Teacher Rating of Student Adjustment
Grade 6 /Year 7

Fast Track Project Technical Report
Clara G. Muschkin
June 16, 2003

Table of Contents

Citations

Instrument
Conduct Problems Prevention Research Group (CPPRG). (1995).
Teacher Rating of Student Adjustment [On-line]. Available: http://www.fasttrackproject.org/

Report
Muschkin, C. G. (2003). *Teacher Rating of Student Adjustment* (Fast Track Project Technical Report) [On-line]. Available: http://www.fasttrackproject.org/

Research Paper
Muschkin, C. G. and Malone, P. S. (2003). Multiple Teacher Ratings: an Evaluation of Measurement Strategies. Manuscript in preparation: Duke University.

Data Sources

Unscored: T7K
Scored: TSA7

I.   Scale Description

The Teacher Rating of Student Adjustment is a 7-item instrument developed by the Fast Track Project to assess dimensions of success in adjusting to middle and high school. The first item queries how well the teacher knows the child; the remaining items target the teacher's perceptions of a student's academic performance, academic motivation, social skills, adult relationships, conduct, and personal maturity. Responses are coded on a five-point scale ranging from 1 to 5, as follows: Poor, unsatisfactory skills (1); Below average skills (2); Average skills (3); Above average skills (4); and Excellent skills (5).

II.   Report Sample

This report contains data collected on Cohort 1, Year 7. The data include a high-risk control sample (n=155) and a normative sample (n = 387) for a total N = 463, including overlap. Of the 463 subjects, 72 are missing all responses for this measure, including 61 from the normative group (13 from Durham, 11 from Washington, 6 from Pennsylvania, and 19 from Nashville), and 23 from the control group (7 from Durham, 6 from Nashville, 6 from Pennsylvania, and 4 from Washington), with overlap between the normative and control groups.

III. **Scaling**

The data from this measure are distinctive in that multiple teacher ratings of the target behaviors were obtained for each student in grades 6, 7, and 8. The goal was to administer the instrument to teachers in each of the student's core classes, as most middle school students move among several classrooms for core academic subjects. The decision to use multiple informants raises methodological issues that impact scaling and analysis of these data. These issues are addressed in the research paper cited above (Muschkin and Malone, 2003), and are briefly referenced below.

One possible problem with the use of these data is the potential lack of concordance among the teacher ratings for each student. Low levels of agreement among the raters may indicate that there is a significant source of unexplained variance, limiting the interpretability of analyses using these data. Thus, our first step in determining a measurement strategy was to evaluate inter-rater reliability. This study used the Intraclass Correlation Coefficient (ICC), appropriate for the condition of varying numbers of raters across classes. In this case, the "class" refers to the individual student in the sample; the ICC is the proportion of variance in the item responses accounted for by the student's identity. The ICC coefficients for the normative sample and control samples are as follows:

	Normative Sample	Control Sample
Academic Performance	0.61	0.58
Academic Motivation	0.62	0.58
Social Skills	0.31	0.29
Relationships with Adults	0.36	0.39
Conduct	0.65	0.58
Personal Maturity	0.55	0.59

These reliability levels indicate that overall, inter-rater disagreement does not preclude the use of multiple ratings. Prior to making that decision, however, we evaluated the validity of a multiple rating score, as compared with the score from a single teacher rating (see the source paper for a complete description of these analyses). The analyses involved structural equation models designed to test the relationship between teacher rating strategies (single vs. multiple) and correlates from other Fast Track measures, including academic performance indicators from school records and items from the Parent Rating of Student Adjustment. The lowest residual variances occurred in the aggregate model, thus indicating that a useful strategy is the use of multiple informant data. The data were aggregated by averaging the teacher ratings available for each student, to create an average scale score for each of the six behavior domains. Caution is advised in using the aggregate scores for Social Skills and Relationships with Adults, as the level of inter-rater reliability is low.

IV. Differences Between Groups

T-tests of means on the aggregate scores between the normative and control samples yielded the following results:

	Normative Sample		Control Sample		DF	t Value	Pr > |t|
	Mean	SD	Mean	SD			
Academic Performance	2.94	1.09	2.41	1.04	380	-4.58	<.0001
Academic Motivation	3.08	1.19	2.49	1.09	380	-4.71	<.0001
Social Skills	3.36	0.92	2.67	1.00	379	-6.70	<.0001
Relationships with Adults	3.23	0.90	2.77	1.03	380	-5.12	<.0001
Conduct	3.40	1.19	2.42	1.10	380	-7.89	<.0001
Personal Maturity	3.10	1.18	2.32	1.16	381	-6.19	<.0001

These results reveal significant differences between the normative and control samples for all of the aggregated scores. For each domain of student adjustment, the normative group received a higher mean rating as compared with the control group. This finding indicates that students in the normative group were, on average, significantly more successful in these dimensions of adjustment to middle school, as compared with the high-risk control group.

V. Recommendations for Use

As noted, the level of inter-rater agreement varies considerably across behavioral domains, with low values of the Intraclass Correlation Coefficient for the "social skills" and "relationships with adults" items. The behavioral characteristics targeted with these items are likely to be the most subject to influence by the particular classroom environment. Researchers should keep in mind that the scores for these domains include considerable variation across teacher ratings for each student, a potentially important source of measurement error in analyses using these data.

When using the aggregate scores, analysts should note that the number of teachers providing ratings varies across students. For the normative and control samples, 45 percent of students had only one teacher rating, 21 percent had two ratings, 26 percent had three ratings, and 10 percent had four or five ratings. In models predicting the TRSA domains, Muschkin and Malone (2003) found that, as expected, the lowest residual variance estimates correspond to the sample of students who received three or more teacher ratings. To the extent that the variation in the number of teachers rating each student is systematic, this could introduce heterogeneity of error variance (heteroscedasticity) in analyses using these data, violating an assumption of many common analytic techniques. An alternative approach is to use a single teacher's rating for each student. This sacrifices the increased precision from the multiple teacher ratings, but avoids the heteroscedasticity problem. Muschkin and Malone found no differences between using reports from a randomly selected teacher and from selection based on the item assessing how well the teacher knows the child; the former is recommended when heteroscedasticity is a concern.

VI. Scale Means and SDs

Means and Standard Deviations for Average Scores,
Teacher Ratings of Student Adjustment
Cohort 1, Year 7 Normative Sample

Variable	Label	Mean	N	Std Dev
tsa7stam	TSA student academic motivation yr 7	2.92	319	1.18
tsa7stap	TSA student academic performance yr 7	2.80	319	1.10
tsa7star	TSA student-adult relationships yr 7	3.14	319	0.96
tsa7stc	TSA student conduct yr 7	3.16	319	1.25
tsa7stpm	TSA student personal maturity yr 7	2.89	320	1.22
tsa7stss	TSA student social skills yr 7	3.22	318	0.96

Means and Standard Deviations for Average Scores,
Teacher Ratings of Student Adjustment
Cohort 1, Year 7 Control Sample

Variable	Label	Mean	N	Std Dev
tsa7stam	TSA student academic motivation yr 7	2.49	131	1.09
tsa7stap	TSA student academic performance yr 7	2.41	131	1.04
tsa7star	TSA student-adult relationships yr 7	2.77	131	1.03
tsa7stc	TSA student conduct yr 7	2.42	131	1.10
tsa7stpm	TSA student personal maturity yr 7	2.32	131	1.16
tsa7stss	TSA student social skills yr 7	2.68	131	1.00

VII.   Scale Correlations

Teacher Ratings of Student Adjustment – Average Score Correlations
Report Sample, Year 7

<table>
<tr><td colspan="7">Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations</td></tr>
<tr><td></td><td>tsa7stam</td><td>tsa7stap</td><td>tsa7star</td><td>tsa7stc</td><td>tsa7stpm</td><td>tsa7stss</td></tr>
<tr><td>tsa7stam
TSA student academic motivation yr 7</td><td>1.000

382.000</td><td>0.758
0.000
382.000</td><td>0.640
0.000
382.000</td><td>0.697
0.000
382.000</td><td>0.813
0.000
382.000</td><td>0.516
0.000
381.000</td></tr>
<tr><td>tsa7stap
TSA student academic performance yr 7</td><td>0.758
0.000
382.000</td><td>1.000

382.000</td><td>0.502
0.000
382.000</td><td>0.561
0.000
382.000</td><td>0.677
0.000
382.000</td><td>0.400
0.000
381.000</td></tr>
<tr><td>tsa7star
TSA student-adult relationships yr 7</td><td>0.640
0.000
382.000</td><td>0.502
0.000
382.000</td><td>1.000

382.000</td><td>0.714
0.000
382.000</td><td>0.656
0.000
382.000</td><td>0.630
0.000
381.000</td></tr>
<tr><td>tsa7stc
TSA student conduct yr 7</td><td>0.697
0.000
382.000</td><td>0.561
0.000
382.000</td><td>0.714
0.000
382.000</td><td>1.000

382.000</td><td>0.747
0.000
382.000</td><td>0.624
0.000
381.000</td></tr>
<tr><td>tsa7stpm
TSA student personal maturity yr 7</td><td>0.813
0.000
382.000</td><td>0.677
0.000
382.000</td><td>0.656
0.000
382.000</td><td>0.747
0.000
382.000</td><td>1.000

383.000</td><td>0.614
0.000
381.000</td></tr>
<tr><td>tsa7stss
TSA student social skills yr 7</td><td>0.516
0.000
381.000</td><td>0.400
0.000
381.000</td><td>0.630
0.000
381.000</td><td>0.624
0.000
381.000</td><td>0.614
0.000
381.000</td><td>1.000

381.000</td></tr>
</table>