRM)

Jennifer Schurer Coldiron, MSW, PhD
April Sather, MPH
Alyssa Hook, BS
DRM assesses practice from documentation in Wraparound records

- Employed by supervisors, coaches, and external evaluators to assess adherence to standards of high-quality Wraparound as documented in the case file
- DRM 1.0 items assessed one of the ten Wraparound principles or one of two additional constructs, access and timeliness
  - Each item was also specific to one of the four phases of wraparound activities
  - Consisted of 33 items scored on a scale of 0 (not met) to 3 (fully met)
- Jim Rast was lead developer of DRM 1.0, along with and other National Wraparound Initiative experts
From the beginning, the DRM was not as highly-rated as other WFAS tools.

User Rating of WFAS Instruments*

“To what extent does the tool adequately capture the strengths and weaknesses of your program?”

<table>
<thead>
<tr>
<th>Answer Options</th>
<th>WFI-1 (n=8)</th>
<th>TOM 1.0 (n=7)</th>
<th>DRM 1.0 (n=6)</th>
<th>CSWI (n=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = Not at all</td>
<td>0.0%</td>
<td>0.0%</td>
<td>20.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>2 = A little bit</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>3 = Somewhat</td>
<td>25.0%</td>
<td>0.0%</td>
<td>40.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>4 = A good deal</td>
<td>62.5%</td>
<td>85.7%</td>
<td>40.0%</td>
<td>66.7%</td>
</tr>
<tr>
<td>5 = Very Much</td>
<td>12.5%</td>
<td>14.3%</td>
<td>0.0%</td>
<td>33.3%</td>
</tr>
</tbody>
</table>

*Based on 2007 development and pilot research funded by the NIH (STTR)
User Rating of WFAS Instruments*

“To what extent did your program or site benefit from use of the tool’s approach?”

<table>
<thead>
<tr>
<th>Answer Options</th>
<th>WFI-1 (n=8)</th>
<th>TOM 1.0 (n=7)</th>
<th>DRM 1.0 (n=6)</th>
<th>CSWI (n=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = Not at all</td>
<td>0.0%</td>
<td>0.0%</td>
<td>33.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>2 = A little bit</td>
<td>12.5%</td>
<td>0.0%</td>
<td>33.3%</td>
<td>16.7%</td>
</tr>
<tr>
<td>3 = Somewhat</td>
<td>37.5%</td>
<td>14.3%</td>
<td>0.0%</td>
<td>16.7%</td>
</tr>
<tr>
<td>4 = A good deal</td>
<td>37.5%</td>
<td>71.4%</td>
<td>16.7%</td>
<td>50.0%</td>
</tr>
<tr>
<td>5 = Very Much</td>
<td>12.5%</td>
<td>14.3%</td>
<td>16.7%</td>
<td>16.7%</td>
</tr>
</tbody>
</table>

*Based on 2007 development and pilot research funded by the NIH (STTR)

“To what extent is the tool feasible for implementation at your Wraparound program or site?”

<table>
<thead>
<tr>
<th>Answer Options</th>
<th>WFI-1 (n=8)</th>
<th>TOM 1.0 (n=7)</th>
<th>DRM 1.0 (n=6)</th>
<th>CSWI (n=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = Not at all</td>
<td>0.0%</td>
<td>0.0%</td>
<td>33.3%</td>
<td>0.0%</td>
</tr>
<tr>
<td>2 = A little bit</td>
<td>14.3%</td>
<td>0.0%</td>
<td>16.7%</td>
<td>16.7%</td>
</tr>
<tr>
<td>3 = Somewhat</td>
<td>28.6%</td>
<td>14.3%</td>
<td>16.7%</td>
<td>0.0%</td>
</tr>
<tr>
<td>4 = A good deal</td>
<td>28.6%</td>
<td>71.4%</td>
<td>33.3%</td>
<td>50.0%</td>
</tr>
<tr>
<td>5 = Very Much</td>
<td>28.6%</td>
<td>14.3%</td>
<td>0.0%</td>
<td>33.3%</td>
</tr>
</tbody>
</table>
Attempts at revising the DRM 1.0 were made in 2010

• Modified using the Delphi process within the NWI members and experts
  – The items were reduced to 22, but the themes and principles remained the same

• Never widely disseminated
  – Was made available to a handful of sites who modified the tool to fit local needs and terminology
DRM has recently been revived to meet needs of sites and evaluators

• Another modified Delphi process with NWI experts
• Goals of 2014 revision included:
  – Make a more comprehensive tool that leverages the rich information a case file may offer
  – Refine the terminology to be generic and/or clear enough that it could be useful, unaltered, in a variety of settings
  – Create a tool that aligns with the National Wraparound Initiative model and other WFAS tools
  – Streamline the tool to make it easier to administer
  – Make the language and terminology clearer and more consistent
  – Strengthen conceptual clarity between subscales
## Tool Comparison by Structure

<table>
<thead>
<tr>
<th></th>
<th>DRM 1.0</th>
<th>DRM 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Subscales</strong></td>
<td>None—just total score</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 Wraparound Key Elements subscales, please one each for overall fidelity, full Meeting Attendance, Timely Engagement, Safety Planning, Crisis Response, and Transition Planning</td>
</tr>
<tr>
<td><strong>Number of scored items/indicators assessing adherence to Wraparound model</strong></td>
<td>33</td>
<td>43</td>
</tr>
<tr>
<td><strong>Optional Sections</strong></td>
<td>None</td>
<td>Outcomes Service planning and receipt</td>
</tr>
<tr>
<td><strong>Aligned with other Wraparound Fidelity Assessment System Tools</strong></td>
<td>No</td>
<td>Yes – assesses fidelity to the same key elements of the Wraparound Fidelity Index (WFI-EZ) and TOM 2.0</td>
</tr>
<tr>
<td><strong>Gathers based youth and team information</strong></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Scoring system</strong></td>
<td>0 (no evidence) to 3 (clear evidence)</td>
<td>0 (no evidence) to 3 (clear evidence)</td>
</tr>
</tbody>
</table>
Next Steps

• WERT in process of using tool in 7 different sites
  – Will revise, if necessary, based on experience
• Also seeking external sites to pilot to assess feasibility and utility in the field
Agenda for Today’s Symposium

• WFAS Overview
• Reviving the DRM
• Refining the TOM
• WrapSTAR
Refining the Team Observation Measure (TOM)

Jennifer Schurer Coldiron, MSW, PhD
Alyssa Hook, BS
April Sather, MPH
Initial TOM Development

• **Initial 78-item TOM was developed in 2007 with other WFAS tools**
  – Item pool was developed by reviewing measures such as the Family Assessment and Planning Team Observation Form (FAPT) and Wraparound Observation Form (WOF)
  – Inter-rater reliability analysis showed mean Cohen’s Kappa of only 0.46, indicating only moderate agreement between raters

• **Tool was revised in 2009**
  – Scoring rules were revised to be more objective and clear
  – 7 items that were difficult for the observers to score reliably were eliminated
  – Yielded the current 71-item version, “TOM 1.0”
    • currently used by 45 collaborators
Despite good reliability and reasonable validity, desire to further refine tool

- **Reliability and validity of the TOM 1.0** (Bruns et al., 2014)
  - High inter-rater reliability with pooled Kappa of 0.73
  - Strong internal consistency with Cronbach’s $\alpha = .80$
  - Program-level mean total TOM 1.0 scores correlated highly with mean total WFI scores for the same programs
  - Agreement with two observer with external roles was near perfect

- Remaining desire to reduce the burden on the observer, clarify concepts, and increase potential variability
Our goals during the revision included:

• Create a more practice-oriented tool that aligns with the National Wraparound Initiative model
• Streamline the tool to make it easier to administer
• Remove redundant items
• Make the language and terminology clearer and more consistent
• Remove items that require follow-up and/or cannot be readily observed within most team meetings
• Remove non-essential items that show little variability on the TOM 1.0
• Separate assessment of facilitation skills from fidelity to the Wraparound model
• Strengthen conceptual clarity between subscales
Revision and Testing Process

• **2014 Revision**
  – Iterative process with multiple rounds of feedback and edits
    • Wraparound experts from The Institute for Innovation & Implementation at the University of Maryland, Baltimore, Portland State University, and the Wraparound Evaluation & Research Team
  – Sought to improve items based on face validity, question clarity, and to provide more variance/specificity

• **Testing**
  – WERT conducted internal pilots
    • 8 inter-rater reliability data points
    • 13 concurrent validity data points
      – Comparing the TOM, TOM 2.0, and WFI-EZ
## Tool Comparison by Structure

<table>
<thead>
<tr>
<th></th>
<th>TOM 1.0</th>
<th>TOM 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organization/Subscales</strong></td>
<td>10 Wraparound Principles</td>
<td>6 Wraparound Key Elements</td>
</tr>
<tr>
<td><strong>Items/Indicators</strong></td>
<td>20 subscales with 3-5 indicators</td>
<td>8 subscales with 5-6 indicators</td>
</tr>
<tr>
<td><strong>Number of scored items/indicators</strong></td>
<td>71</td>
<td>41</td>
</tr>
<tr>
<td><strong>Redundant items</strong></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Follow-up required with facilitator to score certain items</strong></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td><strong>Explicitly assesses completeness of team membership and attendance</strong></td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Aligned with other Wraparound Fidelity Assessment System Tools</strong></td>
<td>No</td>
<td>Yes – assesses fidelity to the same key elements of the Wraparound Fidelity Index (WFI-EZ) and DRM 2.0</td>
</tr>
<tr>
<td><strong>Wording emphasis</strong></td>
<td>On the facilitator’s behavior</td>
<td>On the team’s behavior</td>
</tr>
<tr>
<td><strong>Scoring system</strong></td>
<td>Yes, No, N/A</td>
<td>Yes, No, N/A</td>
</tr>
</tbody>
</table>
## Examples of indicator-level revisions

<table>
<thead>
<tr>
<th>TOM 1.0</th>
<th>TOM 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2b.</strong> The facilitator assists the team to review and prioritize family and youth needs.</td>
<td></td>
</tr>
<tr>
<td><strong>19b.</strong> The team prioritizes services that are community-based.</td>
<td></td>
</tr>
<tr>
<td><strong>19c.</strong> The team prioritizes access to services that are easily accessible to the youth and family.</td>
<td></td>
</tr>
<tr>
<td><strong>13b.</strong> The team assesses goals/strategies using measures of progress.</td>
<td></td>
</tr>
<tr>
<td><strong>3a.</strong> There is a clear agenda or outline for the meeting, which provides an understanding of the overall purpose of the meeting and the major sections of the meeting.</td>
<td></td>
</tr>
<tr>
<td><strong>3b.</strong> The meeting follows an agenda or outline such that team members know the purpose of their activities at a given time.</td>
<td></td>
</tr>
</tbody>
</table>
Examples of indicator-level revisions

<table>
<thead>
<tr>
<th>TOM 1.0</th>
<th>TOM 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2b.</strong> The facilitator assists the team to review and prioritize family and youth needs.</td>
<td><strong>4a.</strong> The team collectively identified, prioritized, and/or reviewed and confirmed the family and youth’s needs.</td>
</tr>
<tr>
<td><strong>19b.</strong> The team prioritizes services that are community-based.</td>
<td><strong>5d.</strong> If accessibility issues were raised, the team prioritized community-based services and supports that are easily accessible to the youth and family.</td>
</tr>
<tr>
<td><strong>19c.</strong> The team prioritizes access to services that are easily accessible to the youth and family.</td>
<td><strong>6c.</strong> The team monitored progress toward meeting needs and achieving outcomes/goals since the last meeting.</td>
</tr>
<tr>
<td><strong>13b.</strong> The team assessed goals/strategies using measures of progress.</td>
<td><strong>8b.</strong> The meeting followed a clear agenda that provided an understanding of the overall purpose of the meeting and the priority agenda items.</td>
</tr>
<tr>
<td><strong>3a.</strong> There is a clear agenda or outline for the meeting, which provides an understanding of the overall purpose of the meeting and the major sections of the meeting.</td>
<td><strong>3b.</strong> The meeting follows an agenda or outline such that team members know the purpose of their activities at a given time.</td>
</tr>
</tbody>
</table>
TOM 2.0 Internal Pilot
Reliability and Validity

• Very strong inter-rater reliability (n=8)
  – Cohen’s Kappa = .93

• Concurrent validity between TOM 1.0, TOM 2.0, and WFI-EZ is mixed; small sample size may have contributed to lower-than-expected correlations (n=13)
  – Concurrent validity is lower at the team-level when compared to correlations of site- or program-level data (Bruns et al., 2014)
    • Internal pilot data currently only available from two sites
  – We are currently collecting more pilot data from several sites
### TOM 2.0 Internal Pilot

#### Pilot Data

- Internal pilot data from two sites in Washington

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Site A (n=8)</th>
<th>Site B (n=13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Meeting Attendance</td>
<td>55%</td>
<td>63%</td>
</tr>
<tr>
<td>Effective Teamwork</td>
<td>87%</td>
<td>87%</td>
</tr>
<tr>
<td>Determined by Families</td>
<td>95%</td>
<td>91%</td>
</tr>
<tr>
<td>Based on Priority Needs</td>
<td>81%</td>
<td>75%</td>
</tr>
<tr>
<td>Natural &amp; Community Supports</td>
<td>64%</td>
<td>74%</td>
</tr>
<tr>
<td>Outcomes-Based Process</td>
<td>47%</td>
<td>70%</td>
</tr>
<tr>
<td>Driven by Strengths</td>
<td>75%</td>
<td>76%</td>
</tr>
<tr>
<td>Skilled Facilitation</td>
<td>92%</td>
<td>91%</td>
</tr>
<tr>
<td><strong>Key Elements Score</strong></td>
<td><strong>75%</strong></td>
<td><strong>79%</strong></td>
</tr>
<tr>
<td><strong>Overall TOM 2.0 Score</strong></td>
<td><strong>74%</strong></td>
<td><strong>77%</strong></td>
</tr>
</tbody>
</table>
TOM 2.0 External Pilot

Status of Pilot Sites

• Based on the internal pilot, modifications were made to 46% of the items.

• 8 sites have signed up to pilot the TOM 2.0
  – 6 are existing TOM 1.0 collaborators
    • Several said that they were eager to pilot a more user-friendly tool
  – 2 are new WFAS collaborators
Improved tool!

• **TOM 2.0 has Increased item variability compared to TOM 1.0**
  – Average non-attendance-related item SD is higher (.33 vs. .20)
  – Number of non-varying items is lower (21.5% vs. 48.6%)

• **Improved end-user experience**
  – Observers universally assessed the TOM 2.0 as being easier to use, resulting in lower cognitive burden than the TOM 1.0
  – End-users also felt that TOM 2.0’s data was conceptually clearer and more useful, especially when viewed alongside data from the WFI-EZ
Next Steps

• Continued refinement, testing, and dissemination is warranted
  – TOM 2.0 shows promising signs of providing the field with a robust instrument to rate activities and behaviors observed in vivo, both for training and coaching purposes, and for quality improvement

• Complete external pilots

• Build site-level data set for further analysis
  – By aggregating data from our Indiana pilot into site-level data rather than simply team-level data, we may be able to better measure concurrent validity

• Build tool into WrapTrack
Agenda for Today’s Symposium

• WFAS Overview
• Reviving the DRM
• Refining the TOM
• WrapSTAR