**Validation of the Mayo High Performance Teamwork Scale and Ottawa Crisis Resource Management Global Rating Scale for Evaluation of Crisis Team Training**

Iwasita, Hiroki1; Emlet, Lillian2; Berg, Benjamin3; Kase, Yoichi2; Uchida, Kotaro2; Ren, Dianxue4; Lutz, John5; Henker Richard6

Dokyo Medical University 1, UPMC2, SimTiki Simulation Center University of Hawaii3, Jikei University4, WISER5, University of Pittsburgh6

University of Pittsburgh Medical Center (UPMC)

Peter M. Winter Institute for Simulation, Education and Research (WISER) Pittsburgh, PA 15213

---

**OBJECTIVE**

Evaluate the reliability and validity of  
- Mayo High Performance Teamwork Scale (MHPTWS)  
- Ottawa Crisis Resource Management Global Rating Scale (Ottawa GRS)  

in a Crisis Team Training (CTT) course at WISER

---

**INTRODUCTION**

- Evaluation is a key component of simulation education  
- Evaluation of team training with simulation is challenging (AMEE Guide No. 82, Med Teach. 35:10 e1511-30, 2013)  
- Development of reliable and validated assessment instruments requires considerable time and effort  
- Several non-technical skill or team performance assessment tools have been developed and used in various settings  
- MHPTWS was developed to assess teamwork skills in CTT training (Malec et al. Sim Health Vol.2 No.1 2007)  
- Ottawa GRS is a reliable and valid assessment tool in team training (Kim et al. Crit Care Med 2006 Vol.34 No.8)  
- Ottawa GRS was developed to assess teamwork skills in CRM

---

**BACKGROUND**

Crisis Team Training Course Overview

---

**METHODS**

- **Assessment Tools**  
  - MHPTWS  
  - Ottawa GRS

- **Standardized Rater Training**  
  - 1.5 hours  
  - Brief orientation to CTT  
  - Review of assigned instrument  
  - Practice evaluation with 3 sessions  
  - Same training for novices and experts

- **Data Collection**  
  - iPad data entry to Form®: Downloaded to Excel  
  - SPSS statistical analysis

---

**RESULTS**

Table 1. MHPTWS inter-rater reliability Kappa Statistic

<table>
<thead>
<tr>
<th>Question</th>
<th>Novice - Novice</th>
<th>Expert - Novice</th>
<th>Expert - Expert</th>
<th>Kappa (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A leader is clearly identified by all team members</td>
<td>0.33 (0.06, 0.60)</td>
<td>0.40 (0.32, 0.48)</td>
<td>0.30 (0.16, 0.43)</td>
<td>0.37 (0.30, 0.43)</td>
</tr>
<tr>
<td>2. The team leader ensures maintenance of an appropriate balance between technical operations and patient care</td>
<td>0.35 (0.09, 0.61)</td>
<td>0.29 (0.09, 0.49)</td>
<td>0.40 (0.26, 0.55)</td>
<td>0.36 (0.13, 0.59)</td>
</tr>
<tr>
<td>3. Each team member demonstrates a client-centered orientation in patient care</td>
<td>0.38 (0.21, 0.56)</td>
<td>0.41 (0.18, 0.64)</td>
<td>0.14 (0.01, 0.28)</td>
<td>0.30 (0.01, 0.50)</td>
</tr>
<tr>
<td>4. A quiet, calm environment is provided for all patients and their families</td>
<td>0.35 (0.10, 0.61)</td>
<td>0.37 (0.01, 0.74)</td>
<td>0.21 (0.00, 0.42)</td>
<td>0.33 (0.05, 0.61)</td>
</tr>
<tr>
<td>5. When team members are acting as their own patient, they verbalize their understanding of their role</td>
<td>0.38 (0.24, 0.53)</td>
<td>0.33 (0.05, 0.61)</td>
<td>0.15 (0.06, 0.25)</td>
<td>0.36 (0.13, 0.59)</td>
</tr>
</tbody>
</table>

Table 2. MHPTWS score 1st session vs. 4th session

<table>
<thead>
<tr>
<th>1st Session Score</th>
<th>4th Session Score</th>
<th>Paired t-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice A</td>
<td>8.69±3.49</td>
<td>11.90±2.15</td>
</tr>
<tr>
<td>Novice B</td>
<td>8.80±3.67</td>
<td>11.55±2.47</td>
</tr>
<tr>
<td>Expert A</td>
<td>7.80±3.64</td>
<td>11.38±2.58</td>
</tr>
<tr>
<td>Expert B</td>
<td>8.20±3.76</td>
<td>11.05±2.37</td>
</tr>
</tbody>
</table>

The expert to expert inter-rater reliability for the MHPTWS was evaluated in 29 of 104 sessions. Due to the small number of sessions evaluated by Expert B, the 1st and 4th session comparisons are pending rating of the remaining videos.

---

**DISCUSSION**

- There is only average inter-rater reliability for both instruments
- Our modest inter-rater reliability in our study was equivalent to the original validation of Ottawa GRS
- Both instruments were able to significantly detect differences in team performance between the 1st and 4th scenarios
- Both tools seem to perform similarly in ability to detect differences between our simulation scenarios of CTT

---

**REFERENCES**