Psychometric Study of the Occupational Self Assessment With Adolescents After Infectious Mononucleosis

Renee Taylor, Sun Wook Lee, Jessica Kramer, Yukiko Shirashi, Gary Kielhofner

KEY WORDS
- activities of daily living
- adolescent behavior
- infectious mononucleosis
- psychometrics
- reproducibility of results

We examined the psychometric characteristics of the Occupational Self Assessment (OSA), which measures clients’ perceptions of their own competence and the value they assign to occupations. Two hundred ninety-six adolescents with acute mononucleosis completed the OSA, the Fatigue Scale, the Checklist of Infectious Symptoms, the Child Health Questionnaire, and the Perceived Stress Scale. OSA items coalesced to capture the intended constructs; the rating scales functioned as intended. More than 90% of adolescents were validly measured. The OSA showed adequate sensitivity and was stable over time. OSA measure of competence was moderately associated with infectious symptoms, fatigue severity, health status, and stress, and the measure of values was not. Neither measure was associated with age, gender, or ethnicity. Finally, adolescents who had not recovered from mononucleosis after 12 mo showed lower competence scores yet attached the same value or importance to occupational participation as adolescents who had recovered.


We examined the reliability, sensitivity, and validity of the Occupational Self Assessment (OSA; Baron, Kielhofner, Iyenger, Goldhammer, & Wolenski, 2006) in a sample of adolescents during their recovery from infectious mononucleosis (mono). Severe fatigue is common during the acute phase of mono (Buchwald, Rea, Katon, Russo, & Ashley, 2000; Hickie et al., 2006; White et al., 2001) and becomes chronic in some people (Buchwald et al., 2000). Occupational consequences of fatigue in adolescents include prolonged absence from school, disruption of family activities, and loss of contact with peers (Buchwald et al., 2000).

The OSA is a client-centered evaluation tool designed to capture people’s perceptions of how illness and disability affect their occupations. Based on the Model of Human Occupation (Kielhofner, 2008), the OSA allows clients to rate their own competence and assign value to 21 areas of performance and participation.

A series of studies using Item Response Theory, through the Rasch measurement model (Rasch, 1960/1980; Smith, Wakely, de Kruif, & Swartz, 2003; Wolfe & Chiu, 1999; Wright & Stone, 1979), has examined the OSA’s internal validity (i.e., whether the items cohered to capture the intended constructs and whether the hierarchy of items representing less or more of the intended trait was meaningful) and instrument sensitivity. The first study of the OSA provided evidence of its internal validity but identified the rating scale as not adequately sensitive (Kielhofner & Forsyth, 2001). Three iterative studies found that a 4-point rating scale resulted in an adequately sensitive instrument (Kielhofner, Forsyth, Kramer, & Iyenger, 2009). Across the studies, items
cohered to capture the two measured constructs, and the item hierarchy showed a meaningful progression that was similar across the studies.

Another study of 112 people demonstrated that the hierarchies of competence and values items were stable across two administrations of the OSA, indicating that the instrument measures the same constructs at both initial and subsequent evaluation (Kielhofner, Dobria, Forsyth, & Kramer, 2010). Moreover, the OSA competence scale detected significant change in 32% of respondents, and the values scale detected significant change in 49% of respondents.

Aim

We continued the psychometric examination of the OSA by focusing on an adolescent population and by combining an Item Response Theory approach via the Rasch Measurement Model (Rasch, 1960/1980; Wright & Stone, 1979) with Classical Test Theory (Allen & Yen, 2002; Lord & Novick, 1968; Novick, 1966) approaches to establishing assessment validity. This combined approach enables a more comprehensive assessment of validity and has been applied to the investigation of other self-report assessments (Asgari & Kramer, 2008; Kyngdon, 2004; Mok, 2004).

Method

Validity refers to an instrument’s capacity to measure the intended construct (Bensen & Schell, 1997). Both Item Response Theory and classical test theory are concerned with validity. However, they approach the problem of providing empirical evidence related to validity in different ways, as noted in the following sections.

Rasch Model Approach

Item Response Theory (Rasch analysis) generates evidence that describes the underlying characteristic, or latent trait, that the scale is intended to capture (Wolfe & Smith, 2007). One type of evidence is whether the items effectively capture the intended trait; each item is analyzed as to whether it “fits” with the rest of the items in targeting the trait. To be considered structurally valid, most of a scale’s items should have adequate fit. In previous studies (Kielhofner & Forsyth, 2001; Kielhofner et al., 2009, 2010), item fit to each scale provided evidence of good structural validity. Therefore, we expected 90% of OSA items to fit the constructs of occupational competence and value.

Rasch analysis also generates evidence of validity by examining the continuum of the latent trait formed by the different items from less to more. Rasch analysis orders (i.e., calibrates) items on this continuum, allowing a determination of whether the item order makes sense from a theoretical perspective, a practical perspective, or both. Because previous research has found that the OSA competence and values continua are theoretically interpretable in an adult population, we expected them to be similarly ordered in this adolescent population.

The Rasch model can also assess substantive validity, which is the extent to which theory explains how and why people completing an assessment differ in their responses (Wolfe & Smith, 2007). As with items, a person fit statistic provides an indication of whether each person’s response pattern fits the underlying item hierarchy (an indication that the person was validly measured). A second concern is whether the participants collectively used the rating scale in the intended manner, that is, as a 4-point continuum measuring less to more competence and value. For instance, OSA respondents rate items in terms of their own competence ("I have a lot of problem doing this," “I have some difficulty doing this,” “I do this well,” “I do this extremely well”). These ratings are intended to be ordered from less to more, and if they are not used as intended, they threaten the scale’s validity. Finally, critical questions can be asked about whether the ordering of items and the rating scale are stable over time. We examined the stability of the rating scales over time in the current study with a subsample for which we had baseline and 1-year follow-up OSA data. We also examined whether the OSA rating scales were able to detect change in this group.

Classical Test Theory Approach

From the perspective of classical test theory, one examines relationships on the basis of theoretical assumptions about the targeted construct and its relationship to other variables. Convergent validity refers to the expectation that measures intended to capture related traits should be correlated. Divergent validity refers to the expectation that measures of unrelated traits should not be correlated (Campbell & Fiske, 1959; Kielhofner, 2006). In this study, we tested whether the correlation of the two OSA variables (competence and value) with several other variables would be as expected on the basis of their theoretical relationships.

Testing of discriminant validity requires groups who are known to differ on some variable relevant to the trait and who are thus expected to show differences in the trait (Campbell & Fiske, 1959; Kielhofner, 2006). A scale’s validity is tested by asking whether the groups will differ significantly on the measures obtained with the scale. In
this study, we compared adolescents who had not recovered from fatiguing symptoms at 12-mo follow-up with matched controls who had recovered. We expected that those adolescents with persistent fatigue would show lower measures on self-reported competence while attaching the same value to occupational participation as the nonfatigued control participants.

Hypotheses

We tested nine hypotheses in this study. Each hypothesis, along with its association with Rasch or classical test theory, the logic of the hypothesis, and the analysis and criteria used to test the hypothesis, is shown in Table 1.

Sample

The Institutional Review Board of the University of Illinois at Chicago approved this study, which was conducted as part of a larger epidemiological study investigating the incidence, correlates, and consequences of severe fatigue in adolescents with mono (Katz, Shiraishi, Mears, Binns, & Taylor, 2009). Participants were 301 adolescents ages 12–18 who had recently been diagnosed with acute mono. They were recruited from schools, hospitals, outpatient clinics, and private physician practices. Of this sample, 296 completed both OSA scales and all the other self-report measures required for this study. The follow-up sample consisted of 31 participants who did not consider themselves to be fully recovered and 59 matched control participants from the group who did consider themselves fully recovered from mono at 12 mo after initial enrollment and after the acute phase of mono. Two matched controls were identified for each participant with persistent fatigue; 3 of the matched controls did not complete the self-report measures necessary for inclusion in this study.

Measures

In addition to the OSA, we used the following measures for this study:

- The Checklist of Infectious Symptoms (Buchwald et al., 2000) is a self-report measure of the presence and severity of mono symptoms. It was developed for and used in a large-scale adult study of chronic fatigue syndrome after mono (Buchwald et al., 2000).

- The Fatigue Scale (Chalder et al., 1993), which evaluates fatigue severity, has been found to be reliable and valid (Chalder et al., 1993) and has been used extensively in a variety of community-based and clinic-based studies of fatigue (e.g., Pawlikowska et al., 1994).

- The Global Health Scale from the Child Health Questionnaire (Landgraf, Abetz, & Ware, 1996) is a measure of health-related quality of life with well-established psychometric properties.

- The Perceived Stress Scale (Cohen, Kamarck, & Mermelstein, 1983) measures self-reported stress in community-based populations; it has well-established psychometric properties and has been used in numerous studies of populations with and without disability.

Data Collection Approach

Participants completed the OSA and the other measures at home in the presence of a research assistant. Baseline data were gathered during the acute stage of mono. Follow-up data were collected 12 mo after the initial infection with mono.

Analysis

Hypotheses derived from Item Response Theory were tested using a Rasch Rating Scale Model (Wright & Masters, 1982) via Winsteps Version 3.49 (Linacre, 2004). The extent to which items captured the underlying trait was evaluated using item fit statistics. In keeping with recommended standards (Bray, Fisher, & Duran, 2001; Velozo et al., 1999; Wright & Linacre, 1994), fit statistics of $MnSq > 1.4$ associated with $Zstd > 2.0$ were taken as indications of item misfit. The OSA’s sensitivity was evaluated by converting the person separation index into strata, or $[4(separation \text{ index}) + 1]/3$, which indicate the number of significant different levels of the construct represented by the items (Wright & Masters, 1982).

The rating scales were examined to determine whether clients were using them as intended, as evidenced by incremental increase in the average client measures associated with each rating category, in the threshold difficulties of each rating scale, and by outfit $MnSq$ fit statistics for each category $< 2.0$ (Linacre, 2002).

We explored the extent to which clients were validly measured by determining the percentage who had acceptable fit statistics at baseline and 12-mo responses to the OSA. Client fit statistics of $MnSq > 1.4$ associated with $Zstd > 2.0$ were taken as indications of a misfit.

When analyzing participants’ responses over time, one must distinguish changes in estimated measures related to actual change in the intended construct from change related to measurement error (Wolfe & Chiu, 1999). In self-report test–retest situations, if participants complete the assessment the same way each time, one can have much higher confidence that changes in measures are the result of actual changes in the underlying trait as opposed to measurement error. To evaluate the stability of the OSA rating scales and items over time, we followed
Table 1. Study Hypotheses

| Hypothesis                                                                 | Logic                                                                 | Approach     | Evidence Criteria                                                                 | Corresponding Results |
|---------------------------------------------------------------------------|                                                                      |             |                                                                                     |                        |
| 1. OSA items will coalesce to capture constructs of competence and values of adolescents with mono. | Findings should replicate previous findings that OSA items cohered in capturing the latent traits of competence and values (Kielhofner et al., 2009, 2010; Kielhofner & Forsyth, 2001). | Rasch analysis | 90% of items have acceptable fit to the Rasch model (i.e., MnSq statistic <1.4 associated with a Zstd < 2). | Table 2               |
| 2. The occupational competence and value rating scales will function as intended among adolescents with mono. | Rating scales should be used by respondents as they were designed to be used, which means that the 4-point scale as used by the OSA should demonstrate the intended ordering from less to more competence and value. | Rasch analysis | Average client measures associated with each rating category and threshold difficulties increase incrementally, and outfit MnSq fit statistics for each category are <2.0 (Linacre, 2002). | Table 3               |
| 3. >90% of all adolescents will be validly measured by the OSA at both baseline and 12 mo. | For a scale to be considered valid, most items should have adequate fit. | Rasch analysis | 90% of people will have acceptable fit to the Rasch model (i.e., MnSq statistic <1.4 associated with a Zstd < 2). | Text in Hypothesis 3 section |
| 4. The OSA rating scales will show adequate sensitivity to discriminate between adolescents with different levels of occupational competence and value for their occupation at baseline. | The revised OSA items measured 6 significantly different levels of competence and value in previous research (Kielhofner et al., 2009). | Rasch analysis | ≥2 strata will be present when the person separation index is converted into strata, using the equation [4(separation index) + 1]/3. | Text in Hypothesis 4 section |
| 5. OSA occupational competence and value rating scales and the hierarchy of each scale will not be substantially different over time. | Item hierarchy should not differ significantly over time and should remain stable (Kielhofner et al., 2009, 2010). | Rasch analysis | Item measures and rating scale threshold difficulties from separate initial and baseline Rasch analysis will not be significantly different. | Supplemental Tables 1, 2, 3, and 4. |
| 6. The OSA competence and values scales will detect significant change in 30% of the adolescents. | In a previous study, 32% and 49% of clients had significantly different measures at Time 2 on the Competence and Values scales, respectively (Kielhofner et al., 2010) | Rasch analysis | OSA values and competence scales will detect statistically significant increases or decreases in 30% of adolescents. Patterns of detected change will make sense for the recovered and chronically fatigued groups. | Text in Hypothesis 6 section |
| 7. Occupational competence and value measures among adolescents with infectious mononucleosis will not be associated with their age, gender, or ethnicity. | Tests designed to measure different traits are expected to show patterns of association that discriminate between the traits. | Independent samples t test, one-way analysis of variance, and Pearson correlation coefficients | No significant difference will be detected in OSA measures between groups for age, gender, and ethnicity. | Supplemental Tables 5 and 6. |
| 8. Client occupational competence measures will show significant associations with measures of symptoms, fatigue, health status, and stress and value measures will not. | Because fatigue affects performance, it should be correlated with clients’ perception of their own competence but not with the value they attach to areas of performance. | Pearson correlation coefficient | Moderate relationships will be detected between competence and symptoms, fatigue, health status, and stress; negligible or nonsignificant relationships will be detected between OSA values and these variables | Table 4               |
| 9. Adolescents who have not recovered will show lower measures on self-reported competence while still attaching the same value to occupational participation as recovered adolescents. | Adolescents with persistent fatigue should show lower self-report of competence but attach similar value to areas of performance when compared with adolescents without fatigue. | t test | Difference in competence measure will be detected between adolescents in nonrecovered group and recovered adolescents at 12 mo. | Text in Hypothesis 9 section |

Note: OSA = Occupational Self Assessment.
the procedure described by Wolfe and Chiu (1999). First, each time point for the 90 participants for whom we had repeated measures was analyzed separately using a Rating Scale Model approach. Then resulting item and rating scale Rasch estimates from baseline and 12 mo were compared for significant differences over time as given by the following formula:

\[
\frac{\text{Time 1 item measure} - \text{Time 2 item measure}}{\sqrt{(\text{Time 1 item measure error})^2 + (\text{Time 2 item measure error})^2}}
\]

Values >±2 are taken as indications of significant differences in assessment parameters between administrations. In this instance, steps must be taken to scale readministrations onto a common metric and adjust person measures accordingly, a process that ensures a rigorous evaluation of reported change over time. In this study, because rating scale steps showed small but significant changes, we obtained initial and 12-mo OSA client measures using a common rating scale structure derived from a combined, simultaneous analysis of both time points. Finally, to examine the ability of the OSA scales to detect change in people, we examined the number of adolescents who showed significant change from baseline to 1-year follow-up. For each adolescent, we derived a change score by means of the following formula, which accounts for measurement error associated with each estimated measure:

\[
\sqrt{(\text{Time 1 adolescent measure} - \text{Time 2 adolescent measure})^2 + (\text{Time 1 adolescent measure error})^2 + (\text{Time 2 adolescent measure error})^2}
\]

Values >±2 indicate significant differences between administrations.

Hypotheses derived from classical psychometric theory were tested using \(t\) tests, Pearson’s correlation coefficients, and analysis of variance. Because of multiple comparisons, we used a conservative \(\alpha\) level (\(p < .01\)) to determine significance.

Results

Hypothesis 1

At baseline, 90.48% of the occupational competence scale items and 90.48% of the values scale items met fit requirements (Table 2). The findings supported the hypothesis that the scale items would coalesce to capture the intended constructs.

Hypothesis 2

At baseline, the occupational competence rating scale was used as intended because average client measures and thresholds were ordered and all rating scales had acceptable

### Table 2. Occupational Self Assessment Competence and Values Item Fit Statistics at Baseline (\(N = 296\))

<table>
<thead>
<tr>
<th>Item</th>
<th>Competence</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\text{Infit}_1)</td>
<td>(\text{MnSq}_1)</td>
</tr>
<tr>
<td>Concentrating on my tasks</td>
<td>0.92</td>
<td>1.00</td>
</tr>
<tr>
<td>Physically doing what I need to do</td>
<td>1.04</td>
<td>0.40</td>
</tr>
<tr>
<td>Taking care of the place where I live</td>
<td>1.18</td>
<td>2.00</td>
</tr>
<tr>
<td>Taking care of myself</td>
<td>0.96</td>
<td>−0.60</td>
</tr>
<tr>
<td>Taking care of others for whom I am responsible</td>
<td>0.77</td>
<td>−3.00</td>
</tr>
<tr>
<td>Getting where I need to go</td>
<td>1.01</td>
<td>0.20</td>
</tr>
<tr>
<td>Managing my finances</td>
<td>1.31</td>
<td>2.80</td>
</tr>
<tr>
<td>Managing my basic needs (food, medicine)</td>
<td>0.91</td>
<td>−1.10</td>
</tr>
<tr>
<td>Expressing myself to others</td>
<td>1.58a</td>
<td>5.80a</td>
</tr>
<tr>
<td>Getting along with others</td>
<td>1.28</td>
<td>3.20</td>
</tr>
<tr>
<td>Identifying and solving problems</td>
<td>0.90</td>
<td>−1.20</td>
</tr>
<tr>
<td>Relaxing and enjoying myself</td>
<td>1.50a</td>
<td>5.20a</td>
</tr>
<tr>
<td>Getting done what I need to do</td>
<td>0.78</td>
<td>−2.80</td>
</tr>
<tr>
<td>Having a satisfying routine</td>
<td>0.82</td>
<td>−2.30</td>
</tr>
<tr>
<td>Handling my responsibilities</td>
<td>0.69</td>
<td>−4.20</td>
</tr>
<tr>
<td>Being involved as a student, worker, volunteer, or family member</td>
<td>1.05</td>
<td>0.60</td>
</tr>
<tr>
<td>Doing activities I like</td>
<td>1.27</td>
<td>3.10</td>
</tr>
<tr>
<td>Working toward my goals</td>
<td>0.72</td>
<td>−3.70</td>
</tr>
<tr>
<td>Making decisions based on what I think is important</td>
<td>0.99</td>
<td>−0.10</td>
</tr>
<tr>
<td>Accomplishing what I set out to do</td>
<td>0.72</td>
<td>−3.60</td>
</tr>
<tr>
<td>Effectively using my abilities</td>
<td>0.72</td>
<td>−3.70</td>
</tr>
</tbody>
</table>

*a*Item exceeds Rasch fit criteria with \(\text{MnSq} > 1.4\) associated with \(\text{Zstd} > 2\).
fit statistics. The values rating scale was also used as intended at baseline (Table 3).

**Hypothesis 3**

Fit statistics could not be calculated on the small numbers of participants who used the highest ratings for all items; percentages are based on the total number of participants for whom fit statistics could be calculated. At baseline, 7.12% of participants misfit on the competence scale and 8.87% misfit on the values scale. At 12-mo follow-up, 8.23% of participants misfit on the competence scale and 6.74% misfit on the values scale. Thus, in all instances >90% of adolescents met criteria for being validly measured.

**Hypothesis 4**

At baseline, the OSA competence items measured almost four significantly different levels (3.99 strata) of competence among adolescents (person separation index = 2.74). The OSA value items also measured four significantly different levels (4.03 strata) of value among adolescents (person separation index = 2.77).

**Hypothesis 5**

The OSA had evidence of minimal variance between baseline and 12 mo. First, the competence item calibrations remained stable over time (Supplemental Table 1\(^1\)). Only one competence item became significantly more challenging at 12 mo. The OSA competence rating scale thresholds were not all stable over time; adolescents felt more competent and had an easier time using the highest rating category of extremely well at baseline (Supplemental Table 2). However, the estimates were separated only by 0.44 logits, and the associated errors were small. That is, the difference was significant, yet small.

Most of the items on the values item hierarchy were the same over time (Supplemental Table 3). Three OSA values items were significantly different at 12 mo. The OSA value rating scale thresholds were not all stable over time; adolescents indicated more value for more items and had an easier time using the highest rating category of most important at baseline (Supplemental Table 4). However, as with the competence rating scale, the baseline and 12-mo parameter estimates differed by a small amount (0.29 logits), and the associated errors were small. That is, the difference was significant, yet small.

\(^1\)Supplemental Tables 1–6 are available online at www.ajot.aotpress.net (navigate to this article, and click on “supplemental materials”).

**Table 3. Occupational Competence and Values Rating Scale Function at Baseline \((N = 296)\)***

<table>
<thead>
<tr>
<th>Measure</th>
<th>Threshold</th>
<th>Difficulties</th>
<th>MnSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competence</td>
<td>I have a lot of problem doing this.</td>
<td>-0.97</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>I have some difficulty doing this.</td>
<td>0.17</td>
<td>-2.35</td>
</tr>
<tr>
<td></td>
<td>I do this well.</td>
<td>1.34</td>
<td>-0.37</td>
</tr>
<tr>
<td></td>
<td>I do this extremely well.</td>
<td>2.61</td>
<td>2.72</td>
</tr>
<tr>
<td>Values</td>
<td>This is not so important to me.</td>
<td>-0.74</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>This is important to me.</td>
<td>-0.01</td>
<td>-2.20</td>
</tr>
<tr>
<td></td>
<td>This is more important to me.</td>
<td>0.86</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>This is most important to me.</td>
<td>1.80</td>
<td>1.63</td>
</tr>
</tbody>
</table>

To account for changes in the rating scale over time, initial and 12-mo OSA client measures were obtained using a common rating scale structure derived from a combined, simultaneous analysis of both time points. Although four items varied over time, no item varied over time on both scales. Therefore, the OSA items were considered invariant over time and considered to also have a shared item estimate at both time points. Deriving person measures from this common rating scale places the baseline and 12-mo assessments on a shared and common metric, or frame of reference, and derives a more conservative but more rigorous estimate of change (Wolfe & Chiu, 1999). These adjusted baseline and 12-mo measures were used to answer Hypotheses 7 and 8.

**Hypothesis 6**

The OSA competence scale detected significant change in 34% of the participants, whereas the values scale detected significant change in 38%. The OSA competence scale detected significant increases in 34% of adolescents who were recovered and in 19% of those with ongoing fatigue. The OSA competence scale detected significant decreases in 3% of those who had recovered and 10% of those with ongoing fatigue. The OSA values scale detected similar rates of significant increases and decreases in both groups. Twenty-four percent of participants who had recovered and 26% of those with ongoing fatigue showed increases, whereas 14% of those who had recovered and 13% of those with ongoing fatigue showed decreases.

**Hypothesis 7**

The adjusted OSA competence and OSA value measures showed no significant associations with gender and ethnicity at baseline, as shown in Supplemental Tables 5 and 6. OSA competence and value baseline measures and age showed weak or no associations (competence: \(R = .05, p = .38\); value: \(R = .17, p < .01\), two-tailed).

\(\text{Average Client} \quad \text{Threshold} \quad \text{Difficulties} \quad \text{MnSq}\)

\(\begin{array}{cccc}
\text{Competence} & \text{Measure} & \text{Difficulties} & \text{MnSq} \\
\text{I have a lot of problem doing this.} & -0.97 & — & 1.09 \\
\text{I have some difficulty doing this.} & 0.17 & -2.35 & 1.00 \\
\text{I do this well.} & 1.34 & -0.37 & 0.92 \\
\text{I do this extremely well.} & 2.61 & 2.72 & 1.00 \\
\text{Values} & \text{This is not so important to me.} & -0.74 & — & 1.09 \\
\text{This is important to me.} & -0.01 & -2.20 & 0.95 \\
\text{This is more important to me.} & 0.86 & 0.56 & 0.91 \\
\text{This is most important to me.} & 1.80 & 1.63 & 1.01 \\
\end{array}\)
Hypothesis 8

Table 4 shows the correlations between the adjusted OSA measures and the measures of infectious symptoms, fatigue, health status, and stress. As anticipated, the OSA competence scale was moderately correlated with infectious symptoms, fatigue, health status, and perceived stress. As expected, the OSA value measures were not correlated with infectious symptoms or fatigue severity. Significant but minimal correlations were found with stress and global health.

Hypothesis 9

t tests comparing adolescents who had not recovered at 12-mo follow-up with matched control adolescents who had recovered revealed that nonrecovered adolescents had lower adjusted competence measures (M = 54.98, SD = 8.79) than those who had recovered (M = 62.21, SD = 8.33, t[88] = −3.84, p = .00). Moreover, as hypothesized, these groups showed no significant difference in the reported values and value measures.

Discussion

This psychometric study provides evidence of the OSA's validity, reliability, sensitivity, and stability in a sample of adolescents. First, OSA items coalesced to capture the intended constructs. Findings support use of the OSA as a valid tool to measure what adolescents perceive as areas of competence and what they perceive as important in terms of their everyday activities. All the competence items meaningfully cohered to measure the construct of competence in the sample, and all the value items that are in the OSA cohered to measure the construct of values.

In addition, findings indicated that the OSA rating scales were used as intended—that is, as continua representing less to more competence and values. The adolescents appropriately used the higher rating categories to indicate more competence and value and the lower rating categories to indicate less competence and less value.

Moreover, >90% of the adolescents were validly measured and gave responses consistent with Rasch model expectations. A typical criterion applied to person misfit is that no more than 5% of the respondents should have poor fit (Smith, Conrad, Chang, & Piazza, 2002). However, 10% has been considered an acceptable target for the OSA because it is completed without the type of extensive training associated with professionally administered assessments (Kielhofner et al., 2009). Identifying clients whose pattern of responses misfit can be done using the key forms found in the OSA manual (Baron et al., 2006) and should be followed with a clarifying discussion to make sure the clients’ responses are understood.

Another indicator of psychometric stability was that the competence and value items had the sensitivity to measure approximately four significantly different levels of competence and value among adolescents. This finding supports the OSA’s use as a tool for identifying group differences in reported competence and value.

Last, this study illustrated the use of a Rasch-based procedure (Wolfe & Chiu, 1999) to examine the stability of OSA items and rating scales over time when used by adolescents with mono. The relative difficulty and value of everyday tasks does not change over time for this group, as indicated by the relative stability of item measures over time. This group, however, is slightly more selective in the extent to which it reports the highest level of competence and value for everyday activities, as shown by the significant differences in the highest rating scale thresholds over time. Although statistically significant, the difference in rating scale use over time was small and may be considered negligible for use in practice.

The OSA competence scale detected significant change in approximately one-third of the adolescents from baseline to 1 year later. As expected, recovered adolescents showed a higher rate of increase and a lower rate of decrease in competence scores than did adolescents with ongoing fatigue. The OSA values scale detected significant change in about one-quarter of both groups who showed increases, whereas 13% with ongoing fatigue and 14% of those recovered had significantly decreased scores. The values change scores likely reflect a combination of developmental changes and the impact of an illness experience on adolescents’ thoughts on what was important. That the values scale detected somewhat more change than the competence scale is consistent with previous research on the OSA (Kielhofner et al., 2010). The findings indicate that the OSA can detect statistically significant and meaningful change in people.

The OSA competence scale was associated with measures of symptoms, fatigue, health status, and stress, as
expected. Moreover, as expected, the OSA values scale was not associated with those measures. Finally, no association was found between adolescents’ perceived competence and values and the sociodemographic variables of age, gender, or ethnicity. These findings provide evidence of the OSA’s convergent and divergent validity.

Twelve months after infection, when nonrecovered adolescents were compared with adolescents who recovered, the former showed lower measures on the competence scale but not on the values scale. This finding was expected, because we hypothesized that although competence would be affected by prolonged fatigue, values would be likely to be stable over time. This finding provides evidence of discriminant validity.

Conclusion

Findings from this study contribute additional evidence of the OSA’s reliability and validity from both an item response and a classical test theory approach. Findings also indicate that the OSA is an adequately sensitive measure capable of discriminating between clients with different competence and values and of detecting meaningful change. Finally, the study provided evidence that the OSA can be used in a valid and reliable manner with adolescents. ▲

Acknowledgments

We thank the Chicago Children’s Memorial Hospital’s Pediatric Practice Research Group and all the physicians and school nurses who referred patients into the study. We also thank student volunteers who contributed to the recruitment of physicians and school nurses. This work was supported by grant R01HD4330101A1 from the National Institute of Child Health and Human Development.

References


