



# Behavior Intervention Monitoring Assessment System

James L. McDougal, Psy.D.,  
Achilles N. Bardos, Ph.D., &  
Scott T. Meier, Ph.D.

## TECHNICAL MANUAL



# **Behavior Intervention Monitoring Assessment System (BIMAS™)**

## **Technical Manual**

**James L. McDougal, Psy.D.**

**Achilles N. Bardos, Ph.D.**

**Scott T. Meier, Ph.D.**



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.  
No part of this manual may be reproduced by any means without permission from the publisher.  
In the U.S.A., P.O. Box 950, North Tonawanda, NY 14120-0950  
In Canada, 3770 Victoria Park Avenue, Toronto, ON M2H 3M6

## Contact MHS

1-800-456-3003 (U.S.)  
1-800-268-6011 (Canada)  
+1-416-492-2627 (International)  
customerservice@mhs.com  
www.mhs.com

This manual was edited and typeset by Maggie Bailey, Heather (Hau) Co, Gurcharn Birdi, and Vanessa Gratsas using InDesign CS3 and Microsoft Word 2003, using the Times New Roman, Arial, Verdana, and Tahoma fonts. The cover was designed by Heather (Hau) Co using Adobe InDesign CS4, Adobe Illustrator CS4 and Adobe Photoshop CS4.

BIMAS, Conners Comprehensive Behavior Rating Scales, and Conners CBRS, are trademarks of Multi-Health Systems Inc.

Devereux Scales of Mental Disorders, DSMD, Devereux Behavior Rating Scale–School Form, and Devereux–School Form are trademarks of the Devereux Foundation.

GAMA is a registered trademark of Pearson Education, Inc.

DSM, DSM-IV, and DSM-IV-TR are trademarks of the American Psychiatric Association.

Apple Safari is a registered trademark of Apple Inc.

Google and Google Chrome are trademarks of Google Inc.

Firefox is a registered trademark of Mozilla.

Microsoft, Internet Explorer, and Windows are registered trademarks of Microsoft Corporation.

Opera is a registered trademark of Opera Software ASA.

Copyright © 2011 Multi-Health Systems Inc. All rights reserved. No part of this manual, the sample forms, or any related materials protected by copyrights are to be printed or otherwise reproduced by any means, including electronic storage within a computer program or database, without the permission of the publisher. These materials may not be translated into a natural or computer language without permission.

This copyright is protected through the laws of the United States, Canada, and other countries. Persons who violate the copyrights on these materials may be liable to prosecution. Ethical codes of various professional associations to which users are likely to belong specifically prohibit both illegal behaviors and actions that would deny other parties fair compensation for their work. Persons who violate professional ethical codes related to inappropriate and unfair use of these materials may be brought before the relevant professional associations to which they belong.

The information included in this manual does not constitute, and shall not be considered, the advice, recommendation, assessment, or endorsement of MHS. To the extent permitted by law, any statutory or implied warranty of merchantability or fitness for a particular purpose is completely denied and disclaimed. MHS shall not be liable for any third party claims, lost profits, lost savings, loss of information, or any other incidental damages or other economic consequential damages resulting from the use of the manual.

Published in Canada by Multi-Health Systems Inc.

Printed in Canada.

September, 2011

# Contents

<b>Chapter 1: Introduction</b> .....	<b>1</b>
<b>Main Features of the BIMAS</b> .....	<b>1</b>
<b>BIMAS Components</b> .....	<b>1</b>
BIMAS Standard.....	1
BIMAS Flex .....	2
<b>Multi-Informant Form Options</b> .....	<b>2</b>
<b>BIMAS Formats</b> .....	<b>2</b>
Online administration and scoring.....	2
Paper-and-pencil administration .....	2
<b>Report Types</b> .....	<b>2</b>
<b>Uses/Applicability</b> .....	<b>3</b>
Screening.....	3
Intervention Planning .....	3
Progress Monitoring.....	3
Outcome Assessment .....	3
Program Evaluation .....	3
<b>Principles of Use</b> .....	<b>4</b>
<b>Users and User Qualifications</b> .....	<b>4</b>
<b>Contents of the Manual</b> .....	<b>5</b>
<b>Chapter 2: Background</b> .....	<b>7</b>
<b>Concerns in School-Aged Youth</b> .....	<b>7</b>
<b>Response to Intervention (RTI)</b> .....	<b>7</b>
The 3-Tier RTI Model .....	7
Tier 1 .....	8
Tier 2.....	8
Tier 3.....	8
RTI and Behavior Progress Monitoring.....	8
<b>Outcome Assessments</b> .....	<b>8</b>
Outcome Assessments in Clinical Settings.....	9
Outcome Assessment in Educational Settings .....	9
Strengths and Limitations of Current Outcome Measures.....	10
<b>The BIMAS: A Change-Sensitive Measure</b> .....	<b>10</b>
<b>Chapter 3: Administration and Scoring</b> .....	<b>13</b>
<b>Choosing the Appropriate Form</b> .....	<b>13</b>
BIMAS Standard .....	14
Purpose of BIMAS Standard .....	14
BIMAS Flex.....	14
Purpose of BIMAS Flex.....	14
BIMAS Flex Progress Monitoring Tools.....	15
Use of the BIMAS Forms within the RTI Framework .....	15
<b>Administration</b> .....	<b>16</b>
General Administration Guidelines.....	16
Age Range .....	16
Reading Levels.....	16
Administration Time.....	16
Appropriate Raters .....	16
Time Frame .....	17
Administration Settings .....	17
Reading the BIMAS Items Aloud for a Rater.....	17
Ethical and Legal Guidelines.....	18
Administration Procedure.....	18
Preparation.....	19

During the Administration .....	19
After the Administration .....	20
<b>Scoring .....</b>	<b>20</b>
Omitted Responses.....	20
Features of Online Scoring.....	20
<b>Hardware and Software Requirements.....</b>	<b>21</b>
<b>Technical Support .....</b>	<b>21</b>
<b>Chapter 4: BIMAS Reports.....</b>	<b>23</b>
<b>Types of Reports .....</b>	<b>23</b>
<b>Assessment Reports .....</b>	<b>23</b>
Multi-student Level.....	23
Individual Level .....	23
<b>Progress Reports .....</b>	<b>27</b>
Multi-student Level.....	27
Individual Level .....	27
<b>Comparative Reports .....</b>	<b>28</b>
Multi-student Level.....	28
Individual Level .....	28
<b>Demographic Reports .....</b>	<b>28</b>
Multi-student Level.....	28
<b>Report Options .....</b>	<b>29</b>
<b>Chapter 5: Understanding and Interpreting BIMAS Scores .....</b>	<b>31</b>
<b>Scale-Level Scores.....</b>	<b>32</b>
Raw Scores.....	32
T-Scores .....	32
Interpreting T-scores .....	32
Behavioral Concern Scales.....	32
Adaptive Scales.....	33
Percentiles .....	33
Confidence Intervals .....	33
<b>Item-Level Scores .....</b>	<b>34</b>
Interpreting Item-Level Scores.....	34
Behavioral Concern Scales.....	34
Adaptive Scales .....	35
<b>Scores for Progress and Outcome Monitoring .....</b>	<b>35</b>
Visual Displays .....	35
Reliable Change Index (RCI).....	36
<b>Interpreting Change with the Reliable Change Index .....</b>	<b>37</b>
Effect Size Estimates .....	37
<b>Interpreting Change with Effect Size Estimates.....</b>	<b>38</b>
Comparison of the Reliable Change Index and Effect Size Estimates .....	38
<b>Scores for Comparing Results Between Raters &amp; Groups .....</b>	<b>38</b>
Statistically Significant Differences Between Raters.....	38
Comparing Between Group Averages .....	39
<b>Chapter 6: Step-by-Step Interpretation of BIMAS Results .....</b>	<b>41</b>
<b>Step-by-Step Interpretation Guidelines: Group Evaluation.....</b>	<b>41</b>
Step 1. Identify General Areas of Concern .....	41
Step 2. Identify Specific Populations in Need of Intervention .....	43
Step 3. Collect Additional Information.....	43
Step 4. Develop an Intervention and Monitoring Plan .....	43
Step 5. Monitor Group Progress .....	44
<b>Step-by-Step Interpretation Guidelines: Individual Evaluation.....</b>	<b>44</b>
Step 1. Interpret the BIMAS Standard Scale Scores to Identify General Areas of Concern.....	45
Step 1a: Interpret the Behavioral Concern Scale Scores.....	45
Step 1b: Interpret the Adaptive Scale Scores.....	45

Step 2. Examine Item-Level Responses for an In-depth Analysis of Concern .....	45
Step 2a: Behavioral Concern Scales .....	45
Step 2b: Adaptive Scales .....	45
Step 3. Collect Additional Information and Compare Results Between Raters .....	46
Step 3a: Scale-level Comparisons .....	46
Step 3b: Item-level Comparisons .....	46
Step 4. Develop an Intervention and Monitoring Plan.....	47
Step 5. Monitor Response to Intervention.....	47
Step 5a: BIMAS Flex .....	47
Step 5b: BIMAS Standard .....	47
<b>Chapter 7: Case Study .....</b>	<b>49</b>
<b>Tier 1 (Universal Level): Group Evaluation .....</b>	<b>49</b>
Step 1. Identify General Areas of Concern .....	49
Step 2. Identify Specific Populations In Need of Intervention .....	50
Step 3. Collect Additional Information .....	52
Step 4. Develop an Intervention and Monitoring Plan .....	52
Step 5. Monitor Group Progress over Time .....	52
<b>Tiers 2 and 3 (Targeted and Intense Levels): Individual Evaluation .....</b>	<b>54</b>
Step 1. Interpret the BIMAS Standard Scale Scores to Identify General Areas of Concern .....	55
Step 1a: Interpret the Behavioral Concern Scale Scores .....	55
Step 1b: Interpret the Adaptive Scale Scores.....	55
Step 2. Examine Item-Level Responses for an In-depth Analysis of Concerns.....	55
Step 2a: Behavioral Concern Scales .....	55
Step 2b: Adaptive Scales .....	56
Step 3. Collect Additional Information and Compare Results Between Raters .....	56
Step 4. Develop an Intervention and Monitoring Plan.....	58
Step 5. Monitor Response to Intervention .....	59
<b>Chapter 8: Development .....</b>	<b>63</b>
<b>Rationale and Goals .....</b>	<b>63</b>
Rationale .....	63
Development Goals .....	63
<b>Preliminary Development .....</b>	<b>64</b>
Preliminary Research: Intervention Item Selection Rules .....	64
Preliminary Structure and Item Development .....	65
Response Schemas.....	66
Creation of the Forms for the Normative Study .....	66
<b>Final Scale Construction .....</b>	<b>66</b>
Data Collection.....	66
Development of the BIMAS Standard .....	67
Development of the Behavioral Concern Scales .....	67
Development of the Adaptive Scales .....	67
Development of the BIMAS Flex Items .....	67
Determining the Scoring Criteria of Item Descriptors for BIMAS Standard Items.....	68
Behavioral Concern Scales .....	68
Adaptive Scales .....	69
Creation of the BIMAS Standard Final Forms .....	69
<b>Chapter 9: Standardization.....</b>	<b>71</b>
<b>Data Collection .....</b>	<b>71</b>
<b>Normative Sample Description .....</b>	<b>71</b>
Teacher Normative Sample (BIMAS–T).....	72
Parent Normative Sample (BIMAS–P).....	72
Self-Report Normative Sample (BIMAS–SR).....	73

<b>Chapter 10: Reliability</b> .....	<b>75</b>
Overview of Results .....	75
Internal Consistency .....	75
Standard Error of Measurement.....	77
Test-Retest Reliability and Standard Error of Prediction .....	78
Consistency between Raters.....	80
<b>Chapter 11: Validity</b> .....	<b>81</b>
Overview of Results .....	81
Content Validity and BIMAS Scale Structure .....	81
Relationship Between the BIMAS and Other Measures.....	84
<b>The BIMAS as a Screening Tool</b> .....	<b>86</b>
The BIMAS as a Screening Tool: Teacher Ratings .....	88
Differences Between Clinical and Non-Clinical Groups .....	88
Differences Between Clinical Groups.....	88
The BIMAS as a Screening Tool: Parent Ratings .....	90
Differences Between Clinical and Non-Clinical Groups .....	90
Differences Between Clinical Groups.....	91
The BIMAS as a Screening Tool: Self-Report Ratings.....	93
Differences Between Clinical and Non-Clinical Groups .....	93
Differences Between Clinical Groups.....	93
<b>Validity of the BIMAS as a Progress Monitoring Tool</b> .....	<b>95</b>
Change-Sensitivity of Early Versions of the BIMAS.....	95
Change-Sensitivity of the BIMAS.....	97
<b>Chapter 12: Concluding Comments</b> .....	<b>99</b>
<b>References</b> .....	<b>101</b>
<b>Appendix A: BIMAS Standard Items by Scale</b> .....	<b>105</b>
<b>Appendix B: BIMAS Flex Items by Scale</b> .....	<b>107</b>
<b>Appendix C: Reliable Change Index Values</b> .....	<b>131</b>
<b>Appendix D: Values for Rater Comparisons</b> .....	<b>133</b>
<b>Appendix E.1: Risk Level Pyramids</b> .....	<b>135</b>
<b>Appendix E.2a: Student by Risk Level (Assessment Report Version)</b> .....	<b>137</b>
<b>Appendix E.2b: Student by Risk Level (Progress Report Version)</b> .....	<b>139</b>
<b>Appendix E.3: Class/Group Student Scores</b> .....	<b>141</b>
<b>Appendix E.4: Standard Individual Assessment Report</b> .....	<b>143</b>
<b>Appendix E.5: Demographics Breakdown</b> .....	<b>147</b>
<b>Appendix E.6: Risk Level by Demographics</b> .....	<b>149</b>
<b>Appendix E.7: Average Score Comparison Report</b> .....	<b>151</b>
<b>Appendix E.8: Standard Individual Comparison Between Raters</b> .....	<b>153</b>
<b>Appendix E.9: Average Score Comparison: Progress Report</b> .....	<b>159</b>
<b>Appendix E.10: Class/Group Student Scores: Progress Report</b> .....	<b>161</b>
<b>Appendix E.11: Risk Level by Demographics: Progress Report</b> .....	<b>167</b>
<b>Appendix E.12: Flex Individual Progress Report</b> .....	<b>169</b>
<b>Appendix E.13: Flex Individual Comparison Between Raters</b> .....	<b>177</b>
<b>Appendix E.14: Standard Individual Progress Report</b> .....	<b>185</b>

<b>Appendix E.15a: Standard Individual Progress Report: Significant Change Over Time (Statistically Significant Change [Based on <i>T</i>-scores] Version)</b> .....	<b>189</b>
<b>Appendix E.15b: Standard Individual Progress Report: Significant Change Over Time (Effect Size [Based on raw scores] Version)</b> .....	<b>193</b>
<b>Appendix E.16: Standard Individual Progress Report: Item Analysis</b> .....	<b>197</b>
<b>Appendix F: Generalizability of the BIMAS Standard Across Race/Ethnic Groups</b> ....	<b>205</b>
<b>Appendix G: Age and Gender Effects</b> .....	<b>207</b>
<b>Appendix H: Raw Score Standard Error of Measurement Values</b> .....	<b>209</b>
<b>Appendix I: Group Membership MANOVAs</b> .....	<b>211</b>
<b>Index</b> .....	<b>215</b>

## Tables

Table 3.1. Maximum Number of Allowable Item Omissions for the BIMAS Standard .....	20
Table 4.1. Description, Special Features, and Function of BIMAS Multi-student Level Reports.....	24
Table 4.2. Description, Special Features, and Function of BIMAS Individual Level Reports .....	25
Table 4.3. BIMAS Report Types & Levels .....	27
Table 5.1. Types of Scores Available on BIMAS Reports.....	31
Table 5.2. Understanding <i>T</i> -scores and Scale Descriptors: Behavioral Concern Scales .....	32
Table 5.3. Common Characteristics of Youth with High Scores on the BIMAS Behavioral Concern Scales.....	32
Table 5.4. Understanding <i>T</i> -scores and Scale Descriptors: Adaptive Scales.....	33
Table 5.5. Common Characteristics of Youth with Low and High Scores on the BIMAS Adaptive Scales.....	33
Table 5.6. Understanding Item Descriptors: Behavioral Concern Scales.....	35
Table 5.7. Understanding Item Descriptors: Adaptive Scales.....	35
Table 5.8. Effect Size Interpretations for Individual Clients on the BIMAS Standard .....	38
Table 8.1. Development Guidelines for Item Descriptors: Behavioral Concern Scales.....	69
Table 8.2. Development Guidelines for Item Descriptors: Adaptive Scales.....	69
Table 9.1. Race/Ethnicity Distribution: BIMAS–T Standard Normative Sample .....	72
Table 9.2. Geographic Region Distribution: BIMAS–T Standard Normative Sample .....	72
Table 9.3. Race/Ethnicity Distribution: BIMAS–P Standard Normative Sample .....	72
Table 9.4. Parental Education Level Distribution: BIMAS–P Standard Normative Sample .....	72
Table 9.5. Geographic Region Distribution: BIMAS–P Standard Normative Sample .....	73
Table 9.6. Race/Ethnicity Distribution: BIMAS–SR Standard Normative Sample .....	73
Table 9.7. Geographic Region Distribution: BIMAS–SR Standard Normative Sample .....	73
Table 10.1. Cronbach’s Alpha: BIMAS–T Standard.....	76
Table 10.2. Cronbach’s Alpha: BIMAS–P Standard .....	76
Table 10.3. Cronbach’s Alpha: BIMAS–SR Standard.....	76
Table 10.4. Standard Error of Measurement: BIMAS–T Standard <i>T</i> -Scores.....	77
Table 10.5. Standard Error of Measurement: BIMAS–P Standard <i>T</i> -Scores.....	78
Table 10.6. Standard Error of Measurement: BIMAS–SR Standard <i>T</i> -Scores .....	78
Table 10.7. Demographic Characteristics of the BIMAS Standard Test-Retest Reliability Samples .....	79
Table 10.8. Test-Retest Reliability Coefficients: BIMAS–T Standard <i>T</i> -scores.....	79
Table 10.9. Test-Retest Reliability Coefficients: BIMAS–P Standard <i>T</i> -scores.....	79
Table 10.10. Test-Retest Reliability Coefficients: BIMAS–SR Standard <i>T</i> -scores.....	79
Table 10.11. Standard Error of Prediction Coefficients: BIMAS Standard <i>T</i> -scores.....	79

# Tables

Table 10.12. Demographic Characteristics of the BIMAS Standard Consistency Between Raters Sample .....	80
Table 10.13. Consistency Between Rater <i>T</i> -scores: Teacher to Self-Report Ratings.....	80
Table 10.14. Consistency Between Rater <i>T</i> -scores: Parent to Self-Report Ratings.....	80
Table 10.15. Consistency Between Rater <i>T</i> -scores: Teacher to Parent Ratings .....	80
Table 11.1. Overview of BIMAS Standard Item Content and Scale Assignment on the BIMAS Standard.....	83
Table 11.2. BIMAS–T Standard Scale Intercorrelations .....	84
Table 11.3. BIMAS–P Standard Scale Intercorrelations.....	84
Table 11.4. BIMAS–SR Standard Scale Intercorrelations .....	84
Table 11.5. Demographic Characteristics of the BIMAS Standard and Conners CBRS Validity Study Samples.....	85
Table 11.6. Correlations Between the BIMAS Standard Conduct Scale <i>T</i> -scores and Relevant Conners CBRS Scales .....	85
Table 11.7. Correlations Between the BIMAS Standard Negative Affect Scale <i>T</i> -scores and Relevant Conners CBRS Scales .....	85
Table 11.8. Correlations Between the BIMAS Standard Cognitive/Attention Scale <i>T</i> -scores and Relevant Conners CBRS Scales .....	86
Table 11.9. Correlations Between the BIMAS Standard Social Scale <i>T</i> -scores and Relevant Conners CBRS Scales .....	86
Table 11.10. Correlations Between the BIMAS Standard Academic Functioning Scale <i>T</i> -scores and Relevant Conners CBRS Scales .....	86
Table 11.11. Clinical Samples: Primary Diagnoses by Rater .....	87
Table 11.12. Demographic Characteristic of the Clinical Samples.....	87
Table 11.13. Clinical Sample Descriptive Statistics ( <i>T</i> -scores) and Effect Sizes (Comparison to the Norm) Across BIMAS–T Standard Scales .....	89
Table 11.14. Standardized Discriminant Function Coefficients of the BIMAS–T Standard Scales .....	89
Table 11.15. Group Classification as Predicted by Scores on the BIMAS–T Standard Scales.....	89
Table 11.16. Classification Accuracy of the BIMAS–T Standard Scales .....	89
Table 11.17. Descriptive Statistics ( <i>T</i> -scores) of the Different Clinical Samples and Effect Sizes (Comparison to the Norm) Across the BIMAS–T Standard Scales .....	89
Table 11.18. Clinical Sample Descriptive Statistics ( <i>T</i> -scores) and Effect Sizes (Comparison to the Norm) Across the BIMAS–P Standard Scales.....	91
Table 11.19. Standardized Discriminant Function Coefficients of the BIMAS–P Standard Scales .....	91
Table 11.20. Group Classification as Predicted by Scores on the BIMAS–P Standard Scales.....	91
Table 11.21. Classification Accuracy of the BIMAS–P Standard Scales .....	91
Table 11.22. Descriptive Statistics ( <i>T</i> -scores) of the Different Clinical Samples and Effect Sizes (Comparison to the Norm) across the BIMAS–P Standard Scales .....	92
Table 11.23. Clinical Sample Descriptive Statistics ( <i>T</i> -scores) and Effect Sizes (Comparison to the Norm) Across the BIMAS–SR Standard Scales .....	93
Table 11.24. Standardized Discriminant Function Coefficients of the BIMAS–SR Standard Scales.....	93
Table 11.25. Group Classification as Predicted by Scores on the BIMAS–SR Standard Scales.....	94
Table 11.26. Classification Accuracy of the BIMAS–SR Standard Scales .....	94
Table 11.27. Descriptive Statistics ( <i>T</i> -scores) of the Different Clinical Samples and Effect Sizes (Comparison to the Norm) Across the BIMAS–SR Standard Scales .....	94
Table 11.28. Pre- to Post-Performance of an Anger Management Treatment Group: BIMAS–T Standard <i>T</i> -Scores .....	98
Table 11.29. Pre- to Post-Performance of an Anger Management Treatment Group: BIMAS–P Standard <i>T</i> -Scores.....	98
Table 11.30. Pre- to Post-Performance of an Anger Management Treatment Group: BIMAS–SR Standard <i>T</i> -Scores .....	98
Table C.1. RCI Values for the BIMAS Standard.....	131
Table D.1. Minimum Values for Comparisons Between Raters on the BIMAS Standard Using Combined Gender Norms (Default Scoring Option).....	133
Table D.2. Minimum Values for Comparisons Between Raters on the BIMAS Standard Using Gender-Specific Norms .....	134
Table F.1. Race/Ethnicity Effects: Multivariate Results .....	205

Table F.2.	Differences Between Race/Ethnic Groups: BIMAS–T Standard .....	205
Table F.3.	Differences Between Race/Ethnic Groups: BIMAS–P Standard .....	206
Table F.4.	Differences Between Race/Ethnic Groups: BIMAS–SR Standard.....	206
Table G.1.	Age and Gender Effects: Multivariate Results .....	207
Table G.2.	Age Effects: Univariate Results .....	207
Table G.3.	Gender Effects: Univariate Results.....	208
Table G.4.	Age × Gender Effects: Univariate Results .....	208
Table H.1.	Standard Error of Measurement: BIMAS–T Standard Raw Scores.....	209
Table H.2.	Standard Error of Measurement: BIMAS–P Standard Raw Scores.....	209
Table H.3.	Standard Error of Measurement: BIMAS–SR Standard Raw Scores .....	210
Table I.1.	Group Membership Effects: Multivariate Results.....	211
Table I.2.	Differences Between Groups: BIMAS–T Standard.....	212
Table I.3.	Differences Between Groups: BIMAS–P Standard.....	212
Table I.4.	Differences Between Groups: BIMAS–SR Standard .....	213

## Figures

Figure 3.1.	Administration and Scoring Options .....	13
Figure 5.1.	Hypothetical Improvement on a BIMAS Behavioral Concern Scale over a Course of Intervention.....	36
Figure 5.2.	Hypothetical Worsening on a BIMAS Behavioral Concern Scale over a Course of Intervention.....	36
Figure 5.3.	Hypothetical Cycles on a BIMAS Behavioral Concern Scale over a Course of Intervention .....	36
Figure 5.4.	Hypothetical Stability on a BIMAS Behavioral Concern Scale over a Course of Intervention .....	36
Figure 6.1.	Application of BIMAS Reports Within the Step-by-Step Interpretation Method .....	42
Figure 7.1.	Risk Level Pyramids .....	49
Figure 7.2.	Class/Group Student Scores .....	50
Figure 7.3.	Risk Level by Demographics .....	51
Figure 7.4.	Average Score Comparison: Progress Report .....	53
Figure 7.5.	Student List by Risk Level .....	54
Figure 7.6.	Standard Individual Assessment Report (Scale Scores) .....	55
Figure 7.7.	Standard Individual Assessment Report (Item-Level Scores) .....	56
Figure 7.8.	Standard Individual Comparison Between Raters (Scale-Level Comparison) .....	57
Figure 7.9.	Standard Individual Comparison Between Raters (Item-Level Comparison) .....	58
Figure 7.10.	Flex Individual Progress Report.....	59
Figure 7.11.	Standard Individual Progress Report.....	60
Figure 7.12.	Standard Individual Progress Report: Significant Change Over Time (RCI–Based on <i>T</i> -score Option).....	61
Figure 11.1.	Mean <i>T</i> -scores by Clinical Group: BIMAS–T Standard.....	90
Figure 11.2.	Mean <i>T</i> -scores by Clinical Group: BIMAS–P Standard.....	92
Figure 11.3.	Mean <i>T</i> -scores by Clinical Group: BIMAS–SR Standard.....	95
Figure 11.4.	Anger Management Treatment Group: Pre- to Post-Treatment BIMAS Standard <i>T</i> -score.....	97

# About the Authors

## **Jim L. McDougal, Psy.D.**

Dr. James McDougal is an Associate Professor and Director of the School Psychology Program at the State University of New York at Oswego. Dr. McDougal teaches courses in assessment and intervention for academics and behavior. He also consults and supervises field-based practicum and internship experiences. He is the former Mental Health Coordinator for the Syracuse City School District, where he integrated mental health services in schools and provided mental health, behavioral, and academic consultation services for more than 40 schools and programs.

Dr. McDougal has 20 years of experience as a practicing school psychologist, and has considerable experience consulting with students with academic and behavioral problems and conducting staff development for educators. He has conducted over 200 training seminars at state and national forums in the areas of academic and behavioral assessment/intervention, and on the implementation of Response to Intervention (RTI) models in schools to improve the academic, behavioral, and socio-emotional functioning of students. Dr. McDougal also has a considerable record of scholarly publications, including two recently released co-authored books entitled *A Practitioner's Guide to Functional Assessment and Behavior Intervention Planning in the Schools* (McDougal, Chafouleas, & Waterman, 2006), and *RTI in Practice: A Practical Guide to Implementing Effective Evidenced-Based Interventions in Your Schools* (McDougal, Graney, Wright, & Ardoin, 2009). In addition, he has published peer-reviewed articles in *School Psychology Review*, *School Psychology Quarterly*, *Psychology in the Schools*, the *Canadian Journal of School Psychology*, and the *Communiqué*.

## **Achilles N. Bardos, Ph.D.**

Dr. Achilles Bardos graduated from Ohio State University and is a Professor in the Department of School Psychology at the University of Northern Colorado. Dr. Bardos' teaching and research interests include educational and psychological assessment, program evaluation, and computer applications in psychology. He has published many articles and book chapters, and presented over 100 papers and workshops at national and international conferences.

Dr. Bardos holds an appointment as Senior Research Scientist with the Devereux Foundation, where he assisted with the development of the Devereux Scales of Mental Disorders™ (DSMD™; Naglieri, LeBuffe, & Pfeiffer, 1994), the Devereux Behavior Rating Scale–School Form™ (Devereux–School Form™; Naglieri, LeBuffe, & Pfeiffer, 1993), the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999), and most recently, the Devereux Early Childhood Assessment–Clinical Form (DECA-C; LeBuffe & Naglieri, 2003).

Dr. Bardos co-authored the Draw A Person: Screening Procedure for Emotional Disturbance (Naglieri, McNeish, & Bardos, 1991), the General Ability Measure for Adults (GAMA®, Naglieri & Bardos, 1997), a non-verbal intelligence test, and the Basic Achievement Skills Inventory–Comprehensive and Survey (BASI; Bardos, 2004), a series of self-administered achievement tests that can be used in schools and other settings such as corrections, public safety, and personnel selection.

Currently, Dr. Bardos is also co-authoring the Reading Competency Inventories (RCI; Bardos & Naglieri, in press), a series of reading vocabulary tests specially designed to assist with a client's reading comprehension skills in a personality evaluation.

## **Scott T. Meier, Ph.D.**

Dr. Scott Meier is a Professor at the State University of New York at Buffalo in the Department of Counseling, School, and Educational Psychology. Dr. Meier is also a licensed psychologist who received his Ph.D. in Counseling Psychology from Southern Illinois University, Carbondale. Dr. Meier focuses his research and teaching on psychological measurement (particularly outcome assessment), research methods (particularly program evaluation), and counseling skills (particularly the integration of case conceptualization and assessment with intervention). Dr. Meier has conducted numerous program evaluations and has worked with organizations such as the Veterans Administration, Child and Adolescent Treatment Services of Western New York, the Society for Prevention of Cruelty to Animals, and numerous school systems in Western and Central New York. Dr. Meier is a member of the American Evaluation Association, the Association for Psychological Science, the Association for Behavioral and Cognitive Therapies, and the Psychological Association of Western New York.

Dr. Meier is the author or co-author of five books (including *Measuring Change in Counseling and Psychotherapy*; Meier, 2008 and *Elements of Counseling*; Meier & Davis, 2010) and has published work in the *American Psychologist*, *Canadian Journal of School Psychology*, *Journal of Counseling Psychology*, *Measurement and Evaluation in Counseling and Development*, and the *American Journal of Evaluation*.

## Publisher's Preface

Multi-Health Systems (MHS) is proud to announce the arrival of the Behavior Intervention Monitoring Assessment System (BIMAS™), a new, web-based assessment and data management system that is specially designed for the screening of behavioral, social, emotional, and academic difficulties as well as progress monitoring response to intervention. The BIMAS was specifically developed for use in schools and districts within the Response to Intervention (RTI) framework to identify behavioral concerns or strengths in youth between the ages of 5 and 18. This unique, change-sensitive tool contains items that function both as a brief screener to identify youth who may require further assessment or intervention, and as a progress monitoring/program evaluation tool. The BIMAS covers a wide range of behaviors that commonly affect student functioning in academic settings (e.g., conduct problems, depressive/anxiety symptoms, cognitive/attention concerns, as well as social and academic functioning).

We are proud to have had the opportunity to work with Drs. James L. McDougal, Achilles N. Bardos, and Scott T. Meier. This author team brings together extensive expertise in School and Counseling Psychology, test development, and change-sensitive assessment which sets the foundation on which the BIMAS is built. All of the authors are highly recognized in the field of psychology and education with numerous contributions and publications over the years.

MHS would like to acknowledge the outstanding contributions of Jane Wong, Anand Nathan, and Noel Demello, who worked extensively on the BIMAS. Together, they worked diligently to develop a tool that will be of the utmost utility to educators and/or clinicians. Their devotion to this project and incredible attention to detail helped ensure that the BIMAS is a product of the highest caliber and meets the quality standards of our other products.

We would like to hear about your experience with the BIMAS; your feedback is helpful for future refinements and enhancements to the assessment. Please send any questions or comments to [r&d@mhs.com](mailto:r&d@mhs.com). You may also wish to visit the MHS website at [www.mhs.com](http://www.mhs.com) to find out about our other innovative products.

*Steven J. Stein, Ph.D.*  
*CEO*  
*ceo@mhs.com*  
*September, 2011*

*Hazel Wheldon, M.A.*  
*Publisher & COO*  
*publisher@mhs.com*  
*September, 2011*

# Author's Preface

Our quest to develop the Behavior Intervention Monitoring Assessment System (BIMAS™) began over a decade ago with an unlikely partnership that has resulted in a product which we are very proud to now release. This Preface is the intertwined story of three authors and the publishing company that agreed to support them.

While employed as a school psychologist in the Syracuse City schools in central New York, Dr. James McDougal served in a district-level position, consulting on young students with challenging behavior, and also working with community and county mental health organizations to integrate therapeutic services into the schools. Over several years, a district-level behavior consultation team was established to support schools to develop behavioral interventions for young, often aggressive students. Additionally, mental health services were established in seven or eight schools with a variety of arrangements and utilizing services from different providers. With the development of these services came a need to evaluate the effectiveness of the services provided and the outcomes of the children and youth receiving them.

A search of the academic literature for an assessment that could monitor the progress of school-aged children receiving therapeutic, psychosocial, and/or behavioral interventions suggested that there was no brief, repeatable behavioral measure that was developed to be “change-sensitive.” The techniques available involved either lengthy behavior rating scales also used for diagnostic purposes or direct observation techniques. Neither of these seemed feasible for widespread use in schools.

At about this time, Dr. McDougal also became extremely interested in the Response to Intervention (RTI) model for providing academic and behavior support for needy students. The new RTI model seemed like a significant improvement over pre-referral intervention team models in that universal screenings were conducted to ensure that all needy students received services, not just those who were referred by their teachers. Further, the creation of levels of interventions of varying intensity based on student need and the requirement for progress monitoring with data-based decision making made this a very exciting educational model. While the evolution of validated progress monitoring techniques for measuring academic skill development was impressive, the same was not true for innovations in measures of behavioral and social-emotional development. The measures available did not seem feasible for universal screenings, were not developed for intervention monitoring, and lacked the utility needed to evaluate the effectiveness of a variety of intervention programs.

While Dr. McDougal's interest in developing a Behavior Intervention Monitoring and Assessment System grew, it was by very good fortune that he met Drs. Scott Meier and Achilles Bardos. Dr. Meier, who is a counseling psychologist and professor at the State University of New York at Buffalo, has conducted considerable work in creating behavior rating scales designed to be sensitive to change. Dr. Meier is a gifted researcher and scholar who, over the past 20 years, has developed, refined, and empirically supported an Intervention Item Selection Rules model for the creation of “change-sensitive” scales to measure behavioral and social-emotional functioning. Dr. Meier's model for scale development constitutes the basis for the BIMAS.

Dr. Meier's connection to the BIMAS relates to his professional research interests in psychological measurement, particularly outcome assessment in counseling and clinical practice. His research in occupational stress in the 1980s made him aware of the limitations of many methods of psychological testing and that awareness led him to research and write a book tracing the history of testing. That effort convinced Dr. Meier that advances in science (and related advances in practice) often depend on advances in measurement methods. Consequently, he refocused much of his research away from substantive areas and toward measurement methodology. In particular, Dr. Meier spent over a decade investigating change-sensitive tests and the different test construction methods that are needed to produce such tests. These methods are very new to most psychologists and clinicians, and we think they fit very well with the new emphasis on RTI in schools. Given that new methods can be unfamiliar, we applaud Multi-Health Systems, Inc. (MHS; the publisher) for taking a chance on developing the BIMAS and devoting substantial resources into this extensive endeavor.

While the idea for a Behavior Intervention Monitoring and Assessment System was promising, it was the connection with Dr. Achilles Bardos that began to make the idea a reality. For over 20 years, Dr. Bardos has been actively involved in developing and researching a variety of psychological measures for school-aged children and youth. He was a statistical

consultant on the Devereux Behavior Rating Scale – School Form (Devereux–School Form; Naglieri, LeBuffe, & Pfeiffer, 1993) and a co-author of the Draw A Person: Screening Procedure for Emotional Disturbance (Naglieri, McNeish, & Bardos, 1991). More recently, Dr. Achilles published the Basic Achievement Skills Inventory (BASI; Bardos, 2004), a standardized achievement test that includes both survey and comprehensive forms and can be administered in either individual or group formats.

Along with his teaching and interests in computer applications in psychology, Dr. Bardos wanted a mechanism to track student behaviors and gauge the effectiveness of interventions—an accountable system that utilizes technology for progress monitoring. He saw the equivalent concepts of managed care practices move into public schools and started building a generic model to assist practitioners in the Individualized Education Plan (IEP) process utilizing software. When Dr. Bardos began the collaboration, he saw that Dr. Meier’s work in community centers and schools offered the empirically-based content for which he was looking. Dr. Bardos also recognized that Dr. McDougal’s work and applications in public school settings gave these efforts more meaning. In this respect, since we have been working on versions of the BIMAS since 2002, Dr. Bardos felt that we were ahead of time and practice as far as RTI is concerned, yet we are at a point where the principles and philosophy of RTI can be satisfied quite well by the BIMAS. Dr. Bardos has been invaluable in building a relationship with our publishing company, MHS, and in guiding the development of the BIMAS. He has dedicated significant effort toward BIMAS norming, standardization, and software development.

Even with the talents of this authorship team, the BIMAS would be little more than intricate scribbling on a cocktail napkin and a series of three-page proposals if not for the right publishing company. Hazel Wheldon, the Publisher and COO of MHS, encouraged us through the proposal process and shared our enthusiasm for developing a different type of assessment, even though that development deviated from procedures used in other traditional behavior rating scales. The entire team at MHS has been excellent to work with and we have benefited from their experience in test development, computer programming, software development, and publishing. While we have benefited from the support of many at MHS, we especially want to acknowledge the efforts of Jane Wong, Research Associate; Jenni Pitkanen, Manager, Product Development; Gill Sitarenios, Chief Scientist; Sara Rzepa, Manager, Data and Psychometrics; and Anand Nathan and Noel Doemello, Software Developers; who toiled with us throughout the development process.

The authorship team would like to extend many thanks to MHS for believing in our idea and for supporting us down the long road toward its fruition. We wish to acknowledge explicitly the equal efforts of each author in the development and publication of the BIMAS—it was truly a team effort. We are hopeful that the BIMAS can contribute significantly to the evolution of outcome assessments and RTI models of service provision for school-aged individuals receiving behavioral and psychosocial intervention.

*James L. McDougal, Psy.D.*

*Achilles N. Bardos, Ph.D.*

*Scott T. Meier, Ph.D.*

*September, 2011*

# Dedications

I would like to dedicate my efforts to this project to my wife Lisa and my two children Ben and Cara.

*James L. McDougal, Psy. D.*

To Katerina and Nicolette...for all the joys and purpose you are bringing to my life.

*Achilles N. Bardos, Ph. D.*

To Susan and Katie.

*Scott T. Meier, Ph. D.*

# Acknowledgments

I feel honored to be given the opportunity to work with Drs. James McDougal, Achilles Bardos, and Scott Meier on the BIMAS. Each of them brought years of experience and expertise in School and Counseling Psychology, psycho-educational assessment development as well as treatment outcome research into the product. Working with the authors has been a truly enriching and rewarding experience.

Thanks go to all of the dedicated individuals here at MHS whose talent and dedication contributed greatly to the development of the BIMAS.

- Hazel Wheldon, Publisher and COO
- Gill Sitarenios, Chief Scientist
- Thomas Gale, Director, Product Development
- Jenni Pitkanen, Manager, Product Development; Maria-Luisa Marocco, Project Associate; Kent Lam and Melissa Fudge, Senior Research Associates; Stephanie Mears, Research Associate
- Sara Rzepa, Manager, Data & Psychometrics; Craig Nathanson and Kevin Williams, Senior Research Associates/Psychometricians; Ronald Tumbas, Data Collection Coordinator
- Rick Walrond, Manager, Software and Development; Anand Nathan, Noel Demello, Richard LeBreton, Nicolei Martin, Jaroslaw Kunicyn, Nadia Danial, and Andrew Cherney, Software Programmers
- Charlene Colella, Manager, Project Management; Maria-Christina Micieli and Christine Teeple, Project Managers; Betty Mangos, Production Coordinator
- Scott Millar, Manager, User Experience; Vanessa Gratsas and Lisa Kucman, Content Editors/Document Designers; David Wiechorek, Technical Writer; Heather Co, Senior Graphic Designer; David Griesz, Senior Content & Graphic Designer; Gurcharn Birdi and Christopher Musgrave, Graphic Designers
- Rob Greenaway, Director, Clinical/Educational Division; Penny Koepsel, Danielle Politi, and Josie Woodson, Clinical and Educational Assessment Consultants

Collaborating with these highly skilled individuals at MHS has made the development process of the BIMAS a true team effort.

*Jane Wong, M.A.  
Research Associate  
September, 2011*

## Data Site Acknowledgments

MHS would like to thank the many helpful individuals who contributed to the BIMAS data collection effort. A significant amount of data needed to be collected from school districts and clinics across the U.S. and Canada. MHS recognizes the constant enthusiasm, positive attitudes, and co-operation these individuals exhibited during the development of this product.

The people listed below contributed significantly to the data collection process. These individuals, who truly care about the children and adolescents they work with, embody so much of what MHS stands for: excellence, constant growth, and the desire to improve lives.

Jeffrey Alpert, M.A., Lynwood Unified School District/Lindbergh Elementary; Alan Bragwell, LPC, Bragwell Services; Jamie Brigham, M.A.; Lisa Brockhuizen, M.A.; Mike Casale, M.A., Brockport High School; Michele Diamond; Jennifer Elkins, M.S., Moore Intermediate School; Barbara Kearney, M.S.; Tracy Larson, Psy.D.; David Leaman, Ph.D.; Annabella Lopez, Ph.D., Andrews University Center for Reading, Learning, and Assessment; Eva R. Markham, Ed.D., University of Louisville; James McCray, M.S.; Jodi Mullen, Ph.D., Integrative Counseling Services; Sheena Oakland, M.A.; Christine A. Offutt, Ph.D., Lock Haven University; Maria E. Rodriguez, Ed.S., Assessments Benefiting Children Inc.; Laurie Schreiber, M.S., Horry County School District; Clarissa E. Steffen, Ph.D.; Tim Truemper, M.A.

## Beta Site Acknowledgments

MHS would also like to thank all the dedicated individuals who contributed to the BIMAS Online beta-testing effort. The BIMAS beta test initiatives required the participants to attend a 6-week webinar program as well as the completion of weekly tasks including the setup of school and student data as well as completion of BIMAS Standard assessments for students in a number of classes for universal screening purposes. MHS appreciates all the time and effort that were put forth in the beta process by the staff members and all invaluable feedback regarding the BIMAS Online user experience.

Staff members at the two sites listed below contributed significantly to the beta-test process.

*Sumner Elementary School  
Claremont Unified School District  
Claremont, CA*

*Van Buren Elementary School  
Baldwinsville Central School District  
Baldwinsville, NY*

# 1 Introduction

This chapter provides an introduction to the Behavior Intervention Monitoring Assessment System (BIMAS™), a behavior rating system empirically developed to monitor children/youth's progress in behavioral and psychosocial intervention using items that are sensitive to change in response to intervention. The main features of the BIMAS are reviewed here, as is the theoretical link to Response to Intervention (RTI). An overview of the model used to identify the change-sensitive items for the BIMAS is also offered. Finally, a brief review of BIMAS uses and testing considerations is provided.

## Main Features of the BIMAS

The BIMAS represents an innovation in the assessment of behavioral, emotional, social, and academic functioning of school-aged children and youth. It is a brief and repeatable measure designed to be used for screening, progress monitoring, outcome assessment, and program evaluation within the RTI framework.

The BIMAS was developed with a new model for test construction, which involves using only items that are reliable and valid in the sense that they are sensitive to change in response to an intervention or treatment. The main features of the BIMAS include:

- multiple rating forms so that information may be collected from teachers, parents, youth, and clinicians
- a brief, standardized, change-sensitive Standard form that can be used for universal screening of behavioral, emotional, social, and academic difficulties as well as for assessing response to intervention
- Flex items that can be used when developing individualized goals and/or for intervention monitoring on a daily or weekly basis
- BIMAS Online offers dynamic web-based data-management, graphing, and reporting options, which allow users to examine data in a variety of ways, in real time, to assist in evidence-based decision making within an RTI framework

- multi-level report options for teachers, special education professionals, school psychologists, and principals as well as district directors or superintendents, to screen and monitor the progress of groups (i.e., class, grade, school, and/or district) and individuals
- utility for screening school-aged children and youth to identify those at risk for emotional/behavioral difficulties
- youth receiving an intervention can be periodically reassessed with a measure sensitive to change, thereby yielding an indication of response to the intervention or therapy
- development of the instrument included collecting more than 4,500 assessments, including the normative sample, clinical cases, and validity study cases
- excellent reliability and validity estimates
- easy administration, scoring, and interpretation of results

## BIMAS Components

The BIMAS Standard comprises five standardized scales. It can be used as a screening device to collect baseline data at the beginning of an intervention and for periodic progress monitoring. The BIMAS also includes a pool of items that can be used to customize behavioral intervention goals for frequent progress monitoring (BIMAS Flex). The BIMAS is available in a multi-informant format (Teacher, Parent, Self-Report, and Clinician).

### BIMAS Standard

The BIMAS Standard can be used as a screening device and/or to collect baseline data at the beginning of an intervention. The Behavioral Concern scales of the BIMAS Standard allow for the assessment of behavioral problems prior to, during, and at the conclusion of an intervention program. In preliminary BIMAS development research, Meier (e.g., 1998, 2000, 2004) found that change-sensitive items could be grouped into two categories: distress/problems and strengths. Based on this original categorization, the BIMAS further evolved into three Behavior Concern scales—Conduct, Negative Affect, and Cognitive/Attention, and two Adaptive Scales—Social and Academic Functioning. The Behavioral Concern

scales, as the name suggests, assess problem behaviors and comprise negatively worded items. The Adaptive scales include mainly items with positive content and are used to assess increases in adaptive functioning that are often the targets of intervention.

### BIMAS Flex

The BIMAS Flex can be used as an extension of the Standard form by allowing assessors to target specific behaviors suggested by *Concern* or *Mild Concern* item-level scores on the Standard form. While the BIMAS Standard provides information on the individual's progress across scales, Flex items provide information on specific intervention targets.

For each of the items on the BIMAS Standard form, there are both positively worded and negatively worded Flex items for the assessor to select from. Assessors can also custom-design Flex items to suit their own needs. Flex items allow assessors to create one- to five-item mini-assessments for frequent progress monitoring to gauge any decrease in behavioral concerns, any increase in adaptive skills, or both, simultaneously.

## Multi-Informant Form Options

Both the BIMAS Standard and the BIMAS Flex have three main rating form options with parallel content, including Teacher, Parent (for rating youth aged 5–18), and Self-Report forms (for youth aged 12–18). (A fourth Clinician form, described in more detail below, is also available.) Teacher ratings reveal the youth's academic, social, and emotional behaviors in the school setting. Parent ratings describe the youth's behavior at home and in other environments. Self-report ratings collect a third source of information that can supplement teacher and parent reports by providing the youth's own insight into his/her functioning (Collet, Ohan, & Myers, 2003). Self-reports can provide valuable information about feelings and thoughts that might not be easily observable by others. There is consistency in the content of the scales and items across the teacher, parent, and self-report versions, thus facilitating the comparison of information between sources. A non-standardized clinician form (for rating youth aged 5–18) is also available with slightly different content to reflect behavioral functioning typically assessed in a mental health clinic.

## BIMAS Formats

This section provides a general overview of the administration and scoring options for the BIMAS. For details about administration and scoring, see chapter 3.

## Online administration and scoring

The BIMAS is usually completed and scored online. (For administration to raters who do not have Internet access, see *Paper-and-Pencil Administration*, below.) The online administration may be completed in a school or clinical setting, or a link can be emailed to the rater(s) who would then be able to open the online assessment on their Internet browser (see *Hardware and Software Requirements* in chapter 3, *Administration and Scoring*, for a list of recommended browsers). Responses are automatically recorded, scored, and saved to a secure server. Raters who have a valid BIMAS Online username and password (e.g., a teacher) can complete BIMAS Standard and/or Flex assessments within the BIMAS Online for any youth to which the rater has been granted access (e.g., that teacher's own students). Paper-and-pencil administrations are entered manually or scanned into the BIMAS Online for scoring.

## Paper-and-pencil administration

Paper-and-pencil forms can be generated in batches from the BIMAS Online for school/district-wide universal screening as well as for individual students. Student demographic information is automatically populated on the form for ease of administration. Upon completion, the forms can be scanned in batches using any regular scanner and uploaded to the web-based data management system for scoring. (Alternatively, a manual score entry option is also available for entering individual or small numbers of responses.) The paper-and-pencil forms cannot be scored by hand.

## Report Types

BIMAS reports can be generated quickly and easily from BIMAS Online. The BIMAS offers a number of report types on a multi-student level as well as individual level. Multi-student-level reports allow personnel involved in the implementation of RTI at schools or districts to assess and monitor behaviors at a group, class, grade, school, and/or district level. Assessing and monitoring at-risk students or those who require intensive behavioral intervention is facilitated by different types of reports offered at the individual student level.

There are four broad categories of BIMAS reports:

1. **Assessment Reports.** Results of a BIMAS Standard administration (i.e., static results from a single point in time) are presented numerically and graphically. Assessment reports are available at the Multi-student level for Universal Screening or Intervention Design for a district, school, grade, class, and/or intervention group. At the Individual level, a student's assessment report from a Universal Assessment can be used to

assist in identifying problem areas for the student and to aid in Intervention Design (e.g., in the formulation of an Individualized Education Plan [IEP]) both at the scale-level and item-level.

2. **Progress Reports.** Two or more BIMAS Standard or Flex assessments can be monitored with Progress reports. At the Multi-student Level, Progress reports can assist in Program Evaluation for a district, school, grade, class, and/or intervention group. When generated at the Individual Level, student Progress Monitoring can be examined at the scale-level and item-level with the aid of time series graphs.
3. **Comparative Reports.** These reports offer comparisons between the results of different groups of students or individuals. At the Multi-student level, Comparative reports compare BIMAS Standard scale scores between different schools, grades, or classes numerically and graphically. At the Individual Level, BIMAS Standard and/or Flex assessments by different raters for the same student can be compared to assist in Student Progress Monitoring.
4. **Demographic Reports.** Percentage/numeric breakdown of students in a district, school, grade, or class in various categories of a particular demographic variable (e.g., age, gender, service code, or race/ethnicity) are presented graphically in demographic reports. They are available at the Multi-student Level only, since they require the aggregation of students from different demographic groups. This report type is ideal for identifying and targeting specific demographic risk factors in Universal Screening or Intervention Design.

Within each of these four categories, multiple report types are available for different screening and/or progress monitoring needs. See chapter 4, *BIMAS Reports*, for a detailed discussion of BIMAS report types and features.

## Uses/Applicability

The BIMAS is a useful tool for screening, intervention planning, progress monitoring, outcome assessment, and program evaluation. These uses are summarized in this section; for a more in-depth discussion, see chapters 4 to 7, *BIMAS Reports, Understanding and Interpreting BIMAS Scores, Step-by-Step Interpretation of BIMAS Results*, and *Case Study*.

## Screening

The BIMAS Standard is well suited for screening school-aged youth for behavioral and social/emotional difficulties in order to detect students at risk and to identify their respective areas of need. The BIMAS Standard can also be used to collect baseline data at the beginning of intervention.

## Intervention Planning

Results from the BIMAS can be used to help create individualized intervention plans. Elevated BIMAS Standard scale scores suggest areas to target in an intervention, and endorsed individual items suggest specific behaviors that may require intervention. The BIMAS Flex items can be used to zero-in on further specific intervention targets for the specific problem areas suggested by elevated scores on the Standard form.

## Progress Monitoring

Progress monitoring generally consists of more frequent assessments for students identified as at-risk and/or those receiving intervention services. The BIMAS Standard can be periodically re-administered to monitor whether students are responding to a particular intervention plan. Results from re-administration may indicate the need for changes or modifications that would make the intervention more effective. Smaller assessments comprised of BIMAS Flex Items can be used more frequently to focus on specific treatment targets and to gather raters' perceptions of student progress.

## Outcome Assessment

Outcome assessment generally refers to the overall effect(s) of an intervention or treatment for the student and is also referred to as summative assessment. Outcome assessment would be the assessed effect at the end of a pre-determined point in time (e.g., at the conclusion of the school year, or at the end of a 6-week group intervention). Because the BIMAS Standard is a standardized change-sensitive measure, it is well suited to serve as a main component of a comprehensive outcome assessment. Changes in youth behavior over time as rated by teachers, parents, or the youth themselves provide a basis for assessing the effects of a treatment or intervention. Although norms are not available on the Clinician form, it can still offer supplemental information regarding a youth's progress.

## Program Evaluation

Program evaluation generally refers to the assessment of progress or change observed in groups of students receiving the same or similar treatments. Results from the BIMAS

can inform decisions about the effectiveness of a particular individual or group intervention. BIMAS results can be collected at the beginning of an intervention and at several points throughout the intervention to evaluate whether a particular program is associated with reduction in social, emotional, and/or behavioral concerns. This methodology allows assessors to determine the most effective interventions for the groups of children and youth being served.

## Principles of Use

The BIMAS is not recommended for use with those who are unwilling or unable to respond honestly to a questionnaire; this may include some adolescents who may under-report socially sensitive behaviors during initial assessments. The BIMAS is not recommended for use with individuals who are disoriented or severely impaired, who possess poor reading abilities, or who are not proficient in English. In some cases, the assessor may wish to read scale items aloud to ensure that the rater understands each item; see *Reading the BIMAS Items Aloud for a Rater* in chapter 3, *Administration and Scoring*, for instructions.

It is particularly important to remember that as a rating scale, the BIMAS is a tool to aid in diagnostic decision making and not a diagnostic instrument. The BIMAS is not intended to be the sole tool used for early identification, diagnostic decision making, assessing a child, or treatment planning. Assessors are cautioned against drawing unsupported interpretations. To obtain a comprehensive view of an individual, the assessor must combine BIMAS score information with information gathered from other psychometrically sound measures, as well as other pertinent information obtained through interviews and discussions with the individual and other informants. Consideration of information from other sources helps the assessor make a more comprehensive and ecologically valid assessment. The BIMAS cannot be used as a substitute for direct clinical observation. It is important to consider any factors that may bias or otherwise compromise results, such as socially desirable responding, misunderstanding of item content, or carelessness in responding.

MHS (the publisher) and the authors are not responsible for the use of the BIMAS assessments in any manner other than that which is outlined in this manual. If the user intends to use the BIMAS in any way other than that which is outlined in this manual, the user must justify its use and interpret results with caution.

The BIMAS was carefully developed and researched to provide the most useful set of items for clinical identification and intervention. Any rating scale has inherent limitations; however, when used appropriately, the BIMAS can be extremely useful in the entire process of defining a problem, eliciting further information from respondents, planning

treatment and intervention, and measuring treatment outcomes. The process is a dynamic one that requires the synthesis of several types of information and sound clinical judgment.

Use of the BIMAS scales for automated diagnoses or actuarial decision making is inappropriate and is therefore strongly discouraged. Raters approach their tasks with varying degrees of observational skill, openness, defensiveness, and candor. The context of assessment is important; the relationship of the respondent to the assessor may vary considerably, and results may vary accordingly. The BIMAS can never be used as a substitute for a trained and experienced assessor's overall judgment regarding treatment planning and treatment outcome.

## Users and User Qualifications

Potential users of the BIMAS include psychologists, counselors, Positive Behavior Support (PBS) specialists, principals, teachers, clinicians, and administrators. The BIMAS can be administered by individuals who do not have advanced formal training in clinical psychology or psychometrics (e.g., teachers), but the procedures outlined in this manual must be followed closely. A test interpreter with access to *T*-scores on BIMAS reports must have MHS B-level qualifications, which require that, at minimum, he/she has completed graduate-level courses in tests and measurement at a university or has received equivalent documented training.

In addition, those interpreting the BIMAS standardized scores must be familiar with this manual and should be familiar with the standards for educational and psychological testing jointly developed by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (AERA, APA, & NCME, 1999). Furthermore, qualified interpreters of this test should also be members of professional associations that endorse a set of standards for the ethical use of psychological or educational tests, or licensed professionals in the areas of psychology, education, medicine, social work, or an allied field. Test users must assume ultimate responsibility for the administration, scoring, and interpretation of any test materials they use. Individuals whose only exposure to testing is gained from this manual will not, in general, be qualified interpreters of the BIMAS. Experienced BIMAS interpreters should periodically review the administration and interpretation portions of this manual to ensure that accepted principles are followed and best practices are regularly used.

## Contents of the Manual

This manual contains information required to use the BIMAS. This chapter introduced the BIMAS. Chapter 2 outlines the theoretical and conceptual framework upon which the BIMAS is based. Chapter 3 provides information about administration and scoring. Chapter 4 provides the description, functions, and main features of BIMAS reports. The types of scores on the BIMAS and interpretation guidelines for each type of score are provided in chapter 5. Step-by-step interpretation guidelines of BIMAS results at both the multi-student and individual levels are provided in chapter 6 and a case study is provided in chapter 7. Chapter 8 gives an in-depth look into the development of the BIMAS. Standardization, reliability, and validity information are provided in chapters 9, 10, and 11, respectively. Finally, chapter 12 presents some concluding comments.

# 2 Background

The Behavior Intervention Monitoring Assessment System (BIMAS™) was designed to be used within the Response to Intervention (RTI) framework and was empirically developed to monitor treatment outcomes. This chapter outlines the theoretical and conceptual framework upon which the BIMAS is based. The chapter begins with a discussion of the kinds of concerns commonly found among school-aged youth and how those concerns relate to assessments within educational and/or clinical settings. Following that, the 3-Tier RTI model is described and the strengths and weaknesses of existing outcome assessments are discussed. Finally, the chapter presents the advantages and utility of the BIMAS, a change-sensitive measure, and describes how it differs from other existing outcome measures.

## Concerns in School-Aged Youth

The BIMAS target population includes youth from a wide age range—between the ages of 5 through 18. Many different types of problems may be encountered in this substantial age range. The types of concerns experienced by school-aged youth are typically classified into three domains: behavioral, affective, and cognitive. Behavioral concern is a broad category that includes such problems as acting impulsively, fighting with others, lying or cheating, and fidgeting. Feelings of anger, depression, and shame are examples of affective issues experienced by children and adolescents. Finally, cognitive problems include issues such as trouble paying attention, following directions, and staying on task.

Most theorists believe that any problem involves a combination of behaviors, cognitions, and affect (Barrett, 2006). In the schools, challenging childhood behavior has been defined as those acts that interfere with skill acquisition, that result in injury to self or others, that result in damage to physical surroundings, or that isolate the student (Doss & Reichle, 1991; Reichle et al., 1996). There has been a significant increase in the number of youth who demonstrate these types of behaviors. Prevalence rates have been noted at around 14% to 20% for typical or at-risk youth and at 13% to 30% for youth with developmental disabilities (Brandenburg, Friedman, & Silver, 1990; Reichle et al., 1996).

Studies have consistently indicated that aggressive, challenging, and noncompliant behavior in children even as young as pre-school age is typically not outgrown and, to the contrary, has a high potential to get worse over time (Kazdin, 1993). In a review of early intervention with children with conduct problems, Reid and Patterson (1991) indicated that in terms of developmental stability, aggressive behavior is comparable to the stability of intellectual performance. Furthermore, they suggest that for a substantial group of children (especially boys), seriously challenging and/or aggressive behaviors that are likely to emerge by first grade will remain relatively stable thereafter. Taken together, these findings suggest a clear need for effective early intervention.

## Response to Intervention (RTI)

In order to counteract problems typically encountered by school-aged youth, many schools and community-based and/or behavioral-health organizations have adopted 3-Tier (or in some cases 4-Tier) models of prevention and intervention. In general, these models incorporate some screening of children/youth and the provision of services of varying intensity, based on need. Individuals receiving services are then monitored to assess progress over time and to inform intervention revision or discontinuation. In the schools, this 3-Tier model has been coined “Response to Intervention” (RTI) and is included in federal legislation, including the Individuals with Disabilities Education Improvement Act (IDEA) of 2004. While most widely adopted for academic skills development, RTI has significant implications for the treatment of behavior and social-emotional skills development. The following section briefly reviews the RTI model and its relation to behavior progress monitoring.

## The 3-Tier RTI Model

Using the 3-Tier RTI model, school officials are able to determine how to match a student’s level of need with the appropriate resource (intervention) intensity. A 3-Tier model is a way to accommodate students within a continuum of three levels of support that addresses the needs of all students, not just those with diagnosable learning or

behavioral problems. In this model, students who respond to the intervention either continue to receive or are faded out of services. The RTI approach fits well with the trend toward data-based decision making and increased accountability across many fields (including behavioral and mental health services) and forms the foundation for preventative 3-Tier models being implemented in the schools (Fiona, 2005; Klotz & Canter, 2007). The following section describes each tier in more detail.

### **Tier 1**

Tier 1 includes approximately 80% to 90% of the student population who will respond to the core curriculum and behavioral supports provided in general education. Tier 1 is the beginning of the RTI process, where the general population is screened to (a) assess the efficacy of universally implemented preventative programming (e.g., character education); and (b) to identify those students who may require additional assessment and/or assistance. Additional assistance takes the form of evidence-based intervention. In Tier 1, these interventions are generally grade- or classroom-based.

### **Tier 2**

Tier 2 includes the approximately 10% to 15% of students who will require supplemental supports along with standard programming in order for positive change to occur. The Tier 2 group includes those students who show some increased risk for emotional or behavioral disturbance but do not meet the criteria for a clinical diagnosis. Supplemental intervention can include group or individual counseling, social skills training, and an added level of behavioral support or some other specialized psychosocial intervention specific to the individual's area of risk. Successfully remediated students are transitioned to Tier 1 supports. If students are not responsive to Tier 2 interventions, additional and more intensive interventions are provided, perhaps within the Tier 3 framework.

### **Tier 3**

Tier 3 includes the approximate 5% or less of students who evidence chronic and persistent emotional and behavioral difficulties, and therefore require intense and individualized interventions which might include a specific treatment plan, specialized classes, training programs, counseling, and/or involvement of inter-agency support plans (Tilly III, 2008; Gresham, 2008).

## **RTI and Behavior Progress Monitoring**

While often considered in the context of academic areas such as reading, the RTI philosophy has been shown to be useful in addressing students' behavioral health concerns.

One example, the School-Wide Positive Behavioral Interventions and Supports (School-Wide PBIS; U.S. Department of Education, Office of Special Education Programs, 2010) program is an empirically based, 3-Tier model being implemented in schools across the United States. The School-Wide PBIS was developed to help schools more effectively promote positive student behavior through a systematic and recursive data-based approach to developing and assessing school-wide behavioral interventions. Schools implementing these practices have reported significant reductions in antisocial behavior (e.g., McCurdy, Mannella, & Eldridge, 2003) and office discipline referrals (ODRs; e.g., Metzler, Biglan, Rusby, & Sprague, 2001; Nakasato, 2000), as well as increases in academic performance and social climate (e.g., McCurdy et al., 2003; Nelson, Martella, & Marchand-Martella, 2002).

Although many promising effects have been reported, there have been some concerns regarding the overreliance on ODRs as the primary source of effectiveness determination data. Limitations of relying solely on ODRs include (a) a lack of sensitivity to students with internalizing difficulties; (b) the potential for teacher bias in the documentation of student behavior; (c) variations in teacher tolerance for misbehavior; and (d) a lack of independent objective data related to the behavior (e.g., Morrison & Skiba, 2001; Nelson et al., 2002; Wright & Dusek, 1998).

What is needed to inform this type of systemic preventative programming is a general screening instrument that can be administered to large groups of children and adolescents to identify those at risk for emotional/behavioral difficulties. This instrument should be brief, reliable, and valid in terms of identifying students with significant behavioral health problems. Equally important, it should be adaptable and ready to be employed as a repeated measurement to identify change resulting from any subsequent intervention(s).

## **Outcome Assessments**

Outcome assessment refers to the process of gauging the effects of counseling and psychotherapy over the course of an intervention. While literature-informed recommendations for outcome assessment typically include the use of a psychological test, in contemporary practice, most clinicians employ only their own judgments or those of their clients about the type and amount of change that has taken place (Clement, 1999). Historically, outcome measures have predominantly been employed in research as a key component of efficacy (i.e., laboratory-based) and effectiveness (community-based) studies that examine the effects of psychosocial interventions (Hill & Lambert, 2004; Strupp, Horowitz, & Lambert, 1997).

## Outcome Assessments in Clinical Settings

Most research in the area of outcome assessment in clinical settings relates to adult clients. However, the same issues are evident in the area of outcome assessments for children and adolescents as well. According to investigators who have conducted psychotherapy research, most outcome measures have significant problems. Because the primary purpose of an outcome measure is to assess client change, any particular instrument's sensitivity to change is directly related to its construct validity (Vermeersch et al., 2004). Yet Hill and Lambert (2004) noted that "most outcome measures have not been developed with an eye toward choosing items that are sensitive to change, and little is known about this aspect of test validity" (p. 117). Outcome measures appear to differ considerably in their ability to detect the types and amount of change resulting from psychotherapy. Measures of depression have been found to produce estimates of change that differ as much as the estimated effect of treatment itself (Lambert, Hatch, Kingston, & Edwards, 1986). In addition, outcome measures based on specific targets of therapy produce greater effects than more distal assessments, therapist and expert judge ratings produce greater effects than client self-reports, and global ratings of symptoms produce greater effects than assessments of specific symptoms (Lambert, 1994). Lambert (p. 85) concluded that "there are reliable differences in the sensitivity of instruments to change."

Psychotherapy outcome research is also the basis upon which recommendations for evidence-based therapies are based. And yet, individuals exhibit significant variation when participating in any psychotherapeutic intervention. Many individuals will demonstrate improvement over time or in comparison to a control group; other individuals will show no change or even worsen with the same intervention. The problem for clinicians in practice is that no method exists for identifying which particular client will improve, stay the same, or worsen even when an evidence-based approach is employed. Given this uncertainty, clinicians need some method of monitoring ongoing progress.

In practice settings, a variety of decision makers, from school districts to managed-care companies, continue to press counselors and clinicians to produce data that demonstrate the effectiveness of clinical work (Botcheva, White, & Huffman, 2002; Brown & Reed, 2002; Gibbs, Napp, Jolly, Westover, & Uhl, 2002; Thayer & Fine, 2001). What Burlingame et al. (2006, p. 77) termed the "robust assessment of patient outcome" has become an essential task, as funding sources for mental health, educational, and social service programs now routinely require evidence of program effectiveness (Botcheva, White, & Huffman, 2002; Erbes et al., 2004; Gibbs, Napp, Jolly, Westover, & Uhl, 2002; Thayer & Fine, 2001). Erbes et al. (2004, p. 31) noted that there has been "a

persistent call within the field of psychology for practitioners to evaluate the outcomes of their treatment," particularly as managed-care organizations "have required the documentation of treatment efficacy in everyday practice."

When faced with decisions about the choice of an outcome measure, clinicians traditionally have employed tests that were not constructed for assessing change (Froyd, Lambert, & Froyd, 1996). Similarly, many of the best-known scales employed for outcome measurement with children and adolescents are lengthy diagnostic and screening instruments. Research participants may be required to spend an hour completing a battery of outcome measures, but in practice settings, most parents, teachers, children, and clinicians will not complete tests or tasks that they perceive as too lengthy. Brevity of an outcome measure may often be the single most important criterion for a measure's selection and actual use in clinical settings. In fact, research suggests that clinicians and clients have a very short tolerance threshold for completing such measures (Miller, Duncan, Sorrell, & Brown, 2005). Because perceived survey length influences completion rates (Dillman, 1978), lengthy, traditional scales become problematic (Meier & Letsch, 2000).

Most clinicians and clients want to spend as little time as possible completing outcome measures, and agencies such as community mental health agencies have relatively small amounts of money to spend on testing. In essence, clinicians are paid to provide interventions, not to collect or analyze the data that document that interventions work. Unlike medicine (Guyatt, Walter, & Norman, 1987) or education (Cross & Angelo, 1988), practitioners in fields like counseling, mental health, psychology, or social work historically have underutilized outcome assessment, thereby losing the benefits of feedback about client progress, evaluation of clinical outcomes and associated practices, and subsequent improvement of services for clients.

## Outcome Assessment in Educational Settings

The RTI approach uses a combination of assessment and intervention procedures designed to identify at-risk students who need additional interventions and to monitor their progress during those interventions. Unlike high-stakes testing strategies where the focus is on summative evaluation, RTI assessment focuses on providing teachers, counselors, parents, and administrators with information useful for assessing short-term student progress (Marston, 2005). In addition to benchmark assessment, a key to the success of the RTI process is repeated testing with students receiving evidence-based interventions. Thus, the RTI process requires tests useful for both benchmarking and repeated assessments of progress.

The implementation of RTI in the schools has largely been facilitated by the development of progress-monitoring techniques that are brief, repeatable, sensitive to short-term change, and valid and reliable for progress monitoring. With respect to academics, Curriculum-Based Measurement (CBM) techniques have allowed schools to universally screen students developing basic academic skills and to intervene with and monitor the progress of those who display noted difficulties. There are research-based as well as commercially available tools widely available for academic screening and intervention progress-monitoring activities. These tools helped in the advancement of school-based implementations of RTI.

With respect to behavior, existing screening and progress-monitoring techniques that have been implemented within the schools include the previously mentioned School-Wide PBIS program that makes use of ODRs as the primary data source. Other behavioral outcome measures used in the schools typically involve student tardiness, attendance, grades, and suspensions (Kingery & Walker, 2002; Tobin & Sugai, 1999; Walker & Severson, 1992; Walker & Shinn, 2002). While all of these data sources are routinely collected and easily available, none is specific to the behavioral difficulties of the student, nor does research exist indicating that these are valid or reliable techniques for monitoring progress.

### Strengths and Limitations of Current Outcome Measures

While the literature contains dozens of measures that could potentially be employed for progress assessment with youth, few of these were developed with the primary purpose of monitoring student progress. While potentially useful for comparing pre- to post-test change, many checklists are too lengthy for repeated assessments of progress required for comprehensive models of RTI, as noted above.

Test developers typically demonstrate strong psychometric properties for their scales, and the supporting evidence often comes in the form of moderate-to-high internal consistency estimates, moderate-to-high correlations with similar scales, and differences in expected directions between clinical and normal samples. A few scales also have evidence indicating that total and subscale scores show improvement after counseling. However, few of these measures have demonstrated that scores remain free from developmental or practice effects and are stable over time in control groups (e.g., free from spontaneous remission).

Given that maturation is most rapid during childhood and adolescence (Brems, 1993; Spiegel, 1987), changes over periods of as short as a few weeks or months may be as attributable to maturation as to psychosocial interventions. For all these reasons, outcome scales that do not include

an examination of item change in the absence of interventions or examine the potential effects of maturation may include items sensitive to practice or developmental effects. Moreover, while scores on some scales have been found to be intervention-sensitive, none of them appear to select individual items during the test construction process on the basis of their intervention sensitivity. Thus, a subset of the items on these tests is likely not to be sensitive to intervention. Change in aggregate total or subscale scores may be the result of a relatively small number of change-sensitive items, and the content of those items may be more specific than the label attached to the scale as a whole.

### The BIMAS: A Change-Sensitive Measure

Because the major purpose of educational and psychological tests historically has been to select persons for entrance to educational, business, and military settings on the basis of individuals' traits, test developers have sought items that discriminate among individuals based on traits of interest (Danziger, 1990). Reviewers of test construction procedures (Collins, 1991; Dawis, 1992; Guyatt, Walter, & Norman, 1987; Lipsey, 1983; Meier, 1994; Tryon, 1991) find that developers of traditional tests, such as measures of intelligence or vocational interests, seek items that maximize variability among individuals and demonstrate stability over time. The BIMAS was developed with the premise that a different set of criteria, however, is more appropriate for creating measures of change.

To identify intervention-sensitive items, either during test construction or during item evaluation, Meier (1997, 1998, 2000, 2004) proposed a set of Intervention Item Selection Rules (IISRs). These IISRs were designed to test two broadly competing claims regarding change at the item level: that such change is the result of an intervention, or that such change results from other factors that constitute error in the context of scale development. Meier has developed and studied scales constructed with both traditional scale development methods and IISR procedures in a variety of clinical and school settings (e.g., Meier 1998, 2000, 2004). Overall, scales constructed with IISRs procedures demonstrate larger treatment effect sizes than traditional scales and adequate reliability estimates.

The central philosophy of the IISRs is that intervention-sensitive items should be able show change in response to an intervention (i.e., a decrease in the maladaptive behaviors being measured or an increase in the adaptive behaviors measured). Moreover, intervention-sensitive items should also behave in a theoretically expected manner in other conditions (i.e., remain stable over time when no intervention is present). The IISRs approach assumes that (a) test

items and tasks differ along a trait–state continuum, and (b) different test construction and item-analysis procedures are necessary to select items with a high state loading that reflect the results of interventions.

The BIMAS is the first measurement tool explicitly designed to be change-sensitive based on the IISRs model. This characteristic allows the BIMAS to work perfectly within the RTI framework. RTI requires high quality, valid, and reliable methods for selection and measurement of change purposes. While traditional tests typically focus on diagnosis, they may be too long and too insensitive to treatment effects to be employed repeatedly for progress monitoring and outcome assessment. The BIMAS is a brief measure that may be completed by multiple informants, and it contains scales designed to perform student screening and progress-monitoring functions. In addition, BIMAS scores may be used for group-level program evaluation to assess student response to specific interventions, intervention providers, and/or intervention programs. Please see chapter 8, *Development*, for more detailed information about IISRs and their application in the context of BIMAS development.

# 3 Administration and Scoring

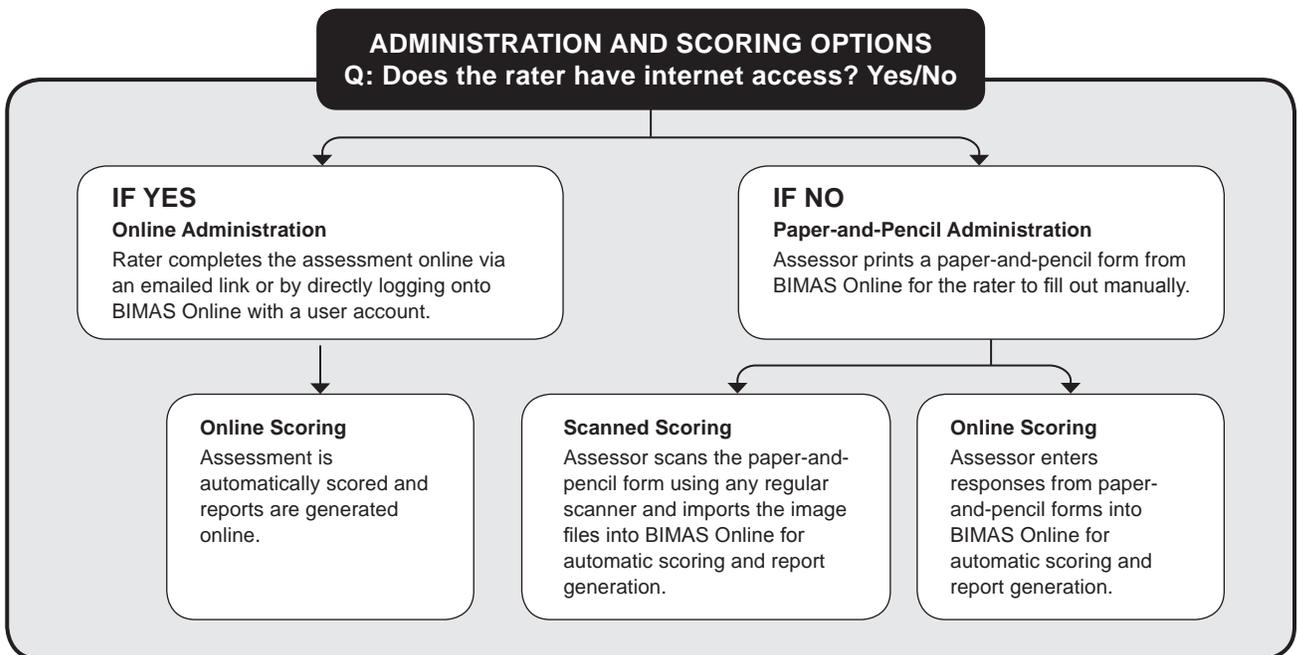
This chapter provides information about administering and scoring the Behavior Intervention Monitoring Assessment System (BIMAS™). Specific details are provided about age ranges, reading levels, administration time frame, and appropriate raters. This chapter also provides guidance that will help increase rater compliance while ensuring that ethical and legal guidelines are followed. Information about informed consent, confidentiality, and debriefing is also included.

The BIMAS can be administered online or via a paper-and-pencil form that the assessor prints through the BIMAS Online (a web-based system for BIMAS administration, scoring, and reporting). This chapter contains a brief guide on how to administer online and paper-and-pencil forms and how to enter ratings to score administrations. Figure 3.1 outlines the various online and paper-and-pencil administration and scoring options.

## Choosing the Appropriate Form

The BIMAS has both a Standard and a Flex version. The BIMAS Standard was developed for both screening and progress monitoring, and the BIMAS Flex was developed entirely for progress monitoring of specific treatment targets. Therefore, prior to any administration of the BIMAS Flex, the BIMAS Standard should be administered in order to identify which behavioral domain(s) are of concern and should be monitored. Thereafter, the decision concerning which version to administer will largely depend on the assessor's needs, intervention progress, and the rater's availability.

Figure 3.1. Administration and Scoring Options



## BIMAS Standard

The BIMAS Standard consists of 34 change-sensitive items that can be used for both screening and progress monitoring changes in response to intervention. There are four rater forms: the BIMAS–Teacher Standard (BIMAS–T Standard); BIMAS–Parent Standard (BIMAS–P Standard); BIMAS–Self-Report Standard (BIMAS–SR Standard); and BIMAS–Clinician Standard (BIMAS–C Standard). The BIMAS–T Standard, BIMAS–P Standard, and BIMAS–SR Standard contain 34 items parallel across forms, and the BIMAS–C Standard has a similar set of 31 items with an extra item related to therapy appointments attendance that replaces four items assessing academic functioning on the other three forms. For a list of items by scale across the different forms, consult appendix A.

All BIMAS Standard rater forms utilize the same response format. The form requires the respondent to indicate how often a particular behavior manifested during the past week: 0 = Never, 1 = Rarely, 2 = Sometimes, 3 = Often, and 4 = Very Often. Each of these response categories is paired with descriptors:

- *Never* is observed 0 times or not observed
- *Rarely* is observed 1–2 times or to a minimum extent
- *Sometimes* is observed 3–4 times or to a moderate extent
- *Often* is observed 5–6 times or to a significant extent
- *Very Often* is observed 7 or more times or to an extreme extent

## Purpose of BIMAS Standard

Ideally, the BIMAS Standard would be administered at the beginning of the school year as a school-wide screening tool. Following a period of getting familiar with their classes, teachers would be asked to provide BIMAS–T Standard ratings for their students. Additionally, the BIMAS–P, BIMAS–SR, and BIMAS–C are available to supplement the teachers’ ratings to provide information on how students fare in different settings. For instance, parents can be asked to provide BIMAS–P Standard ratings for their child in order to facilitate communication between parents and the school regarding the youth’s behavioral/emotional status following the summer break. Screening results on the BIMAS Standard at the beginning of the school year provides a baseline measure for each of the domains assessed on the BIMAS—Conduct, Negative Affect, Cognitive/Attention, Social, and Academic Functioning. The BIMAS Standard helps to identify and pinpoint area(s) of concern for individual students who might require further assessment or intervention. The school’s Behavioral Response-to-Intervention (RTI) Team can then develop intervention plans and establish a frequency of progress

monitoring to suit the youth’s behavioral needs. Any scales or items identified on the BIMAS Standard as concern can then be targeted and monitored frequently using the shorter, more individualized BIMAS Flex (see *BIMAS Flex*, in the following section). Since the BIMAS Standard is the norm referenced form<sup>1</sup>, it can be used for periodic progress monitoring (e.g., following a 6-week intervention program) and to verify any changes in behavior evidenced on the BIMAS Flex. The BIMAS Standard can also be used for follow-up school-wide screenings in the Winter, Spring, or even Summer term. See *Use of the BIMAS Forms within the RTI Framework* later in this section for more detailed discussion on screening and progress monitoring across the RTI tiers.

## BIMAS Flex

The BIMAS Flex items can be used as an extension of the Standard form for monitoring individual students with behavioral needs. For each of the 34 items on the BIMAS–T Standard, BIMAS–P Standard, and BIMAS–SR Standard (31 items on the BIMAS–C Standard), there is a pool of both positively and negatively worded Flex items for the assessor to choose from to create individualized BIMAS Flex forms. For a list of Flex items by scale across the different forms, consult appendix B.

Features of the BIMAS Flex include:

- a bank of positively and negatively worded items linked to each BIMAS Standard item.
- flexibility to create custom Flex items.
- flexibility to adjust the frequency of observation.
- the ability to assign the scoring criteria for each Flex item.
- the ability to assign the target item score for each Flex item.
- the ability to set the date by which a goal is to be achieved.
- time series graphs and goal lines.
- Assessment reminders, emailed to assessors to remind them of upcoming assessments.

## Purpose of BIMAS Flex

The rationale for Flex item development was three-fold. Firstly, while the BIMAS Standard is feasible for administration every 6 to 10 weeks, often student monitoring is desired more frequently. In the case of medication monitoring or monitoring the student’s response to intervention, weekly, daily, or even hourly data might be important. Flex items allow assessors the flexibility to create one to five item mini-

<sup>1</sup> Norms are available on the teacher, parent, and self-report forms only; not the clinician form.

assessments<sup>2</sup> for frequent progress monitoring. Since they are brief, compliance can easily be obtained from teachers, parents, or other invested parties to complete the forms on a fairly frequent basis.

Secondly, Flex items offer the flexibility of measuring either positive or negative behaviors (or both). Some items are worded positively where an increase in such behaviors are desirable, while others are worded negatively where a decrease in such behaviors would signify that the intervention is effective. Therefore, for each youth, assessors can evaluate the decrease in behavioral concerns and/or any increase in adaptive skills.

Lastly, Flex items allow assessors to create specific targets or treatment goals pertaining to a behavior that has been identified as a potential problem with the BIMAS Standard form. Standard form items flagged as a concern by one or more raters can provide an anchor for selecting Flex items. A Flex item can be selected from a global list of items, or a custom item can be created as deemed appropriate by the RTI Team. For example, more specific target behaviors such as “worked out problems with peers by himself/herself” or “used a strategy to resolve a conflict” might be selected for the more general Standard anchor item “worked out problems with others.” Using these Flex items, progress in the attainment of specific targets in an intervention (e.g., learning a specific conflict-solving strategy) can be monitored.

### BIMAS Flex Progress Monitoring Tools

Upon selecting/creating the Flex items, assessors can set the scoring criteria (i.e., Item Descriptor: *Concern, Mild Concern, Fair, or Positive*) for any Flex item. That is, a particular item response (i.e., 0 = Never to 4 = Very Often) can be determined as one that is desirable or one that should be flagged as a concern. For example, at an early stage of intervention, the assessor can decide whether a response of 2 (Sometimes) for a Flex item such as “took part in group activities” can be considered *Fair* for a youth who started off with a response of 0 (Never) at the beginning of the intervention. As the treatment progresses, the scoring criteria for any Flex item response can be adjusted at any time depending on the student’s response to intervention.

Moreover, a particular item response can be predetermined as the goal to be achieved within a specified time frame. For example, for the week prior to intervention, if the teacher’s ratings of a student’s on-task behavior were 1 (Rarely), a reasonable goal might be to increase teacher’s ratings of on-task behavior to 2 (Sometimes) within 4 weeks. Following weekly ratings, progress on item scores can then be compared to the Goal Line to determine whether the intervention is on target. This Goal Line is based on three things: (1) the item

score at the beginning of the intervention or assessment period (the baseline); (2) the specified target score (e.g., 4 = Very Often); and (3) the anticipated achievement date.

Assessment Reminders can be scheduled to remind the assessor of upcoming assessments. Upon receipt of an assessment reminder, the assessor can generate an online assessment link to be emailed to a rater. The rater then completes the Standard/Flex assessment remotely on their own time. Raters who have a BIMAS Online user account (e.g., teachers or clinicians affiliated with the school) and have been given access to the student by the system administrator can log on using their BIMAS Online username and password to complete the assessment. Refer to the BIMAS Online Tutorial for details on the assignment of Flex items and assessment reminders.

When the Flex assessment has shown improvement in the targeted behavior, the BIMAS Standard may be re-administered to provide a standardized assessment of change. Alternatively, if repeated Flex assessments are not illustrating positive change, the intervention can quickly be revised so that ineffective treatment components need not be continually implemented until the next scheduled Standard assessment.

## Use of the BIMAS Forms within the RTI Framework

The BIMAS Standard should be administered at Tier 1 (Universal Level) during Universal Assessments (UAs), typically in the Fall, Winter, and Spring. UAs are school- or district-wide behavioral screenings for the early identification of at-risk students and for progress monitoring over the school year. They are considered “Universal” because everyone in the school/district/organization is assessed. UA can be conducted in a homeroom class or a different period class as long as the teacher is well acquainted with the student to assess the student’s behavioral functioning. UAs can also take place alongside academic benchmarks. Examination of UA data at a specific point in time allows for: (a) the assessment of the overall functioning or performance of the population compared to a benchmark criterion; (b) the identification of areas in need of improvement (e.g., identify targets of Tier 1 intervention); and (c) the identification of specific individuals in need of additional evaluation or intervention. UAs across time allow for the evaluation of Tier 1 programming effectiveness. The BIMAS Online can report on up to four UAs in one academic year, allowing for progress monitoring at the group level across school terms within the year.

In addition to school-wide UAs, Group Assessments (GAs) can be conducted to assess and progress monitor small

<sup>2</sup> The use of one to five Flex items is recommended for brevity, which allows for frequent progress monitoring.

groups of students participating in an intervention group. For example, students from different classes or grade levels may be assembled together in a social skills training group or an anger management group based on the behavioral difficulties identified in a UA. The teacher/instructor facilitating the group can complete the BIMAS Standard for each student in the group in a GA to obtain baseline data for the group at the beginning of the intervention program. Follow-up GAs can then be used to track the group's progress. Examination of GA data at a specific point in time and across time serves the same functions as outlined for UA above, but in a smaller scope, specific to the group of students in question. Note that GA results are not provided in all BIMAS multi-student level reports. Refer to the *Types of Reports* section in chapter 4, *BIMAS Reports*, for the list of reports that provide GA results.

At Tiers 2 and 3 (Targeted and Intensive Levels), the scope of assessment moves to those individuals identified as at-risk or having intense need and the interventions or supports that they receive. Both the BIMAS Standard and BIMAS Flex can be used at these two higher tiers for more frequent progress monitoring purposes. Each assessment at Tier 2 and 3 can be used: (a) to assess the gap between a student's level of functioning against age-based expectations; (b) to identify specific students' needs so that they can be grouped for intervention; and (c) to establish baseline levels of functioning to inform goal setting and later program evaluation. Repeated assessments in these tiers allow for an individual's response to an intervention to be assessed, and with follow-up group assessments, the effectiveness of a given intervention can be assessed. Based on results, the intensity of the intervention for students might be modified, or an intervention program might be formatively revised in order to increase intervention efficacy. It is important to remember, however, that placement in specific tiers within the RTI model and the resulting services/intervention is determined by the decision of the school-based professionals managing the behavioral RTI program, and such decisions should be based on multiple sources of pertinent information, and not limited to BIMAS scores alone.

## Administration

There are two administration options for the BIMAS assessments: paper-and-pencil and online. The information included in this section is applicable to both methods.

### General Administration Guidelines

#### Age Range

Teachers, parents, and clinicians can rate youth who are 5 to 18 years old. Youth between the ages of 12 and 18 can

rate themselves on the BIMAS. The BIMAS should not be used to rate anyone beyond the age range of the standardization sample.

Note that the BIMAS Online will not prevent data entry and scoring for a youth outside of the recommended age range, but will issue a caution that the normative data (and therefore the *T*-scores and percentiles) are not applicable when one generates a *BIMAS Standard Individual Assessment Report* (see chapter 4, *BIMAS Reports*, for information on BIMAS reports).

### Reading Levels

The BIMAS can accommodate a low reading level. BIMAS reading level scores are expressed as a grade score (e.g., a reading level score of 6 indicates that the items can be read by the average sixth-grade student). The reading level values were determined using the Flesch-Kincaid Grade Level Formula (Flesch, 1948; Kincaid, Fishburne, Rogers, & Chissom, 1975):

$$\text{Readability Score} = .39 \times \left[ \frac{\text{total words}}{\text{total sentences}} \right] + \left[ \frac{\text{total syllables}}{\text{total words}} \right] - 15.59$$

The reading level for the BIMAS–T Standard and BIMAS–P Standard is 5.7 and for the BIMAS–SR Standard is 4.6. The average reading level for the BIMAS–T Flex and BIMAS–P Flex is 4.8, and BIMAS–SR Flex is 4.6. (Note that users can create custom Flex items and so the reading level on any Flex form may change, depending on the words used in the items). Alternative methods of administration may be needed (see *Reading the BIMAS Items Aloud for a Rater* later in this chapter) for youth with very low reading skills or reading comprehension abilities.

### Administration Time

The BIMAS Standard can be completed in approximately 5 to 10 minutes, while the BIMAS Flex administration time depends on the number of items included in the assessment, but any one to five item mini-assessment should be able to be completed in less than 5 minutes.

### Appropriate Raters

The BIMAS can be completed by teachers, parents, or clinicians for youth between the ages of 5 and 18 and by the youth himself/herself between the ages of 12 and 18. Teachers who provided ratings for youth in the normative study (see chapter 9, *Standardization*) included those who teach standard academic and elective courses and those who provide special educational services, remedial, general education, and enrichment/gifted programs. Parent data collected for the BIMAS normative study came from biological, adoptive, and foster parents. The most useful

parent data is usually provided by the primary caregiver, even if that caregiver is not the biological mother or father. Self-reports provide information about a youth's own perceptions and feelings about his/her behavior that parents and teachers may not be aware of. Clinicians who know the youth well can also provide invaluable information regarding the youth's mental health or progress in treatment. Information obtained across settings and by different raters can be used to develop a more comprehensive intervention plan for progress monitoring.

Whenever feasible, multiple respondents should be asked to complete the BIMAS for a given youth. Both convergent and divergent reports convey useful information that can help the assessor to formulate a more complete picture of the youth's behavior. It is also extremely valuable if several teachers who are familiar with the youth are able to complete BIMAS Teacher forms, especially for students who have been identified as belonging to Tier 2 or 3 within the RTI model or those who may require more intensive monitoring of behaviors.

### Time Frame

The rater should consider behavior observed *during the past week* when completing the BIMAS Standard. This time frame was chosen for several reasons. The BIMAS is meant to provide information on current/recent functioning. Raters generally find it easier to respond and the responses tend to be more reliable when raters are given a recent and fixed time period to consider. In situations where changes in response to intervention occur at a fast pace, responses based on the past week also allow for sensitivity to change over time.

The BIMAS Flex time frame is chosen by the assessor because it may be administered daily, every other day, weekly, biweekly, etc. With frequent administrations, however, the assessor should consider the increased likelihood of measurement error such as practice effects (e.g., changes in responses due to frequent, repeated administrations) taking place, and adjust decisions about interventions appropriately.

### Administration Settings

The BIMAS assessments can be administered on an individual basis, in a group setting, in person, or remotely. It is recommended that the BIMAS assessments be completed in a quiet setting with as few distractions as possible, and the assessment should be completed in one sitting. The *Administration Procedure* described later in this chapter is applicable to all settings.

When administered **remotely** (e.g., given to a rater to be completed at another time or sent to a rater via an emailed link for online completion), it is recommended that the assessor make personal contact with the rater (e.g., by scheduling a brief meeting) or include a cover letter (or email) that specifies the key information for accurate completion of the assessment, as outlined in the *Administration Procedure* section of this chapter. In addition to relating the information contained in the *Administration Procedure* section, encourage the rater to complete the assessment in one sitting and to complete it independently without the aid of others. For raters who will be completing a paper-and-pencil administration on their own, it is recommended that a stamped, addressed envelope in which to return the completed form be provided. If the rater is under the legal age of 18, before administering a BIMAS assessment, always obtain informed consent from the appropriate guardian of the child or adolescent, as well as from the youth. It is also recommended that the assessor provide general debriefing or feedback regarding the results following an administration.

When the BIMAS assessments are being administered in a **group setting**, which would most often occur in the case of class- or school-wide screening using the BIMAS–SR Standard, make sure each respondent has sufficient physical space to ensure privacy while responding. Instruct the respondents to complete the assessment to the best of their ability and to save questions until the end; allowing questions and/or comments during the response period may bias all youths in the room. In small groups (i.e., fewer than 10 people), one administrator, or “proctor,” is usually sufficient to ensure that all answers are provided independently and that the room is free from distractions. Several administrators may be required if the BIMAS is being administered to 10 or more people. Otherwise, group administrations can proceed in the same manner as individual administrations. In some cases, such as Tier 2 or 3 individual student assessments, scheduling an individual, in-person meeting with each youth before and/or after a BIMAS administration may be helpful for the youth to understand why he/she was asked to complete the ratings and how the findings from the assessment might help him/her. Any method of administration that deviates from the method outlined in the *Administration Procedure* section of this chapter should be noted by the assessor and caution must be taken when interpreting results.

### Reading the BIMAS Items Aloud for a Rater

If a rater cannot read the BIMAS (e.g., due to poor eyesight or difficulty comprehending written English), it is acceptable to read the instructions and items aloud. If possible, the rater should be given a separate form to follow during the reading, and should mark his/her own responses. If a rater

cannot record his/her own responses (e.g., due to limited motor skills), it is acceptable for the assessor to record the rater's responses on the form or online.

In addition to following the standardized administration procedures (see *Administration Procedure*, in this chapter), the assessor should:

- prevent inadvertent reinforcement of any responses (e.g., avoid saying “good answer” to a given response)
- avoid emphasizing the importance of any particular item
- ensure that the rater completes the items independently, without input from others
- take caution when interpreting results of a modified administration (deviation from the standardized procedure should be noted by the assessor and caution must be taken when interpreting results)

### Ethical and Legal Guidelines

Most assessors and practitioners are guided by certain ethical and legal principles relevant to their professional affiliation. The standards described in this section represent good practice, even if the practitioner is not required to follow them according to a given professional association or laws governing licensure/certification. These guidelines include obtaining informed consent, ensuring confidentiality, debriefing, and complying with copyright laws.

The process of obtaining **informed consent** begins by informing the rater about the general purpose of the assessment, including the reason he/she is being asked to complete the BIMAS. In most settings, the rater must be informed of any limits to confidentiality and of the efforts the administrator will take to protect confidential information. The rater must also be instructed that he/she has the right to stop the assessment at any time for any reason (consistent with the ethical standard of “freedom to withdraw”), but must be informed of the consequences of this choice and of any alternative options. “Informed consent” means that the rater has agreed to complete the rating scale without being forced to do so. Legally, consent can only be obtained from an adult. In the school setting, federal law requires the school to obtain consent from the youth's legal guardian before conducting any individual assessment; however, rules and regulations regarding informed consent for group assessments (e.g., school-wide) may differ between districts, schools, jurisdictions, etc. It is recommended that local laws and regulations regarding informed consent be obtained prior to any school- or district-wide screening. **Assent** is agreement from a youth to participate; although assent has little legal basis, it is important from an ethical perspective.

**Confidentiality** is often addressed as part of the informed consent discussion and should include information about two main points: protecting information about the child/adolescent, and protecting his/her responses. The assessor must be honest about who will have access to the information, how the information will be shared, and with whom the information will be shared. The youth and the youth's parents/guardians should be informed about the steps that will be taken to keep the information confidential (e.g., in a locked file cabinet in a locked office). Assessors are also advised to visit [www.mhs.com](http://www.mhs.com) to access the *MHS Security Policy*, which includes information about the advanced security methods used by the BIMAS Online for the protection of sensitive and confidential data.

Raters should be **debriefed** after responding. Debriefing includes a general explanation of how the collected information will be used. General debriefing can be completed by the assessor immediately after completion of the BIMAS. Detailed debriefing or feedback about the results requires knowledge of BIMAS scoring and interpretation and should be conducted by a professional who is qualified in these areas.

The BIMAS Standard and BIMAS Flex are **copyrighted assessments**; therefore, they cannot be copied without the explicit written permission of the publisher. The BIMAS assessments are also regulated tests; in accordance with the ethical and professional standards of the American Psychological Association and the Standards for Educational and Psychological Testing, the use and interpretation of results is restricted to qualified professionals (see *Users and User Qualifications* in chapter 1, *Introductions*, for more information). Completed BIMAS paper-and-pencil forms should be retained in a confidential and secure location as long as required by any applicable laws or regulations.

Recent laws require that individuals have access to the personal information contained in their files. Additionally, a file may be subpoenaed for a legal proceeding. It can be difficult to know how to provide the legally required information while maintaining copyright protection and security of test materials. It is recommended that local laws and regulations be followed for the appropriate storage and release of personal information, test materials, and results. Please visit [www.mhs.com](http://www.mhs.com) to view the *MHS Test Disclosure Policy*, which includes recommendations for best practices regarding client access to test results and the release of test materials under legal and ethical obligations.

### Administration Procedure

This administration procedure is applicable to the BIMAS Standard and BIMAS Flex, both paper-and-pencil and online versions. Administration of the BIMAS requires minimal

training. The administrator should follow this protocol closely to help increase compliance, reduce bias, and ensure that ethical and legal guidelines are followed.

## Preparation

- Ensure that informed consent has been obtained from the youth's parent or legal guardian (see *Ethical and Legal Guidelines* in this chapter).
- Ensure that enough time is available to complete the assessment in one sitting (see *Administration Time* earlier in this chapter).
- Provide a quiet setting with minimal distractions. If the rater is completing the BIMAS remotely (either paper-and-pencil or online), encourage him/her to choose a place and time that will be conducive to concentration.
- If administering the paper-and-pencil version, give the rater the appropriate form, a pen or pencil, and a comfortable place to work. If no table is available, provide a clipboard.
- Since all paper-and-pencil BIMAS forms are generated via the BIMAS Online as a downloadable PDF file, each printed PDF form contains the youth's demographic information (school name, student's name, grade level, class, date of birth, student unique ID, rater's name [if other than self], and the date the form was printed.). If the BIMAS–SR is administered in a classroom setting, the teacher should distribute the correct form to the individual student according to the name printed on the form. The teacher should also verify that the forms have the correct class name as well as teacher's name printed on them, otherwise the forms will be scored under the wrong class and/or teacher's name in multi-student reports. If the forms were generated for a group, the group name and group instructor's name would appear on the forms, and the group instructor should verify the accuracy of the group's information as well.
- If administering the online version locally, provide the rater with a computer that has Internet access.
- If administering the online version remotely, ensure that the rater has Internet access and a valid email address.
- Explain to the rater why he or she is being asked to complete the assessment. Use general language (e.g., do not use diagnostic labels) to avoid biasing the rater's responses.
- Emphasize that there are no right or wrong answers and that answers should be based on the rater's own opinions, observations, and in the case of BIMAS–SR, self-knowledge (i.e., not based on what other people think or say).

- Provide blank paper for the rater to write any additional comments or questions that may arise during the completion of the BIMAS assessment. Encourage raters who are completing the assessment remotely to keep track of any such comments or questions.
- Explain confidentiality. Inform the rater if the answers and scores will be shared with anyone. Doing so further increases the chances of honest responding and is part of the ethical/legal requirements of administration (see *Ethical and Legal Guidelines* in this chapter).

## During the Administration

- As mentioned in the *Preparation* section, BIMAS–SR paper-and-pencil forms generated through BIMAS Online have the youth's demographic information as well as the date on which the form was generated printed on them. The rater's name is also printed on BIMAS–T, BIMAS–P, and BIMAS–C paper-and-pencil forms along with the student demographic information.
- Note that the names of parents and clinicians are printed on the Standard and Flex forms in the Rater Legend area provided that such information had been set up in students' demographic information on the BIMAS Online<sup>3</sup>. The system allows for the information of up to four parents/guardians and four clinicians to be entered for every student. Before completing an assessment, the rater simply has to identify himself/herself by selecting the rater code next to his/her name in the Rater Legend area and bubbling in the appropriate Rater Code of the paper-and-pencil form. (A teacher is not required to identify himself/herself by selecting a rater code on BIMAS–T Standard or Flex paper-and-pencil forms because the selection of class or group name and teacher's name take place during the generation of forms.)
- The same selection of rater codes for parents or clinicians applies during online administrations where the rater is sent an online assessment link to complete the BIMAS on his/her own Internet browser. The list of parent/clinician raters will appear on the screen before the online BIMAS assessment begins and the rater must select his/her own name by clicking the corresponding radio button on the screen.
- Any rater completing a paper-and-pencil form will need to indicate the date of assessment. If necessary, help the rater to do so.
- Ensure that the rater reads and understands all of the instructions on the form.
- Encourage the rater to read each item carefully. The rater should decide about each item separately.

<sup>3</sup> Detailed instructions on how to set up student demographic information are available in the BIMAS Online Help (including instructions on how to add/edit parents'/guardians' and/or clinicians' names).

- The BIMAS assessment must be completed independently. Ask the rater to choose only one response for each item. Even though it may be hard to decide, the one response that the rater thinks describes the youth or himself/herself the best should be picked. If the rater expresses indecision about an item, the administrator can say, “I know that it may be hard to answer some of the questions, but please try your best and respond to each question.”
- Avoid biasing the rater’s response to any item. Even a simple clarification of one item can have an impact on responses to subsequent items. Refrain from any interpretation of responses. Make note of any explanations provided for specific items or general concepts.

### After the Administration

- For paper-and-pencil administrations, quickly review the form, including the date of assessment and rater information. If such information were missing, it may not be possible to score the form as those are required fields when entering responses on the BIMAS Online. Ensure also that all of the items have been completed. (See *Omitted Responses* in this chapter for information on how to handle missing responses.) It is particularly important to review the form when scanned scoring is the method of choice for importing responses onto the BIMAS Online. Bubbles that are not fully shaded may easily be recognized as omitted responses. (Note: The BIMAS Online allows for the correction or verification of imported assessments; see the BIMAS Online Help for detailed instructions).
- Ask the rater to correct any missing or unclear information.
- Debrief the rater by reiterating the reasons for doing the assessment, and provide a timeline for when results and recommendations will be provided. Communicate appreciation to the rater for taking the time to complete the assessment and that the information he/she provided is helpful
- Ask the rater if he/she has any additional comments or questions.

## Scoring

Practitioners can quickly and easily score and generate BIMAS reports using a computer that meets the hardware and software requirements with Internet access (see *Hardware and Software Requirements*, later in this chapter for details). BIMAS assessments can be scored over the Internet via the BIMAS Online or scanned in batches and uploaded onto the BIMAS Online. (See the BIMAS Online Help for the specific steps to take to enter and score responses.)

## Omitted Responses

The assessment results will be less valuable and less accurate if responses are omitted. Responses are treated as omitted if the rater skips an item. In the case of paper-and-pencil administration, items are also treated as omitted if the rater gives multiple responses to a single item (e.g., could not decide between two responses so endorsed both, or accidentally marked more than one response), or uses non-standard responses (e.g., the rater did not choose a numeric response but wrote a text explanation, or the rater marked an “in-between” response, such as “2.5”). The best solution is to quickly review the responses and obtain a proper response from the rater before the assessment period is complete. If this is not possible, the recommendation is to treat these missing or questionable responses as omitted items. The scoring function on the BIMAS Online displays a warning when more than 10% of the items on the assessment have been omitted on the BIMAS Standard.

Statistically, any scale can become unstable if too many responses are omitted. Table 3.1 presents the allowable number of omitted responses for each scale. If the maximum number of allowable omissions has been exceeded for a given scale, the score for that scale will not be calculated, and a note appears on the report indicating that there were too many omissions to calculate a score for the affected scale. If items have been omitted but the maximum number of allowable omissions has not been exceeded, the online scoring system will prorate scores to obtain an estimation of the raw score using the following formula:

$$\text{Prorated Score} = \frac{(\text{Obtained raw score for scale}) \times (\text{Total \# of items on scale})}{\text{Total \# items on scale with responses}}$$

The assessor should pay attention to which scales had missing response data as indicated in the report and interpret these scales with caution.

**Table 3.1. Maximum Number of Allowable Item Omissions for the BIMAS Standard**

Scale	Maximum Number of Omitted Items
<b>Behavioral Concern Scales</b>	
Conduct	1
Negative Affect	1
Cognitive/Attention	1
<b>Adaptive Scales</b>	
Social	1
Academic Functioning	1
<b>All Scales Combined</b>	<b>3</b>

## Features of Online Scoring

The BIMAS Online is accessible from any computer with Internet access meeting online requirements. It allows for different levels of reporting—district, school, grade, class/

intervention group, and individual student level. Responses from paper-and-pencil forms can either be entered into the online system individually for report generation, or scanned in batches using any regular scanner with automatic paper feeding function. Respondents with access to the BIMAS Online (i.e., a person with a valid username and password given by the administrator of the system) can complete assessments online. Reports can be generated instantaneously once assessments are completed. Respondents who are not authorized users on the BIMAS Online (e.g., parents, students, other mental health providers) can take the BIMAS assessment directly via the Internet by following a unique, emailed link. Online administration allows the respondent the flexibility to complete the assessment at a convenient location (e.g., home, school). It also significantly reduces administration and data entry time.

In addition, the BIMAS Online scoring provides:

- Data verification of scanned assessments.
  - A unique Form ID associated with each form ensures that student and rater information are linked to each individual form generated from the BIMAS Online, eliminating the need to re-enter the information when entering responses.
  - Optional report features that can be enabled or disabled according to the assessor's needs (e.g., 95% confidence interval).
  - Scores are sent directly to the web server which allows for immediate generation of reports.
  - Dynamic, web-based reports allow users to generate different reports in real time by changing any of the report variable menus (e.g., school, grade, rater form, etc.).
  - Direct linkage between related report types allows for easy access to pertinent information regarding specific students' level of risk/functioning using various drill-down options.
- Reports can be printed directly from the screen or in a printer-friendly format. Sections of the web-based report can be integrated (by dragging the report element[s] or using copy and paste functions) into the assessor's own full evaluation report.
  - Unlimited number of follow-up assessments can be administered and reports can be regenerated an unlimited number of times.
  - Records are saved for future access and for comparison of progress over time (e.g., year-to-year).
  - Records can be sorted and searched using various criteria (e.g., Student Name/ID, grade, class, Standard/Flex form, rater type).
  - Optional double-entry feature to verify data entry accuracy.

## Hardware and Software Requirements

The minimum requirements to use the BIMAS Online include Internet access and a current version of one of the following Internet browsers: Microsoft® Internet Explorer®, Mozilla Firefox®, Apple Safari®, Opera® or Google Chrome™. A 1024 x 768 pixels screen size is recommended. It is highly recommended that the latest version of any of the above browsers is installed when using the BIMAS Online.

## Technical Support

Multi-Health Systems Inc. maintains toll-free technical support numbers in both the U.S. and Canada. Subscribers to the BIMAS Online can call MHS in the U.S. at 1-800-496-TECH (8324) and in Canada at 1-800-268-6011 between the hours of 9:00 a.m. and 5:00 p.m., Eastern Standard Time. International customers can call +1-416-492-2627. Fax and email support is available at 1-416-492-3343 and support@mhs.com, respectively.

For more information about the BIMAS, call MHS in the U.S. at 1-800-456-3003 and in Canada at 1-800-268-6011. Email inquiries can be sent to customerservice@mhs.com.

# 4 BIMAS Reports

This chapter provides information about the types of reports available on the Behavior Intervention Monitoring Assessment System (BIMAS™). Information about the function and main features of specific reports as well as the level in which various reports are available at (multi-student vs. individual student level) are discussed. Report options are also described in this chapter.

## Types of Reports

The BIMAS has four basic types of reports: (1) Assessment reports, (2) Progress reports, (3) Comparative reports, and (4) Demographic reports. Assessment, Progress, and Comparative reports are available at both multi-student and individual levels. Demographic reports can be generated at the multi-student level. The reports are aligned with the Response-to-Intervention (RTI) focus on assessing functioning at one point in time as well response to intervention over time. Table 4.1 provides a list of all the BIMAS multi-student level reports and a description of the content and main features of each report. Individual reports are presented in Table 4.2. For ease of reference, each of the 16 reports is provided with a Report ID which is listed within the tables along with the appendix number where a sample of the report can be accessed. (*Note:* Report IDs can also be used to reference reports in the *Report* menu on the BIMAS Online). Table 4.3 presents the available level of each BIMAS report.

## Assessment Reports

Assessment reports provide information on student functioning at one single point in time.

### Multi-student Level

Assessment reports generated at the **district, school, or grade** level allow assessors to review the overall level of risk and adaptive functioning of students across the Behavioral Concern and Adaptive scales on the BIMAS. A review of such information may help district or school level assessors and/or administrators to pinpoint areas of need for psychosocial or behavioral programs and to identify groups

of students (i.e., schools, grades, or classes) most in need of such programming. Screening data from Universal Assessment (UA) at the beginning of each year may also serve as a baseline for comparison to later assessments. In particular, Assessment reports at the **class/group** level can be used to examine how a specific class is faring overall in a UA or how a group of students participating in an intervention program are doing in a Group Assessment (GA). (See *Use of the BIMAS Forms within the RTI Framework* in chapter 3, *Administration and Scoring*, for the definition of UA and GA.) The reports also allow for the identification of students within a class or intervention group who may be at risk for developing behavioral problems. These reports can also be used to identify students who possess adaptive strengths. Students with strengths in Social/Academic Functioning can potentially be paired-up with at-risk students in the class or act as positive role models for others.

The following Assessment reports are available at the multi-student level:

- *Risk Level Pyramids* (Report ID: 1; available for reporting a Class's UA results only)
- *Student List by Risk Level* (Report ID: 2; available for reporting a Class's UA results only)
- *Class/Group Student Scores* (Report ID: 3; available for reporting both a Class's UA results and a Group's GA results)

### Individual Level

At the **individual** level, the *BIMAS Standard Individual Assessment Report* (Report ID: 4) can be used to identify behavioral concerns and strengths of an individual student in one single assessment. The raters' responses to each item can also be assessed by the item analysis provided in the report to inform treatment/intervention goal formation, and in particular, the selection of Flex items that can be used to monitor the student's response to intervention.

**Table 4.1. Description, Special Features, and Function of BIMAS Multi-student Level Reports**

Type of Report	Report	Report ID	Description and Special Features	Function	Appendix
Assessment	Risk Level Pyramids	1	<ul style="list-style-type: none"> <li>Pyramid graphs showing risk level percentage distribution across the five BIMAS scales in one UA.</li> <li>Table showing number and percentage of students at each risk level.</li> <li>Breakdown by school (district level report), grade (school level report), or class (grade level report).</li> </ul>	<ul style="list-style-type: none"> <li>Identify general areas of concern in a district, school, grade, or class.</li> </ul>	E.1
Assessment/ Progress	Student List by Risk Level	2	<ul style="list-style-type: none"> <li>List of student names in alphabetical order sorted by risk level in a UA (Assessment report) or across several UAs (Progress report).</li> <li>Individual student names provide direct web links to the student's <i>Standard Individual Assessment Report</i> (Report ID: 4)/<i>Standard Individual Progress Report</i> (Report ID: 14) for the selected UA(s).</li> <li>"In Prog Monitor" column to show students who are currently in progress monitoring (● = Yes; ○ = No), if the information has been updated in the student Demographic Profile on the BIMAS Online.</li> <li>Student names can be sorted by first or last name, or by student scores, within each risk level.</li> </ul>	<ul style="list-style-type: none"> <li>Identify at-risk students.</li> <li>Monitor at-risk students' scale score/risk level progress over UAs.</li> <li>Track progress monitoring status of at-risk students.</li> </ul>	E.2a (Assessment) E.2b (Progress)
Assessment	Class/Group Student Scores	3	<ul style="list-style-type: none"> <li>Alphabetical list of student names and their scores across all five BIMAS scales in a class/group.</li> <li>Individual student names provide direct links to the student's <i>Standard Individual Assessment Report</i> (Report ID: 4) for the selected UA/GA.</li> <li>Student names can be sorted by first or last name.</li> </ul>	<ul style="list-style-type: none"> <li>Identify general areas of concern in a class or a group.</li> <li>Identify at-risk students.</li> <li>Identify students with adaptive strengths.</li> </ul>	E.3
Demographic	Demographics Breakdown	5	<ul style="list-style-type: none"> <li>Percentage/number of students by demographic variable in a UA.</li> <li>Offers two options of graphical display: pie chart (default) and histogram.</li> <li>Two display options for histogram graph: percent of students or number of students.</li> <li>Table showing number and percentage of students for each demographic category.</li> </ul>	<ul style="list-style-type: none"> <li>Identify percentage distribution of demographic variables within a district, school, grade, or class in a UA.</li> </ul>	E.5
	Risk Level by Demographics	6	<ul style="list-style-type: none"> <li>Stacked column graphs with percentage of students within each risk level by demographic variable in one UA.</li> <li>Three display options to show <i>High Risk</i> only (default), <i>High Risk</i> and <i>Some Risk</i> only, or all risk levels.</li> <li>Table showing number and percentage of students at each risk level.</li> <li>Direct link to the <i>Student List by Risk Level</i> (Report ID: 2) report for the selected UA, filtered by the selected demographic category.</li> </ul>	<ul style="list-style-type: none"> <li>Identify demographic risk factors within a district, school, grade, or class.</li> <li>Identify specific populations in need of intervention.</li> </ul>	E.6
Comparative	Average Score Comparison Report	7	<ul style="list-style-type: none"> <li>Comparison of group average scale scores between schools within a district, grades within a school, or classes within a grade in one UA in a bar graph and table format.</li> <li>Option to filter out certain Service Codes.</li> <li>Bar graph indicates the overall average of a district, school, or a grade stands against the group average <i>T</i>-scores.</li> <li>Table indicates the average level of risk for each group average <i>T</i>-score.</li> </ul>	<ul style="list-style-type: none"> <li>Determine the overall functioning of a group by obtaining group average <i>T</i>-score.</li> <li>Identify specific populations in need of intervention and those that are performing better than others.</li> <li>Determine group average score goals for subsequent UAs.</li> </ul>	E.7

Note. UA = Universal Assessment; GA = Group Assessment.

Continued...

Table 4.1. (Continued) Description, Special Features, and Function of BIMAS Multi-student Level Reports

Type of Report	Report	Report ID	Description and Special Features	Function	Appendix
Progress	Average Score Comparison: Progress Report	9	<ul style="list-style-type: none"> <li>• Comparison of group average scale scores between schools within a district, grades within a school, or classes within a grade over several UAs or across school years in a clustered column graph and table format.</li> <li>• Option to filter out certain Service Codes.</li> </ul>	<ul style="list-style-type: none"> <li>• Monitor progress in group average <i>T</i>-scores across time (UAs or school years).</li> <li>• Determine whether group-based intervention goals have been met.</li> <li>• Examine the need to revise group-level interventions.</li> </ul>	E.9
Progress	Class/Group Student Scores: Progress Report	10	<ul style="list-style-type: none"> <li>• List of student names and their scores across selected UAs/GAs on individual BIMAS scales in alphabetical order.</li> <li>• Individual student names provide direct links to the student's <i>Standard Individual Progress Report</i> (Report ID: 14) for the selected UAs/GAs.</li> <li>• Student names can be sorted by first or last name.</li> </ul>	<ul style="list-style-type: none"> <li>• Monitor scale score and risk level progress of students in a class or a group.</li> </ul>	E.10
Progress/ Demographic	Risk Level by Demographics: Progress Report	11	<ul style="list-style-type: none"> <li>• Comparison of the percentage of students within each risk level by demographic categories across several UAs or school years in a stacked column graph and table format.</li> <li>• Three display options to show <i>High Risk</i> only (default), <i>Some Risk</i> only, <i>Low Risk</i> only, or all risk levels.</li> <li>• Table showing number or percentage of students at each risk level.</li> <li>• Direct link to the progress version of the <i>Student List by Risk Level</i> (Report ID: 2) for the selected UAs, filtered by the selected demographic category.</li> </ul>	<ul style="list-style-type: none"> <li>• Monitor progress in risk level distribution of various demographic categories over time.</li> <li>• Determine the need to revise intervention programs and/or allocation of resources for specific populations.</li> </ul>	E.11

Note. UA = Universal Assessment; GA = Group Assessment.

Table 4.2. Description, Special Features, and Function of BIMAS Individual Level Reports

Type of Report	Report	Report ID	Description and Special Features	Function	Appendix
Assessment	Standard Individual Assessment Report	4	<ul style="list-style-type: none"> <li>• Student's Standard form assessment results across the five BIMAS scales in one administration.</li> <li>• Scale-level <i>T</i>-score graph and table (item responses for the Clinician form).</li> <li>• Item-level analysis.</li> <li>• Text summary of results and description of behaviors assessed by each BIMAS scale.</li> </ul>	<ul style="list-style-type: none"> <li>• Identify general areas of concern and/or strength.</li> <li>• Identify/prioritize behavioral items for intervention.</li> <li>• Items rated as <i>Concern/Mild Concern</i>, or <i>Positive</i> can be used as anchor items for the selection of Flex items for progress monitoring.</li> </ul>	E.4
Comparative	Standard Individual Comparison Between Raters	8	<ul style="list-style-type: none"> <li>• Comparison of Standard form scale- and item-level scores between a maximum of five raters (e.g., Teacher 1, Teacher 2, Parent 1, Parent 2, and Self) in a bar graph and table format (item responses for the Clinician form).</li> <li>• Lists all pair-wise statistically significant differences in scale scores between raters.</li> <li>• Item-level analysis of different raters' responses and the corresponding risk levels.</li> </ul>	<ul style="list-style-type: none"> <li>• Compare results between raters at both the scale and item-level.</li> <li>• Obtain information about a youth across different settings (e.g., at school vs. at home).</li> <li>• Identify/prioritize behavioral items for intervention targets.</li> <li>• Items rated as <i>Concern/Mild Concern</i>, or <i>Positive</i> by all raters can be used as anchor items for the selection of Flex items for progress monitoring.</li> </ul>	E.8

Continued...

**Table 4.2. (Continued) Description, Special Features, and Function of BIMAS Individual Level Reports**

Type of Report	Report	Report ID	Description and Special Features	Function	Appendix
Progress	Flex Individual Progress Report	12	<ul style="list-style-type: none"> <li>• Comparison of a student’s Flex item scores across administrations by the same rater in a time series graph and table format.</li> <li>• Displays comments by raters at each Flex assessment.</li> <li>• Scoring criteria determined by the assessor are incorporated into the score tables and graphs using risk level color zones.</li> <li>• Target item response and target date displayed in the graph as a Goal Line for easy comparison with the item score line for assessing rate of improvement.</li> </ul>	<ul style="list-style-type: none"> <li>• Monitor progress on specific behavioral goals.</li> <li>• Evaluate the effectiveness of intervention program.</li> </ul>	E.12
	Progress/Comparative	Flex Individual Comparison Between Raters	13	<ul style="list-style-type: none"> <li>• Comparison of Flex item scores between a maximum of five raters (e.g., Teacher 1, Teacher 2, Parent 1, Parent 2, and Self) in a graph and table format.</li> <li>• Displays comments by raters for each Flex item.</li> <li>• Scoring criteria determined by the assessor are incorporated into the score tables and graphs using risk level color zones.</li> </ul>	<ul style="list-style-type: none"> <li>• Compare results between raters to assess student’s response to intervention on specific intervention goals</li> <li>• Evaluate the effectiveness of an intervention program.</li> </ul>
Progress	Standard Individual Progress Report	14	<ul style="list-style-type: none"> <li>• Comparison of a student’s Standard form scale scores across administrations.</li> <li>• Phase lines displayed within time series graphs (if Intervention Phase Dates have been updated in the <b>Intervention Dates/Notes</b> section on the BIMAS Online).</li> <li>• Intervention Notes appear below the graph of the corresponding scale.</li> <li>• Direct link to the <b>Edit Intervention Dates/Notes</b> page on the BIMAS Online.</li> <li>• Direct links to the student’s <b>Standard Individual Progress Report: Significant Change Over Time</b> (Report ID: 15) and <b>Standard Individual Progress Report: Item Analysis</b> (Report ID: 16).</li> </ul>	<ul style="list-style-type: none"> <li>• Monitor student’s response to intervention at the scale level.</li> </ul>	E.14
Progress	Standard Individual Progress Report: Significant Change Over Time	15	<ul style="list-style-type: none"> <li>• Time series graph and table with a comparison of a student’s Standard form scale scores across administrations with the indication of statistically significant change based on <i>T</i>-scores (Reliable Change Index; RCI) or magnitude of change based on raw scores (effect size estimates).</li> <li>• Assessors with <i>T</i>-score rights have the option to assess change using either metric of change.</li> <li>• Assessors without <i>T</i>-score rights can assess changes in Levels of Risk/Functioning based on RCI.</li> <li>• Interpretive guidelines for both metrics of change.</li> <li>• Direct links to the student’s <b>Standard Individual Progress Report</b> (Report ID: 14) and <b>Standard Individual Progress Report: Item Analysis</b> (Report ID: 16).</li> </ul>	<ul style="list-style-type: none"> <li>• Monitor student’s response to intervention at the scale level.</li> </ul>	E.15a (statistically significant change based on <i>T</i> -scores)  E.15b (Effect size based on raw scores changes)
	Standard Individual Progress Report: Item Analysis	16	<ul style="list-style-type: none"> <li>• Comparison of a student’s Standard form item-level scores (item responses for the Clinician form) across administrations.</li> <li>• Analysis of item-level progress in a time series graph and table format.</li> <li>• Direct links to the student’s <b>Standard Individual Progress Report</b> (Report ID: 14).</li> </ul>	<ul style="list-style-type: none"> <li>• Monitor student’s response to intervention at the item level.</li> </ul>	E.16

Table 4.3. BIMAS Report Types &amp; Levels

Type of Report	Report	Report ID	Level Report Available At					
			District	School	Grade	Class	Group	Student
Assessment	Risk Level Pyramids	1	✓	✓	✓	✓		
Assessment/ Progress	Student List by Risk Level	2	✓	✓	✓	✓		
Assessment	Class/Group Student Scores	3				✓	✓	
	Standard Individual Assessment Report	4						✓
Demographic	Demographics Breakdown	5	✓	✓	✓	✓		
	Risk Level by Demographics	6	✓	✓	✓	✓		
Comparative	Average Score Comparison Report	7	✓	✓	✓			
	Standard Individual Comparison Between Raters	8						✓
Progress	Average Score Comparison: Progress Report	9	✓	✓	✓			
	Class/Group Student Scores: Progress Report	10				✓	✓	
Progress/ Demographic	Risk Level by Demographics: Progress Report	11	✓	✓	✓	✓		
Progress	Flex Individual Progress Report	12						✓
Progress/ Comparative	Flex Individual Comparison Between Raters	13						✓
Progress	Standard Individual Progress Report	14						✓
	Standard Individual Progress Report: Significant Change Over Time	15						✓
	Standard Individual Progress Report: Item Analysis	16						✓

## Progress Reports

Progress reports illustrate the effects of intervention over time for individuals or groups of students.

### Multi-student Level

Progress reports at the **district** level allow assessors to measure student response to district programming and can inform decisions regarding formative revisions of such programs. The efficacy of **school/grade** level intervention programs (e.g., school-wide character education or grade-level positive behavior initiatives) can also be shown through the use of Progress reports. These reports can either illustrate (a) changes in the percent and number of students within each risk level; or (b) changes in group average scores for the behavioral domains assessed by the BIMAS across UAs or school years. The efficacy of interventions implemented at the **class/group level** can also be examined with information from the progress report. For example, the results help identify areas that have improved or suggest emerging problems that may require new intervention planning. The use of progress reports at this class/group level also allows for identification of students who may need change in the level of services (e.g., elevated to higher Tier[s] if progress is not evident, or vice versa).

The following Progress reports are available at the multi-student level:

- *Student List by Risk Level* (Report ID: 2; this report can function as a Progress report for a Class when more than one UA is selected)
- *Average Score Comparison: Progress Report* (Report ID: 9)
- *Class/Group Student Scores: Progress Report* (Report ID: 10; available for reporting a Class's progress over UAs or a Group's progress over GAs)

### Individual Level

An **individual** student's response to intervention can be examined in multiple ways: time series graphs of scale or item scores across time as well as the evaluation of statistically significant changes in *T*-scores based on the Reliable Change Index (RCI) or the magnitude of raw score changes determined from effect size estimates. Refer to the *Scores for Progress and Outcome Monitoring* section in chapter 5, *Understanding and Interpreting BIMAS Scores*, for details and interpretative information on the BIMAS progress monitoring scores.

The following Progress reports are available at the individual student-level:

- *Flex Individual Progress Report* (Report ID: 12)
- *Standard Individual Progress Report* (Report ID: 14)
- *Standard Individual Progress Report: Significant Change Over Time* (Report ID: 15)
- *Standard Individual Progress Report: Item Analysis* (Report ID: 16)

## Comparative Reports

Comparative reports—offer comparisons between the results of different raters or groups of students.

### Multi-student Level

Comparative report at the **district** level can assist in examining how the district is performing on the whole in terms of an overall district average score on any of the behavioral concern or adaptive behavior domains. The *Average Score Comparison Report* (Report ID: 7) at the district level also displays the average scores across all schools selected in the report.

By comparing each school’s average score against one another, administrators at the district level can pinpoint which school(s) is/are performing better or worse than other schools within the district. Using the graph in the Comparative report, district level administrators can also visually compare the difference between each school’s average score against the overall district average. This information can be used to identify schools that are in need of behavioral or psychosocial programming. The ability to identify schools that demonstrate low levels of behavioral concerns and high levels of adaptive functioning may also be useful—investigation into the reasons for a particular school’s successful functioning may enhance the overall success of the entire district by providing guidance or direction for schools that may not be performing at the same level.

**School/grade** level Comparative report allows for grade-to-grade or class-to-class comparisons of average scores. For a detailed description on how to interpret comparative results at the group level, consult the *Scores for Comparing Results Between Raters and Groups: Comparing Between Group Averages* section in chapter 5, *Understanding and Interpreting BIMAS Scores*.

Note that the multi-student-level Comparative report is not available at the class/group level. Within classes/groups, comparisons between individuals can be done by using the *Class/Group Student Scores* (Report ID: 3) listed in the *Assessment Reports* section.

## Individual Level

Comparative reports at the **individual** level compare ratings from different raters for the same individual. By comparing scale scores as well as item scores across different raters, the assessor can identify discrepancies in scores to determine whether the student behaves differently in different settings (e.g., home vs. school) and if the student shares the same opinions about his/her own behaviors with his/her observers (e.g., parents or teachers). The BIMAS also offers a Comparative/Progress report (for BIMAS Flex assessments only), which enables the assessor to examine how different raters’ Flex items ratings change over time. This unique report type addresses questions such as: 1) Have there been parallel improvements made following intervention across settings and informants? and 2) Which rater perceived the most/least change? With the ability for raters to insert comments for each item on BIMAS Flex assessments, raters can comment on their responses or observations during intervention implementation and monitoring. Such comments may help the assessor to gain additional insights as to what worked or what might not be as effective for the youth in question. Individual-level Comparative reports include the following:

- *Standard Individual Comparison Between Raters* (Report ID: 8)
- *Flex Individual Comparison Between Raters* (Report ID: 13)

## Demographic Reports

Demographic reports provide graphical displays of the percentage and numeric breakdown of students along a dimension of demographic risk factor categories (e.g., Service Code, which includes categories of General Education, Special Education, and Title 1<sup>1</sup>).

### Multi-student Level

Demographic reports facilitate the examination of demographic risk factors which may be associated with certain levels of risk or adaptive functioning within the **district**, **school**, **grade**, or **class/group**. Demographic reports help provide a visual dissection of the percentage or number of students at each level of risk or adaptive functioning across demographic variables such as age, gender, Service Code, Meal Plan, or English as a Second Language (ESL). This information may be helpful for resource and program planning purposes by allowing administrators to identify

<sup>1</sup> Title 1 is a United States federal program that provides funds for additional instructional resources for students who are not performing at grade-level and are of a lower socioeconomic background.

particular demographic categories as risk-factors and consider funding or resources to mitigate those areas of need (e.g., increasing the funding within a district for ESL programs across schools). When progress on percentage/number of students within each risk level for each demographic category are compared across UAs or school years, assessors can also evaluate the effectiveness of intervention programs targeted to aid specific populations. The following Demographic reports are available at the group level:

- *Demographics Breakdown* (Report ID: 5)
- *Risk Level by Demographics* (Report ID: 6)
- *Risk Level by Demographics: Progress Report* (Report ID: 11)

## Report Options

The BIMAS Online provides several report options. Item responses on the five BIMAS scales are scored using standardized *T*-scores and each *T*-score is categorized into a specific level of risk or functioning (see chapter 5, *Understanding and Interpreting BIMAS Scores*, for a detailed discussion of BIMAS scores). All users on the BIMAS Online system can view risk/functioning level information on all the BIMAS reports. However, information pertaining to *T*-scores, percentile ranking, and confidence interval are only visible to users with *T*-score rights. This special right is given by the **Data Manager** (the BIMAS Online system administrator) who sets up other users on the system. Users with *T*-score rights can also change the following default settings on student-level reports:

**Confidence Intervals:** This option inserts confidence intervals in the table of detailed scores and in the Summary of Results section in the *Standard Individual Assessment Report* (Report ID: 4; see sample of the report in appendix E.4). The user can choose either the 90% (default) or 95% level of confidence (see *Scale Level-Scores: Confidence Interval* section in chapter 5, *Understanding and Interpreting BIMAS Scores*, for a discussion on how to interpret confidence intervals).

**Norms:** The default setting is to score using combined-gender norms. The assessor may choose to score using gender-specific norms (see *Norming Procedure & Derivation of Standardized Scores* in chapter 9, *Standardization*, for a discussion of combined-gender norms as the default scoring option).

**Summary of Results:** This option inserts a *Summary of Results* section in the *Standard Individual Assessment Report* in text format. Scores reported in the summary include the obtained *T*-score along with the 90% (or 95%) confidence interval and percentile ranking, as well as the level of risk/functioning associated with the *T*-score. Behaviors tapped by each scale are also provided in this summary section. The assessor may instead choose to turn the summary section off (i.e., it will not appear in the report).

# 5 Understanding and Interpreting BIMAS Scores

This chapter provides interpretation guidelines for scores on the Behavior Intervention Monitoring Assessment System (BIMAS™). Discussion on how to interpret scale- and item-level scores, how to compare results obtained over time (including how to determine both statistically significant and clinically meaningful change), and how to compare results across raters and groups of students are provided. As with any assessment tool, only individuals with the appropriate

background and experience can accurately interpret results from the BIMAS. See *Users and User Qualifications* in chapter 1, *Introduction*, for further guidance.

The BIMAS Standard provides both scale-level and item-level scores, as well as scores for progress and outcome monitoring and comparing results across raters and groups.<sup>1</sup> Table 5.1 presents the types of scores available on each BIMAS report.

**Table 5.1. Types of Scores Available on BIMAS Reports**

Report	Report ID	Types of Scores												
		Raw Score	T-score	Percentile	Confidence Interval	Scale Descriptor	Item Score	Item Descriptor	Reliable Change Index	Effect Size	Significant Difference in Scale Scores between Raters	Group Average T-score	Number/Percentage of Students at each Risk/Functioning Level	Number/Percentage of Students for each Demographic Category
Risk Level Pyramids	1												✓	
Student List by Risk Level	2		✓			✓								
Class/Group Student Scores	3		✓			✓							✓	
Standard Individual Assessment Report	4		✓	✓	✓	✓	✓	✓					✓	
Demographics Breakdown	5													✓
Risk Level by Demographics	6												✓	✓
Average Score Comparison Report	7											✓		
Standard Individual Comparison Between Raters	8		✓	✓	✓	✓	✓	✓		✓				
Average Score Comparison: Progress Report	9											✓		
Class/Group Student Scores: Progress Report	10		✓			✓							✓	
Risk Level by Demographics: Progress Report	11												✓	✓
Flex Individual Progress Report	12						✓	✓						
Flex Individual Comparison Between Raters	13						✓	✓						
Standard Individual Progress Report	14		✓			✓								
Standard Individual Progress Report: Significant Change Over Time	15	✓	✓	✓	✓	✓			✓	✓				
Standard Individual Progress Report: Item Analysis	16						✓	✓						

<b>LEGEND</b>	<span style="display:inline-block; width:15px; height:15px; background-color:purple; border:1px solid black;"></span> Assessment Report	<span style="display:inline-block; width:15px; height:15px; background-color:blue; border:1px solid black;"></span> Progress Report	<span style="display:inline-block; width:15px; height:15px; background-color:green; border:1px solid black;"></span> Comparative Report	<span style="display:inline-block; width:15px; height:15px; background-color:purple; border:1px solid black;"></span> Demographic Report
---------------	---	---	---	--

<sup>1</sup> The BIMAS–Clinician Standard is not norm-referenced; only item responses are available in Standard individual reports (Report IDs: 4, 8, and 16).

## Scale-Level Scores

### Raw Scores

The raw scores for each scale on the BIMAS Standard were converted to standardized scores to facilitate interpretation of the results. Raw scores are hard to interpret because the scales have different numbers of items as well as different means and standard deviations. All BIMAS raw scores are converted to standardized scores (*T*-scores) in order to obtain values that are more easily interpretable and to allow for comparisons across different forms, scales, and raters (see *Norming Procedures and Derivation of Standardized Scores* section in chapter 9, *Standardization*, for a detailed description of how raw scores were converted to *T*-scores).

### *T*-Scores

Standardized scores are based upon a comparison to the normative sample, which is intended to represent a typical population. The BIMAS–Teacher Standard (BIMAS–T Standard), BIMAS–Parent Standard (BIMAS–P Standard), and BIMAS–Self-Report Standard (BIMAS–SR Standard) raw scores are converted to *T*-scores, which have a mean of 50 and a standard deviation of 10. The BIMAS *T*-scores are based on the relationship between an individual’s raw scores and typical scores for that individual’s age group. For example, a raw score of 4 on the Conduct scale (a Behavioral Concern scale) on the BIMAS Parent form for a 5-year-old youth is equal to a *T*-score of 50, which represents the mean of that comparison group (children aged 5 to 6 years). The more the rater reports observing behavioral concerns, the higher the *T*-scores. The reverse is true on the Adaptive scales. Since higher *T*-scores indicate a higher level of strength in social or academic functioning, the more the rater observes a lack of adaptive skills, the lower the *T*-scores.

## Interpreting *T*-scores

### Behavioral Concern Scales

Table 5.2 presents the Scale Descriptors (or Levels of Risk) of the *T*-score ranges on the BIMAS Standard Behavioral Concern Scales (i.e., Conduct, Negative Affect, and Cognitive/Attention). This table also provides interpretive guidelines that can be used to describe *T*-scores in a general manner. Within the three Behavioral Concern scales, higher scores indicate more concerns. A *T*-score of 75, for example, can be described as *High Risk* (all *T*-scores of 70 or higher fall into this category). This high *T*-score suggests a very high number of concerns were reported on items in a specific domain in relation to the normative group (i.e., to other youth within the same age group). *T*-scores within this category warrants a high level of concern. A *T*-score of 65, on the other

hand, can be described as *Some Risk* (all *T*-scores between 60–69). A score within the *Some Risk* category means that the rater gave higher ratings to the items on a scale than what were found in the normative group. Scores within this category warrant a moderate level of concern. Assessors are recommended to review the *T*-scores within the *Some Risk* range (in conjunction with other information and knowledge of the youth) in order to decide whether or not the concerns in the associated area warrant follow-up and/or intervention. Generally speaking, the higher the *T*-score is within this range, the more likely it is that follow-up and/or intervention is necessary. Finally, a *T*-score of 50 can be described as *Low Risk* (all *T*-scores below 60 fall into this category). A *Low Risk* score means that the youth’s behaviors were rated at a level that is typical when compared to the normative group and indicates a very low level of concern.

Scores that fall near the juncture of two categories warrant special consideration. For example, *T*-scores of 59 and 60 fall into different categories (*Low Risk* and *Some Risk*, respectively), even though the difference between them may not be statistically significant. It may be more helpful to describe borderline *T*-scores such as 59 or 60 as falling at the juncture of the *Low Risk* and *Some Risk* categories. Table 5.2 provides the common characteristics of high scorers on each of the Behavioral Concern scales.

**Table 5.2. Understanding *T*-scores and Scale Descriptors: Behavioral Concern Scales**

<i>T</i> -score	Scale Descriptor	Interpretive Guideline
<i>T</i> = 70+	<b>High Risk</b>	Many more concerns than are typically reported.
<i>T</i> = 60–69	<b>Some Risk</b>	More concerns than are typically reported.
<i>T</i> < 60	<b>Low Risk</b>	Fewer concerns than are typically reported.

**Table 5.3. Common Characteristics of Youth with High Scores on the BIMAS Behavioral Concern Scales**

Behavioral Concern Scale	Common Characteristics of High Scorers
<b>Conduct</b>	Problems with physical and/or verbal aggression; may fight or bully. May be defiant and/or manipulative. May have problems controlling temper. May be dishonest. May use alcohol, drugs, or tobacco. May engage in risky behavior.
<b>Negative Affect</b>	Shows symptoms of depression and/or anxiety. Mood problems may include sadness, negativity, anhedonia, shamefulness, and nervousness. May be tearful. Ideation about hurting self may be present. May be fearful or worry a lot. Feelings may be easily hurt.
<b>Cognitive/Attention</b>	Difficulty with control of attention and/or behavior. May have poor concentration or be easily distracted. May lose interest quickly or have difficulty finishing things. May have high activity levels, impulsivity, and difficulty staying seated. May talk too much. May be fidgety or easily excited. May have difficulty organizing and planning.

## Adaptive Scales

Table 5.4 presents the Scale Descriptors (or Levels of Functioning) of the *T*-score ranges on the BIMAS Standard Adaptive scales (i.e., Social and Academic Functioning). This table provides interpretive guidelines that can be used to describe *T*-scores in a general manner. The BIMAS Adaptive Scales are strength-based—they can be used to identify areas of concern as well as strengths in relation to social and academic functioning. As opposed to the Behavioral Concern scales, where higher *T*-scores indicate more concerns, on the two Adaptive Scales, higher *T*-scores indicate a greater number of positive/desirable behaviors. For example, a *T*-score of 65 falls within the *Strength* level of functioning (all *T*-scores above 60 fall into this category), and it means that the rater gave higher ratings on items on the scale compared to the normative group (i.e., to other youth within the same age group). A score that is categorized as a *Strength* may indicate a high level of adaptive skills and/or positive behaviors within the social and academic functioning domains by the youth. On the other hand, a *T*-score of 50 can be described as *Typical* (*T*-scores between 41 and 59 fall into this category). A *Typical* *T*-score on the Adaptive scales means that the youth's adaptive behaviors were rated at a level that is typical when compared to the normative group. In contrast, a *T*-score of 30 would be described as *Concern* (all *T*-scores of 40 or less), and it means that the rater gave lower ratings to the items than those found in the normative group. A *T*-score in the *Concern* range suggests that the frequency of adaptive behaviors observed was quite low, and further assessment/intervention may be required. Assessors are recommended to review the *T*-scores within the *Concern* range (in conjunction with other information and knowledge of the youth) in order to decide whether or not the concerns in the associated area warrant follow-up and/or intervention. In general, the lower the *T*-score is within this range, the more likely it is that follow-up and/or intervention is necessary. Table 5.5 provides the common characteristics of low and high scorers on each of the Adaptive scales.

**Table 5.4. Understanding *T*-scores and Scale Descriptors: Adaptive Scales**

<i>T</i> -score	Scale Descriptor	Interpretive Guideline
<i>T</i> = 60+	<b>Strength</b>	Higher than typical level of adaptive behaviors reported.
<i>T</i> = 41–59	<b>Typical</b>	Typical level of adaptive behaviors reported.
<i>T</i> ≤ 40	<b>Concern</b>	Many more concerns than are typically reported.

**Table 5.5. Common Characteristics of Youth with Low and High Scores on the BIMAS Adaptive Scales**

Adaptive Scale	Common Characteristics of Low Scorers	Common Characteristics of High Scorers
<b>Social</b>	Poor social skills. May have difficulty with friendships, body language, social cues, or emotions. May be socially awkward. May appear disinterested in social interactions. May seem rude or unfriendly. May have difficulty with expressing thoughts or emotions. May have no or very few friends; may be unpopular, unaccepted, or ignored by peers. May be teased by other youth.	Good social skills. May excel at making friends, reading body language and social cues, and/or understanding emotions. Friendly and comfortable in expressing thoughts or emotions. May have many friends.
<b>Academic Functioning</b>	May have difficulty following directions. May receive failing grades at school. May not work up to academic potential and may not be well-prepared for class. May be absent from school on a frequent basis.	Follows directions well. Receives good grades at school. Works up to academic potential and comes to class prepared. Is rarely absent from class.

## Percentiles

Scores on the BIMAS can also be described as percentile ranks. A percentile rank indicates the percentage of youth in the normative sample who earned a raw score that is the same as or lower than that earned by the youth who was rated. For example, if an individual's percentile rank is 65, this means that 65% of the youth in the standardization sample earned a raw score that is the same or lower.

BIMAS percentile ranks, which range from 1 to 99, are provided for every possible raw score on a BIMAS scale. Percentile ranks are most useful for describing the relative standing of an individual in the normative sample distribution. *T*-scores are better for comparing a youth's scores across scales on the BIMAS.

## Confidence Intervals

All measurements contain error (see Naglieri & Chambers, 2009, for a discussion of measurement error). The degree of measurement error found in the BIMAS is provided in chapter 10, *Reliability* (see *Standard Error of Measurement*). Measurement error must be taken into account whenever the scores are presented and compared. Measurement error in BIMAS Standard scores is described in terms of the confidence interval. Confidence intervals take measurement error into account by providing, at a specific level of probability, a range of scores within which the true score (i.e., the score

that would be obtained if there were no errors in measurement) is expected to fall. Both the 90% and 95% confidence intervals were computed for all scales (see *Standard Error of Measurement* section in chapter 10 for more information on the computation of confidence intervals). BIMAS Standard individual (student) level reports include the confidence interval for every *T*-score (available only for reports generated by users given *T*-score Rights); the user can choose whether to use 90% or 95% confidence intervals scoring options on the BIMAS Online. The 90% level of confidence is recommended, and is the default setting for BIMAS reports.

## Item-Level Scores

The BIMAS Standard also offers Item Descriptors for the interpretation of item-level scores. This feature offers assessors versatility in goal-setting and customized intervention design/monitoring (for an illustration of how to interpret and use item-level results to assist in intervention design, please refer to chapter 6, *Step-by-Step Interpretation of BIMAS Results*).

Raters respond to each item on a Likert scale of 0 to 4 (0 = Never or observed 0 times/not observed during the past week; 1 = Rarely or observed 1 – 2 times/to a minimum extent; 2 = Sometimes or observed 3 – 4 times/to a moderate extent; 3 = Often or observed 5 – 6 times/to a moderate extent; 4 = Very Often or observed 7 or more times/to an extreme extent). Each item score (0 to 4) is given an Item Descriptor (*Concern, Mild Concern, No Concern/Fair, or Positive*) to indicate the Level of Risk/Functioning associated with the score.<sup>2</sup> At the item level, the approach used to designate the risk levels of a BIMAS Behavioral Concern scale item score is similar to the technique used by Naglieri, McNeish, and Bardos (1991), Naglieri, LeBuffe, and Pfeiffer (1994), and LeBuffe and Naglieri (2003); they all suggested that an individual item score that falls in the top 15% of the normative group distribution (e.g., exceeds the mean normative item score plus one standard deviation) can be considered problematic. In order to designate Behavioral Concern item scores as *Concern* or *Mild Concern*, the normative sample cumulative frequencies, as well as means and standard deviations for every item score on each BIMAS

item, were examined across various BIMAS Standard rater forms where norms are available (i.e., BIMAS–T, BIMAS–P, and BIMAS–SR). The same logic, though reversed, was applied to the Adaptive Scales—an item score that falls in the top 15% of the normative group distribution (i.e., exceeds the mean normative item score plus one standard deviation) was considered *Positive*. *Concern* and *Mild Concern* categories denote item scores that fall in the lowest 10% or between the lowest 11% to 20%, respectively. See chapter 8, *Development*, for a more detailed discussion of the development of Item Descriptors.<sup>3</sup>

As with item analysis on any psychological measure, a good deal of caution and clinical judgment is warranted in interpretation as each item assesses a very specific behavior. Scale-level scores are generally more stable and reliable than individual item scores; therefore, in general, item-level information should be used to supplement scale scores. For instance, it may be reasonable to provide an individual with interventions for attention problems if he/she had a high score on the Cognitive/Attention scale (i.e., scoring within the *High Risk/Some Risk* ranges); however, it would likely not be appropriate to establish a large-scale intervention plan based on a single item rating without any corresponding scale elevations. Other times, a high endorsement of an individual item (e.g., “expressed thoughts of hurting self”) can be critical enough to warrant immediate attention.

In addition to examining specific item content, assessors should look for a convergence of data supporting an item-level analysis. Items with elevated scores across multiple raters as indicated by Item Descriptors of *Concern* or *Mild Concern* should be considered in association with supporting evidence from the student’s background/history. In particular, flagged items within behavioral domains considered by the assessor as having the most interference with daily functioning and/or success in the academic setting may also inform intervention planning.

## Interpreting Item-Level Scores

### Behavioral Concern Scales

Table 5.6 illustrates the Item Descriptors or interpretive guidelines associated with item scores on any given item on the Behavioral Concern scales (i.e., Conduct, Negative Affect, or Cognitive/Attention). Whether a response to an item is a *Concern, Mild Concern, or No Concern* depends on the item content, type of rater (Teacher/Parent/Self), and the youth’s normative age group (e.g., ratings for a 5-year-old child vs. an 18-year-old youth). For example, an item score

<sup>2</sup> The same item response schemas (0 to 4) and Item Descriptors (*Concern, Mild Concern, or No Concern* for negatively worded items and *Concern, Mild Concern, Fair, or Positive* for positively worded items) are used on the BIMAS Flex. However, since the BIMAS Flex is not norm-referenced, assessors must set up their own item-level scoring criteria based on what item response is considered to be “*Concern*” or “*Typical*” for an individual youth. Interpretation of Item Descriptors on the BIMAS Flex is similar to interpretation of the item-level scores on the BIMAS Standard. A Flex item scored in the *Concern* or *Mild Concern* category may suggest areas of need, while an item scored in the *Positive* range could be used to complement any intervention that incorporates a student’s adaptive behaviors.

<sup>3</sup> BIMAS Standard Item Descriptors for all items across age groups (and gender, if gender-specific norms are preferred) are available upon request for research purposes. See chapter 12, *Concluding Comments*, for the publisher’s contact information.

of (1 = Rarely) can trigger a *Concern* for the item “expressed thoughts of hurting him/herself,” whereas the same item score of (1 = Rarely) may produce an Item Descriptor of *No Concern* for the item “acted without thinking.”

In general, if an item from any of the Behavioral Concern scales produced a *Concern* item score, it indicates that the behavior in question was rated as occurring much more frequently than observed amongst most youth in a comparable age group by the same rater. In such cases, further investigation and/or intervention in association with the scale-level scores may be necessary. However, a *Mild Concern* item score indicates that the youth’s behavior was rated as occurring slightly more frequently than the normative group, and a mild level of concern is warranted. On the other hand, a *No Concern* item score means that the youth is displaying the behavior at a frequency comparable to the normative group.

**Table 5.6. Understanding Item Descriptors: Behavioral Concern Scales**

Item Descriptor	Interpretive Guideline
<b>Concern</b>	Behavior was rated as occurring much more frequently than is typical.
<b>Mild Concern</b>	Behavior was rated as occurring more frequently than is typical.
<b>No Concern</b>	Behavior was rated as occurring at a typical level.

### Adaptive Scales

Since higher scores indicate fewer concerns on BIMAS Adaptive Scales, there is a fourth Item Descriptor—*Positive*—added to reflect adaptive behaviors which are considered positive or beyond the expected level of functioning. Table 5.7 illustrates the Item Descriptors or interpretive guidelines associated with item scores on the Adaptive scales (i.e., Social or Academic Functioning). Similarly to the Behavioral Concern scales, whether a response to an item is considered a *Concern*, *Mild Concern*, *Fair*, or *Positive* depends on the item content, type of rater (Teacher/Parent/Self), and the normative age group.

A *Positive* descriptor for item scores on any of the Adaptive scales may indicate that the youth is displaying the adaptive behavior more frequently than is typically observed amongst youth in a comparable age group, or that the youth is performing an adaptive skill at a level that is beyond age-level expectations. A *Fair* item score means that the youth is displaying the behavior at a frequency comparable to the normative group. A *Mild Concern* item score indicates that the youth’s adaptive behavior was rated as occurring slightly less frequently than in the normative group, and a mild level of concern is warranted. On the other hand, a *Concern* item score indicates that the behavior was rated as occurring much less frequently than observed amongst

most youth in a comparable age group by the same rater type. Further investigation and/or intervention in association with the scale-level scores may be required.

**Table 5.7. Understanding Item Descriptors: Adaptive Scales**

Item Descriptor	Interpretive Guideline
<b>Positive</b>	Behavior was rated as occurring more frequently than is typical or performed at a level that is beyond expectation.
<b>Fair</b>	Behavior was rated as occurring at a typical level.
<b>Mild Concern</b>	Behavior was rated as occurring less frequently than is typical.
<b>Concern</b>	Behavior was rated as occurring much less frequently than is typical.

## Scores for Progress and Outcome Monitoring

The focus of progress monitoring is on change in an individual as she or he participates in an intervention. Outcome monitoring, in contrast, focuses on at least a tentative final estimate of the improvement (or lack thereof) experienced by the individual after completion of the intervention. From an educational perspective, progress monitoring corresponds with formative assessment, while outcome monitoring is similar to summative assessment.

Repeated measurement with the BIMAS forms provides invaluable progress and outcome monitoring data. The literature offers several methods for estimating the type and extent of change over time, but currently no consensus about the best method exists. Instead, an experienced assessor who can combine the BIMAS change estimates with other clinical information is required in order to determine the effectiveness of the intervention. Three types of change information are provided with repeated BIMAS measurement: visual displays, the reliable change index (RCI), and effect size (ES) estimates.

### Visual Displays

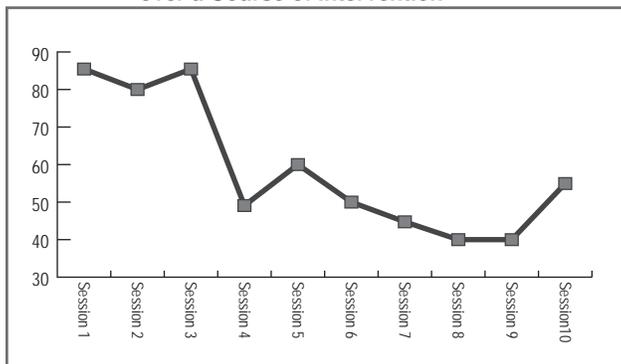
Visual displays are graphics and tables employed to illustrate, reason about, and make decisions about relationships among important quantitative and qualitative information (Tuft, 1983). The ideal graphic induces the viewer to think about the substance of the results (rather than the methodology employed) and has the potential to reveal multiple dimensions of the data. Equally important, visual displays can also aid in the retention of important information.

One of the most popular visual displays for use in estimating change is a time series graph, an illustration of a single variable recorded over a period of time (Tuft, 1983). Best employed with larger sets of data that show (or have the potential to show) substantial variability, time series graphics can be applied to BIMAS data for the purposes of examining progress and outcome with the same individual over time. Behavioral assessors, for example, employ time series analyses to gauge the effects of interventions, beginning with a baseline period and then searching for patterns of change in one or more variables once an intervention is introduced.

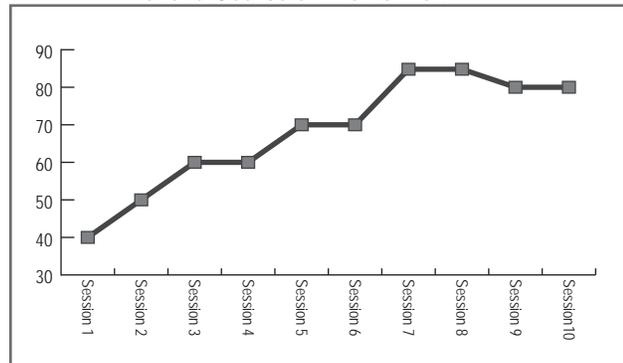
One may ask, what kind of data patterns are likely to be apparent? The counseling and psychotherapy outcome literature typically focuses on improvement vs. worsening. For example, as illustrated in Figure 5.1, many individuals receiving psychosocial interventions will evidence an improvement on constructs such as conduct, attention, or negative affect. Conversely, some individuals, particularly those not receiving treatment, may worsen over the course of time, including over the course of the academic year (see Figure 5.2). In addition, some adolescents may initially under-report problems and upon retest, report more accurate information after they have formed a relationship with a therapist (i.e., establishment of a rapport). In such cases it may appear that the individual has worsened over time, when in fact, the second administration represents more valid data and should be employed as the proper starting point (baseline) against which to evaluate subsequent data.

In actual practice, it is also common to see two other data patterns: Cycles of variability (Figure 5.3) and stability (Figure 5.4). Potentially tied to environmental or dispositional factors, regular variability or cycles are apparent for many clients. Others demonstrate only small, seemingly random changes on a certain variable over time; this stability becomes problematic when the behavior, thought process, or affect in question is dysfunctional.

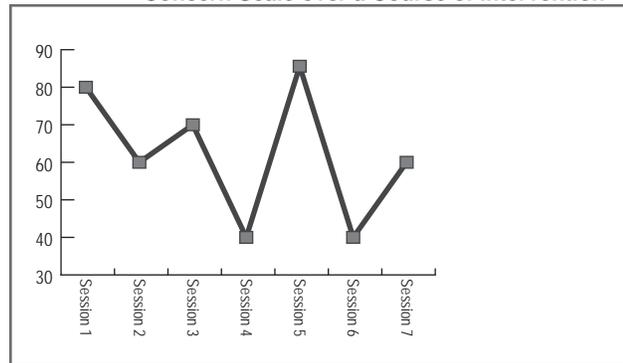
**Figure 5.1. Hypothetical Improvement on a BIMAS Behavioral Concern Scale over a Course of Intervention**



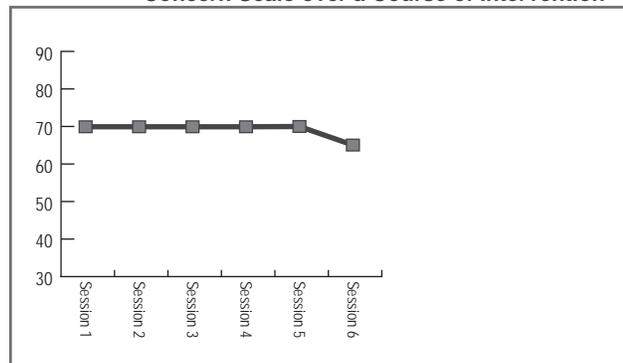
**Figure 5.2. Hypothetical Worsening on a BIMAS Behavioral Concern Scale over a Course of Intervention**



**Figure 5.3. Hypothetical Cycles on a BIMAS Behavioral Concern Scale over a Course of Intervention**



**Figure 5.4. Hypothetical Stability on a BIMAS Behavioral Concern Scale over a Course of Intervention**



## Reliable Change Index (RCI)

A method for gauging statistically significant change in *T*-scores was outlined by Jacobson and Truax (1991). The Jacobson and Truax (1991) method involves calculating a Reliable Change Index (RCI) that is used to determine whether a change in scores between test administrations is statistically significant. The RCI takes into account the difference in test scores between two administrations as well as the standard error of difference between them (the standard error of difference is computed with standard

error of prediction estimates; see the *Test-Retest Reliability and Standard Error of Prediction* section in chapter 10, *Reliability*, for more information).

The RCI values needed to establish statistical significance, when comparing Time 1 to Time 2 scores, were calculated for all scales on the BIMAS Standard, at the  $p < .10$  level of significance. The  $p < .10$  criterion was used in order to ensure that important changes (both increases and decreases in scores) are not missed. If the absolute difference between two administrations is equal to or greater than the corresponding RCI value, then the amount of change in that area of functioning is statistically significant. This means that the measured change is a function of an actual difference between test scores and not the result of random fluctuations in behavior or error in measurement. This calculation method is integrated into the BIMAS *Standard Individual Progress Report: Significant Change Over Time report* (Report ID: 15 in Table 5.1; see report sample in appendix E.15a). Appendix C provides the *T*-score values needed for statistically significant difference between pairs of administrations by the same rater.

### Interpreting Change with the Reliable Change Index

A statistically significant result means that the measured change can be attributed to reliable differences between the scores rather than random fluctuations in behavior, or error in measurement. If scores have decreased significantly from Time 1 to Time 2 on a Behavioral Concern scale, then one can conclude that the youth's scores show significant improvement. If the scores have significantly increased, then the scores show significant worsening. The reverse is true for the Adaptive scales, since higher scores indicate *fewer* concern. If a change in scores is not statistically significant, then the amount of change should be attributed to random fluctuation or measurement error. If, however, the change in scores is statistically significant, then judging the *clinical importance* of the change is the next step.

### Effect Size Estimates

Effect Size (ES) is a measure of the strength of change, based on an individual's raw scores. Typically employed to determine if groups (e.g., intervention versus control) show differences over time, ES has been adapted for examining whether individuals exhibit change over time (see Clement,

1999). To examine change in an individual's scores over time, we can compute ES for each BIMAS scale or total score using the following formula<sup>4</sup>:

$$ES = (\text{Time 2 Mean} - \text{Time 1 Mean}) / \text{Standard Deviation at Time 1}$$

where *Time 1 Mean* denotes the mean of all item scores for an individual on a BIMAS scale at time 1, which is typically the baseline BIMAS assessment prior to intervention, and *Time 2 Mean* denotes the mean of all item scores for the same individual on the same scale at time 2 (a follow-up assessment during a course of intervention or a follow-up assessment at the termination of intervention). For example, Student A's scores on the five BIMAS scales at Time 1 and 2 are as follows:

	Time 1	Time 2
Conduct	75	62
Negative Affect	60	45
Cognitive/Attention	65	60
Social	55	50
Academic Functioning	40	52

When ES is calculated per scale, per individual using Clement's (1999) formula, Student A's self-reported item scores on the Conduct scale for Time 1 and Time 2 are as follows:

	Time 1	Time 2
Angry	4	2
Risky behaviors	4	2
Fought (verbally/physically)	4	2
Lied/cheated	0	1
Lost temper	4	1
Aggressive	4	0
Alcohol and/or drug use	2	1
Disciplinary referral	3	1
Tobacco use	4	1
<i>Ms:</i>	3.22	1.22
<i>SD at Time 1:</i>	1.39	

So,  $ES = (\text{Time 2 } M - \text{Time 1 } M) / SD \text{ at Time 1} = (1.22 - 3.22) / 1.39 = -1.44$ , or "*Improved*" as indicated by Clement's (1999) interpretive guideline (see Table 5.8).

<sup>4</sup> Minor modifications were made to Clement's (1999) formula in calculating ES: (a) Instead of subtracting Post-test (Time 2) Mean from the Pre-test (Time 1) Mean as suggested by Clement (1999), the BIMAS *Standard Individual Progress Report: Significant Change Over Time* (Report ID: 15) was programmed to subtract the Time 1 Mean from the Time 2 Mean so that a negative ES value suggests improvements on Behavioral Concern Scales (a decrease in scores) and a positive ES value suggests improvement on Adaptive Scales (an increase in scores). This is to allow for a more intuitive interpretation of change; (b) In cases where no variability is observed at Time 1 (e.g., all item scores = 1), *SD at Time 1* = 0, and ES cannot be calculated with 0 as the denominator, the formula for calculating ES is modified to use *SD at both Time 1 and Time 2* as the denominator (i.e., *SD* is calculated on all item scores across Time 1 and Time 2); and (c) in cases where no variability is observed at both Time 1 and Time 2 (e.g., all item scores = 4 on the scale at both administrations), a ES of 0 (*No Change*) is indicated in the report.

For progress monitoring, ES can be calculated for any two administrations of the BIMAS Standard (e.g., after Sessions 1 and 5). ES can function as an estimate of the amount of change demonstrated by an individual over time.

### Interpreting Change with Effect Size Estimates

Thus far, no tests of statistical significance exist to suggest that a certain level of ES indicates that a definitive change has occurred, especially when it comes to change within an individual as opposed to interpreting change at the group level. Instead, various authors have proposed criteria to gauge the amount of change, such as the one created by Clement (1999) as presented in Table 5.8.

**Table 5.8. Effect Size Interpretations for Individual Clients on the BIMAS Standard**

Effect Size	Interpretation for Behavioral Concern Scales	Interpretation for Adaptive Scales
≤ -1.50	Much Improved	Much Worse
-.51 to -1.49	Improved	Worse
-.50 to +.50	No Change	No Change
.51 to 1.49	Worse	Improved
≥ 1.50	Much Worse	Much Improved

The BIMAS *Standard Individual Progress Report: Significant Change Over Time* (Report ID: 15 in Table 5.1; see appendix E.15b for a sample report) provides the option to display ES. This report also provides the same interpretive guidelines as presented in Table 5.8 to assist in the interpretation of obtained ES values. In general, higher negative ES values indicate that improvements have been made on Behavioral Concern scales where a decrease in Conduct, Negative Affect, or Cognitive/Attention types of concerns is desirable. Conversely, higher positive ES values indicate improvements have been made on Adaptive scales where an increase in Social or Academic Functioning types of behaviors is desirable.

### Comparison of the Reliable Change Index and Effect Size Estimates

RCI and ES are two different change statistics. RCI allows the assessor to assess the statistical significance of change, whereas ES does not. On the other hand, ES allows for the evaluation of the magnitude of change, which one cannot obtain from the RCI. Both change statistics are available in the BIMAS *Standard Individual Progress Report: Significant Change Over Time* report (Report ID: 15). Assessors can choose between the two options to suit their progress monitoring needs.

## Scores for Comparing Results Between Raters & Groups

The BIMAS provides comparative types of scores at both the individual level and multi-student level. At the individual student level, it is worthwhile to compare how different raters (e.g., teachers vs. parents) rate the same individual because it may provide insight as to how the youth behaves in different settings (e.g., at school vs. at home). Moreover, by comparing the youth's self-report to an observer, one can gain invaluable information about the youth's perception of his/her own behaviors.

At the multi-student level, comparisons can be made between different groups (e.g., Class A to Class B at the Grade level, Grade X to Grade Y at the School level, or School A to School B at the District level). BIMAS scale scores can be averaged across individuals within different student groups and comparisons between group averages can be made to provide assessors with a general picture of how a district, school, or grade is functioning. Such group comparisons are useful in an evidence-based practice whereby success in group-based programs can be shown, or changes to existing services or programs can be considered.

### Statistically Significant Differences Between Raters

The BIMAS *Standard Individual Comparison Between Raters* (Report ID: 8 in Table 5.1) compares *T*-scores obtained from multiple raters of one youth. Assessors can choose up to five raters to be included in a report, but the *T*-score comparisons are done in pairs (e.g., Teacher 1 to Parent 1; Parent 1 to Parent 2; Self to Teacher 1; Self to Teacher 2; etc.). Because a comparison of scores obtained from different raters must take measurement error into consideration, the Comparative report indicates whether or not differences in scale scores across raters are *statistically* significant. The values needed to establish statistical significance were calculated using the formula provided by Anastasi and Urbina (1997). This formula takes into account the standard error of measurement (*SEM*) for each of the scales (see chapter 10, *Reliability*), and  $p < .10$  is used as the level of significance (see appendix D for the values needed to determine significance between a pair of ratings).

When *T*-scores obtained by different raters are indicated on the BIMAS comparative report as being similar (i.e., difference between *T*-scores did not reach statistical significance), they should be interpreted as similar, despite their possible differing scale descriptors (e.g., *High Risk* from Rater 1 vs. *Some Risk* from Rater 2). In contrast, when a rater's *T*-score

is indicated to be significantly higher than another rater's  $T$ -score, denoted by Rater X > Rater Y, it means that the difference in  $T$ -scores between that pair of raters reached statistical significance, and should be interpreted as different from one another, even if the same Scale Descriptors were obtained (e.g., *Some Risk* from both Rater 1 and Rater 2).

For example, ratings by Parent 1, Parent 2, and Student A (Self) produced the following  $T$ -scores on the Negative Attention scale, respectively: 64, 76, and 69. Based on *SEM* established from the normative sample, the amount of difference in  $T$ -scores required between two parents to be considered statistically significant is 11 for Student A's age group (12–13 years) on the Negative Affect scale, while the amount of difference required between a parent and self-report is 10. In this case, the amount of difference in  $T$ -scores between Parent 1 and Parent 2 is 12 ( $|64 - 76| = 12$ ). Therefore, Parent 1's ratings on the Negative Affect scale for Student A is considered significantly different from that of Parent 2's. However, both Parent 1 and Student A ( $|64 - 69| = 5$ ) and Parent 2 and Student A ( $|76 - 69| = 7$ ) provided ratings that were not statistically different from one another.

## Comparing Between Group Averages

At the District level, an assessor may wish to for example, obtain an overall average  $T$ -score on the Conduct Scale across teacher ratings at a particular screening or Universal Assessment (UA) to see which school is performing below (or above) other schools within the district. The *BIMAS Average Score Comparison Report* (Report ID: 7 in Table 5.1) aggregates each individual student's Conduct scale score in a particular UA in order to compute an average  $T$ -score for each of the schools selected by the user within the district. A visual display by way of a bar graph showcases how each individual school's average Conduct  $T$ -score compares to one another and an overall average Conduct  $T$ -score across all selected schools within the district is also provided on the Comparative report. (Note that the district's overall average  $T$ -score is the mean score of all students in the selected schools rather than the arithmetic mean of each school's average  $T$ -score.)

# 6 Step-by-Step Interpretation of BIMAS Results

The Behavior Intervention Monitoring Assessment System (BIMAS™) was developed for individual student screening and progress monitoring, but it is also designed for group screening and the evaluation of group-based intervention programs. In order to systematically describe the interpretation of BIMAS results, this chapter begins by focusing on step-by-step interpretations of the BIMAS findings at the group level, as schools or districts would generally analyze results from a Tier 1 universal screening in order to identify students who might require Tier 2 or 3 levels of services in behavioral progress monitoring. The chapter will conclude with a step-by-step discussion of the interpretation of BIMAS results for individual evaluation. At the individual level, concepts relating to the analysis of scale scores and item-level responses, comparing results between raters, and how to monitor individual response to intervention results will be discussed. Since the interpretive steps at both the group and individual levels involve reviewing different BIMAS reports at every step, the BIMAS reports are organized into a flowchart in Figure 6.1 to follow the step-by-step interpretation sequence. In chapter 7, *Case Study*, the same step-by-step interpretive sequence is also followed.

## Step-by-Step Interpretation Guidelines: Group Evaluation

When assessing multiple students in a district, school, grade, class, or group, there are a number of reports that assist assessors with screening and progress monitoring. The step-by-step interpretive method outlined below is useful for group-level (Tier 1) universal screening, which might involve district- or school-wide Universal Assessments (UAs).

### Step 1. Identify General Areas of Concern

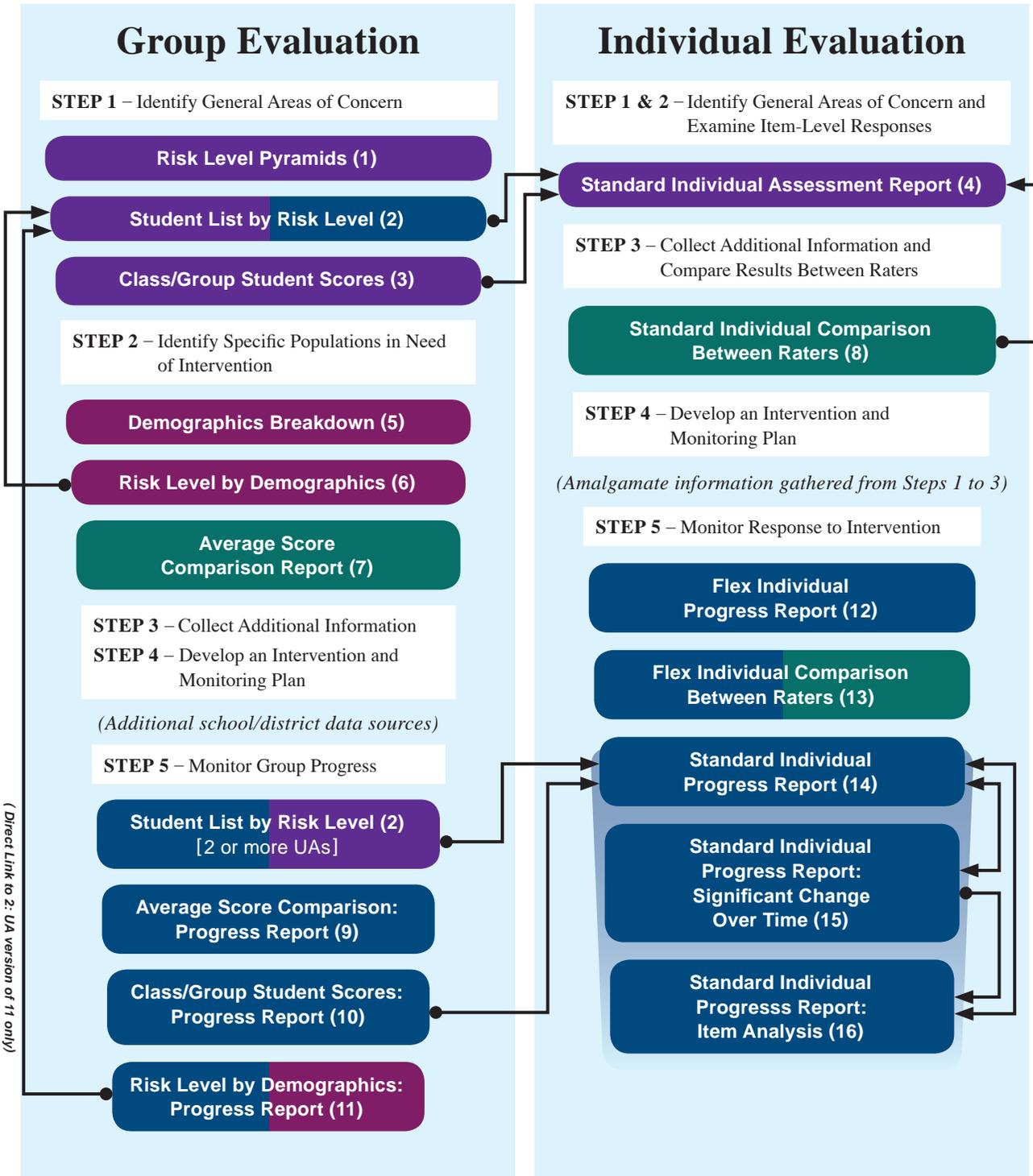
At the **district** level, UA data entered in the BIMAS Online provide levels of risk across the BIMAS scales for the district. Using the *Risk Level Pyramids* report (Report ID: 1; see appendix E.1 for a sample of this report generated at the district level) district-level assessors can start by identifying

the BIMAS domains with the highest percentage of students in the *Some Risk/High Risk/Concern* categories for the district and for each individual school. **School** level assessors can also use the same report (generated at the school level) to identify the school's general areas of concern as well as the functioning of individual grades.

At the **grade** level, besides using the above report to identify general areas of concern within a particular grade as well as individual classes belonging to the grade, assessors can also use the *Student List by Risk Level* report (Report ID: 2) to obtain a list of students in each grade who have been identified by their BIMAS scale scores to be at-risk for a certain behavioral domain (appendix E.2a shows the Assessment report version; a Progress report version is shown in appendix E.2b; see *Step 5. Monitor Group Progress* later in this section for discussion on the use of the Progress version). Each student name within these reports is a direct link to the *Standard Individual Assessment Report* (Report ID: 3; see appendix E.3 for a sample student report), which facilitates the examination of individual functioning of students in the *High Risk/Some Risk* groups on any of the Behavioral Concern scales (or the *Concern* category on the Adaptive scales).

At the **class/group** level, assessors can use the *Class/Group Student Scores* report (Report ID: 3; see appendix E.3) to review students' scores within a classroom or an intervention group across all five BIMAS scales. Students scoring in the *High Risk* category across the three Behavioral Concern scales or those scoring in the *Concern* category across the two Adaptive scales may need to be referred for follow-up assessment. Other pertinent information about individual students might be required to decide whether or not some or all of the students scoring in the *Some Risk* category require follow-up assessment and/or intervention. On the other hand, students who are in the *Strength* category on the Adaptive scales can perhaps be paired with individuals who are not performing as well behaviorally to provide assistance or to model good classroom behaviors.

Figure 6.1. Application of BIMAS Reports Within the Step-by-Step Interpretation Method



(Direct Link to 2: UA version of 11 only)

**LEGEND**

<span style="display: inline-block; width: 15px; height: 15px; background-color: purple; border: 1px solid black; margin-right: 5px;"></span> Assessment Report	<span style="display: inline-block; width: 15px; height: 15px; background-color: teal; border: 1px solid black; margin-right: 5px;"></span> Comparative Report
<span style="display: inline-block; width: 15px; height: 15px; background-color: blue; border: 1px solid black; margin-right: 5px;"></span> Progress Report	<span style="display: inline-block; width: 15px; height: 15px; background-color: purple; border: 1px solid black; margin-right: 5px;"></span> Demographic Report

→ Direct web-link *from* Reports A to B (one-way)  
 ↔ Direct web-link *between* Reports A & B (two-way)

## Step 2. Identify Specific Populations in Need of Intervention

After general areas of concern have been identified in Step 1, the task turns to identifying populations of students most in need of intervention. To begin this task, assessors can review Demographic reports for the identification of specific risk factors. Following that, Comparative Report at the multi-student level can be used to identify populations in need across various levels within the district. The *Demographics Breakdown* (Report ID: 5; sample report shown with Service Code as the selected demographic variable in appendix E.5) might reveal a high percentage of students within certain demographic categories, indicating higher needs for services within those demographic categories. The *Risk Level by Demographics* (Report ID: 6; see appendix E.6 with Service Code as the demographic variable selected) can be used to examine behavioral functioning across demographic categories. District-level assessors should be cognizant of demographic categories that are disproportionately identified with behavioral or adaptive concerns because this type of information may suggest populations in need of intervention. At the same time, it could also indicate a need to sensitize raters to cultural and socioeconomic issues related to behavior.

When interpreting the *Average Score Comparison Report* (Report ID: 7; see appendix E.7 for a sample report generated at the district level), on a Behavioral Concern scale, schools with higher average scores than other schools in the district, or ones that exceed the district's overall average score, would indicate intervention needs in that domain. On the other hand, schools with much lower average scores compared to other schools on Adaptive scales would indicate a need for positive behavioral programs in the Social or Academic Functioning domains. Moreover, it is equally important to identify schools that are performing well within the district. For example, these well-performing schools can shed light on specific programs that might have contributed to their success.

At the **school**, **grade**, and **class** level, the same reports can be reviewed to identify specific populations in need of intervention. Targeted programs or additional services may be required for specific populations identified in this step. Alternatively, preventive services or programs can be designed at the Tier 1 (Universal Level) as a proactive initiative.

By using the reports identified in Steps 1 and 2 in the flowchart presented in Figure 6.1, risk levels across all schools in a district can be compared to each other in order to funnel intervention resources to the neediest schools. At the school level, BIMAS reports can be used to target the

specific grade levels, classrooms, and groups of students that require the most intervention. This screening information may be used as baseline data for program evaluation and to inform the design of group-level interventions. Existing programs catered to Tier 1 (Universal Level) can also be revised based on data to improve school- or district-wide behavioral functioning.

## Step 3. Collect Additional Information

This step requires that the information collected by the administrations of the BIMAS Standard in district-, school-, grade-, class-, or group-wide Universal/Group Assessments be validated and contextualized with other sources of data. These additional data sources might include State test performance, standardized academic or curriculum-based testing, attendance, office referrals, and demographic data. Collection of these additional data allow the assessor to substantiate areas of need illustrated by BIMAS ratings and to explore the scope of the presenting problem or need. All of this is used in the following step to develop an intervention and monitoring plan.

## Step 4. Develop an Intervention and Monitoring Plan

With the information collected in the steps prior, the assessor would have identified specific areas of concern/need, specific schools and/or groups of students most in need, and the scope of the presenting problem. At this step, research-based interventions are reviewed and selected for implementation. While a complete discussion of intervention and monitoring plan development is beyond the scope of this discussion, in general these plans have the following components: a) clearly identified intervention targets; b) baseline data to which subsequent data may be compared; c) an analysis of the scope of the problem and conditions that influence the intervention target(s); d) specific intervention components selected based on that analysis; e) clear goals; f) clearly identified and valid techniques for monitoring progress; and g) a timetable for evaluation. More information on effective intervention and progress monitoring techniques can be accessed on websites such as the US Department of Education's What Works Clearinghouse, and the National Center on Student Progress Monitoring.

## Step 5. Monitor Group Progress

In addition to progress monitoring with individuals, the BIMAS Standard is also designed to be used to assess change over time in groups of students. Schools can use Progress reports at the multi-student level to examine Universal Assessment (UA) results across a school year. In the RTI process, typically three benchmarks, or Universal Assessments, are conducted at each grade level to identify at-risk students and to assess the progress of all students over the school year. Similarly, Group Assessments (GAs) can be conducted on a group of students in an intervention group several times over a term or semester (or over the course of the intervention program). Over several UAs or GAs, the number and/or percentage of students within each level of concern (as identified with *T*-scores) can be examined with the following group-level Progress reports:

- **Student List by Risk Level** (Report ID: 2; see appendix E.2b for the Progress report version of this report with UAs 1 and 2 selected). When many of the students who have been identified in the *Some Risk/High Risk/Concern* groups have follow-up scores in the *Low Risk/Typical* range of risk/functioning, the intervention can be deemed to be effective. When improvements are noted in scores but many still remain in the *Some Risk/High Risk* ranges (i.e., change in scores are not large enough to produce a change in risk categories), the intervention program appears to be effective, but the need for more intervention is indicated.
- **Average Score Comparison: Progress Report** (Report ID: 9 see appendix E.9, generated at the district level). When group average scores for individual schools, grades, or classes decrease on Behavioral Concern scales and increase on Adaptive scales across UAs or school years, the assessor can conclude that improvements have been made. For example, if a school in the district shows progressively lower average scores on a Behavioral Concern scale at every UA (or on an Adaptive scale, progressively higher average scores over UAs), the assessor can conclude that the intervention program introduced at this school is making a positive impact on the behavioral functioning of students.

- **Class/Group Student Scores: Progress Report** (Report ID: 10; see appendix E.10). When the scale scores of students progress from *High Risk/Some Risk* to *Low Risk* on Behavioral Concern scales at subsequent UAs, and/or from *Concern* to *Typical*, or even *Strength*, on Adaptive Scales, the assessor can conclude that improvements have been made in the classroom. Similarly, the effectiveness of two or more interventions can be compared over time across intervention groups by generating two intervention group Progress reports.
- **Risk Level by Demographics: Progress Report** (Report ID: 11; a grade-level report shown in appendix E.11). A decrease in the percentage and number of students with previously identified areas of concern at subsequent UAs would denote a positive change. Assessing the progress of students in specific demographic categories allows educators to ensure that existing programming or any intervention programs introduced as a result of UA 1 baseline data are effective for the majority of students involved in them.

In sum, BIMAS multi-student level Assessment, Progress, Comparative, and Demographic reports provide assessors with different ways to examine group-based assessment data. The main goal at each interpretive step for the assessor working within the RTI framework is to find ways to better allocate resources where needed. The use of different BIMAS reports can assist in such endeavors.

## Step-by-Step Interpretation Guidelines: Individual Evaluation

Assessment reports at the multi-student level help identify students who may be at-risk in developing behavioral problems. Once students requiring intervention and progress monitoring have been identified, the assessor may progress into the evaluation of specific individuals. The following is a suggested strategy for interpreting BIMAS results from an individual evaluation. This systematic, five-step strategy can be used when interpreting scores obtained from one rater or multiple raters providing observations about a single youth.

Before beginning the step-by-step interpretation, the assessor should always consider whether the respondent provided usable data, or if he/she may have provided misleading responses. The BIMAS results will only be accurate if the information provided by the respondent is a reasonable reflection of the rated youth.

## Step 1. Interpret the BIMAS Standard Scale Scores to Identify General Areas of Concern

Interpretation Step 1 consists of identifying general areas of behavioral concern and adaptive functioning using the BIMAS *Standard Individual Assessment Report* (Report ID: 4; see sample report in appendix E.4). This process begins with the examination of *T*-scores for the Behavioral Concern scales and the Adaptive scales and noting the corresponding Scale Descriptors for risk/strength level classifications in the scale-level section of the report.

### Step 1a: Interpret the Behavioral Concern Scale Scores

The BIMAS Behavioral Concern scale results are reported as *T*-scores or Levels of Risk (see Table 5.2 for *T*-score cut-offs and the corresponding Level of Risk/Scale Descriptors). Table 5.3 summarizes the common characteristics of high scorers for each Behavioral Concern scale (see appendix A for items included on the scales). Note that users who were set up on BIMAS Online without *T*-score rights cannot access *T*-scores, percentiles, or confidence interval information on any BIMAS reports, but rather, Levels of Risk information (i.e., Scale Descriptors) only. The *T*-score graphs of the *Standard Individual Assessment Report* provide a visual display of score elevations within the various risk categories. *T*-scores falling within the *High Risk* range indicate potential problems in specific areas (Conduct, Negative Affect, and/or Cognitive/Attention), and can suggest key targets for intervention. *T*-scores within the *Some Risk* range should be considered in conjunction with other information and knowledge of the youth for decisions about whether or not the concerns in the associated area warrant follow-up and/or intervention. Generally speaking, the higher the *T*-score is within this range, the more likely it is that follow-up and/or intervention is necessary.

### Step 1b: Interpret the Adaptive Scale Scores

The BIMAS Adaptive scale results are reported as *T*-scores or Levels of Functioning in both graphical and table display on the individual assessment report (see Table 5.4 for *T*-score cut-offs and the corresponding Level of Functioning/Scale Descriptors). Table 5.5 summarizes the common characteristics of high vs. low scorers for each Adaptive scale (see appendix A for items included on the scales). Again, users who are not authorized to access *T*-scores, percentiles, and confidence intervals can only obtain Levels of Functioning (i.e., Scale Descriptors) on BIMAS reports. *T*-scores falling within the *Concern* range indicate potential problems in specific areas (Social and/or Academic Functioning), and

can suggest key targets for intervention. In general, the lower the *T*-score is within this range, the more likely it is that follow-up and/or intervention is necessary. On the other hand, scores on the Adaptive scales falling within the *Typical* and *Strength* levels of functioning can be used as part of a strength-based assessment to identify a youth's personal strengths and incorporate them into an individualized intervention to overcome areas of concern.

## Step 2. Examine Item-Level Responses for an In-depth Analysis of Concern

Each scale can be better understood by determining which items from each BIMAS Standard scale contributed to a low or high scale score. One way to accomplish this goal is to determine if any item scores were substantially higher or lower than those of the normative group.

### Step 2a: Behavioral Concern Scales

See *Interpreting Item-Level Scores: Behavioral Concern Scales* and Table 5.6 in chapter 5, *Understanding and Interpreting BIMAS Scores*, for a description of Behavioral Concern item-level scores and interpretive guidelines. At the scale level, a Behavioral Concern scale *T*-score can be elevated when a few items are rated as *Concern*, and/or when many items are rated as *Mild Concern*. The BIMAS *Standard Individual Assessment Report* provides item-level scores as well as the norm-referenced interpretations of the item scores in the form of Item Descriptors. Item analysis begins with reviewing the Behavioral Concern scales for items flagged as a *Concern* or *Mild Concern*. These flagged items should be reviewed for specific content and to see if patterns can be ascertained. Reviewing the norm-based levels of risk by item can be helpful in identifying particular areas of concern and in designing and selecting interventions. Items with elevated scores help pinpoint which behaviors are in immediate need of intervention. It is also important to review the Item Descriptors for the potentially critical item “expressed thoughts of hurting self” (item number 24) across the Teacher, Parent, Self-Report, and Clinician forms, if available. Immediate follow-up may be required should the descriptor for this item be flagged as a *Concern* (see *Step 3. Collect Additional Information and Compare Results Between Raters* in the next section for a discussion on rater results comparisons).

### Step 2b: Adaptive Scales

See *Interpreting Item-Level Scores: Adaptive Scales* and Table 5.7 in chapter 5, *Understanding and Interpreting BIMAS Scores* for a description of Adaptive Scale item-level scores and interpretative guidelines. When reviewing the

Adaptive Scales (i.e., Social and Academic Functioning), it is useful to review not only the items flagged as *Concern* and *Mild Concern* but also those item ratings indicating *Fair* and *Positive* levels of functioning. Items on the Adaptive scales flagged as *Concern* or *Mild Concern* highlight specific areas in need of further assessment and ultimate intervention, while items rated at *Fair* to *Positive* levels of functioning might suggest personal strengths that could be capitalized upon during intervention design.

### Step 3. Collect Additional Information and Compare Results Between Raters

While the reports from each BIMAS administration are informative, they should never be used in isolation for the purpose of making diagnoses, intervention design/selection, or making high-stakes placement decisions. Additional information collected to inform the intervention/treatment efforts for students identified by the BIMAS report as having significant levels of concern might include a review of the specific student's school records, interviews with parents and teachers, classroom observations, and additional testing including BIMAS administrations to additional raters.

Comparisons of BIMAS scores obtained from different combinations of raters can be useful in obtaining a multi-rater perspective of the youth's problems or adaptive functioning. Moreover, comparing results from multiple raters can also help determine if there is consistency in behaviors across environments. If agreements between raters are observed on most Standard form items (i.e., item scores between raters do not show large discrepancies), it means that the youth's behaviors are consistent across different settings. However, when large discrepancies are observed between raters' item scores, the assessor may need to identify reasons why the ratings might differ by gathering additional information (e.g., interviewing different raters). For example, a youth may be viewed as being able to follow directions in the home by a parent, but not so much in the classroom by the teacher. Identifying discrepancies between raters' ratings can also help inform an effective intervention plan.

#### Step 3a: Scale-level Comparisons

A comparison of scale scores obtained from different raters must take measurement error into consideration. This is done by examining the statistical differences between raters' scores. The BIMAS *Standard Individual Comparison Between Raters* (Report ID: 8; see sample report in appendix E.8) provides the results of such comparisons between up to five raters automatically. For example, a *T*-score of 68 on the BIMAS Conduct scale on the Teacher report is higher than the score of 60 for the same scale on the Parent report. When

the measurement error associated with these scales is taken into account, however, the difference may not be statistically significant. A statistically significant difference between the scores is one that exceeds the measurement error associated with the scores (see *Statistically Significant Differences Between Rater* section in chapter 5, *Understanding and Interpreting BIMAS Scores*, for a detailed discussion on the topic. See chapter 7, *Case Study*, for a case study example on interpreting differences between raters using the report shown in appendix E.8). When a statistically significant difference of scale scores between any two given raters for the same individual student is found, the next step is to evaluate this difference in the context of individual behavioral items. For example, if the difference in the parent's and teacher's scale scores were found to be statistically significant, the assessor may want to find out the reason for the disparity by examining the item-level differences (outlined in the next step) or collect additional information relating to how the youth behaves in different environmental settings. On the other hand, when convergence in scale scores is found between raters on a particular scale, the assessor should review items that were rated as behavioral concerns for the youth by multiple raters.

#### Step 3b: Item-level Comparisons

When comparing items with elevated scores, assessors should consider three basic considerations: item content, convergence of evidence, and relevance to intervention. As previously discussed, the content of items can help to prioritize them as targets for intervention. For example, an item rated as *Concern/Mild Concern* related to self harm must be prioritized over one related to fidgeting. The "convergence of evidence" consideration requires the assessor to review all the items rated in the *Concern/Mild Concern* ranges across multiple raters and in the context of available background and history information. This review is done to further prioritize targets for intervention based on the convergence of evidence suggesting that the item taps an area of concern across multiple settings. The third consideration in item-level comparisons is to take the *Concern/Mild Concern* items with prioritized content and a convergence of evidence and to review them for relevance to intervention. This consideration is somewhat setting-dependent, as the targets for a clinical treatment plan would likely differ from those of a school-based behavioral intervention plan. The idea here is to identify the items that are most functionally related to the student's true area of need and those most likely to be related to targets of the ensuing intervention.

Since the Clinician form is not norm-referenced, item-level comparison between raters involving Clinician ratings can be done through a comparison between raters' item raw

scores. The same three basic considerations relating to item content, convergence of evidence, and relevance to intervention apply.

## Step 4. Develop an Intervention and Monitoring Plan

During this step of the process, information obtained from the *Standard Individual Assessment Report* (Report ID: 4, shown in appendix E.4) at Steps 1 and 2 is combined with information gathered in Step 3 to develop the appropriate intervention for the student. This process involves creating an intervention to address the behavioral and adaptive areas of concern while incorporating areas of adaptive functioning in which the student is perceived to be performing at typical levels or higher. School-based professionals can use the gathered information to place the student into existing intervention(s) or to individualize an intervention based on need.

In addition to developing an intervention, this step involves the creation of a student progress monitoring plan. Here BIMAS Standard items of particular concern can be used for goal development and to create brief BIMAS Flex assessments. For example, during a meeting for a student who has been identified as at-risk for developing behavioral concerns, the assessor can examine which Standard form items were rated as *Concern* or *Mild Concern* from Step 2. In particular, heed should be paid to those items that have been rated as *Concern* or *Mild Concern* by multiple raters (from Step 3). Items believed to be the most relevant to the youth's intervention should be considered as the anchor items for BIMAS Flex. The assessor can select one to five Flex items on the BIMAS Online (see the *BIMAS Flex* section in chapter 3, *Administration and Scoring*, for a discussion on the use of the BIMAS Flex). The BIMAS Standard, then, can be used for longer-term progress monitoring (e.g., quarterly).

## Step 5. Monitor Response to Intervention

Once an intervention and monitoring plan has been established for the student, routine assessments should be conducted over time to assess the student's response on both frequent BIMAS Flex assessments and more detailed/norm-referenced assessments using the BIMAS Standard. The results of both data sources can be compared to the baseline and to previously determined goals to evaluate the student's response and to revise the intervention accordingly.

## Step 5a: BIMAS Flex

The following Progress reports are available for evaluating students' RTI on the BIMAS Flex:

- ***Flex Individual Progress Report*** (Report ID: 12; see appendix E.12). The time series graphs can be used to evaluate progress in terms of changes in item scores on each Flex item. For instance, when tracking a positively worded Flex item, where item scores are expected to increase as intervention progresses (i.e., a positive change), if the slope of the item score line is steeper than that of the Goal Line, it means that the change in behavior is happening at a faster rate than expected. This would indicate that the intervention is making a positive impact. In such cases, the assessor can consider moving onto other behaviors to be progress-monitored. On the other hand, if the item score line is progressing in the opposite direction (i.e., item scores are dropping) or advancing at a much slower pace than the Goal Line, the assessor may want to review the intervention and/or the goal that was set previously.
- ***Flex Individual Comparison Between Raters*** (Report ID: 13; see sample report in appendix E.13). Since BIMAS Flex assessments can be completed by teachers, parents, the youth, and clinicians (e.g., the school psychologist), it is useful to compare Flex item ratings between raters with the column graphs provided in this report. If agreements between raters are obtained on most Flex items (i.e., item scores between raters do not show large discrepancies), it means that the effects of intervention are apparent across different settings. However, when large discrepancies are observed between raters in Flex item scores, comments from different raters (if any) during the course of intervention should be reviewed to identify the reasons why the ratings might differ. This analysis may help shape intervention strategies.

## Step 5b: BIMAS Standard

Given the change-sensitivity of the BIMAS Standard items, they are ideal for monitoring changes over time. The following Progress reports are available for evaluating students' RTI on the BIMAS Standard:

- ***Standard Individual Progress Report*** (Report ID: 14; see appendix E.14). The assessor should first review the time series graphs in this report to pinpoint general patterns of change as well as any changes in scores that can be associated with intervention program changes (as shown by Intervention Phase Lines in the graphs).

- **Standard Individual Progress Report: Significant Change Over Time** (Report ID: 15; see appendix E.15a). This report offers assessors the tool to identify statistically significant changes based on *T*-scores using the Reliable Change Index (RCI; Jacobson & Truax, 1991) between baseline and follow-up test scores. The magnitude of change across raw scores can be examined with effect size (ES) estimates (see appendix E.15b). Refer to *Scores for Progress and Outcome Monitoring* in chapter 5, *Understanding and Interpreting BIMAS Scores*, for more interpretive information about these metrics of change.
- **Standard Individual Progress Report: Item Analysis** (Report ID: 16; see appendix E.16). Changes on an item-level on the BIMAS Standard can be tracked using the time series graphs. Assessor can compare how specific behaviors progress within certain domains during the course of intervention, even in circumstances where scale-level changes are not observed.

Note that when a youth moves from one normative age group to another between two administrations (e.g., from the 5–6 age group to the 7–9 age group), *T*-scores, Levels of Risk/Functioning, and Item Descriptors may change just because the norm group has changed. For example, a raw score of 13 on the BIMAS–Teacher Conduct scale converts to a *T*-score of 71 (*High Risk*) for a 6-year-old youth (using the default combined gender norms), but a *T*-score of 68 (*Some Risk*) for a 7-year-old youth. By the same token, the item response of 2 = Sometimes on the item “acted sad or withdrawn” produces an Item Descriptor of *Concern* for a 6-year-old youth on the BIMAS–Teacher Negative Affect Scale, but of *Mild Concern* for a 7-year-old youth. In cases where a change in normative age groups has occurred, the assessor should review raw scores as well as item responses when interpreting results.

# 7 Case Study

The case study described in this chapter illustrates the use of the Behavior Intervention Monitoring Assessment System (BIMAS™) in a typical 3-Tier assessment of a school district. However, assessors can easily apply BIMAS results to a 4-tier RTI program as the principles described in this manual apply readily to any system that involves stages of broad screening to more targeted/intensive individual interventions. The case illustrates one of the many possible functions and uses of the BIMAS. It follows the same interpretation steps discussed in chapter 6, *Step-by-Step Interpretation of BIMAS Results*. Sample reports are presented in appendix E. Certain part(s) of a few selected reports are shown within the chapter to help illustrate the text.

The Springfield Central School District is a small-sized district of 500 students (fictional district name). The district has two elementary schools (Kindergarten to Grade 5) and one middle school (Grade 6 to Grade 8). Springfield implemented academic RTI programming several years ago, initially targeting literacy in the primary grades and gradually expanding to other subject areas and into the middle school. With demonstrated success in improving students' academic performance, the district has moved to implementing a model of RTI for monitoring student behavior.

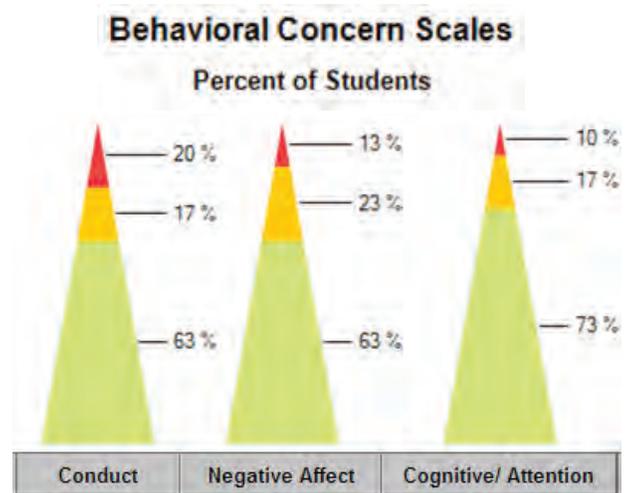
## Tier 1 (Universal Level): Group Evaluation

Springfield began the implementation of behavioral RTI by targeting the elementary schools in the district. After having some time to get familiar with their classes, teachers were asked to provide BIMAS Standard ratings for their students. These ratings were entered into the BIMAS Online, and reports were generated to assist with universal screening at a district, school, grade, and class level.

### Step 1. Identify General Areas of Concern

At the **district** level, a preliminary analysis involves reviewing district-level data with Assessment reports so that general areas of concern can be identified. The **Risk**

Figure 7.1. Risk Level Pyramids



*Level Pyramids* report (Report ID: 1; see appendix E.1 and Figure 7.1) showed that across the three Behavioral Concern scales, Conduct showed the highest percentage of students (20%) scoring in the *High Risk* category within the district, compared to 13% on the Negative Affect scale and 10% on Cognitive/Attention. District level decision-makers learned that Conduct was also an area of concern at both Lincoln Middle School (20% in *High Risk*) and MacDonald Elementary School (21% in *High Risk*). A review of office referrals and school discipline data confirmed that these two schools struggled most with students who have conduct-related issues. All of the above information suggested a need for systemic programming in Character Education and/or Positive Behavioral Supports for Springfield, especially for the identified schools.

District level assessors also identified concerns surrounding Social and Academic Functioning in the district. For the simplicity of discussion, the following sections will focus on reviewing data on the Conduct scale. However, it is important to note that even when a particular BIMAS scale has been identified as the most problematic at the group level, results pertaining to all other scales should be considered simultaneously. The same interpretation steps can be followed for all of the BIMAS scales.

At the **school** level, the same report was generated for individual schools within Springfield to review each school's

overall functioning as well as a breakdown by grade. Lincoln’s report revealed that the 6<sup>th</sup> Grade (ages 11 to 12) had the highest percentage of *High Risk* ratings on the Conduct scale across all grades, while the 4<sup>th</sup> Grade (ages 9 to 10) at MacDonald school showed the most concern in that domain. These data further clarified the grades within each school in which behavioral programs targeting conduct issues were most vital.

At each **grade** level, the list of students within the *Some Risk* or *High Risk* groups on the **Student List by Risk Level** report (Report ID: 2; sample report shown in appendix E.2a contains a portion of the output only; a selected number of students within the grade are included at each risk level for demonstration purposes) were considered at school RTI meetings to see if any were identified as requiring elevation to Tier 2 or Tier 3 services. In order to identify the general areas of concern within individual classes, the *Risk Level Pyramids* report (Report ID: 1) was generated at the grade level. It revealed that at MacDonald, Class 4A showed the highest percentage of at-risk students on the Conduct scale among all Grade 4 classes.

At the **class** level, assessors at Lincoln and MacDonald used the **Class/Group Student Scores** report (Report ID: 3) to review the BIMAS scale scores for each student in specific classrooms. Using the 4A classroom from MacDonald as an example (refer to appendix E.3 and Figure 7.2), school assessors were able to visually pinpoint the six students who were scoring in the *Some Risk* (yellow) and *High Risk* (red) zones by focusing on the Conduct scale column in the report. Looking across each row, students who were scoring within the *Some Risk/High Risk* ranges on Conduct were also showing concerns across multiple BIMAS scales. This helped to identify students who might require Tier 2 or Tier 3 intervention in the classroom. The total number and percentage of students scoring at each risk level obtained from the classroom report aided in the development and prioritization of class-based intervention programs based on an overview of students’ needs across the five BIMAS scales.

In addition, students who scored at the *Strength* level (scale scores highlighted blue) across the two Adaptive scales were identified. For instance, even though Joey Deppe scored in the *Some Risk* range on both Conduct and Negative Affect, and *High Risk* on the Cognitive/Attention scale, his teacher’s ratings resulted in a classification of *Strength* on the Social scale. Joey’s abilities in the social domain were used to help compensate for his weaknesses in the design of his intervention program. Other students who were identified with strengths on the Adaptive scales were paired up with other at-risk students in the class to provide assistance and to model good classroom behaviors. Lastly, the easy access to individual student’s *Standard Individual Assessment Report* (Report ID: 4; see appendix E.4 for a sample) offered by the hyperlinks in the classroom report helped assessors at both schools to prioritize students within each class for possible elevation into Tiers 2 or 3, based on the severity of students’ behavioral concerns. Refer to *Tiers 2 and 3 (Targeted and Intense Levels): Individual Evaluation* for related discussion on the topic.

## Step 2. Identify Specific Populations In Need of Intervention

In this step, multi-student-level Demographic and Comparative reports were reviewed to pinpoint specific populations in need and to identify any demographic risk factors in order to address possible inadequacies in existing services for various demographic groups.

**District** level assessors at Springfield began by examining how demographic categories were distributed within the district. **Demographics Breakdown** reports were generated at different administrative levels (Report ID: 5; with Service Code shown as an example in appendix E.5), and assessors at Springfield identified that 57% of students were in General Education in the district in UA 1, 23% were in Special Education, and 20% of students were receiving Title 1 funding. In particular, the same report generated at the

Figure 7.2. Class/Group Student Scores

Click on student name to view the student’s BIMAS Standard Individual Assessment Report for the selected Universal Assessment.

Student Name	Behavioral Concern Scales Higher T-scores indicate MORE concerns.			Adaptive Scales Higher T-scores indicate LESS concerns.	
	Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
<a href="#">Black, Juliana</a>	57	54	52	47	54
<a href="#">Chan, Christina</a>	55	56	54	53	66
<a href="#">Deppe, Joey</a>	63	68	75	73	54
<a href="#">Farrah, Colleen</a>	83	81	59	39	31
<a href="#">Fernandez, Jose</a>	57	50	52	47	51

school level revealed that a higher percentage of students were in Special Education at Lincoln (25%) compared to the other schools in the district (21% at MacDonald and 17% at Rouge).

To examine how the demographic variable interacts with behavioral risk levels, *Risk Level by Demographics* (Report ID: 6; see appendix E.6 and Figure 7.3) was generated with Service Code selected as the demographic variable. It was found that across the district, students receiving Title I funding had a disproportionate percentage of students scoring in the *High Risk* level on Conduct (67%), and students in the Special Education category had a high percentage of at-risk students (29% in *High Risk* and 43% in *Some Risk*). Report data were used to support discussions about the ways in which psychosocial service delivery could be improved for specific populations in need.

In addition, the school-to-school comparison of average *T*-scores on the *Average Score Comparison Report* (Report ID: 7; see appendix E.7) showed that MacDonald and Lincoln had school average scores above the overall district average of 61 on the Conduct Scale (62 at both MacDonald and Lincoln). These average *T*-scores at UA 1 became the baseline for the district and for individual schools against which progress across subsequent UAs can be compared across the academic year.

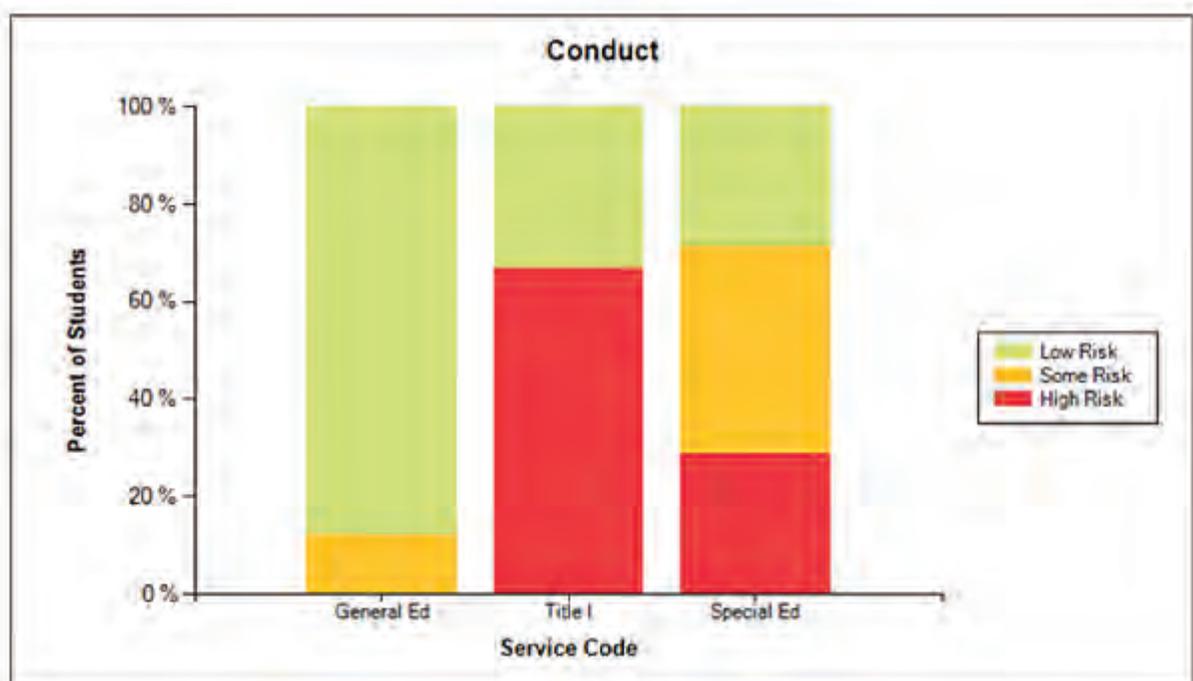
By pairing the information obtained from the *Average Score Comparison Report* with the *Risk Level by Demographics* report, assessors at Springfield identified demographic risk factors that were associated with concerns at schools that required the most attention in the district. The average *T*-scores for specific demographic categories were also obtained at each school. This information was used to set the appropriate goals for each of the Service Categories for subsequent UAs at each school. Assessors at Springfield also examined the programs and demographic factors within Rouge Elementary that might have contributed to the school's lower average score on Conduct to see if some of those programs can be adapted in the other two schools within the district.

Using the same reports generated at the **school/grade/class** levels, school officials at individual schools identified the same types of risk factors that were identified at the district level. Strategies to improve service delivery for the identified populations were examined. This information was funneled back to the district for consideration of increased funding to mitigate any inefficacy in current programming.

Intervention teams at individual schools also used the drill-down function from *Risk Level by Demographics* report (Report ID: 6) to *Student List by Risk Level* (Report ID: 2) to prioritize individual cases for follow-up assessments, consultation, and/or recommendation for psychosocial services specific to the high-risk demographic group to which each student belongs.

**Figure 7.3. Risk Level by Demographics**

Click on any of the columns in the graph to obtain a **Student List by Risk Level** report filtered by that specific demographic variable category.



### Step 3. Collect Additional Information

Once the areas of concern were identified and specific schools, grade levels, classes, and groups of students were specified as “at-risk,” additional information was collected to validate BIMAS ratings and to inform selection of students to be elevated into Tiers 2 or 3 for progress monitoring and intervention design at various group levels. At the district level, office referral and school discipline data confirmed that Lincoln and MacDonald were the schools with the highest levels of concern related to student conduct. For instance, the principal of MacDonald was aware that students, especially those in the 4<sup>th</sup> Grade, were in serious need of behavioral intervention, as they had been a challenging group for teachers and administrators during the year prior. A review of the 4<sup>th</sup> Grade class-level data helped to highlight a particular class with more students experiencing conduct problems. The school psychologist, along with the principal and teachers in the school, discussed how to tailor instructions within classes that have a higher number of students with behavioral needs.

At the grade and class levels, BIMAS data were used to generate a list of students rated with concerns in the *Some Risk* (yellow) to *High Risk* (red) levels on the Conduct scale. Additional information was collected for these students, including academic and behavioral records, ratings from parents and the youth, and additional assessments when deemed necessary to more fully understand the concerns for the student and to inform intervention design.

District and school officials also reviewed data on other BIMAS scales and discussed what programs can be implemented to improve students’ functioning in respective behavioral domains.

### Step 4. Develop an Intervention and Monitoring Plan

Based on district-level BIMAS data from Step 1 through 3, Springfield District, in consultation with the targeted schools, decided to implement a school-wide positive behavioral support project at both Lincoln and MacDonald. The project involved the whole school at both sites as well as staff and student training. Clear behavioral school-wide expectations were developed, student motivators were provided for positive behavior, and discipline referrals continued to be monitored. Additional funding for resources was allocated to Lincoln to address the specific needs of students within the Special Education classes. Use of BIMAS data at MacDonald resulted in additional grade-level interventions implemented in all 4<sup>th</sup> Grade classrooms, including additional behavioral skill instruction and increased use of incentives to motivate

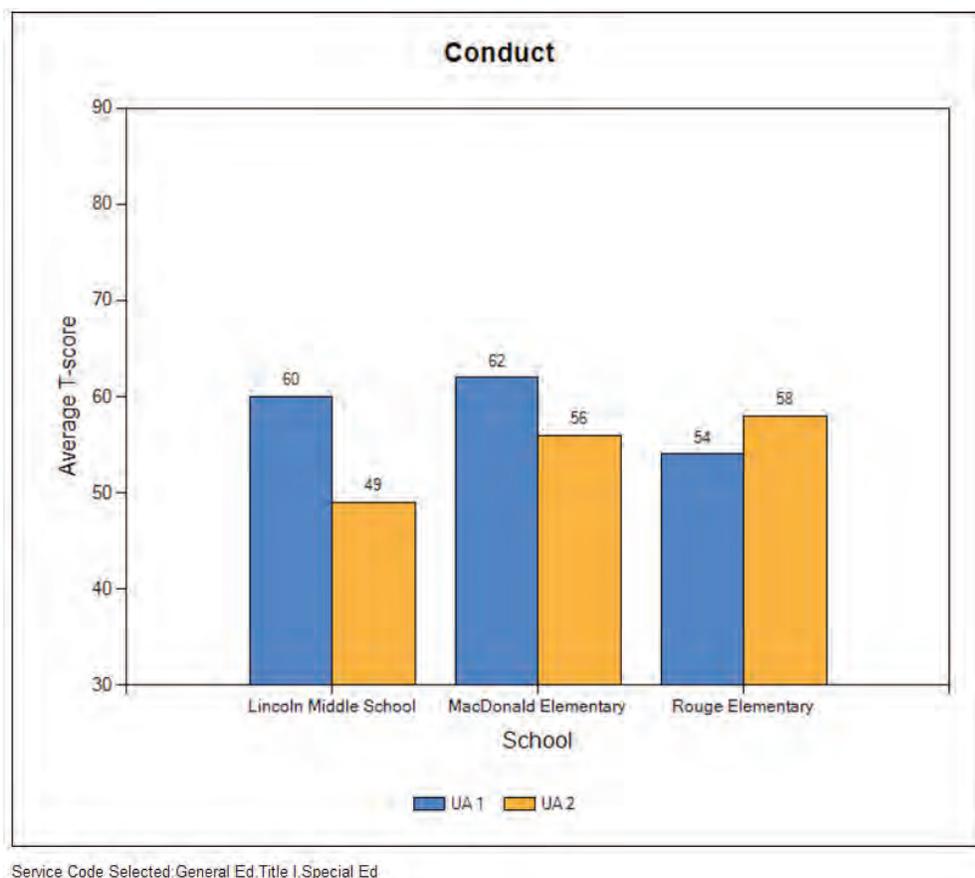
students to exhibit desired behaviors. In the 4A classroom where there was a high number of students with behavioral concerns, an experienced teacher mentor was assigned to the class teacher to help her develop a more effective classroom management system. Students who had received *High Risk* or *Some Risk* ratings were reviewed for possible inclusion in Tier 2 or 3 intervention programming.

The monitoring plan for Springfield District included three UAs, with BIMAS screening to be conducted again in Winter and in Spring. After the first UA in the Fall term, a goal was created to have 80% of students fall into the *Low Risk* level across all BIMAS Behavioral Concern scales by the end of the year at both Lincoln and MacDonald. Additionally, a goal was set to have average *T*-scores on the Conduct scale at the two schools drop below 60 by the Winter UA at all grade levels. In addition to these goals, other more specific targets were established with respect to attendance, discipline referrals, grades, and curriculum-based test scores.

### Step 5. Monitor Group Progress over Time

To assess progress in meeting these goals, UA 2 BIMAS data were collected for all students at the beginning of the Winter term alongside curriculum-based testing at the Winter term benchmark. Following every UA within the school year, district- and school-level assessors examined results specific to the UA in question using the previously described BIMAS Assessment reports types. For example, at the **district** level in UA 2, most schools were at or close to having 80% of students achieving *Low Risk* levels of concern on the Conduct scale as revealed by the *Risk Level Pyramids* report (Report ID: 1). Although MacDonald had not yet reached the goal of having 80% students scoring in the *Low Risk* range on the Conduct scale at UA 2, results suggested that progress had been made as the percentage did show a decrease. On the other hand, Lincoln seemed to have achieved the goal, with 80% of the school’s students receiving teacher ratings within the *Low Risk* range. Moreover, BIMAS Progress reports such as the **Average Score Comparison: Progress Report** (Report ID: 9; see appendix E.9 and Figure 7.4) indicated encouraging results in UA 2. The average *T*-scores on the Conduct scale was well below 60 for all three schools within the Springfield District. In particular, the school average *T*-score at Lincoln showed big improvement across the two UAs, from 60 at UA 1 to 49 at UA 2. The situation also improved for MacDonald—from the Fall UA school average *T*-score of 62 to Winter UA’s average *T*-score of 56. Although the school average score for Rouge Elementary had gone up from 54 to 58, it was still well within the district-wide goal of 60.

Figure 7.4. Average Score Comparison: Progress Report



The impact of the district’s psychosocial service delivery to the level of risks in behavioral functioning among students within the Title 1 category was examined with the *Risk Level by Demographics: Progress Report* (Report ID: 11; see appendix E.11). The percentage of students within the *High Risk* group had dropped from 67% in UA 1 to 20% at UA 2. Most of the students from the *High Risk* group in UA 1 had moved into the *Some Risk* category (40% at UA 2) or the *Low Risk* category. In other words, the *Low Risk* group showed an increase of 7% across the two UAs. Students in the Special Education group also showed great improvement. There were no more *High Risk* scores evident in UA 2 and the percentage of students in the *Some Risk* group had dramatically decreased as well for this demographic category. These progress data showed that the reallocation of resources within the district to address the behavioral needs of at-risk students in specific demographic populations during the Fall term was very effective.

At the **school** level, the same report was generated to review progress made in each grade for various Service Code categories. Students within the Title 1 and Special Education categories who were identified as either *High Risk* or *Some Risk* on the Conduct scale in UA 1 showed much improvement in UA 2. In fact, within each Service Code category,

the pre-determined goals with respect of risk level distribution were reached. These changes provided good evidence that the school-wide behavioral programs that were put into effect since the Fall UA had brought about some improvements, particularly within those grade levels that required the most urgent change.

Furthermore, at the **grade** level, individual students who were identified by the RTI teams as requiring behavioral progress monitoring were noted as such on the BIMAS Online in the students’ demographic profile. Thus, assessors at Lincoln or MacDonald were able to keep track of services provided to students in need. The progress version of the *Student List by Risk Level* report (Report ID: 2; sample report shown in appendix E.2b includes a portion of the output and does not include all students in the grade. The same students within each risk level in appendix E.2a [assessment version] are included in this progress version. See also Figure 7.5 for a portion of the report reproduced here) was then generated to include both UA 1 and 2 data. The report indicated that all students scoring within the *Some Risk/High Risk* groups from UA 1 were in progress monitoring with a closed circle shown in the “In Prog Monitor” column, reflecting the information entered in the BIMAS Online student demographic information profile page. Students who had a *High*

Figure 7.5. Student List by Risk Level

**Conduct Scale**

In Prog Monitor = In Progress Monitoring at time of assessment      N/A = No Score Available      ● = Yes ○ = No

High Risk Group (T-score = 70+) in Universal Assessment 1				
	UA 1		UA 2	
Student Name	T-score	In Prog Monitor	T-score	In Prog Monitor
<a href="#">Farrah, Colleen</a>	83	●	65	●
<a href="#">Klumming, Heather</a>	81	●	62	●
<a href="#">Nathanson, Andy</a>	71	●	65	●
Some Risk Group (T-score = 60-69) in Universal Assessment 1				
	UA 1		UA 2	
Student Name	T-score	In Prog Monitor	T-score	In Prog Monitor
<a href="#">Deppe, Joey</a>	63	○	55	●
<a href="#">Lola, Jojo</a>	69	●	46	●
<a href="#">Taylor, Dudley</a>	63	○	61	●
Low Risk Group (T-score < 60) in Universal Assessment 1				
	UA 1		UA 2	
Student Name	T-score	In Prog Monitor	T-score	In Prog Monitor
<a href="#">Black, Juliana</a>	57	○	57	○
<a href="#">Chan, Christina</a>	55	○	53	○
<a href="#">Fernandez, Insa</a>	57	○	53	○

Risk (red) score in UA 1 on the Conduct scale had either a *Some Risk* (yellow) or *Low Risk* (green) score by UA 2. Assessors were also able to quickly review each student’s *Standard Individual Progress Report* (Report ID: 14; see sample report in appendix E.14) by using the student name hyperlinks provided in this report. (See *Step 5: Monitor Response to Intervention within Tier 2 and 3 [Targeted and Intense Levels]: Individual Evaluation* section later in this chapter for a detailed discussion on how to monitor response to intervention at the individual level).

At the **class** level, the ***Class/Group Student Scores: Progress Report*** (Report ID: 10; see appendix E.10 for a sample report generated for the 4A class at Macdonald) showed that even though some improvements have been made (i.e., many *High Risk* scores have disappeared in UA 2) in the targeted classes, there continued to be a small number of students rated in the *Some Risk/High Risk* level in relation to conduct problems. As a result, a meeting was scheduled to collect more information and to consider continual Tier 2 or Tier 3 programming for these students.

The aforementioned steps were planned to be repeated in the Spring UA to ascertain whether further improvements were made.

## Tiers 2 and 3 (Targeted and Intense Levels): Individual Evaluation

Alongside the district-/school-/or grade-wide programs targeting universal conduct-based issues identified through UAs at Tier 1 (Universal level), individual students were identified as *Some Risk* or *High Risk* within other behavioral or adaptive domains assessed on the BIMAS. Some of the students were elevated into more targeted or intensive levels of progress monitoring at Tiers 2 or 3. To illustrate this process, the case study of John Oakes is presented. John is a student from the 6<sup>th</sup> Grade at Lincoln who was identified as a student in need of intervention. While many of the students were identified as being at-risk for developing conduct-related issues, John was identified from the multi-student level drill-down reports, such as the *Class/Group Student Scores* report (Report ID: 3), as scoring in the *High Risk* range on Negative Affect. Further, the teacher team referred John for Tier 2 (Targeted Level) intervention due to the fact that some ideation related to themes of self-harm was reported. This portion of the case study illustrates the *Step-by-Step Interpretation Guidelines: Individual Evaluation* section described in chapter 6.

## Step 1. Interpret the BIMAS Standard Scale Scores to Identify General Areas of Concern

### Step 1a: Interpret the Behavioral Concern Scale Scores

To begin this process, John's *Standard Individual Assessment Report* (Report ID: 4, see appendix E.4 and Figure 7.6) was reviewed to identify scales that were scoring in the *Some Risk* ( $T$ -score = 60 to 69) to *High Risk* ( $T$ -score = 70 to 85) range, as these would be areas for additional assessment and potential intervention. As illustrated by John's report, teacher ratings indicate a borderline score on the Conduct scale ( $T = 58$ ; *Low Risk*), a high level of concern related to Negative Affect ( $T = 78$ ; *High Risk*), and a borderline score of  $T = 59$  (*Low Risk*) on Cognitive/Attention.

### Step 1b: Interpret the Adaptive Scale Scores

The *Adaptive Scales: T-scores* portion of John's individual assessment report was reviewed to identify Adaptive scales that were rated in the *Concern* range ( $T \leq 40$ ) as well as the *Typical* ( $41 \leq T \leq 59$ ) or *Strength* ( $T = 60+$ ) ranges. Adaptive scales in the *Concern* level are areas for additional assessment and potential intervention, whereas adaptive areas rated in the *Typical* or *Strength* levels of functioning can be incorporated into an individualized intervention to overcome areas of concern. As illustrated on John's report

(presented in appendix E.4), his Academic Functioning score was a significant *Concern* (within the red zone) within the Adaptive scales ( $T = 23$ ). This finding led to the collection of additional information related to John's academic motivation, performance, and skill development. John was rated well within the *Typical* (green) range on Social ( $T = 42$ ), indicating that an intervention that uses group interventions and social rewards might be more powerful for John than other types of positive reinforcements.

## Step 2. Examine Item-Level Responses for an In-depth Analysis of Concerns

Moving to the *Items Score by Scale* portion of the report, the assessor proceeded to review items contributing to John's *High Risk* level score on Negative Affect and *Concern* level score on Academic Functioning to aid in the design or selection of the type of intervention.

### Step 2a: Behavioral Concern Scales

As illustrated in the report in appendix E.4 and Figure 7.7, John's Negative Affect problems might include being sad/withdrawn, depressed, and experiencing thoughts of self-harm. This last item by itself at a *Concern* level rating would warrant immediate follow-up. Other items rated at a *Concern* level included loss of temper and difficulties staying on task. Further, a mild level of concern was indicated with respect to anger, fighting, sleepiness, being

Figure 7.6. Standard Individual Assessment Report (Scale Scores)

#### Behavioral Concern Scales: T-scores

Higher scores indicate MORE concerns.

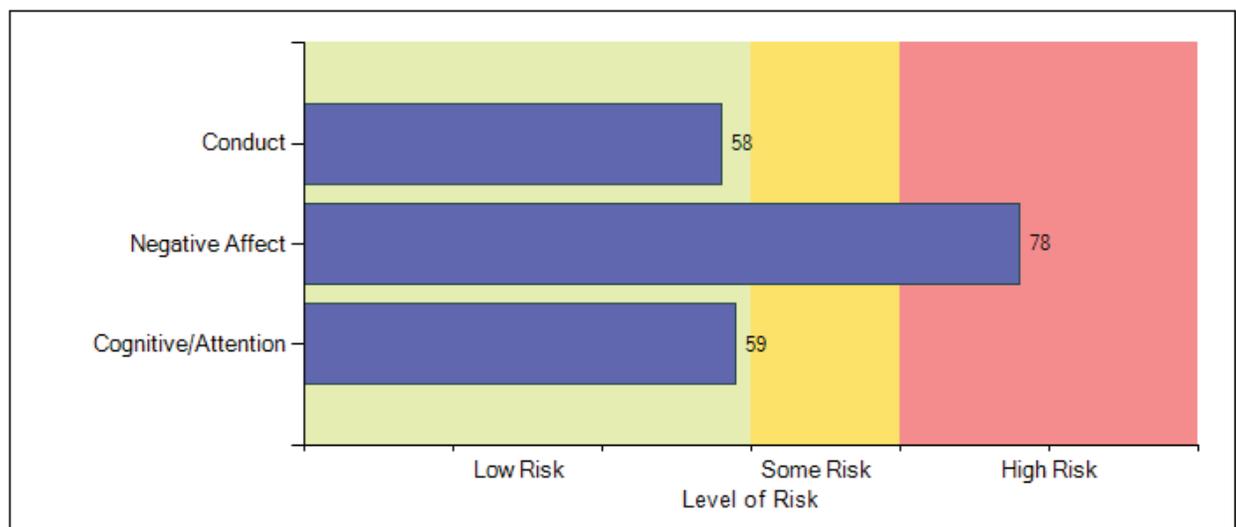


Figure 7.7. Standard Individual Assessment Report (Item-Level Scores)

Negative Affect				
T-score	Percentile	90% CI	Scale Descriptor	
78	99	72-84	High Risk	
Item			Item Score	Item Descriptor
12	sad/withdrawn		4	Concern
24	thoughts of hurting self		3	Concern
8	depressed		2	Concern
5	sleepy/tired		2	Mild Concern
27	emotional or easily upset		2	Mild Concern
20	anxious		1	No Concern
16	embarrassed/ashamed		0	No Concern

emotional or easily upset, fidgeting, and trouble in organization and planning. Given John's item-level scores, one emerging clinical hypothesis was that he was experiencing sadness and depression, and he might be externalizing those emotions in the form of conduct problems at times. He also encountered negative emotions in the form of negative thoughts, in particular those related to self-harm. Lastly, he was manifesting cognitive/attentional concerns related to fidgeting behavior and diminished ability to stay on-task as well as to plan and organize.

This case study example illustrates the importance of reviewing item-level scores in each student's individual assessment report as soon as any of the Behavioral Concern scales have been flagged as *Some Risk* (yellow) or *High Risk* (red). It is possible that for John, once the negative affective concerns have been alleviated, any conduct-type behaviors that were related to emotional adjustment will improve as the root of the problems is addressed.

## Step 2b: Adaptive Scales

A review of John's Adaptive scale item scores suggested that on the Academic Functioning scale, items at the *Concern* (red) level included preparing for class, failing grades, working up to potential, and school absences. Further, although at the scale level, John was rated in the *Typical* range on the Social scale, a breakdown of item scores within the scale revealed that one item on the scale, "relating with others," was rated at a *Concern* level while another item, "shared thoughts," was rated at the *Mild Concern* (yellow) level. The rest of the items on the Social scale were rated as *Fair* (green). In relation to his academic functioning, the item "followed directions" was rated in the *Mild Concern* range, while the rest of the items on the scale were rated in the *Concern* range. Based on these item-level results, the assessor theorized that John's academic functioning might

be related to his cognitive and attentional issues (e.g., diminished ability to stay on task and difficulty with organization and planning). As well, his discomfort with social interactions might have been related to the emotional issues he was experiencing.

While the picture of John's areas of concern was becoming clearer, a review of his Adaptive scale scores in the *Typical* or *Concern* range was undertaken to inform the intervention design/selection process. While John was perceived to have some discomfort with social interactions, his report also suggested that he was friendly, able to work out problems with others, and able to speak clearly. These items suggest that John might do well in group interventions to address his depression, sadness, thoughts of self-harm, and problems staying on task. Further, his Adaptive Functioning ratings suggest that academic interventions might need to focus on class preparation, homework, and attendance. The intervention team also considered academic supports that involved supervised peer mentoring, given that John's overall social functioning at the scale level is in the *Typical* range.

## Step 3. Collect Additional Information and Compare Results Between Raters

While the BIMAS *Standard Individual Assessment Report* from each BIMAS administration is informative, it should never be used in isolation for diagnoses, intervention design/selection, or high stakes placement decisions. Additional information that was collected for John included a review of his school records, interviews with his parents and teachers, classroom observations, and additional testing, including BIMAS administrations to additional raters (e.g., parent, self). These additional data sources did confirm that he was a student in need of additional support.

As illustrated on the *Standard Individual Comparison Between Raters* (Report ID: 8; see appendix E.8 and Figures 7.8 and 7.9), the ratings from John and his mother provided additional evidence that John did struggle with Negative Affect. Specifically, John provided a *Some Risk* ( $T = 68$ ) level self-rating and Mrs. Oakes gave a *High Risk* ( $T = 75$ ) rating for her son, which converged with the teacher's *High Risk* rating ( $T = 78$ ). In relation to other Behavioral Concern scales, all three raters provided similar *Low Risk* ratings on Conduct and Cognitive/Attention. With regards to Adaptive scales, all three raters provided *Concern* level ratings for John on Academic Functioning ( $T = 23$ ,  $T = 28$ , and  $T = 34$  for Teacher, Parent, and Self-Report, respectively). These results provided converging evidence that John is struggling academically. Further, the teacher's ratings produced a borderline *Typical* score on the Social scale ( $T = 42$ ), whereas both the parent and self-ratings identified on the Social scale as an area of concern ( $T = 39$  and  $33$ , for parent and self-report, respectively).

Turning to the item-level analysis portion of the rater comparative report and focusing on the Negative Affect scale (see Figure 7.9), all three raters provided *Concern* level ratings on the self-harm item. Other items that received

*Concern* or *Mild Concern* level ratings across all three raters included "sleepy/tired," "depressed," "sad/withdrawn," and "emotional or upset." The assessor team further explored the issue of self-harm ideation and made referrals for the school psychologist to have interviews with John alone and also with his mother. It became clear from the school psychologist's assessment that John did not intend to act on his thoughts.

On the other two Behavioral Concern scales (identified by all three raters as *Low Risk* at the scale level), both John and his mother endorsed very few items related to conduct behaviors, even though loss of temper, fighting (verbally/physically/both), and anger were identified by the teacher as *Concern/Mild Concern*. However, trouble in organization and planning were identified by all three raters as a *Mild Concern*. Problems staying on task seemed to be exhibited in the classroom, as rated by the teacher as a *Concern*, but not by Mrs. Oakes in the home or by John himself. On the other hand, item-level scores on the Social scale showed that speaking clearly was an adaptive skill that John possessed since he was rated as *Positive* by his mother on this item, and both John and the teacher had provided a *Fair* item rating. Sharing thoughts and relating with others seemed

Figure 7.8. Standard Individual Comparison Between Raters (Scale-Level Comparison)

Rater	Teacher 1	Parent 1	Self
Rater Name	Mr. Steve Murr	Peggy Oakes	John Oakes
Date of Assessment	12/09/2010	12/10/2010	12/10/2010

### Behavioral Concern Scales: T-scores Comparison Between Raters

Higher scores indicate MORE concerns

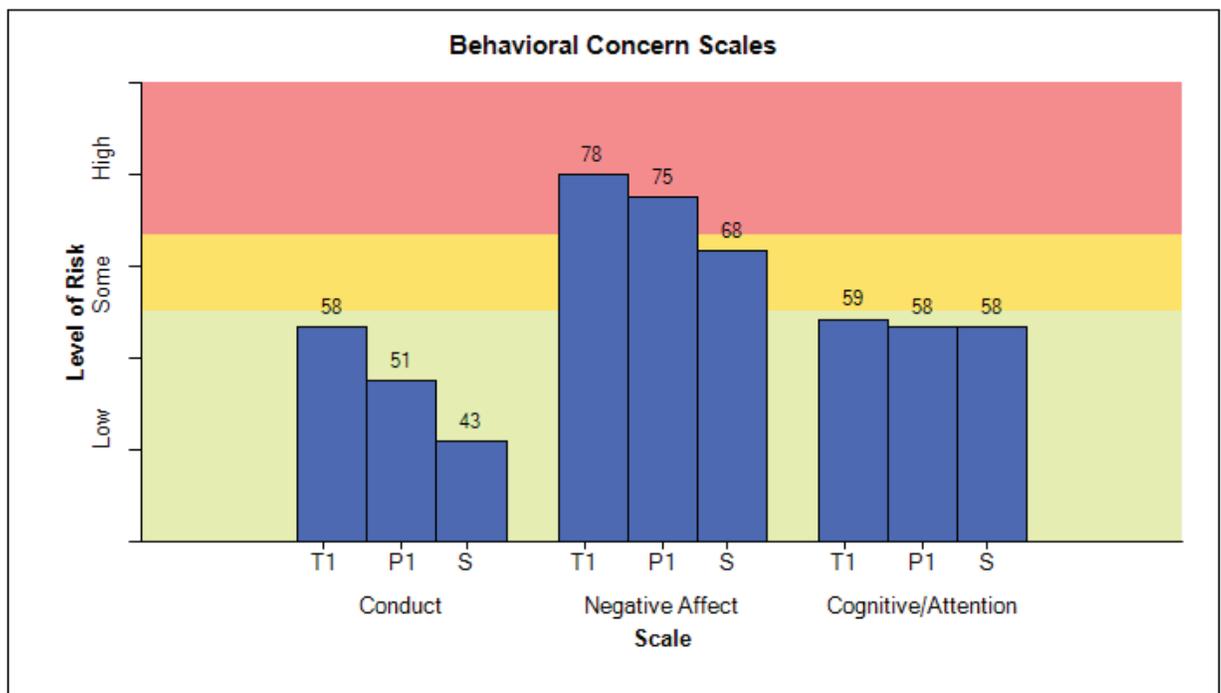


Figure 7.9. Standard Individual Comparison Between Raters (Item-Level Comparison)

Item No.	Item Stem	Rater's Item Raw Score		
		T1	P1	S
<b>Negative Affect</b>				
5	sleepy/tired	2 Mild Concern	4 Concern	3 Mild Concern
8	depressed	2 Concern	3 Concern	3 Concern
12	sad/withdrawn	4 Concern	3 Concern	2 Mild Concern
16	embarrassed/ashamed	0 No Concern	1 No Concern	2 Mild Concern
20	anxious/worried/nervous	1 No Concern	1 No Concern	1 No Concern
24	thoughts of hurting self	3 Concern	3 Concern	3 Concern
27	emotional/easily upset	2 Mild Concern	3 Concern	2 Mild Concern
<b>Scale Score Scale Descriptor</b>		<b>T = 78 High Risk</b>	<b>T = 75 High Risk</b>	<b>T = 68 Some Risk</b>

to be common concerns shared between all raters. On the Academic Functioning scale, which was rated as *Concern* across all three raters at the scale level, failing grades, working up to academic potential, being prepared for class, and absence from school were rated in the *Mild Concern* or *Concern* range by all three raters.

Taken together, all the rater comparative information suggested that John displayed a high level of concern in observable, mood-related behavioral concerns, both at home and at school. Moreover, the lower scores in John’s self-rating might suggest John’s lack of insight into his own affective functioning, or it might suggest John’s reluctance to share/communicate about his emotional difficulties. John’s lower self-reported score on Negative Affect did not necessarily mean that John was not experiencing difficulties in affective functioning: it could suggest that he might require help in recognizing his own emotional states and their connection with outward, observable behaviors.

### Step 4. Develop an Intervention and Monitoring Plan

During this step of the process, the interpretative information from the BIMAS is taken in the context of the additional information gained in Step 3 in order to create or select an appropriate intervention for the student of concern. In John’s case, he was provided with daily academic intervention services to support him in school. He was also provided with an intervention plan to address negative affect with a placement in the school psychologist’s weekly support group. Lastly, John was enlisted as a peer tutor to read books to younger students twice a week.

To monitor John’s progress, a weekly BIMAS Flex assessment was developed and a follow-up BIMAS Standard assessment was scheduled to be administered to the same teacher, the parent, and John in 10 weeks. During a meeting for John, the team managing John’s case at school selected five Standard form items that were rated as *Concern* from UA 1 screening to be used as the anchor in selecting Flex items for John out of the global list available on BIMAS Online. Standard anchor items “depressed,” “sad/withdrawn,” and “thoughts of hurting self” were tracked with Flex items “had a negative self image,” “took part in group activities,” and “expressed thoughts of hurting himself/herself.” On the Adaptive scales, Standard anchor items “prepared for class” and “absent from school” were tracked with the Flex items “turned in completed work on time” and “attended classes.” Particular attention was given to choosing positively worded Flex items to encourage the increase of positive/adaptive behaviors at this early stage of intervention, since in John’s case, reinforcing a positive self-image may help mitigate depressive mood and/or ideation of self-harm.

After selecting the five Flex items, the team proceeded to assign the target item score for each Flex item, as well as the date for the goal to be achieved, so that a Goal Line would appear on each Flex item’s time series graph. The team also assigned the scoring criteria for each Flex item. Weekly Flex assessment reminders were scheduled on the BIMAS Online to remind the assessor managing John’s case to send Mrs. Oakes and John the Flex forms via emailed links. The teacher logged onto the BIMAS Online to complete the weekly Flex assessments. At the end of the 10-week period, the teacher again completed a Standard form to see if any improvements evident on the Flex form could be verified by the norm-referenced Standard form.

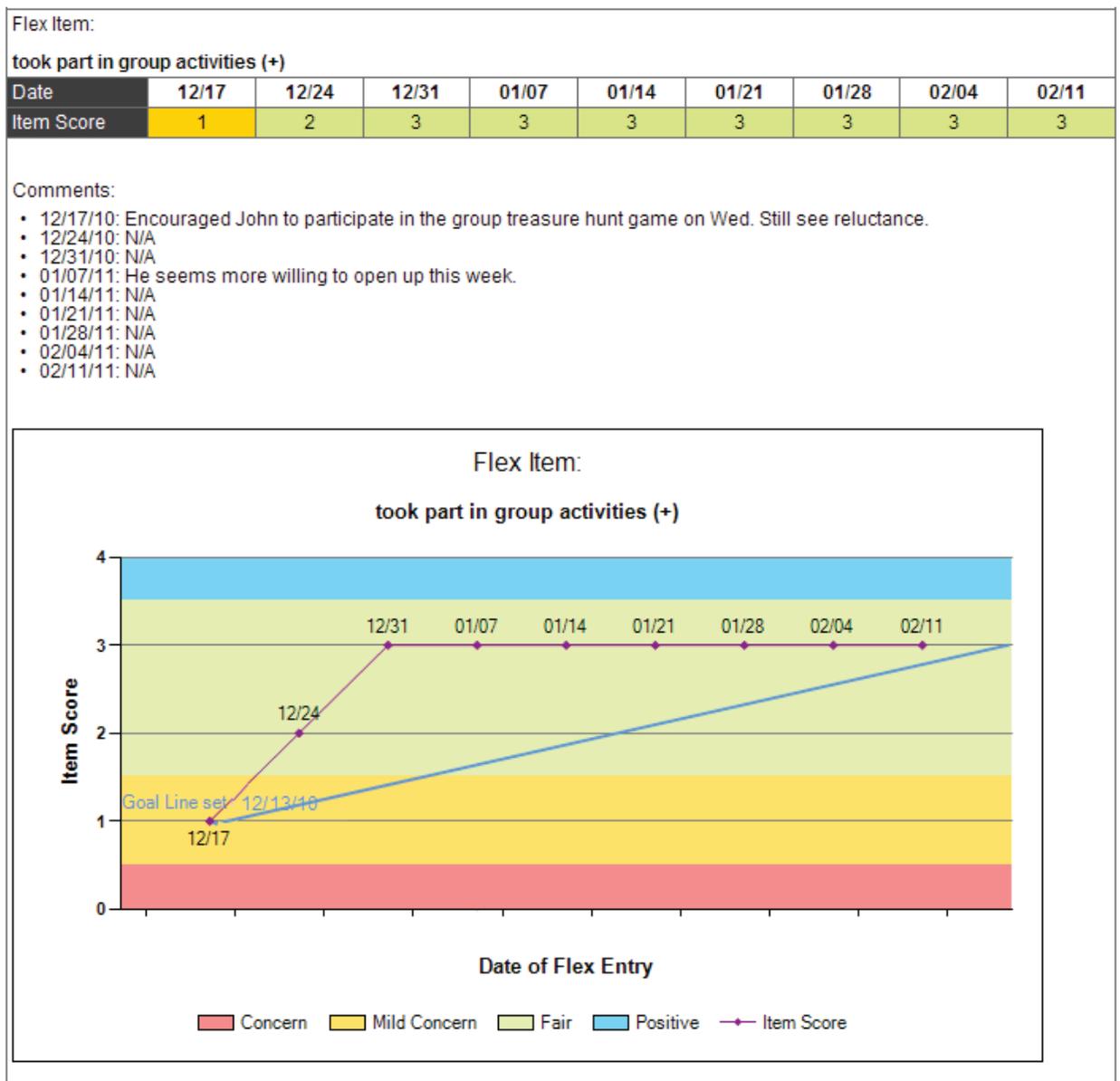
## Step 5. Monitor Response to Intervention

John’s progress throughout the intervention was monitored weekly with the BIMAS Flex. As illustrated in the *Flex Individual Progress Report* (Report ID: 12; see appendix E.12 and Figure 7.10), John’s progress based on weekly teacher ratings indicated that he was making progress on all five specific behaviors being tracked. The time series graph of the Flex items indicated change in the desirable direction (i.e., negative self-image and thoughts of hurting self had decreased, whereas participation in group activities, turning in completed work on time, and class attendance had increased). Progress for all the items was also on target (or

for certain Flex items even exceeded expectation) since the item score lines and the blue Goal Lines ran almost parallel in terms of the slope.

Each rater’s weekly Flex assessments were also compared using the *Flex Individual Comparison Between Raters* (Report ID: 13; see appendix E.13 for a sample of one of the weekly assessments). The raters differed sometimes on certain Flex items, but most of the times, their ratings on the Flex items converged. The assessor also noted the comments from different raters on an item-to-item basis at each weekly Flex assessment and incorporated them into the intervention along the way.

Figure 7.10. Flex Individual Progress Report



In addition to the Flex reports, the follow-up BIMAS Standard completed by the teacher after 10 weeks of intervention indicated that John’s *Concern* level ratings had generally decreased. The **Standard Individual Progress Report** (Report ID: 14; see appendix E.14 and Figure 7.11; sample report shown in the appendix and Figure 7.11 include John’s BIMAS Standard UA 2 assessment past the 10-week mark in order to demonstrate the Intervention Phase Dates/Notes feature of the report) revealed that the intervention designed to address John’s Negative Affect and Academic Functioning issues were effective since the time series graphs showed a decrease on Negative Affect and an increase in Academic Functioning scale scores. (The Intervention Phase Dates/Notes on the Academic Functioning scale showed that intervention required some adjustment to accelerate change.)

The **Standard Individual Progress Report: Significant Change Over Time** (Report ID: 15) was reviewed to further examine the above changes. The assessor first identified any statistically significant change based on *T*-scores using the Reliable Change Index (RCI; see appendix E.15a and Figure 7.12; sample report shown in the appendix also includes John’s BIMAS Standard UA 2 assessment past the 10-week mark to demonstrate the Overall Change column). The report indicated that the 10-point change in *T*-Scores on the Negative Affect scale between UA 1 and the subsequent 10-week follow-up assessment was a statistically significant change. This was a very positive result, as Negative Affect was John’s primary concern based on the first assessment. However, this report also indicated that the increase in

Academic Functioning ratings as well as the decline in scores on the other Behavioral Concern scales did not reach statistical levels of significance.

As such, the assessor proceeded to assess change over time using effect size (ES; based on changes in raw scores; see appendix E.15b) in order to evaluate the magnitude of change across scales. As the report shows, contrary to the results based on RCI or changes in *T*-scores, a meaningful change was only observed on the Academic Functioning scale (ES = 1.41, *Improved*). No meaningful changes were observed on other BIMAS scales, including the Negative Affect scale (ES = -.44, *No Change*) between UA 1 and the subsequent 10-week follow-up assessment. Results differed on the two metrics of change because each depends on a different set of parameters (e.g., RCI is affected by the norm group’s standard error of prediction estimates and ES is affected by item score fluctuations between assessments; refer to *Scores for Progress and Outcome Monitoring* section in chapter 5, *Understanding and Interpreting BIMAS Scores*, for detailed discussion on the two changes metrics). Information from the two versions of the report suggested that while John’s Academic Functioning ratings did not produce a change in *T*-scores large enough to reach statistical significance, he likely made noticeable positive changes as rated by the teacher. On the other hand, on the Negative Affect scale, since John was scoring in the *High Risk* level, the progress made in the domain overall in terms of *T*-scores had reached statistical significance, putting John in the *Some Risk* category. On an item-level, there might still be room for improvement to trigger a bigger magnitude of change

Figure 7.11. Standard Individual Progress Report

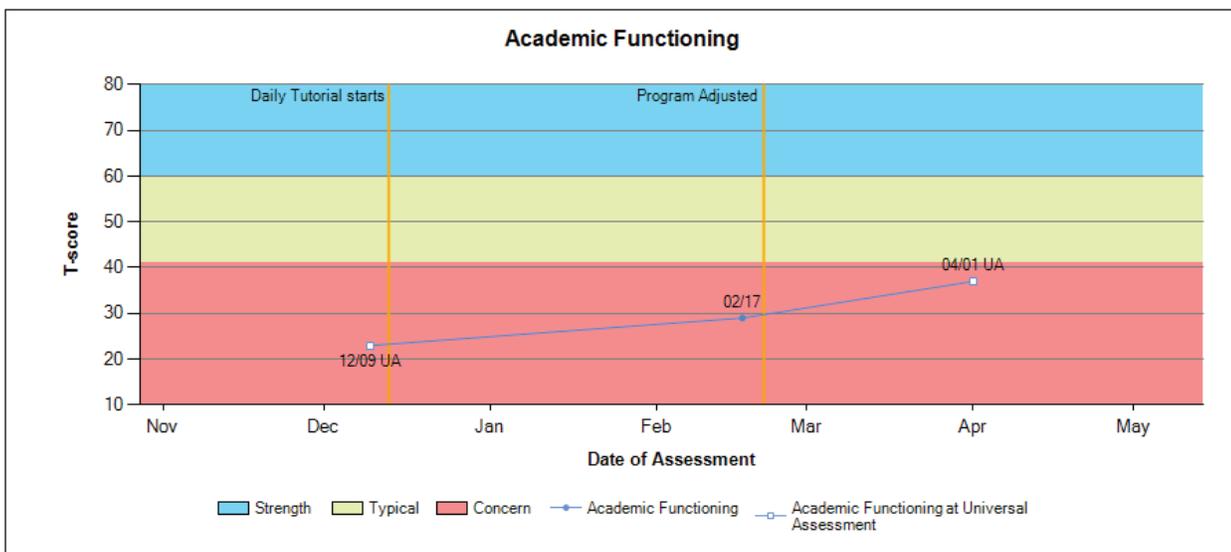


Figure 7.12. Standard Individual Progress Report: Significant Change Over Time (RCI–Based on T-score Option)

Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Negative Affect	T-score	78	68	63	Improved
	Level of Risk	High	Some	Some	
	Significant Change ( $p < .10$ )	12/09 <sup>UA</sup> -02/17 (Improved) 02/17-04/01 <sup>UA</sup> (No Change)			

in order for it to be detected by ES. The assessor, therefore, used the direct link provided on the *Standard Individual Progress Report: Significant Change Over Time* report to obtain time series graphs of item responses on the *Standard Individual Progress Report: Item Analysis* (Report ID: 16; see appendix E.16).

Altogether, it appeared from the individual level Progress reports that John was responding in a positive direction to the interventions being implemented at school. Data indicated that the current interventions should continue. The school's RTI team proceeded to schedule a meeting with John's parents to share the results. In addition, 10 additional weekly Flex assessments were scheduled with the teacher to continue progress monitoring John's RTI. Another follow-up assessment using the norm-referenced BIMAS Standard would take place should further positive changes be observed in this second round of Flex assessments.

# 8 Development

The Behavior Intervention Monitoring Assessment System (BIMAS™) development project began in 1997 with a series of empirical studies aimed at identifying items that are sensitive for monitoring changes in response to intervention. The project encompassed three years of data collection wherein thousands of ratings from teachers, parents, and clinicians as well as students' self-ratings were gathered from multiple data collection sites. Intensive research and sophisticated statistical analyses informed each step of the instrument's development, as described in this chapter.

First, the rationale and goals for BIMAS development are outlined. The preliminary research forming the basis for the development of the BIMAS is summarized, including a description of: (1) the Intervention Item Selection Rules (IISRs) used for BIMAS item selection, (2) the preliminary structure and item development process, and (3) the development of the BIMAS response schemas. Next, final scale construction during the normative phase of development is described. The normative sample is briefly described, followed by a description of the BIMAS Standard and BIMAS Flex development. Next, the criteria for developing BIMAS Standard Item Descriptors scores are presented. Finally, the creation of the final forms is described.

## Rationale and Goals

The BIMAS development process began with the identification of a need for a specific kind of test and a set of test-related goals and purposes. The following section describes the rationale and development goals for the BIMAS.

### Rationale

Initial discussions regarding the development of the BIMAS occurred in the context of two major professional trends that highlighted the need for a measure like the BIMAS. First, efforts towards cost containment of health services, including mental health services, led to the creation and rapid spread of managed care companies. Managed care initially slowed health care expenditures, but its creation also brought along a host of negative effects on the provision of behavioral health services (Davis & Meier, 2001); for example, many counselors and counseling agencies, in settings from private practice to college counseling centers, now routinely

employ brief therapy. This trend towards brief therapy has put pressure on mental health practitioners to deliver the most cost-effective treatment. Often times, costs (from the managed care company's standpoint) and benefits (from the client's perspective) is a difficult balancing act. As a result, managed care's emphasis on increased accountability has led to a greater focus on more objective measurements of client progress and outcomes that has the potential to improve both science and practice efforts.

In the schools, the 2004 reauthorization of the Individuals with Disabilities Education Improvement Act (IDEA) further emphasized the importance of prevention, early intervention, and accountability as related to special education and intervention services. As a result, educators have been employing Response-to-Intervention (RTI) approaches in the development and monitoring of their programs. As described in chapter 2, *Background*, RTI is a 3-Tier model that integrates assessment and intervention; schools identify students at risk for poor learning outcomes through universal screening and provide evidence-based interventions to maximize student achievement and to reduce behavior problems. The intensity and nature of interventions can be adjusted depending on students' responsiveness to intervention. As a result, the movement of RTI has led to a greater need for measures that can be employed to screen students for behavioral health problems as well as monitor student progress and outcomes resulting from behavioral and psychosocial interventions.

Whether in schools, clinics, or other settings, intervention approaches like RTI require valid instruments that can track a youth's status over time. The main rationale for the development of BIMAS was to produce a measurement system that could be easily administered multiple times within the RTI system for screening, progress monitoring, program evaluation, and data-based decision making.

### Development Goals

The following goals were established for the BIMAS development project:

1. Develop a brief, repeatable, psychometrically sound measure for universal screening, progress monitoring, and program evaluation.

2. Provide a change-sensitive measure based on an empirically supported model for selecting change-sensitive items, specifically, Meier's (1997, 1998, 2000, 2004) Intervention Item Selection Rules (IISRs; see *Preliminary Research: Intervention Item Selection Rules* in the next section).
3. Include content to assess both adaptive behaviors as well as the most common behavioral and social-emotional concerns observed in childhood and adolescence.
4. Validate forms for use with multiple informants so that information may be collected across the various contexts and from different perspectives.
5. Create a user-friendly yet sophisticated web-based graphing and reporting system for easy data entry, automated reports, and graphic displays of data for screening, progress monitoring, and program evaluation.

## Preliminary Development

### Preliminary Research: Intervention Item Selection Rules

In an effort to create a measure that is change-sensitive, a series of studies were conducted (Meier, 1997, 1998, 2000, 2004; Meier, McDougal, & Bardos, 2008; Weinstock, & Meier, 2003), and this body of empirical work formed the basis for the development of the BIMAS. In the Meier et al. (2008) study, for example, parents of 896 elementary school-aged children receiving psychotherapy interventions from community mental health agencies completed a preliminary outcome measure assessing the youth's symptoms and functioning. Pre- and post-intervention data were examined to compare sensitivity to change between item groupings in scales. Results indicated that scales formed with change-sensitive items evidenced larger effect sizes than scales composed of the original item pool, and demonstrate adequate reliability estimates.

The process employed in the development of the BIMAS focused on identifying constructs that change as a result of emotional and behavioral interventions. To identify intervention-sensitive items either during test construction or item evaluation, Meier (1997, 1998, 2000, 2004) proposed a set of Intervention Item Selection Rules (IISRs). The central philosophy of the IISRs is that intervention-sensitive items should change in response to an intervention and behave in a theoretically expected manner in other conditions (e.g., remain stable over time when no intervention is present). This approach assumes that (a) test items and tasks differ along a

trait-state continuum, and (b) different test construction and item analysis procedures are necessary to select items with a high state loading that reflect the results of interventions. IISRs are designed to test two broadly competing claims regarding change at the item level: such change is the result of an intervention, or the change results from other factors that constitute error in the context of scale development. Meier has developed and studied scales constructed with both traditional and IISRs procedures in a variety of clinical and school settings (e.g., Meier 1998, 2000, 2004). Overall, scales constructed with IISRs procedures demonstrated larger treatment effect sizes than traditional scales, and produced adequate reliability estimates (Meier, McDougal, & Bardos, 2008). Items on change-sensitive measures will share some characteristics with traditional, trait-sensitive tests. Intervention-sensitive items, for example, should also be theoretically based, reliable, and unrelated to systematic error sources. However, intervention-sensitive items should possess additional properties, foremost of which is that they change in response to an intervention.

The nine IISRs are described below. The first two describe assumptions about the nature of intervention-sensitive items, and the remaining seven IISRs describe procedures for collecting and evaluating empirical data.

The IISRs are as follows:

1. Items are initially identified in a literature review such that they are theoretically grounded and related to relevant research.
2. Intervention-based items are aggregated across individuals, but not across items or occasions (as with trait-based tests). Aggregation across individuals decreases random error (Messick, 1989) and increases the possibility of detecting item scores that are responsive to intervention effects (Epstein, 1979, 1980).
3. The range of scores at pre-test is reviewed so that items demonstrating obvious ceiling or floor effects may be removed.
4. Scores on intervention-sensitive items over time must demonstrate change in intervention groups.
5. Scores on intervention-sensitive items over time must exhibit change in the expected direction. This ensures that the items are indeed change sensitive and are able to monitor change in the expected direction.
6. Change observed in an intervention group is investigated relative to a non-treated comparison group.
7. The items must not demonstrate differential effects between these groups prior to intervention.
8. The results of the item change must not be related to systematic error sources such as social desirability.

9. Steps 3 through 8 should be cross validated with repeated studies of new samples from the population of interest.

As indicated in step 9 above, the development of an evaluation instrument based on IISRs requires a recursive process of cross validating items both for screening and progress monitoring purposes with different samples of populations of interest. Towards this end, preliminary versions of the BIMAS were developed using both clinical and school samples in field settings, a combination rarely studied in psychotherapy research (Kazdin, 2000). In fact, Weisz, Huey, and Weersing (1998) found only nine studies (over a period of 50 years) that examined the effects of child treatment in applied settings. This is important because meta-analyses examining the effects of psychotherapy provided in laboratory settings versus field settings indicate that the former produced effect sizes around .75 and the latter around 0 (Weisz, Weiss, & Donenberg, 1992). Thus, these results imply that the psychometric properties of tests developed with non-clients in non-clinical settings may not be generalizable for use with children and adolescents receiving psychosocial interventions in applied settings. Since the BIMAS was developed to be used in school districts or organizational (applied) settings for the purpose of behavioral screening or progress monitoring, field studies utilizing both clinical and school samples add weight to the external validity of the tool.

In an unpublished intervention program evaluation study conducted in 1998, McDougal and Meier utilized an initial item set to evaluate outcomes of children who were receiving school-based mental health services in a number of schools in Syracuse, NY. Over the course of the evaluation, teacher ratings were obtained for 338 children receiving services as well as for 47 children without any behavioral or mental health concerns who served as a control group. Results of the evaluation served as impetus for further research, as children receiving services overall demonstrated improvement in teacher ratings over time while teacher ratings of the control group remained stable. Further, the composite means of children in the clinical group were significantly different ( $p > .001$ ) from the means of non-clinical children, indicating that this item set might have some utility as a screening measure.

Initial published studies have found differences in the psychometric properties of items developed with IISRs and other more traditional item selection rules (Meier, 1998, 2000; Weinstock & Meier, 2003). For example, in an intervention study aimed at changing college students' attitudes toward alcohol use, Meier (1998) contrasted the IISRs approach with traditional item evaluation methods using a 16-item alcohol attitude scale. Application of the more traditional type of intervention item selection criteria versus the IISRs approach resulted in the creation of two sets of items with differing psychometric properties. The intervention-sensitive items demonstrated greater

pre-intervention to post-intervention change. In another study, 116 parents completed a social skills scale at intake and follow-up periods as part of treatment for children in a community health center, it was also found that scales composed of items that met IISRs criteria had larger effect sizes (Meier, 2000). Furthermore, in an early study investigating preliminary change-sensitive items to be included on the BIMAS, Weinstock and Meier (2003) subjected both intake and follow-up item scores on a 56-item self-report checklist completed by 615 university counseling center clients to principal component analysis (PCA) as well as an IISRs evaluation. As predicted, items selected using IISRs demonstrated larger effect sizes (when combined into scales) than items selected through PCA. Taken together, the results from these three empirical studies provide evidence that the more traditional use of trait-based methods of item selection and evaluation that focuses on detecting individual differences may be problematic for measuring constructs that should theoretically change in response to intervention.

More recently, Meier, McDougal, and Bardos (2008) studied a preliminary version of the BIMAS using parent ratings of 896 youth clients (Kindergarten through Grade 10) receiving psychotherapy at a community clinic. Results indicated that this early version of the BIMAS was a change-sensitive scale which demonstrated adequate reliability, larger effect sizes between pre- and post-intervention group means than most assessment tools created for other purposes (e.g., ones that diagnose a stable trait-like construct), and that it was brief enough to be administered repeatedly for the purpose of progress monitoring.

## Preliminary Structure and Item Development

The change-sensitive items found in Meier's (1998, 2000, 2004) research were used as a preliminary pool of items for the BIMAS. The items from these studies could be grouped into two categories: Distress/Problems and Strengths. Items within the Distress/Problems category included feelings of depression, behaving differently than peers, impulsivity, fidgeting, fighting with others, and failing grades. Items within the Strengths category included content related to interpersonal communication, paying attention to others when engaged in a conversation, ability to make friends, and helping with household tasks. More items were developed following a review of the literature on children's behavioral problems (Stiffman, Orme, Evans, Feldman, & Keeney, 1984), reviews of other measures of children's distress and functioning (Meier, 1998), and suggestions from mental health professionals (e.g., school counselors, psychologists, and social workers).

Using clinical judgment and rationale, the initial item pool was examined to (a) categorize items by similar content, (b) determine if item content was consistent for each rater

type (e.g., teacher, parent, or student) or if there was a need to revise or delete certain items on a particular rater form, and (c) examine if item content made sense for different age groups. Additional items were also created as the need arose to meet the goals of (a), (b), and (c).

Next, the initial item pool was categorized on the basis of face validity, into four domains. *Externalizing* behaviors referred to items assessing conduct problems, substance abuse, and deviant behaviors, while *Internalizing* behaviors focused on negative affect related to anxiety and depression. *Cognitive Processing* items focused on themes related to attention, focus, bizarre thoughts as well as behaviors. Finally, *Adaptive Functioning* items consisted of a broad category of content that included academic, social, and interpersonal functioning. With all items, item content was revised or created so that they were observable or behavioral in nature (e.g., “appeared sad” was employed instead of “felt sad”).

Another goal for the BIMAS was to create forms, with matching content and roughly equal numbers of items, for use by different raters. A final number of about 30 to 35 items per rater form was desired to keep the BIMAS brief and suitable for a one-page format. Similar to other scales developed for screening and monitoring purposes, comparable forms that can be completed by teachers, parents, clinicians, and youth were created. Given the desire to create parallel forms for these rater types, the appropriateness and wording of item contents for each informant were reviewed. For content related to school (e.g., received failing grades, came prepared to class), clinicians who work with the youth in a community mental health setting may not have access to this information; consequently, these items were dropped from the Clinician form. Some of these school-related items were retained but others were dropped from the Parent form as well. Similarly, the Clinician form contains an item relevant to therapy content (i.e., attended therapy session) that was irrelevant for other raters. The vast majority of items, however, assessed information deemed to be accessible to all four informants.

Item content was also reviewed in relation to age appropriateness. While the item content domains remained the same across rater forms, attention was paid to ensure that item wording was appropriate for the age span across Kindergarten through Grade 12 (Ages 5–18). In addition, a review of item content and previous research indicated that a self-report version for youth below Grade 7 was unlikely to evidence sufficient reliability and validity estimates; therefore, this version of the self-report was dropped.

## Response Schemas

Instructions for raters on the BIMAS forms ask the respondent to assess how often a child or adolescent manifested each of the behaviors during the past week. Response choices include 0 = Never, 1 = Rarely, 2 = Sometimes, 3 = Often,

and 4 = Very Often. Each of these descriptions is paired with a frequency of occurrence during the past week. Never (the response of 0) is paired with 0 times or never observed; Rarely (response = 1) with 1 or 2 times or to a minimal extent; Sometimes (response = 2) with 3 to 4 times or to a moderate extent; Often (response = 3) is 5 to 6 instances or to a significant extent; and Very Often (response = 4) is 7 or more times or to an extreme extent. In previous research with a preliminary version of the scale (see Meier et al., 2008), the pairing of behavioral frequency ratings with the category descriptors was found to increase the mean and standard deviation of the resulting scale scores compared to data collected with a version using the category descriptors alone (Meier, 2008). By pairing the number of occurrence of a behavior (frequency count), a more objective measure, with the category descriptors (an overall impression of the extent of behavior), the authors believe that respondents would be better able to gauge behaviors using different dimensions, thus increasing the change-sensitivity of the BIMAS.

## Creation of the Forms for the Normative Study

Results from the preliminary research and rational analysis of the items guided item selection for the standardization version of the tool. This process resulted in 36 items for the normative version of the BIMAS on the Teacher, Parent, and Self forms and 34 items on the BIMAS Clinician form. The order of the items for each form was randomized using random-number-generating functions (StatSoft, 2001), resulting in new forms for the normative study.

## Final Scale Construction

Development of the final scales involved collection of data for normative and clinical samples, factor analyses to determine the factor structure of the forms, and the development of the Behavioral Concern scales, Adaptive scales, and Flex items.

## Data Collection

Data for the normative versions of the BIMAS forms were collected from the general population and from selected clinical groups. The extensive data collection project resulted in normative samples that included ratings from 1,400 teachers and 1,400 parents (on youth aged 5 to 18 years), and 700 youth (aged 12 to 18 years). The normative samples were representative of the general U.S. population in terms of age, gender, race/ethnicity, geographic location, and parental education level (parent version only) in accordance with the 2000 U.S. Census (see chapter 9, *Standardization*, for a full description of the normative samples). Additionally, 538

teacher ratings, 467 parent ratings, and 350 youth self-ratings were collected for youth with a clinical diagnosis (diagnoses included Disruptive Behavior Disorders, Attention-Deficit/Hyperactivity Disorder, Anxiety Disorders, Depression, Pervasive Developmental Disorders, Learning Disorders, and Developmental Delays; see *Clinical Samples* in chapter 11, *Validity*, for a description of the clinical groups). More detailed descriptions of the normative sample, the standardization process, and the psychometric properties of the BIMAS are provided in chapters 9 to 11, *Standardization*, *Reliability*, and *Validity*.

## Development of the BIMAS Standard

Data from the normative and clinical samples were analyzed to confirm the item content of the BIMAS Standard. Item-level analyses conducted on the 36 BIMAS Standard items revealed that two items (i.e., “behaved differently than his/her peers” and “expressed strange or bizarre thoughts”) showed extreme floor effects (i.e., they had very little variability and were almost always rated as 0 or 1), and were therefore cut from the BIMAS Standard. As a result, the final number of items on the Teacher, Parent, and Self forms is 34 and that for the Clinician form is 31 (four items tapping academic functioning were replaced by one item on the attendance of therapy appointments). In order to verify the proposed scale structure of the BIMAS Standard (see *Preliminary Structure and Item Development*, earlier in this chapter), these final items were subjected to confirmatory factor analyses (CFA using Maximum Likelihood, generalized least squares estimation). Results indicated adequate model fit (see *Content Validity and BIMAS Scale Structure* in chapter 11, *Validity*, for more information on the confirmatory factor analysis procedure and results). Decisions regarding the final scale structure were made based on the authors’ clinical experience, the research literature, as well as results from the CFA.

## Development of the Behavioral Concern Scales

The statistical and rational analysis of the items resulted in three main areas pertaining to behavioral distress/problems: externalizing, internalizing, and cognitive/information processing. Further review of the item content (all negatively worded) contributed to the renaming of the BIMAS Behavioral Concern scales. The Conduct scale (9 items) taps issues related to anger management, bullying behaviors, substance abuse, and deviance (e.g., “appeared angry”). The items within the Negative Affect scale (7 items) assess anxiety and/or depressive types of symptoms (e.g., “acted sad/withdrawn”). The Cognitive/Attention scale (7 items) looks at issues related to attention, focus, organization,

planning, and memory (e.g., “had trouble paying attention”). Refer to appendix A for a full list of items on each BIMAS Behavioral Concern scale.

## Development of the Adaptive Scales

The remaining 11 items on the BIMAS forms had been previously identified as adaptive functioning types of items (see the *Preliminary Structure and Item Development* section earlier in this chapter). These are mostly positively-worded items that tap adaptive skills pertaining to social and academic functioning (note that on the Clinician form, two items that may contribute to therapeutic gains replace the academic functioning items). Having strength-based scales alongside the Behavioral Concern scales may assist in the development of an intervention plan by utilizing the youth’s strengths to overcome potential areas of difficulty. As a result, rather than reverse-scoring all the positively worded items to calculate a scale score that would indicate a behavioral concern or a lack of adaptive skills, all positively worded items are scored as they are worded—higher scores mean that the behaviors in question are more frequently observed. (To this end, two negatively worded items “received failing grades at school” and “was absent from school” were determined to require reverse-scoring.) The Adaptive Scales include: the Social scale (6 items), which taps communication, friendship maintenance, and interpersonal skills (e.g., “shared what he/she was thinking about”) and the Academic Functioning scale (5 items), which assesses a youth’s academic performance, attendance, and attitude in learning (e.g., “worked up to his/her academic potential”). As mentioned before, four of the items on the Academic Functioning scale, with the exception of “followed directions,” were items that pertain more to an academic setting rather than a clinical setting. Instead, the “followed directions” item as well as the item “attended his/her scheduled therapy appointments” composed two standalone Clinician Adaptive items on BIMAS–Clinician. Refer to appendix A for a full list of items on the BIMAS Adaptive scales.

## Development of the BIMAS Flex Items

The BIMAS Flex items were designed to target specific behavioral concerns or skills as measured by the standardized BIMAS Standard items. While the BIMAS Standard was designed to provide information on the individual’s progress across standardized scales, the Flex items were designed to provide information on specific intervention targets, such as “interacted well with friends” or “worked out problems with peers by him/herself,” serving as the basis for Individualized Education Program (IEP) or treatment plan goals. On average, each BIMAS Standard item across the five BIMAS scales has a pool of 10–15 corresponding positively or negatively worded Flex items for assessors

to choose from. Altogether, a total of 655 Flex items were developed for each of the BIMAS Standard Teacher, Parent, and Self-Report forms; and for the Clinician form, a total of 586 Flex items were developed.

Because the Flex items were conceived as a tool to help assessors develop and monitor a set of behavioral goals for a youth, several goal setting guides in the literature, such as the SMART Goals Setting Guide proposed by Nikitina (2006), guided the development of the Flex items. This model suggests that each intervention goal should be:

1. **Specific.** Goals need to be straightforward, specific, clear and easy, and emphasize what needs to happen by answering “who, what, why, and how.”
2. **Measurable.** Establish concrete criteria for measuring progress so you can see the change occur. “If you can’t measure it, you can’t manage it.”
3. **Attainable.** Goals need to challenge the individual, and the individual must believe that he/she will be able to meet the challenge.
4. **Realistic.** Goals that are unrealistic and set too far out of reach will not motivate an individual to commit to intervention and will set the individual up for failure. Change does not usually occur overnight, so devise a plan or way of getting to the goal that is realistic.
5. **Timely.** Set a timeframe for the goal that gives the individual a clear target to work towards.

Appendix B presents the Flex items for each BIMAS Standard item.

## Determining the Scoring Criteria of Item Descriptors for BIMAS Standard Items

Both the BIMAS Standard and BIMAS Flex provide Item Descriptors to guide the interpretation of item-level results: *Concern*, *Mild Concern*, and *No Concern* on Behavioral Concern Scales; *Concern*, *Mild Concern*, *Fair*, and *Positive* on Adaptive Scales, since only the BIMAS Standard is norm-referenced, this section regarding the development of the Item Descriptors pertains only to BIMAS Standard Item Descriptors<sup>1</sup>.

Each item on the BIMAS Standard has been standardized on a U.S. national sample closely matching the most recent U.S. Census in terms of demographics in much the same

<sup>1</sup> The same Item Descriptors are used on BIMAS Flex reports where assessors set up their own item-level scoring criteria based on what item response is considered to be “concern” or “typical” for an individual youth. Assessors are advised to consult the information on how the item-level scoring criteria were developed for the norm-based BIMAS Standard form to help inform the setup of Flex item scoring criteria in accordance to their own school-wide expectations or the individual youth’s behavioral status.

way as the scale-level norms were developed. This standardization process allows assessors to compare a youth to his/her normative group on an item-to-item basis. At the item level, the approach used to designate the risk levels of a BIMAS Behavioral Concern scale item score is similar to the technique used by Naglieri, McNeish, and Bardos (1991), Naglieri, LeBuffe and Pfeiffer (1994), and LeBuffe and Naglieri (2003); all suggested that an individual item score that falls in the top 15% of the normative group distribution (e.g., exceeds the mean normative item score plus one standard deviation) can be considered problematic. To this end, means and standard deviations were calculated for each normative age group (5–6, 7–9, 10–11, 12–13, 14–16, and 17–18 on BIMAS–T & BIMAS–P; 12–13, 14–16, and 17–18 on BIMAS–SR) for each of the 34 BIMAS Standard items for both combined gender (default scoring option) and gender-specific groups. After that, cumulative percentiles were calculated for the various age groups for each Standard item using both combined and gender-specific groups. An Item Descriptor was assigned for each possible item score (0 to 4) on each of the Standard items, separately for each rater form. This descriptor was based on the mean, standard deviation, cumulative frequency, and percentile (see Tables 8.1 and 8.2). In some cases, clinical judgment on item content in relation to gender, age, and observer- vs. self-report differences also assisted in the final Item Descriptors. The following rational rules were considered:

1. *Mild Concern/Concern* should never be assigned for a rating of 1 = Rarely unless the behavior in question is a serious one (e.g., “thoughts of hurting self”).
2. If the response frequency for an item score is more than 25% in the normative sample (e.g., more than 25% of respondents answered a rating of 2 = Sometimes), then that item score should not be flagged as a *Mild Concern/Concern*.
3. For borderline instances, consistency should be maintained as much as possible across normative age groups and across raters (i.e., BIMAS–T, BIMAS–P, and BIMAS–SR).

## Behavioral Concern Scales

As mentioned above, a combination of rationale and data, based on individual item means, standard deviation, cumulative frequency, and percentile across age groups, contributed to the development of the Item Descriptors. Table 8.1 illustrates the data-driven guidelines for the Item Descriptors on the Behavioral Concern scales. The Item Descriptor (e.g., *Concern/Mild Concern/No Concern*) obtained by a rater’s item score (0 = Never, 1 = Rarely, 2 = Sometimes, 3 = Often, or 4 = Very Often) for any given item depends on the item content, type of rater (Teacher/Parent/Self) as well as the youth’s age group (and gender if gender-specific norms were used as the scoring option instead of the default combined-gender norms).

**Table 8.1. Development Guidelines for Item Descriptors: Behavioral Concern Scales**

Item Descriptor	Development Guideline
<b>Concern</b>	$> M + 1 SD$ or $\geq 85^{\text{th}}$ percentile
<b>Mild Concern</b>	$= M + 1 SD$ or between $75^{\text{th}}$ – $84^{\text{th}}$ percentile
<b>No Concern</b>	$< M + 1 SD$ or $< 75^{\text{th}}$ percentile

Note. Development guidelines are in reference to the normative sample's distribution

In general, on any of the Behavioral Concern scales (i.e., Conduct, Negative Affect, or Cognitive/Attention) a *Concern* item score indicates that the behavior was rated as occurring much more frequently than observed amongst most youth in a comparable age group. A *Mild Concern* item score indicates that the youth's behavior was rated as occurring slightly more frequently than in the normative group. On the other hand, a *No Concern* item score means that the behavior occurs at a frequency comparable to the youth's normative age group.

## Adaptive Scales

Table 8.2 illustrates the data-based guidelines for the item-level norms on the Adaptive scales. Since higher scores on the BIMAS Adaptive scales indicate fewer concerns (because they consist of mostly positively worded items), there is an extra Item Descriptor added, *Positive*, to reflect adaptive behaviors which are considered positive or performing beyond the expected level of functioning. Thus, there are four Item Descriptors associated with the Adaptive Scales: *Concern*, *Mild Concern*, *Fair*, and *Positive*. Again, item content, type of rater (Teacher/Parent/Self) as well as the age group (and gender if gender-specific norms was selected as the scoring option) determines the specific Item Descriptor that can be associated with a rater's score (0 to 4) on any item.

A *Concern* item score on any of the Adaptive scales (i.e., Social or Academic Functioning) indicates that the adaptive behavior in question was rated as occurring *much less* frequently than observed amongst most youth in a comparable age group by the same rater. A *Mild Concern* score indicates that the youth's adaptive behavior was rated as occurring *slightly less* frequently than in the normative group. A *Fair* item score means that the youth is displaying the behavior at a frequency comparable to the normative group within the youth's age group. Lastly, a *Positive* item score may indicate that the youth displays the adaptive behavior more frequently than observed amongst most youth in a comparable age group by the same rater.

**Table 8.2. Development Guidelines for Item Descriptors: Adaptive Scales**

Item Descriptor	Development Guideline
<b>Concern</b>	$< M - 1 SD$ or $\leq 10^{\text{th}}$ percentile
<b>Mild Concern</b>	$= M - 1 SD$ or between $11^{\text{th}}$ – $20^{\text{th}}$ percentile
<b>Fair</b>	$> M - 1 SD$ & $< M + .67 SD$ or between $21^{\text{st}}$ – $74^{\text{th}}$ percentile
<b>Positive</b>	$\geq M + .67 SD$ or $\geq 75^{\text{th}}$ percentile

Note. Development guidelines are in reference to the normative sample's distribution

## Creation of the BIMAS Standard Final Forms

After a review of the normative data, the item counts were finalized: 34 items for the BIMAS–T Standard (5–18 Years), BIMAS–P Standard (5–18 Years), BIMAS–SR Standard (12–18 Years) and 31 items for the BIMAS–C Standard (5–18 Years). Appendix A presents the final list of items by scale for each BIMAS form.

# 9 Standardization

This chapter describes the process and methods used to develop the norms for the Behavior Intervention Monitoring Assessment System (BIMAS™) Standard teacher, parent, and self-report versions<sup>1</sup>. As a norm-referenced test, the BIMAS required a nationally standardized sample, which is essential to establish its psychometric qualities. This chapter is organized in the following sections: data collection, description of the normative samples, followed by a discussion on the norming procedures and the derivation of the standardized scores.

## Data Collection

Data collection took place between March 2007 and May 2009. During the standardization and research phase of instrument development, thousands of BIMAS forms were completed; data from these forms were included in the standardization, reliability, and validity research studies (see chapter 10, *Reliability*, for a description of the reliability samples, and chapter 11, *Validity*, for a description of the validity samples). Twenty-five site coordinators throughout the U.S. assisted with the data collection. Site coordinators were recruited by contacting data collection sites known to the publisher, through author contacts, and through several mailing and email campaigns. Each site coordinator was instructed to follow standardized procedures that included obtaining informed consent, having raters follow specific written instructions, and debriefing the raters as needed upon completion of the assessment(s). All participants were aware that forms were being completed as part of the process of BIMAS test development. Site coordinators were compensated for taking part in the data collection process. Once the rating forms were returned to the publisher, they were subjected to visual inspections followed by some initial statistical analyses. Rating forms that had more than 10% of responses missing were excluded from the dataset.

For all assessments, in addition to ratings on the BIMAS items, demographic information about the rated youth was collected, including age, gender, parental education

level (PEL; on the parent forms only), race/ethnicity, and geographic region. Race/ethnicity categories were: Asian/Pacific Islander, Black/African American, Hispanic, Native/Aboriginal, White, Multiracial, and Other. For ease of presentation, these groups are referred to herein in the following manner: Asian, African American, Hispanic, White, and Other (including Native/Aboriginal and Multiracial due to the small sample sizes of these groups). On the parent reports, data were collected on PEL for both parents, and the higher of the two was used to classify the PEL of the child.

Additionally, 538 teacher ratings, 467 parent ratings, and 350 youth self-ratings were collected from youth with a clinical diagnosis. In order to ensure the accuracy of all diagnoses, for every child classified as a clinical case, the data collection site coordinator completed a Clinical Diagnostic Information Form. Clinical cases were accepted only if: (a) a single primary diagnosis was indicated, (b) a qualified professional (e.g., psychiatrist, psychologist) had made the diagnosis, (c) the proper criteria were assessed using either the Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition: Text Revision (DSM-IV-TR; APA, 2000) or the International Statistical Classification of Diseases and Health Related Problems 10th revision (ICD-10; WHO, 2004), and (d) appropriate methods (e.g., record review, rating scales, observation, interview) were used to make the diagnosis. These cases were used for the establishment of the validity of the BIMAS. Refer to *The BIMAS Standard as a Screening Tool* section in chapter 11, *Validity*, for a detailed description of the various clinical samples and their performance on the BIMAS.

## Normative Sample Description

The BIMAS normative samples include 1,400 ratings from teachers on the BIMAS-Teacher (BIMAS-T), 1,400 ratings from parents on the BIMAS-Parent (BIMAS-P), and 700 ratings from adolescents on the BIMAS-Self-Report (BIMAS-SR).

<sup>1</sup> Neither the BIMAS Flex nor the BIMAS Standard Clinician versions are discussed in this chapter, as these versions are not norm-referenced. Since only the norm-referenced Standard form is discussed in this chapter, "BIMAS" is used to denote the BIMAS Standard throughout the chapter.

## Teacher Normative Sample (BIMAS–T)

Teachers completed the BIMAS–T for a normative sample of 1,400 youth. All of the teachers had known the students they were rating for at least 1 month (specifically, the duration of teacher to student familiarity, in months, was: 1–3 = 7.5%, 4–6 = 23.6%, 7–11 = 44.4%,  $\geq 12$  = 23.7%; this data was missing for 0.7% of teachers), thereby meeting the minimum acquaintance requirement for completing the BIMAS.

The normative sample included ratings of 50 males and 50 females at each age (from 5 through 18 years). The sample characteristics were compared to the U.S. population (based on the 2000 U.S. Census report) on race/ethnicity and geographic region. The collected data were very similar to the U.S. Census in terms of race/ethnicity; however, some discrepancies existed between the actual collected data and Census targets for geographic region. To address these discrepancies, the sample was weighted through statistical procedures so that the weighted sample closely matched the U.S. Census statistics both in terms of race/ethnicity (see Table 9.1) and geographic region distribution (see Table 9.2).

**Table 9.1. Race/Ethnicity Distribution:  
BIMAS–T Standard Normative Sample**

Race/Ethnicity of the Rated Youth	Normative Sample (Weighted)		U.S. Census
	N	%	%
Asian	55	4.0	3.8
African American	217	16.0	15.7
Hispanic	203	14.9	15.1
White	836	61.4	61.9
Other	50	3.7	3.5
<b>Total</b>	<b>1,361</b>	<b>100.0</b>	<b>100.0</b>

**Table 9.2. Geographic Region Distribution:  
BIMAS–T Standard Normative Sample**

U.S. Region of the Rated Youth	Normative Sample (Weighted)		U.S. Census
	N	%	%
Northeast	251	18.4	18.1
Midwest	299	22.0	21.9
West	325	23.9	23.3
South	486	35.7	36.7
<b>Total</b>	<b>1,361</b>	<b>100.0</b>	<b>100.0</b>

## Parent Normative Sample (BIMAS–P)

The BIMAS–P rating form was completed for a normative sample of 1,400 children and adolescents. The majority ( $n = 1,116$ ; 79.7%) of the BIMAS–P normative sample comprised assessments completed by the youth's biological mother, while the remaining assessments were completed by the youth's biological father ( $n = 164$ ; 11.7%) or by other significant adults (including non-biological parents and other relatives;  $n = 120$ ; 8.5%). The 1,400 rated youth included 50 males and 50 females at each age (for ages 5 through 18 years). The sample characteristics were compared to the U.S. population (based on the 2000 U.S. Census report) on race/ethnicity, PEL, and geographic region. While race/ethnicity very closely matched the Census targets, a similar statistical weighting procedure described in the Teacher normative sample section was applied to the BIMAS–P sample to correct for discrepancies in PEL and region. The resulting weighted sample therefore closely matched the U.S. Census statistics in terms of race/ethnicity (see Table 9.3), PEL (see Table 9.4), and geographic region distribution (see Table 9.5).

**Table 9.3. Race/Ethnicity Distribution:  
BIMAS–P Standard Normative Sample**

Race/Ethnicity of the Rated Youth	Normative Sample (Weighted)		U.S. Census
	N	%	%
Asian	30	2.2	3.8
African American	214	15.3	15.7
Hispanic	207	14.8	15.1
White	873	62.4	61.9
Other	75	5.4	3.5
<b>Total</b>	<b>1,400</b>	<b>100.0</b>	<b>100.0</b>

**Table 9.4. Parental Education Level Distribution:  
BIMAS–P Standard Normative Sample**

Parental Education Level of the Rated Youth	Normative Sample (Weighted)		U.S. Census
	N	%	%
High school diploma or lower	646	46.2	46.6
Apprenticeship/ Vocational training/ 2-year college/ some university	385	27.5	27.2
4-year college/ university or higher	369	26.4	26.2
<b>Total</b>	<b>1,400</b>	<b>100.0</b>	<b>100.0</b>

**Table 9.5. Geographic Region Distribution:  
BIMAS–P Standard Normative Sample**

U.S. Region of the Rated Youth	Normative Sample (Weighted)		U.S. Census
	<i>N</i>	%	%
Northeast	272	19.4	18.1
Midwest	265	18.9	21.9
West	333	23.8	23.3
South	530	37.9	36.7
<b>Total</b>	<b>1,400</b>	<b>100.0</b>	<b>100.0</b>

## Self-Report Normative Sample (BIMAS–SR)

The BIMAS–SR normative sample consisted of 700 youth aged 12 to 18 years old (350 males, 350 females, 100 youth in each age group by year). Table 9.6 describes the sample’s racial/ethnic distribution, which very closely approximated the U.S. Census. A similar weighting statistical procedure was applied to the BIMAS–SR sample so that the normative sample regional representation would be a close match to U.S. Census data (see Table 9.7).

**Table 9.6. Race/Ethnicity Distribution:  
BIMAS–SR Standard Normative Sample**

Race/Ethnicity	Normative Sample (Weighted)		U.S. Census
	<i>N</i>	%	%
Asian	28	4.0	3.8
African American	110	15.6	15.7
Hispanic	107	15.2	15.1
White	433	61.6	61.9
Other	25	3.5	3.5
<b>Total</b>	<b>703</b>	<b>100.0</b>	<b>100.0</b>

**Table 9.7. Geographic Region Distribution:  
BIMAS–SR Standard Normative Sample**

U.S. Region	Normative Sample (Weighted)		U.S. Census
	<i>N</i>	%	%
Northeast	128	18.3	18.1
Midwest	159	22.6	21.9
West	157	22.4	23.3
South	259	36.8	36.7
<b>Total</b>	<b>703</b>	<b>100.0</b>	<b>100.0</b>

## Norming Procedures and Derivation of Standardized Scores

Following the gathering of teacher, parent, and self-report normative data, the raw score means, standard deviations, and frequency distribution statistics for each of the BIMAS scales were analyzed on the weighted normative samples. The follow-up analyses included an examination of the sample’s performance across the youths’ chronological age and gender<sup>2</sup>. Multivariate Analyses of Variance (MANOVAs) were employed to examine the relationships between gender and age with the BIMAS raw scale scores (see appendix G for results). Results indicated significant main effects for age with small to moderate effect sizes. On the teacher and parent reports, in general, scores on the Behavioral Concern scales decreased slightly from age 5 to age 12, and then increased again until age 18, while the reverse pattern was true for the Adaptive scales. On the self-report, scores on the Behavioral Concern scales tend to increase slightly from age 12 to age 15, and decreased again until age 18; the reverse pattern was observed on the Adaptive scales. Results also indicated some significant (though very small) gender effects. In general, where significant differences were found, boys were rated slightly higher than females on the Behavioral Concern scales, while females were rated slightly higher than males on the Adaptive scales. The main effects were qualified by significant Age × Gender interactions for the majority of the teacher and parent scales. Decomposition of these interactions generally indicated that the gender effects were significant only at certain ages. These results informed the construction of norms broken into six age groups for the BIMAS–T and BIMAS–P (i.e., ages 5–6, 7–9, 10–11, 12–13, 14–16, and 17–18), and three age groups for the BIMAS–SR (i.e., 12–13, 14–16, and 17–18).

The gender differences that were found are consistent with the differences that are reported in the literature (e.g., higher externalizing type of behaviors for males than females). To reflect these differences, the default setting in the BIMAS scoring and reporting software uses norms that are based on a sample comprising males and females (i.e., combined gender norms). Although the combined sample norms reflect the actual differences in the general population, some users might prefer to use gender-specific norms for certain settings and purposes. Therefore, gender-specific norms were also calculated and are available as a scoring option.

<sup>2</sup> Analyses were also conducted to examine the sample’s performance across the youth’s race/ethnicity. Results revealed that there were few meaningful differences between scores on the BIMAS across the races/ethnicities (see appendix F for details).

In the process of developing the BIMAS norms, a cumulative frequency distribution of raw scores was developed for each of the five BIMAS scales across the various age groups. As expected in the case of behavior rating scales, these distributions were skewed. Thorndike (1982) states that in general, “the tendency to psychopathology may not be normal in shape.” Percentiles were computed for BIMAS raw scores to retain the actual shape of the original distribution. Data points that diverged significantly from a smooth curve partly reflect true differences and partly reflect sampling variability (Zachary & Gorsuch, 1985). To mitigate the effect of sampling variability, smoothed percentiles were also obtained using regression analysis. At each age, the predicted percentile score from the regression was used in conjunction with the original (unsmoothed) percentiles score to produce the final set of percentile scores. Specifically, the final “smoothed” percentiles were derived from the original, unsmoothed percentiles (given a 70% weighting) and a regression generated value (given a 30% weighting). Use of this smoothed normative value allows for irregular but real differences between age groups to have an effect, while reducing the impact of random fluctuation. The final smoothed percentile scores were then converted to standard *T*-scores with a mean of 50 and standard deviation of 10. For a discussion of non-linear *T*-score transformations, and smoothing procedures, the reader is referred to educational and psychological measurement textbooks (Urbina, 2004; Crocker & Algina, 1986).

# 10 Reliability

Reliability is defined as “the consistency of scores obtained by the same person when re-examined with the same test on different occasions, or with different sets of equivalent items, or under other variable examining conditions” (Anastasi, 1988, p. 102). This chapter presents results of reliability analyses of the Behavior Intervention Monitoring Assessment System (BIMAS™) Standard<sup>1</sup> (i.e., the BIMAS–Teacher [BIMAS–T], BIMAS–Parent [BIMAS–P], and BIMAS–Self-Report [BIMAS–SR]), including the results of internal consistency, standard error of measurement (*SEM*), test-retest reliability, and consistency between raters analyses.

## Overview of Results

Results of the reliability analyses revealed that the BIMAS forms have good levels of internal consistency, with Cronbach’s alpha values from the total sample ranging from .81 to .91 on the BIMAS–T, from .77 to .90 on the BIMAS–P, and from .75 to .88 on the BIMAS–SR. Good levels of temporal stability (test-retest reliability) were found when the BIMAS was taken twice within a 2- to 4-week period (without any intervention), with correlation coefficients (Pearson’s *r*) ranging from .85 to .91 on the BIMAS–T, from .79 to .96 on the BIMAS–P, and from .81 to .90 on the BIMAS–SR (all *r*s significant,  $p < .001$ ). A good level of consistency between raters (i.e., teacher and self-report; parent and self-report; teacher and parent) was found on ratings of the same child with Pearson’s *r* ranging from .54 to .86 across all scales (all *r*s significant,  $p < .001$ ).

## Internal Consistency

One measure of a test’s reliability is internal consistency, which is assessed using Cronbach’s alpha (Cronbach, 1951). Cronbach’s alpha ranges from 0.0 to 1.0, and is a function of the following parameters: “(a) the interrelatedness of the items in a test or scale, and (b) the length of the test” (John & Benet-Martinez, 2000, p. 343). There is no one level or cut-off of Cronbach’s alpha that universally denotes satisfactory reliability. Adequate alpha depends on how many items are on the scale (the greater the number of items, the higher alpha tends to be), the scale’s purpose, and the construct being measured. Internal consistency estimates (based on Cronbach’s alpha) for the BIMAS were computed on a weighted sample comprising 85% normative cases and 15% clinical cases. For the internal consistency analyses, clinical cases were added to the normative sample to ensure an adequate level of variability in the data and to reflect real-world applications where there is a mix of youth with and without clinical diagnoses. Tables 10.1 to 10.3 present Cronbach’s alpha values for this final weighted sample. The BIMAS–T, BIMAS–P, and BIMAS–SR were all found to demonstrate high levels of internal consistency for the majority of the scales. Specifically, for the total samples, BIMAS–T alpha values ranged from .81 to .91, BIMAS–P values ranged from .77 to .90, and BIMAS–SR values ranged from .75 to .88. With few exceptions, adequate to excellent levels of internal consistency were found across age and gender groups.

---

<sup>1</sup> Since only the norm-referenced standard form is discussed in this chapter, “BIMAS” is used to denote the BIMAS Standard throughout the chapter.

**Table 10.1. Cronbach's Alpha: BIMAS–T Standard**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
<b>Total Sample</b>		<b>.91</b>	<b>.85</b>	<b>.91</b>	<b>.85</b>	<b>.81</b>
<b>Combined Gender</b>	5–6	.85	.79	.91	.76	.66
	7–9	.81	.75	.87	.81	.70
	10–11	.86	.84	.91	.84	.80
	12–13	.92	.87	.93	.89	.86
	14–16	.93	.87	.92	.88	.81
	17–18	.93	.86	.89	.88	.83
<b>Male</b>	5–6	.87	.75	.92	.83	.58
	7–9	.85	.76	.88	.83	.71
	10–11	.87	.82	.90	.83	.71
	12–13	.93	.91	.93	.89	.88
	14–16	.94	.89	.92	.89	.76
	17–18	.94	.87	.86	.86	.81
<b>Female</b>	5–6	.81	.84	.87	.58	.72
	7–9	.65	.74	.83	.79	.69
	10–11	.84	.86	.91	.84	.87
	12–13	.92	.81	.92	.88	.84
	14–16	.92	.84	.93	.86	.86
	17–18	.93	.86	.92	.90	.84

**Table 10.2. Cronbach's Alpha: BIMAS–P Standard**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
<b>Total Sample</b>		<b>.87</b>	<b>.82</b>	<b>.90</b>	<b>.84</b>	<b>.77</b>
<b>Combined Gender</b>	5–6	.79*	.69	.90	.78	.49
	7–9	.83	.77	.87	.76	.61
	10–11	.82	.80	.90	.75	.71
	12–13	.89	.85	.93	.90	.82
	14–16	.90	.86	.89	.86	.84
	17–18	.86	.82	.88	.88	.81
<b>Male</b>	5–6	.78*	.59	.91	.82	.49
	7–9	.82	.81	.87	.72	.66
	10–11	.86	.85	.90	.71	.61
	12–13	.88	.87	.92	.88	.81
	14–16	.90	.86	.90	.81	.83
	17–18	.87	.81	.85	.89	.78
<b>Female</b>	5–6	.80†	.76	.87	.69	.47
	7–9	.82	.74	.86	.77	.55
	10–11	.73*	.67	.88	.82	.81
	12–13	.90	.82	.94	.91	.83
	14–16	.91	.86	.86	.91	.84
	17–18	.87	.83	.91	.88	.85

Note. \*Calculated with item 32 (was suspected of smoking or chewing tobacco) taken out due to null variance. †Calculated with item 29 (was suspected of using alcohol and/or drugs) taken out due to null variance.

**Table 10.3. Cronbach's Alpha: BIMAS–SR Standard**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
<b>Total Sample</b>		<b>.88</b>	<b>.85</b>	<b>.87</b>	<b>.83</b>	<b>.75</b>
<b>Combined Gender</b>	12–13	.87	.85	.85	.79	.70
	14–16	.88	.85	.87	.83	.75
	17–18	.89	.85	.87	.86	.78
<b>Male</b>	12–13	.88	.83	.84	.79	.71
	14–16	.86	.83	.86	.81	.77
	17–18	.89	.83	.86	.87	.74
<b>Female</b>	12–13	.84	.86	.87	.80	.67
	14–16	.90	.86	.88	.84	.72
	17–18	.90	.86	.88	.84	.80

## Standard Error of Measurement

All measurements contain some error, which can be estimated with the standard error of measurement (*SEM*). To obtain a basic understanding of *SEM*, consider an individual's obtained score on the BIMAS as a reflection of the individual's "true" score on the test. Numerous factors may cause the obtained score to differ from, and fail to match exactly, with the individual's true BIMAS score. *SEM* provides an estimate of how much an individual's obtained score might vary from his/her true score. *SEM* is an estimate of the amount of error in the obtained scores. *SEM* values were calculated for each of the BIMAS scale *T*-scores using the *internal consistency* reliability estimates that were derived from the weighted normative and clinical samples (see *Internal Consistency* in this chapter). Tables 10.4 to 10.6 present *SEM* values for the BIMAS *T*-scores (see appendix H for *SEM* values for the BIMAS raw scores).

The *T*-score *SEM* values have been built into Confidence Intervals for the BIMAS *T*-scores (see *Confidence Intervals* in chapter 5, *Understanding and Interpreting BIMAS Scores*, for a description of how Confidence Intervals can be used in the interpretation process). Confidence Intervals surrounding the obtained *T*-score for every BIMAS scale are provided as an option in the computerized reports.

The Confidence Intervals were calculated by obtaining: (1) the standard error of measurement (*SEM*; see formula 1), (2) the Lower Bound of the Confidence Interval (see formula 2), and (3) the Upper Bound of the Confidence Interval (see formula 3).

### Formula 1.

$$SEM = s_x \sqrt{1 - r_x}$$

where  $s_x$  = the standard deviation of the *T*-scores earned by the weighted sample, and reliability = internal consistency estimate (Cronbach's alpha).

### Formula 2.

$$Lower\ Bound = Obtained\ Score - (z \times SEM)$$

where  $z = 1.64$  for the 90% level of confidence, and  $z = 1.96$  for the 95% level of confidence.

### Formula 3.

$$Upper\ Bound = Obtained\ Score + (z \times SEM)$$

where  $z = 1.64$  for the 90% level of confidence, and  $z = 1.96$  for the 95% level of confidence.

The *SEM* values were also used to determine the values needed for significance when comparing BIMAS results between raters (see *Statistically Significant Differences between Raters* in chapter 5, *Understanding and Interpreting BIMAS Scores*, for more information). For further information on the use and interpretation of the *SEM*, refer to McDonald (1999).

**Table 10.4. Standard Error of Measurement: BIMAS–T Standard T-Scores**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
Total Sample		2.58	3.94	3.42	4.20	4.36
Combined Gender	5–6	2.88	3.83	3.13	5.07	5.54
	7–9	3.47	4.30	3.69	4.34	5.29
	10–11	2.79	3.65	2.97	4.20	4.32
	12–13	2.27	3.76	3.23	3.78	3.70
	14–16	2.73	4.22	3.30	3.95	4.52
	17–18	2.49	4.07	3.92	3.81	4.02
Male	5–6	3.03	4.26	2.92	4.41	6.38
	7–9	3.10	4.19	3.59	4.16	5.33
	10–11	3.10	4.05	3.35	4.13	5.28
	12–13	2.32	3.27	3.32	3.89	3.45
	14–16	2.37	4.29	3.39	3.69	5.27
	17–18	2.36	3.83	4.23	4.28	4.06
Female	5–6	2.60	3.08	3.54	6.17	4.66
	7–9	4.32	4.64	3.99	4.47	5.12
	10–11	2.55	3.26	2.68	3.93	3.38
	12–13	2.21	4.33	3.10	3.82	3.92
	14–16	3.04	4.59	3.48	4.32	3.78
	17–18	2.87	4.26	3.39	3.46	4.14

**Table 10.5. Standard Error of Measurement: BIMAS–P Standard T-Scores**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/ Attention	Social	Academic Functioning
Total Sample		3.41	4.28	3.33	4.05	4.48
Combined Gender	5–6	4.02	5.22	3.46	4.53	6.41
	7–9	3.62	4.55	4.01	4.59	5.68
	10–11	3.73	4.37	3.30	4.93	4.97
	12–13	2.98	3.95	2.56	3.41	3.99
	14–16	2.99	3.82	3.32	3.92	3.58
Male	17–18	4.08	4.47	3.38	3.76	4.61
	5–6	4.14	5.87	3.37	4.27	6.54
	7–9	4.00	4.40	4.29	5.17	5.76
	10–11	3.39	3.80	3.25	4.99	5.93
	12–13	3.22	3.75	2.91	3.85	4.14
Female	14–16	3.14	3.78	3.04	4.72	3.82
	17–18	4.04	4.64	4.17	3.62	4.95
	5–6	3.79	4.34	3.74	5.08	6.25
	7–9	3.39	4.77	3.82	4.23	5.77
	10–11	4.33	5.30	3.47	4.12	3.94
	12–13	2.67	4.56	2.40	3.05	4.12
	14–16	2.88	3.93	3.86	3.06	3.75
	17–18	4.10	4.45	3.34	3.86	4.13

**Table 10.6. Standard Error of Measurement: BIMAS–SR Standard T-Scores**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/ Attention	Social	Academic Functioning
Total Sample		3.41	4.28	3.33	4.05	4.48
Combined Gender	12–13	3.28	3.94	4.02	4.62	4.88
	14–16	3.36	4.05	3.63	4.61	4.90
	17–18	3.33	4.07	3.76	4.13	4.16
Male	12–13	3.25	4.13	4.32	4.60	5.30
	14–16	3.58	4.36	3.58	4.45	4.67
	17–18	3.37	4.15	3.82	3.77	4.96
Female	12–13	3.08	4.16	3.76	3.95	5.00
	14–16	3.20	3.76	3.57	4.60	5.38
	17–18	3.29	3.90	3.69	4.80	3.59

## Test-Retest Reliability and Standard Error of Prediction

Test-retest reliability refers to the correlation of scores obtained from two separate administrations for the same youth by the same rater over a specified period of time. This type of reliability was assessed over a 2- to 4-week interval by obtaining *T*-score correlations for the BIMAS with a sample of 112 teachers, 83 parents, and 53 youth who completed the BIMAS twice (no interventions took place between the Time 1 and Time 2 administrations; see Table 10.7 for demographic characteristics of the test-retest samples). The correlations, as well as the means and standard deviations from Time 1 and Time 2 administrations, are provided in Tables 10.8 to 10.10. Across the three forms, *r* ranged from .79 to .96 (all *p* < .001), indicating that the BIMAS has high test-retest reliability.

These test-retest values were then incorporated into the calculation of the standard error of prediction with the following formula:

$$\text{Standard Error of Prediction} = S_x \sqrt{1 - r_{xy}}$$

where  $S_x$  = the standard deviation of the normative sample, and  $r_{xy}$  = test-retest reliability.

The standard error of prediction pertains to outcome assessment and is a method of determining how much scores may be expected to fluctuate over time due to random error when no intervention occurs between administrations. Table 10.11 presents standard error of prediction values for the BIMAS T-scores. These values provide an effective way to assess change over time (see *Scores for Progress and Outcome Monitoring* in chapter 5, *Understanding and Interpreting BIMAS Scores*).

**Table 10.7. Demographic Characteristics of the BIMAS Standard Test-Retest Reliability Samples**

Demographic Characteristic of the Rated Youth		Teacher		Parent		Self-Report	
		<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Population	Non-Clinical	112	100.0	83	100.0	53	100.0
	Clinical	0	0.0	0	0.0	0	0.0
Gender	Male	58	51.8	35	42.2	30	56.6
	Female	24	41.2	48	57.8	23	43.4
Race/Ethnicity	Asian	0	0.0	2	2.4	0	0.0
	African American	26	23.2	15	18.1	7	13.2
	Hispanic	17	15.2	15	18.1	20	37.7
	White	59	52.7	49	59.0	25	47.2
	Other	10	8.9	2	2.4	1	1.9
<b>Total</b>		<b>112</b>	<b>100.0</b>	<b>83</b>	<b>100.0</b>	<b>53</b>	<b>100.0</b>
Age <i>M (SD)</i>		7.14 (3.4)		11.5 (3.7)		15.4 (1.9)	
Days between administrations <i>M (SD)</i>		21.6 (6.5)		21.9 (7.2)		20.2 (7.9)	

**Table 10.8. Test-Retest Reliability Coefficients: BIMAS–T Standard T-scores**

Scale		<i>r</i>	<i>N</i>	Time 1		Time 2	
				<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Behavioral Concern Scales	Conduct	.89	112	54.1	9.1	54.4	9.1
	Negative Affect	.85	112	54.6	10.0	54.0	9.8
	Cognitive/Attention	.91	112	51.8	13.0	51.7	14.0
Adaptive Scales	Social	.91	112	46.1	13.0	46.3	13.0
	Academic Functioning	.91	108	49.8	12.0	49.0	13.0

Note. All *r*s significant,  $p < .001$ . Sample sizes vary due to missing data.

**Table 10.9. Test-Retest Reliability Coefficients: BIMAS–P Standard T-scores**

Scale		<i>r</i>	<i>N</i>	Time 1		Time 2	
				<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Behavioral Concern Scales	Conduct	.79	82	49.6	7.9	49.6	7.4
	Negative Affect	.91	79	54.4	12.0	53.9	11.0
	Cognitive/Attention	.84	83	52.1	10.0	52.2	11.0
Adaptive Scales	Social	.96	79	43.7	16.0	44.1	16.0
	Academic Functioning	.80	83	52.3	8.4	53.2	8.1

Note. All *r*s significant,  $p < .001$ . Sample sizes vary due to missing data.

**Table 10.10. Test-Retest Reliability Coefficients: BIMAS–SR Standard T-scores**

Scale		<i>r</i>	<i>N</i>	Time 1		Time 2	
				<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Behavioral Concern Scales	Conduct	.81	52	45.8	7.0	45.6	7.5
	Negative Affect	.87	52	45.4	9.6	46.2	10.8
	Cognitive/Attention	.82	52	44.7	9.8	44.7	9.4
Adaptive Scales	Social	.90	52	54.5	10.6	53.6	12.1
	Academic Functioning	.85	49	52.9	8.4	52.3	9.2

Note. All *r*s significant,  $p < .001$ . Sample sizes vary due to missing data.

**Table 10.11. Standard Error of Prediction Coefficients: BIMAS Standard T-scores**

Scale		Rater		
		Teacher	Parent	Self-Report
Behavioral Concern Scales	Conduct	3.25	4.64	4.39
	Negative Affect	3.86	3.05	3.58
	Cognitive/Attention	2.94	4.00	4.22
Adaptive Scales	Social	3.01	2.10	3.17
	Academic Functioning	3.03	4.53	3.85

## Consistency between Raters

Because the BIMAS–T, BIMAS–P, and BIMAS–SR measure similar constructs, a degree of consistency is expected across rater types. The level of consistency was assessed by correlating scores from the different raters. Although *some* degree of similarity is expected between raters, it is nonetheless expected that a certain degree of incongruence will exist (i.e., the correlations should be moderate in size). This incongruence occurs because the various raters may have different opinions about, and different experiences with, the

youth's behavior. This incongruence can also occur because the raters see the youth in different contexts. A sample of 162 youth who provided self-report ratings were also rated by a teacher and a parent (see Table 10.12 for a sample description). Correlation coefficients (Pearson's  $r$ ) were calculated between each pair of raters (see Tables 10.13 to 10.15). As anticipated, the correlations were found to be moderate in size (all  $p < .001$ ) for the teacher to self-report comparisons ( $r = .54$  to  $.69$ ) and the parent to self-report comparisons ( $r = .59$  to  $.69$ ). A great deal of consistency was found between teacher and parent ratings, with strong correlations between the two rater types on all scales ( $r = .79$  to  $.86$ ).

**Table 10.12. Demographic Characteristics of the BIMAS Standard Consistency Between Raters Sample**

Demographic Characteristic of the Rated Youth	Group	<i>N</i>	%
Population	Non-Clinical	70	43.2
	Clinical	92	56.8
Gender	Male	80	49.4
	Female	82	50.6
Race/Ethnicity	Asian	2	1.2
	African American	23	14.2
	Hispanic	22	13.6
	White	101	62.3
	Other	14	8.6
<b>Total</b>		<b>162</b>	<b>100.0</b>
<b>Age <i>M</i> (<i>SD</i>)</b>		<b>13.9 (1.8)</b>	

**Table 10.13. Consistency Between Rater *T*-scores: Teacher to Self-Report Ratings**

Scale		<i>r</i>	Teacher		Self-Report	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Behavioral Concern Scales	Conduct	.54	56.5	9.7	49.2	10.1
	Negative Affect	.64	59.0	12.4	52.2	13.4
	Cognitive/Attention	.69	54.7	12.9	48.6	10.0
Adaptive Scales	Social	.59	41.0	11.8	45.0	8.0
	Academic Functioning	.59	47.8	11.6	49.4	8.7

Note.  $N = 162$ . All  $r$ s significant,  $p < .001$ .

**Table 10.14. Consistency Between Rater *T*-scores: Parent to Self-Report Ratings**

Scale		<i>r</i>	Parent		Self-Report	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Behavioral Concern Scales	Conduct	.62	54.4	11.2	49.2	10.1
	Negative Affect	.69	56.8	13.3	52.2	13.4
	Cognitive/Attention	.67	53.8	10.4	48.6	10.0
Adaptive Scales	Social	.59	44.2	10.4	45.0	8.0
	Academic Functioning	.65	46.1	9.9	49.4	8.7

Note.  $N = 162$ . All  $r$ s significant,  $p < .001$ .

**Table 10.15. Consistency Between Rater *T*-scores: Teacher to Parent Ratings**

Scale		<i>r</i>	Teacher		Parent	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Behavioral Concern Scales	Conduct	.82	56.5	9.7	54.4	11.2
	Negative Affect	.86	59.0	12.4	56.8	13.3
	Cognitive/Attention	.84	54.7	12.9	53.8	10.4
Adaptive Scales	Social	.79	41.0	11.8	44.2	10.4
	Academic Functioning	.80	47.8	11.6	46.1	9.9

Note.  $N = 162$ . All  $r$ s significant,  $p < .001$ .

# 11 Validity

The validity of a test refers to the quality of inferences that can be made by the test's scores. That is, how well does the test measure the construct(s) it was designed to measure, and how well are the claims regarding its use and applications supported by empirical evidence? This chapter provides evidence for the validity of the norm-referenced versions of the Behavior Intervention Monitoring Assessment System (BIMAS™) Standard<sup>1</sup> (i.e., BIMAS–Teacher [BIMAS–T], the BIMAS–Parent [BIMAS–P], and the BIMAS–Self-Report [BIMAS–SR]).

Evidence that the BIMAS measures the constructs it was designed to measure include:

- a description of the content validity of the assessments, as well as the appropriateness of the scale structure
- an exploration of the relationship between BIMAS scores and scores from other instruments measuring similar constructs

In terms of the use and applications of the BIMAS, two claims are made. Within a 3-tier Response to Intervention (RTI) model of behavioral health service delivery, the BIMAS can be effectively used:

- in Tier 1 as a screening measure to identify students who might need further attention and support, and as an outcome measure to evaluate (through universal screenings) the effectiveness of existing programs
- in Tiers 2 and 3 as a progress monitoring tool

Validation support is provided for each of these uses.

## Overview of Results

The content validity and appropriateness of the BIMAS scale structure was determined based on the validity of the item to scale assignment, as well as on adequate confirmatory factor analysis fit indicators. Further evidence that the BIMAS measures the constructs it was intended to measure was assessed by correlating BIMAS scores with scores from the Conners Comprehensive Behavior Rating Scales™ (CBRS™; Conners, 2008). Scores that assess similar constructs were

moderately to strongly correlated (all  $p < .001$ ), with the following ranges: Conduct  $r_s = .49$  to  $.78$ ; Negative Affect  $r_s = .38$  to  $.70$ , Cognitive/Attention  $r_s = .46$  to  $.69$ , Social  $r_s = -.47$  to  $-.71$ , Academic Functioning  $r_s = -.31$  to  $-.49$  (the negative correlations for the Social and Academic Functioning scales are in the expected direction). Results of discriminant function analyses (predicting clinical versus non-clinical group membership) provided evidence that the BIMAS can effectively be used as a screening measure to identify students who might need further attention and support, with good overall correct classification rates (BIMAS–T: 82.5% to 85.2%; BIMAS–P: 78.3% to 78.6%; and BIMAS–SR: 71.5% to 71.8%). Results from a progress monitoring study revealed that the BIMAS can effectively be used as a progress monitoring tool. Significant differences in pre- to post-treatment scores were found, with predominantly large effects sizes (Cohen's  $d$ ). From pre- to post-treatment, BIMAS scores were able to capture significant decreases in the Behavioral Concern scale scores with large effect sizes (Conduct  $d_s = 1.5$  to  $2.8$ ; Negative Affect  $d_s = 1.0$  to  $1.8$ ; Cognitive/Attention  $d_s = 1.2$  to  $2.4$ ), as well as significant increases in the Adaptive Skills scale scores, again, with large effect sizes (Social  $d_s = -0.7$  to  $-1.0$ ; Academic Functioning  $d_s = -0.8$  to  $-1.8$ ).

## Content Validity and BIMAS Scale Structure

Often, evidence of the validity of any test begins with its theoretical framework and the procedures used to formulate its structure and create its content (i.e., items, scales, and constructs). The content validity evidence of the BIMAS has an extensive history. It began with early studies by Meier (1997, 1998, 2000, 2004) that included the examination of the relationships of specific items to external criteria (see Meier, 2000), further item/content reviews by the authors, and feedback by colleagues working in public schools and community mental health centers. These efforts resulted in a pool of items that were pilot tested in small scale studies. Following these earlier studies (e.g., Lerew, 2004) and reviews of the literature, a set of items were developed and proposed to represent behaviors that can be classified into: (a) externalizing behaviors; (b) internalizing behaviors; (c)

<sup>1</sup> Since only the norm-referenced Standard form is discussed in this chapter, "BIMAS" is used to denote the BIMAS Standard throughout the chapter.

behaviors that are related to attentive skills; and (d) items related to adaptive skills that can be further divided into areas of socialization and academic performance and/or functioning. Externalizing behaviors are related to problems that are manifested in outward behavior and reflect a child's negative reactions to his or her environment, which may include aggression, delinquency, and hyperactivity. Internalizing behaviors involve problems that are manifested in an inward fashion, typically of an inhibited style that could be described as withdrawn, lonely, depressed, and anxious. See Table 11.1 for the item and scale structure of the BIMAS (see chapter 8, *Development*, for a detailed discussion on early studies of the BIMAS items).

As described above, the scale structure and item assignment detailed in Table 11.1 was arrived at through the use of clinical judgment and the results of empirical studies (Meier, 1997; 1998; 2000; 2004). In order to provide empirical support for the scale structure, a series of confirmatory factor analyses (CFAs) was conducted on the normative and clinical samples (see chapter 9, *Standardization* for a description of the normative sample; see *The BIMAS Standard as a Screening Tool* later in this chapter for a description of the clinical samples). The CFAs were conducted with Maximum Likelihood, generalized least squares estimation. Items were grouped into "parcels" on the basis of intercorrelations; items that correlated most highly with one another were assigned to the same parcel (see Cattell & Burdsal, 1975, as well as Hughey & Burdsal, 1982, for more information on the parceling approach to factor analysis). The item parcels were loaded onto their respective scales, and the

scales were allowed to correlate with each other. As recommended by Hu and Bentler (1999) as well as Cole (1987), and March, Herbert, Balla, and McDonald (1988), multiple criteria were used to assess the fit of the model. Adequate fit of the model was defined as meeting minimal criteria for several fit indices, including the Normed Fit Index (NFI; Bentler & Bonett, 1980), the Non-Normed Fit Index (NNFI; Bentler & Bonett, 1980), the Comparative Fit Index (CFI; Bentler, 1990), and the Steiger-Lind Root Mean Square Error of Approximation Index (RMSEA; Steiger & Lind, 1980, as cited in Browne & Cudeck, 1993). Adequate fit of the model was defined as  $NFI > .90$ ,  $NNFI > .90$ ,  $CFI > .90$ , and  $RMSEA < .10$ . Results indicated that the models approached adequate fit across all three forms. BIMAS-T fit indices were  $NFI = .91$ ,  $NNFI = .87$ ,  $CFI = .91$ ,  $RMSEA = .13$ . BIMAS-P fit indices were  $NFI = .89$ ,  $NNFI = .85$ ,  $CFI = .90$ ,  $RMSEA = .13$ . BIMAS-SR fit indices were  $NFI = .92$ ,  $NNFI = .90$ ,  $CFI = .93$ ,  $RMSEA = .10$ .

The final step in determining the appropriateness of the BIMAS scale structure was to intercorrelate all of the BIMAS scales (with data from the normative sample). This procedure was undertaken in order to provide evidence of the multidimensionality of the BIMAS; theoretical expectations would be met if all of the correlations were significant (indicating that the constructs are related), but not high enough to suggest redundancy in the scales. Tables 11.2 to 11.4 display the results. For all three rater-types, theoretical expectations were met, with significant, but moderate, correlations being found.

Table 11.1. Overview of BIMAS Standard Item Content and Scale Assignment on the BIMAS Standard

Item	Item Wording		Behavioral Concern Scales			Adaptive Scales	
	Teacher/Parent	Self-Report	Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
1	shared what he/she was thinking about.	shared my thoughts with others.				✓	
2	appeared angry.	felt angry.	✓				
3	had trouble paying attention.	had trouble paying attention.			✓		
4	followed directions.	followed directions.					✓
5	appeared sleepy or tired.	felt sleepy or tired.		✓			
6	was impulsive.	was impulsive.			✓		
7	spoke clearly with others.	communicated clearly with others.				✓	
8	appeared depressed.	was depressed.		✓			
9	engaged in risk-taking behavior.	did something risky.	✓				
10	had problems staying on task.	had problems staying on task.			✓		
11	maintained friendships.	maintained friendships.				✓	
12	acted sad or withdrawn.	was sad or withdrawn.		✓			
13	fought with others (verbally, physically, or both).	fought with others (verbally, physically, or both).	✓				
14	acted without thinking.	acted without thinking.			✓		
15	appeared comfortable when relating to others.	felt relaxed interacting with others.				✓	
16	was easily embarrassed or felt ashamed.	was easily embarrassed or felt ashamed.		✓			
17	lied or cheated.	lied or cheated.	✓				
18	had trouble remembering.	had trouble remembering things.			✓		
19	was generally friendly with others.	was friendly with others.				✓	
20	appeared anxious (worried or nervous).	was anxious (worried or nervous).		✓			
21	lost his/her temper when upset.	lost my temper when I was upset.	✓				
22	had trouble with organizing and planning.	had trouble with organizing and planning.			✓		
23	worked out problems with others.	worked out problems with others.				✓	
24	expressed thoughts of hurting himself/herself.	had thoughts of hurting myself.		✓			
25	was aggressive (threatened or bullied others).	threatened or bullied others.	✓				
26	received failing grades at school.	received failing grades at school.					✓
27	was emotional or upset.	felt emotional or upset.		✓			
28	idgeted.	idgeted.			✓		
29	was suspected of using alcohol and/or drugs.	used alcohol and/or drugs.	✓				
30	worked up to his/her academic potential.	tried my hardest when it came to schoolwork.					✓
31	was sent to an authority for discipline.	was sent to an authority for discipline.	✓				
32	was suspected of smoking or chewing tobacco.	smoked or chewed tobacco.	✓				
33	was prepared for class.	went prepared to class.					✓
34	was absent from school.	was absent from school.					✓

**Table 11.2. BIMAS–T Standard Scale Intercorrelations**

BIMAS–T Scale		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
Behavioral Concern Scales	Conduct	–				
	Negative Affect	.55	–			
	Cognitive/Attention	.52	.49	–		
Adaptive Scales	Social	–.36	–.36	–.25	–	
	Academic Functioning	–.46	–.44	–.65	.41	–

Note. Listwise deletion of missing data was used. *N* = 1,361. All correlations significant at *p* < .01 (2-tailed).

**Table 11.3. BIMAS–P Standard Scale Intercorrelations**

BIMAS–P Scale		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
Behavioral Concern Scales	Conduct	–				
	Negative Affect	.66	–			
	Cognitive/Attention	.65	.56	–		
Adaptive Scales	Social	–.41	–.45	–.35	–	
	Academic Functioning	–.55	–.49	–.58	.52	–

Note. Listwise deletion of missing data was used. *N* = 1,396. All correlations significant at *p* < .01 (2-tailed).

**Table 11.4. BIMAS–SR Standard Scale Intercorrelations**

BIMAS–SR Scale		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
Behavioral Concern Scales	Conduct	–				
	Negative Affect	.68	–			
	Cognitive/Attention	.71	.68	–		
Adaptive Scales	Social	–.47	–.40	–.35	–	
	Academic Functioning	–.70	–.53	–.58	.53	–

Note. Listwise deletion of missing data was used. *N* = 702. All correlations significant at *p* < .01 (2-tailed).

## Relationship Between the BIMAS and Other Measures

The validity of the BIMAS was assessed by examining its relationship with the Conners Comprehensive Behavior Rating Scales (Conners CBRS; Conners, 2008). The Conners CBRS “is a comprehensive assessment tool which assesses a wide range of behavioral, emotional, social, and academic concerns and disorders in children and adolescents” (p. 1). Ratings were obtained for three samples of non-clinical youth. Teachers (*N* = 112), parents (*N* = 127), and youth (*N* = 108) completed the respective BIMAS and Conners CBRS rating forms (i.e., Conners CBRS–Teacher [Conners CBRS–T]; Conners CBRS–Parent [Conners CBRS–P]; Conners CBRS–Self-Report [Conners CBRS–SR]). Table 11.5 describes the composition of this study’s sample in terms of the rated youth’s gender, age, and race/ethnicity.

The correlations between the BIMAS and the Conners CBRS are presented in Tables 11.6 through 11.10. The mean scores were similar across the two tests (all means were close to 50, as would be expected from a non-clinical sample). Furthermore, significant correlations (*p* < .01) that were moderate to

large in size were found for all relevant scale comparisons, providing strong evidence that the BIMAS is measuring the constructs it was designed to measure.

**Conduct Scale.** Scores on the BIMAS Conduct scale were compared to scores on two Conners CBRS scales that assess defiant, aggressive, and potentially violent behaviors (i.e., Defiance/Aggression and the Violence Potential Indicator), as well as the DSM-IV-TR Conduct Disorder and DSM-IV-TR Oppositional Defiant Disorder scales (i.e., scales that measure the DSM-IV-TR criteria for these disorders; see Table 11.6). Across the three rater-types, the scale correlations ranged from .49 to .78 (all *p* < .01), indicating a strong level of agreement between the BIMAS and Conners CBRS scales.

**Negative Affect Scale.** Scores on the BIMAS Negative Affect scale were compared to scores on two Conners CBRS scales that assess negative mood (i.e., the empirically derived Emotional Distress scale as well as the DSM-IV-TR Major Depressive Episode scale; see Table 11.7). The scale correlations were moderate to strong; across the three rater-types, correlations ranged from .38 to .70 (all *p* < .01), with the highest convergence occurring for the parent ratings.

**Cognitive/Attention Scale.** Scores on the BIMAS Cognitive/Attention scale were compared to scores on three Conners CBRS scales (see Table 11.8), including the empirically derived Hyperactivity/Impulsivity scale (Hyperactivity on the teacher form), as well as the two DSM-IV-TR ADHD scales (i.e., scales that measure the DSM-IV-TR criteria for ADHD Predominantly Inattentive Type and ADHD Predominantly Hyperactive-Impulsive Type). The scale correlations were moderate to strong; across the three rater-types, correlations ranged from .46 to .69 (all  $p < .01$ ), with the highest convergence occurring for the teacher ratings.

**Social Scale.** Scores on the BIMAS Social scale were compared to scores on the Conners CBRS DSM-IV-TR Autistic Disorder and Asperger's Disorder scales (see Table 11.9; note that these DSM-IV-TR scales are not available on the Conners CBRS-SR, therefore, no self-report comparisons are provided). Strong correlations were found for the

teacher ratings ( $r_s = -.69$  to  $-.71$ ; all  $p < .01$ ), and moderate correlations were found for the parent ratings ( $r_s = -.47$  to  $-.52$ ; all  $p < .01$ ). Note that the negative correlations are in the expected direction, because while higher scores on the BIMAS Social scale indicate better functioning, higher scores on the Conners CBRS scales indicate more problematic functioning.

**Academic Functioning Scale.** Scores on the BIMAS Academic Functioning scale were compared to scores on the Conners CBRS Academic Difficulties scale (as well as the Language and Math subscales that are available only on the teacher and parent versions of the Conners CBRS; see Table 11.10). Across the three rater types, moderate scale correlations were found, ranging from  $-.31$  to  $-.49$  (all  $p < .01$ ). As with the Social scale, the negative correlations are in the expected direction.

**Table 11.5. Demographic Characteristics of the BIMAS Standard and Conners CBRS Validity Study Samples**

Demographic Characteristic of the Rated Youth		Teacher		Parent		Self-Report	
		<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Population	Non-Clinical	112	100.0	127	100.0	108	100.0
	Clinical	0	0.0	0	0.0	0	0.0
Gender	Male	59	52.7	62	48.8	64	59.3
	Female	53	47.3	65	51.2	44	40.7
Race/Ethnicity	Asian	7	6.3	4	3.1	4	3.7
	African American	21	18.8	39	30.7	19	17.6
	Hispanic	5	4.5	14	11.0	27	25.0
	White	67	59.8	62	48.8	53	49.1
	Other	12	10.7	8	6.3	5	4.6
<b>Total</b>		<b>112</b>	<b>100.0</b>	<b>127</b>	<b>100.0</b>	<b>108</b>	<b>100.0</b>
Age <i>M</i> ( <i>SD</i> )		12.14 (2.87)		12.06 (3.15)		15.30 (2.02)	

**Table 11.6. Correlations Between the BIMAS Standard Conduct Scale *T*-scores and Relevant Conners CBRS Scales**

Conners CBRS Scale	<i>r</i>			Conners CBRS-T		Conners CBRS-P		Conners CBRS-SR	
	Teacher	Parent	Self-Report	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Defiant/Aggressive Behaviors	.52	.73	.64	51.1	12.0	51.7	13.7	53.3	14.6
Violence Potential Indicator	.54	.74	.64	50.1	10.1	52.7	15.0	52.4	12.8
DSM-IV-TR Conduct Disorder	.50	.68	.62	50.0	10.3	51.1	13.9	53.2	14.1
DSM-IV-TR Oppositional Defiant Disorder	.49	.78	.62	53.7	13.9	50.6	12.5	51.0	11.0
BIMAS Conduct Scale <i>M</i>	51.9	50.9	49.0						
BIMAS Conduct Scale <i>SD</i>	6.0	9.6	9.5						

Note. Teacher,  $N = 112$ ; Parent,  $N = 127$ ; Self-Report,  $N = 108$ . All correlations significant at  $p < .01$  (2-tailed).

**Table 11.7. Correlations Between the BIMAS Standard Negative Affect Scale *T*-scores and Relevant Conners CBRS Scales**

Conners CBRS Scale	<i>r</i>			Conners CBRS-T		Conners CBRS-P		Conners CBRS-SR	
	Teacher	Parent	Self-Report	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Emotional Distress	.47	.70	.54	48.7	9.4	50.6	13.7	51.0	10.6
DSM-IV-TR Major Depressive Episode	.38	.62	.56	49.8	9.1	50.0	13.1	52.0	11.5
BIMAS Negative Affect Scale <i>M</i>	48.5	49.2	50.8						
BIMAS Negative Affect Scale <i>SD</i>	7.9	10.1	10.2						

Note. Teacher,  $N = 112$ ; Parent,  $N = 127$ ; Self-Report,  $N = 108$ . All correlations significant at  $p < .01$  (2-tailed).

**Table 11.8. Correlations Between the BIMAS Standard Cognitive/Attention Scale T-scores and Relevant Conners CBRS Scales**

Conners CBRS Scale	<i>r</i>			Conners CBRS–T		Conners CBRS–P		Conners CBRS–SR	
	Teacher	Parent	Self-Report	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Hyperactivity/Impulsivity <sup>1</sup>	.63	.56	.46	53.0	13.8	50.3	12.0	53.0	11.9
DSM-IV-TR ADHD Inattentive	.69	.53	.53	51.8	12.9	50.8	11.6	51.9	9.6
DSM-IV-TR ADHD Hyperactive-Impulsive	.62	.56	.56	52.7	12.4	50.6	12.8	52.6	12.7
BIMAS Cognitive/Attention Scale <i>M</i>	49.9	47.5	49.3						
BIMAS Cognitive/Attention Scale <i>SD</i>	11.5	9.6	10.3						

<sup>1</sup> Hyperactivity on the Conners CBRS–T.

Note. Teacher, *N* = 112; Parent, *N* = 127; Self-Report, *N* = 108. All correlations significant at *p* < .01 (2-tailed).

**Table 11.9. Correlations Between the BIMAS Standard Social Scale T-scores and Relevant Conners CBRS Scales**

Conners CBRS Scale	<i>r</i>		Conners CBRS–T		Conners CBRS–P	
	Teacher	Parent	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
DSM-IV-TR Autistic Disorder	–.69	–.47	51.4	12.7	52.9	12.3
DSM-IV-TR Asperger’s Disorder	–.71	–.52	50	11.3	52.7	11.9
BIMAS Social Scale <i>M</i>	50.3	47.0				
BIMAS Social Scale <i>SD</i>	9.7	9.7				

Note. Teacher, *N* = 112; Parent, *N* = 127. All correlations significant at *p* < .01 (2-tailed). Self-Report comparisons are not made as the relevant scales are not available on the Conners CBRS–SR.

**Table 11.10. Correlations Between the BIMAS Standard Academic Functioning Scale T-scores and Relevant Conners CBRS Scales**

Conners CBRS Scale	<i>r</i>			Conners CBRS–T		Conners CBRS–P		Conners CBRS–SR	
	Teacher	Parent	Self-Report	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Academic Difficulties	–.49	–.40	–.49	48.9	9.1	50.4	12.5	50.6	10.5
Academic Difficulties: Language	–.44	–.31	n/a	48.9	8.7	49.9	11.5	n/a	n/a
Academic Difficulties: Math	–.39	–.35	n/a	48.5	7.6	52.2	13.5	n/a	n/a
BIMAS Academic Functioning Scale <i>M</i>	47.2	48.7	50.4						
BIMAS Academic Functioning Scale <i>SD</i>	9.1	9.2	9.3						

Note. Teacher, *N* = 112; Parent, *N* = 127; Self-Report, *N* = 108. All correlations significant at *p* < .01 (2-tailed). n/a = not available (the Language and Math subscales are not available on the Conners CBRS–SR).

## The BIMAS as a Screening Tool

This section presents evidence for the use of the BIMAS as a screening tool for emotional and behavioral difficulties. Several analyses were conducted in order to determine if the BIMAS could differentiate students with no behavior concerns (hereafter referred to as the non-clinical or normative group) from those who experience emotional/behavioral difficulties (hereafter referred to as the clinical groups). If results indicated that the BIMAS scores could produce this level of differentiation, then this would provide support for the instrument’s use as a screening tool (i.e., in Tier 1 for Universal Screening).

In order to conduct these analyses, ratings of youth previously identified with a clinical condition were collected during the standardization of the BIMAS. The ratings were completed by teachers, parents, and students themselves. There were 538 teacher ratings (300 males and 238 females), 467 parent ratings (274 males and 193 females), as well as 350 youth self-report ratings (180 males and 170

females). Ratings of youth with the following diagnoses were collected: Disruptive Behavior Disorders (DBD; includes Conduct Disorder, Oppositional Defiant Disorder); Attention-Deficit/Hyperactivity Disorder (ADHD); Anxiety (includes Generalized Anxiety Disorder, Obsessive Compulsive Disorder, and Social Phobia); Depression (includes Major Depressive Disorder and Dysthymia); Pervasive Developmental Disorders (PDD; includes Autistic Disorder and Asperger’s Disorder); Learning Disorders (LD; teacher report only); and Developmental Delay (DD; teacher report only). See Table 11.11 for the breakdown of clinical diagnostic groups by rater, and Table 11.12 for demographic characteristics of the rated youth (see *Data Collection* in chapter 9, *Standardization*, for more information on data collection procedures).

Ratings from the clinical groups were compared to ratings of the normative sample (see chapter 9, *Standardization*, for a description of the normative sample). Specifically, analyses were conducted to determine if BIMAS ratings could (a) differentiate clinical from non-clinical group membership, and (b) differentiate between various clinical groups.

Table 11.11. Clinical Samples: Primary Diagnoses by Rater

Clinical Group	Rater						
	Teacher		Parent		Self-Report		Total
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>
Disruptive Behavioral Disorders	123	22.9	70	15.0	65	18.6	258
Attention Deficit/Hyperactivity Disorder	109	20.3	117	25.1	89	25.4	315
Anxiety	55	10.2	67	14.3	56	16.0	178
Depression	60	11.2	73	15.6	62	17.7	195
Pervasive Developmental Disorders	95	17.7	86	18.4	65	18.6	246
Learning Disorders	45	8.4	n/a	n/a	n/a	n/a	45
Developmental Delay	30	5.6	n/a	n/a	n/a	n/a	30
Other Clinical	21	3.9	54	11.6	13	3.7	88
<b>Total</b>	<b>538</b>	<b>100.0</b>	<b>467</b>	<b>100.0</b>	<b>350</b>	<b>100.0</b>	<b>1,355</b>

Note. n/a = not applicable.

Table 11.12. Demographic Characteristic of the Clinical Samples

Demographic Characteristic of the Rated Youth		Rater						
		Teacher		Parent		Self-Report		Total
		<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>
Gender	Male	300	55.8	274	58.7	180	51.4	754
	Female	238	44.2	193	41.3	170	48.6	601
Race/Ethnicity	Asian	2	0.4	3	0.6	2	0.6	7
	African American	103	19.1	102	21.8	86	24.6	291
	Hispanic	49	9.1	49	10.5	44	12.6	142
	White	342	63.6	280	60.0	203	58.0	825
	Other	42	7.8	33	7.1	15	4.3	90
<b>Total</b>		<b>538</b>	<b>100.0</b>	<b>467</b>	<b>100.0</b>	<b>350</b>	<b>100.0</b>	<b>1,355</b>
<b>Age <i>M</i> (<i>SD</i>)</b>		<b>13.93 (3.31)</b>		<b>12.78 (3.44)</b>		<b>14.71 (1.86)</b>		<b>–</b>

In order to assess the ability of scores on the BIMAS to differentiate clinical youth from non-clinical youth, two types of analyses were conducted. First, mean *T*-scores from the clinical sample were compared to the mean *T*-scores from the normative sample. To determine the size of the difference between the groups, Cohen's *d*, an estimate of effect size, was calculated for each scale. Guidelines for interpreting Cohen's *d* effect size are provided by Cohen (1988): small effect size = |0.2|; medium effect size = |0.5|; large effect size = |0.8|. Cohen's *d* values can be interpreted as the size of the difference between two scores as a proportion of their pooled standard deviation. For example, a Cohen's *d* value of 0.5 is interpreted as a difference between two groups that is the size of one-half of their pooled standard deviation.

Second, two discriminant function analyses (DFA) were run per rater type (i.e., teacher, parent, self-report) in order to determine if BIMAS scores could predict group membership into the non-clinical or clinical groups. For both DFAs, the five BIMAS scale scores (i.e., Conduct, Negative Affect, Cognitive/Attention, Social, and Academic Functioning) were used as predictors to predict group membership into the non-clinical or clinical sample. The first DFA used the full-range of BIMAS scale scores in order to determine overall classification accuracy. The second DFA used the *T*-scores cut-off and classification criteria (i.e., greater than 60 for the Behavioral Concerns scales and/or a score of less than 40 for the Adaptive Skills scales; see *Interpreting T-scores* in

chapter 5, *Understanding and Interpreting BIMAS Scores*, for more details) were used to identify students who might require some level of intervention and/or assistance (i.e., to simulate the use of the BIMAS as a screening tool). Several classification statistics (see Kessel & Zimmerman, 1993) were calculated to demonstrate the efficacy of the BIMAS as a screening tool.

- **Overall Correct Classification Rate.** The percentage of correct group classifications made using the BIMAS *T*-scores.
- **Sensitivity.** The percentage of clinical cases correctly predicted by the BIMAS *T*-scores to belong to the clinical group.
- **Specificity.** The percentage of normative cases correctly predicted by the BIMAS *T*-scores to belong to the normative group.
- **Positive Predictive Power.** The percentage of youth identified by the BIMAS *T*-scores as *having* a clinical condition who, based on previous diagnosis, actually have a clinical condition.
- **Negative Predictive Power.** The percentage of youth identified by the BIMAS *T*-scores as *not having* a clinical condition who actually *do not* have a clinical condition.

Finally, analyses were also conducted to determine if BIMAS ratings could differentiate between various clinical groups. Because different scales on the BIMAS are more relevant to specific disorders (e.g., Conduct to DBD, Negative Affect to Anxiety and Depression, Cognitive Attention to ADHD, Social to PDD), separate analyses were conducted in order to determine if different clinical groups scored differently on each of the BIMAS scales. Effect sizes and graphical presentation of results are provided. MANOVAs were conducted to assess the statistical significance of the effects (see appendix I for results).

## The BIMAS as a Screening Tool: Teacher Ratings

Ratings on the BIMAS–T form were completed by 538 teachers of students with a reported clinical diagnosis (see Tables 11.11 and 11.12 for sample descriptions). The following sections present results regarding the ability of the BIMAS–T ratings to differentiate clinical from non-clinical youth, and to differentiate between clinical groups.

### Differences Between Clinical and Non-Clinical Groups

Table 11.13 presents the performance of the entire clinical sample as rated by teachers across the BIMAS–T scales. Since BIMAS scale scores are presented as standardized *T*-scores with a normative mean of 50 ( $SD = 10$ ), results indicate that on the Behavioral Concerns scales, teacher ratings of students from the clinical sample were over 1.5 *SDs* above the normative mean on the Negative Affect and Cognitive/Attention scales, and over 1 *SD* on the Conduct scale. Similarly, on the Adaptive scales, teacher ratings of students from the clinical sample were close to 1.5 *SDs* below the normative mean on the Social scale and close to 1 *SD* below the Academic Functioning scale. Effect sizes, as assessed with Cohen's *d* (in calculating effect size, the clinical *M*s and *SD*s presented in Table 11.13 were compared to the full normative sample, where  $N = 1,361$ ,  $M = 50$  and  $SD = 10$ ), revealed large effects for all scales and were all in the expected direction (positive *d* values on a Behavioral Concern scale indicate that the clinical sample mean was *greater* than the normative mean, while negative *d* values on an Adaptive scale indicate that the clinical sample mean was *lower* than normative mean). These results offer some initial evidence for the BIMAS–T's ability to differentiate students with a clinical diagnosis from non-clinical students in the normative sample.

Two DFAs were performed using the five BIMAS–T scale scores as predictors of group membership into the normative sample (i.e., non-clinical) or clinical sample using either the:

1) full-range of BIMAS scores, or 2) BIMAS cut-scores (i.e., greater than 60 for the Behavioral Concerns scales and/or a score of less than 40 for the Adaptive Skills scales).

The Wilks' Lambda values (Full Range of Scores, Wilks' Lambda = .54,  $\chi^2 [5] = 1168.1$ ,  $p < .001$ ; BIMAS Cut-Scores, Wilks' Lambda = .55,  $\chi^2 [5] = 1146.3$ ,  $p < .001$ ) indicate that the five BIMAS predictor variables significantly discriminated between the clinical and non-clinical groups across all comparisons performed. The standardized discriminant function coefficients indicate how heavily each BIMAS–T variable contributed to the discrimination between the two groups. As the values in Table 11.14 reveal, the Negative Affect and Cognitive/Attention scales were the most significantly contributing variables (for both analyses). The overall correct classification rate was 85.2% when the full-range of scores were employed, and was 82.5% when the cut-scores were employed, offering strong support for the use of the BIMAS–T as a screening tool (the classification table can be found in Table 11.15; see Table 11.16 for the full list of classification accuracy results).

### Differences Between Clinical Groups

In addition to differentiating overall clinical versus non-clinical group membership, Table 11.17 and Figure 11.1 provide evidence that teacher ratings on the BIMAS–T can discriminate youth with various specific clinical disorders from non-clinical youth. The mean scores for the Behavioral Concern scales were well above the normative mean (normative  $M = 50$ ,  $SD = 10$ ), and for the Adaptive scales, the mean scores were well below the normative mean. All effect sizes were large. (In calculating Cohen's *d*, the clinical *M*s and *SD*s presented in Table 11.17 were compared to the full normative sample, where  $N = 1,361$ ,  $M = 50$  and  $SD = 10$ .) Moreover, an examination of the means for each clinical group reveals that each clinical group has its own pattern of scale elevations on the BIMAS–T, and these patterns are in the expected direction. For example, while all scales are good at differentiating the DBD group from the normative sample, the scale with the biggest difference from the norm (i.e., the largest effect size) is the Conduct scale ( $d = 2.1$ ). Similarly, Negative Affect was the most discriminating scale for both the Anxiety and Depression groups ( $d = 2.1$  and  $2.2$ , respectively), the Cognitive/Attention scale was the most discriminating scale for the both the ADHD and LD groups ( $d = 1.9$  and  $1.8$ , respectively), and the Social scale was the most discriminating for the PDD group ( $d = -1.8$ ).

**Table 11.13. Clinical Sample Descriptive Statistics (T-scores) and Effect Sizes (Comparison to the Norm) Across BIMAS–T Standard Scales**

BIMAS–T Standard Scales	Clinical Sample			Cohen's <i>d</i>
	<i>N</i>	<i>M</i>	<i>SD</i>	
Conduct	516	63.5	10.9	1.3
Negative Affect	537	66.4	10.4	1.6
Cognitive/Attention	538	66.6	9.8	1.7
Social	538	35.6	10.3	–1.4
Academic Functioning	538	40.2	9.8	–1.0

Note. Clinical *M*s (*SD*s) compared to values from the normative sample ( $N = 1,361$ ,  $M = 50$ ,  $SD = 10$ ). Cohen's *d* values of  $|0.2|$  = small effect,  $|0.5|$  = medium effect, and  $|0.8|$  = large effect. Positive *d* values indicate that the clinical sample score was higher than the normative sample score; negative *d* values indicate that the normative sample score was higher than the clinical sample.

**Table 11.14. Standardized Discriminant Function Coefficients of the BIMAS–T Standard Scales**

DFA	Full Range of Scores	BIMAS Cut-Scores
Conduct	.195	.172
Negative Affect	.505	.528
Cognitive/Attention	.565	.508
Social	–.399	.278
Academic Functioning	.421	–.120

**Table 11.15. Group Classification as Predicted by Scores on the BIMAS–T Standard Scales**

		Predicted Group Membership					
		Full Range of Scores			BIMAS Cut-Scores		
		Normative	Clinical	Total	Normative	Clinical	Total
Actual Group Membership	Normative	<b>1,200</b>	199	1,399	<b>1,167</b>	233	1,400
	Clinical	85	<b>430</b>	515	107	<b>431</b>	538
	Total	1,285	629	1,914	1,274	664	1,938

Note. **Bolded** values indicate correctly classified cases.

**Table 11.16. Classification Accuracy of the BIMAS–T Standard Scales**

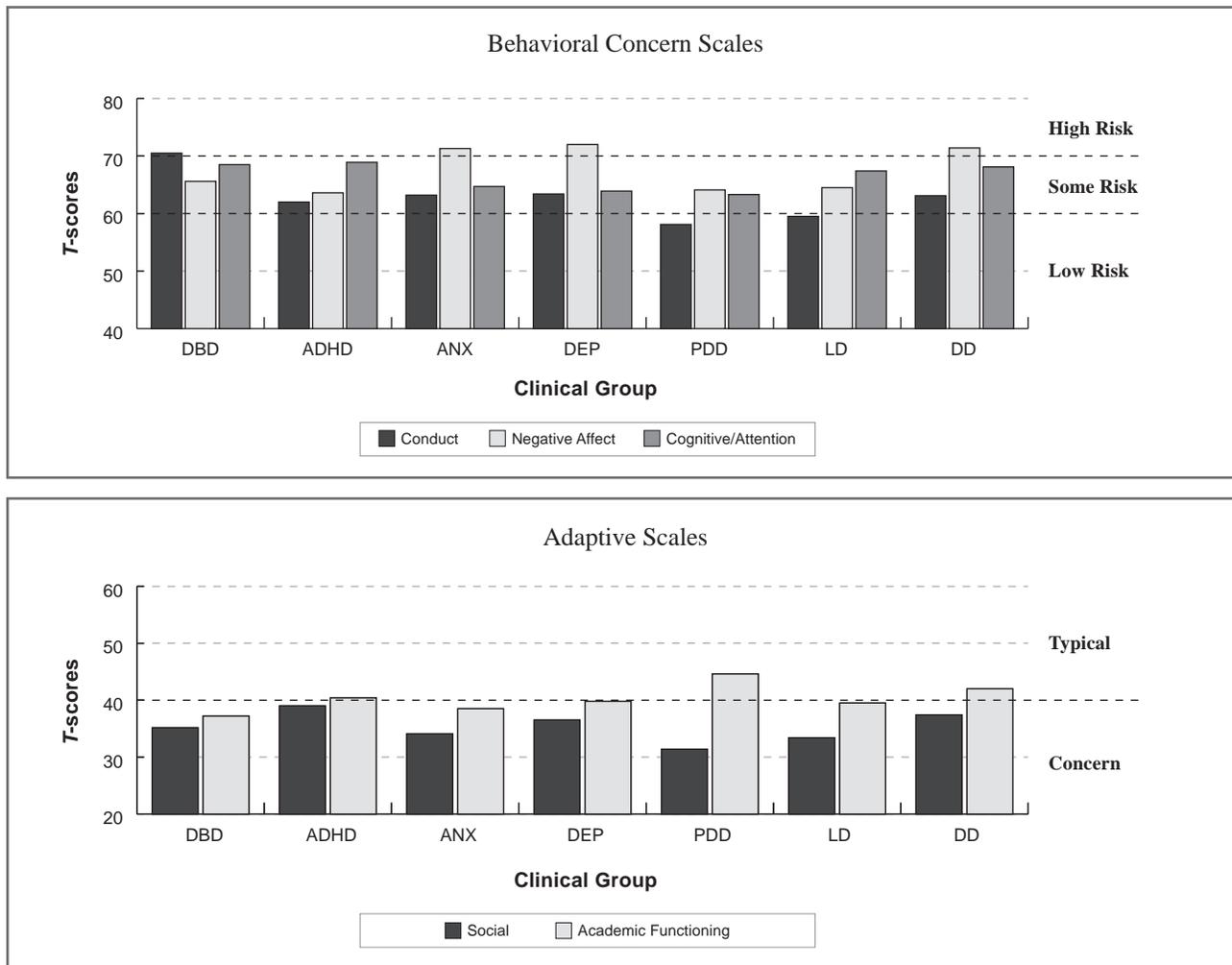
Classification Accuracy Statistic	Full Range of Scores	BIMAS Cut-Scores
Overall Correct Classification	85.2%	82.5%
Sensitivity	83.5%	80.1%
Specificity	85.8%	83.4%
Positive Predictive Power	68.4%	64.9%
Negative Predictive Power	93.4%	91.6%

**Table 11.17. Descriptive Statistics (T-scores) of the Different Clinical Samples and Effect Sizes (Comparison to the Norm) Across the BIMAS–T Standard Scales**

Type of Scale	Scale	<i>N</i>	DBD	ADHD	ANX	DEP	PDD	LD	DD
			123	109	55	60	95	45	30
Behavioral Concern Scales	Conduct	<i>M</i>	<b>70.5</b>	62.0	63.2	63.4	58.1	59.5	63.1
		<i>SD</i>	<b>9.0</b>	9.5	12.0	10.0	9.8	11.3	11.2
		Cohen's <i>d</i>	<b>2.1</b>	1.2	1.3	1.3	0.8	0.9	1.3
	Negative Affect	<i>M</i>	65.6	63.6	<b>71.3</b>	<b>72.0</b>	64.1	64.5	<b>71.4</b>
		<i>SD</i>	10.2	8.4	<b>10.4</b>	<b>8.9</b>	11.6	10.2	<b>10.0</b>
		Cohen's <i>d</i>	1.6	1.4	<b>2.1</b>	<b>2.2</b>	1.4	1.4	<b>2.1</b>
	Cognitive/Attention	<i>M</i>	68.5	<b>68.9</b>	64.7	63.9	63.3	<b>67.4</b>	68.1
		<i>SD</i>	9.6	<b>8.4</b>	11.6	9.0	11.2	<b>6.8</b>	9.3
		Cohen's <i>d</i>	1.9	<b>1.9</b>	1.5	1.4	1.3	<b>1.8</b>	1.8
Adaptive Scales	Social	<i>M</i>	35.2	39.0	34.1	36.5	<b>31.4</b>	33.4	37.4
		<i>SD</i>	8.4	10.7	11.9	9.3	<b>11.4</b>	9.1	8.1
		Cohen's <i>d</i>	–1.5	–1.1	–1.6	–1.4	<b>–1.8</b>	–1.7	–1.3
	Academic Functioning	<i>M</i>	37.2	40.4	38.5	39.8	44.6	39.5	42.0
		<i>SD</i>	8.3	8.8	11.8	9.9	9.8	10.3	10.0
		Cohen's <i>d</i>	–1.3	–1.0	–1.1	–1.0	–0.5	–1.0	–0.8

Note. Clinical *M*s (*SD*s) compared to values from the normative sample ( $N = 1,361$ ,  $M = 50$ ,  $SD = 10$ ). Cohen's *d* values of  $|0.2|$  = small effect,  $|0.5|$  = medium effect, and  $|0.8|$  = large effect. **Bolded** cells indicate the BIMAS–T scale with the largest effect size for each clinical group. Positive *d* values indicate that the clinical sample score was higher than the normative sample score; negative *d* values indicate that the normative sample score was higher than the clinical sample score. DBD = Disruptive Behavior Disorders, ADHD = Attention Deficit/Hyperactivity Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorders, LD = Learning Disorders, DD = Developmental Delay.

Figure 11.1. Mean T-scores by Clinical Group: BIMAS–T Standard



Note. DBD = Disruptive Behavior Disorders, ADHD = Attention Deficit/Hyperactivity Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorder, LD = Learning Disorders, DD = Developmental Delay.

## The BIMAS as a Screening Tool: Parent Ratings

Ratings on the BIMAS–P form were completed by 467 parents of children with a reported clinical diagnosis (see Tables 11.11 and 11.12 for sample descriptions). The following sections present results regarding the ability of the BIMAS–P ratings to differentiate clinical versus non-clinical group membership and to differentiate between clinical groups.

### Differences Between Clinical and Non-Clinical Groups

Table 11.18 presents the performance of the entire clinical sample as rated by parents across the BIMAS–P scales. Results indicate that on the Behavioral Concerns scales, parent ratings of children from the clinical sample were approximately 1 SD above the normative mean (normative  $M = 50$ ,  $SD = 10$ ), while on the Adaptive scales, parent ratings

of children from the clinical sample were close to 1 SD below the normative mean. Effect sizes, as assessed with Cohen’s  $d$  (in calculating effect size, the clinical  $M$ s and  $SD$ s presented in Table 11.18 were compared to the full normative sample, where  $N = 1,400$ ,  $M = 50$  and  $SD = 10$ ), revealed large effects for all scales and were all in the expected direction (positive  $d$  values on a Behavioral Concern scale indicate that the clinical sample mean was *greater* than the normative mean, while negative  $d$  values on an Adaptive scale indicate that the clinical sample mean was *lower* than normative mean). These results offer some initial evidence for the ability of parent’s ratings on the BIMAS–P to differentiate children with a clinical diagnosis from non-clinical children in the normative sample.

Two DFAs were performed using the five BIMAS–P scale scores as predictors to predict group membership into the normative sample or the clinical sample using either the: 1) full-range of BIMAS scores, or 2) BIMAS cut-scores (i.e., greater than 60 for the Behavioral Concerns scales and/or a score of less than 40 for the Adaptive Skills scales).

Similar to the results found from the teacher ratings, Wilks' Lambda values (Full Range of Scores, Wilks' Lambda = .69,  $\chi^2 [5] = 688.3, p < .001$ ; BIMAS Cut-Scores, Wilks' Lambda = .67,  $\chi^2 [5] = 742.0, p < .001$ ) indicate that the five BIMAS predictor variables significantly discriminated between the clinical and non-clinical groups across all comparisons performed. The standardized discriminant function coefficients indicate how heavily each BIMAS-P variable contributed to the discrimination between the two groups (see Table 11.19). The Social scale was the most significantly contributing variable for both DFAs. The overall correct classification rate was 78.3% when the full-range of scores was employed and was 78.6% when the cut-scores were employed (the classification table can be found in Table 11.20; see Table 11.21 for the full list of classification accuracy results).

### Differences Between Clinical Groups

In addition to predicting overall clinical versus non-clinical group membership, Table 11.22 and Figure 11.2 provide evidence that parent ratings on the BIMAS-P can discriminate youth with various specific clinical disorders from non-clinical youth. The mean scores for the Behavioral Concern scales ranged from 0.5 *SDs* to just under 2 *SDs* above the normative mean ( $M = 50, SD = 10$ ), and for the Adaptive scales, the mean scores ranged from 0.5 *SDs* to 1.5 *SDs* below the normative mean. The majority of effect sizes were large with some moderate effect sizes for the ADHD and PDD groups. (In calculating Cohen's *d*, the clinical *M*s and *SD*s presented in Table 11.22 were compared to the full normative sample, where  $N = 1,400, M = 50$  and  $SD = 10$ ). Furthermore, exactly like the results found in the teacher ratings, an examination of the means for each clinical group

revealed expected scale-level differences. Specifically, Conduct is the most discriminating scale for the DBD group ( $d = 1.9$ ), Negative Affect is the most discriminating scale for both the Anxiety and Depression groups ( $d = 1.0$  and  $1.4$ , respectively), Cognitive/Attention is the most discriminating scale for the ADHD group ( $d = 1.4$ ), and the Social scale is the most discriminating for the PDD group ( $d = -1.5$ ).

**Table 11.18. Clinical Sample Descriptive Statistics (T-scores) and Effect Sizes (Comparison to the Norm) Across the BIMAS-P Standard Scales**

BIMAS-P Standard Scales	Clinical Sample			Cohen's <i>d</i>
	<i>N</i>	<i>M</i>	<i>SD</i>	
Conduct	467	60.3	10.5	1.0
Negative Affect	467	61.5	10.3	1.1
Cognitive/Attention	467	60.7	9.9	1.1
Social	467	38.4	9.9	-1.2
Academic Functioning	467	40.4	7.9	-1.0

*Note.* Clinical *M*s (*SD*s) compared to values from the normative sample ( $N = 1,400, M = 50, SD = 10$ ). Cohen's *d* values of  $|0.2|$  = small effect,  $|0.5|$  = medium effect, and  $|0.8|$  = large effect. Positive *d* values indicate that the clinical sample score was higher than the normative sample score; negative *d* values indicate that the normative sample score was higher than the clinical sample.

**Table 11.19. Standardized Discriminant Function Coefficients of the BIMAS-P Standard Scales**

DFA	Full Range of Scores	BIMAS Cut-Scores
Conduct	-.156	.128
Negative Affect	.388	.297
Cognitive/Attention	.437	.348
Social	-.463	.550
Academic Functioning	-.168	.130

**Table 11.20. Group Classification as Predicted by Scores on the BIMAS-P Standard Scales**

		Predicted Group Membership					
		Full Range of Scores			BIMAS Cut-Scores		
		Normative	Clinical	Total	Normative	Clinical	Total
Actual Group Membership	Normative	<b>1,085</b>	311	1,396	<b>1,124</b>	276	1,400
	Clinical	93	<b>374</b>	467	124	<b>343</b>	467
	Total	1,178	685	1,863	1,248	619	1,867

*Note.* **Bolded** values indicate correctly classified cases.

**Table 11.21. Classification Accuracy of the BIMAS-P Standard Scales**

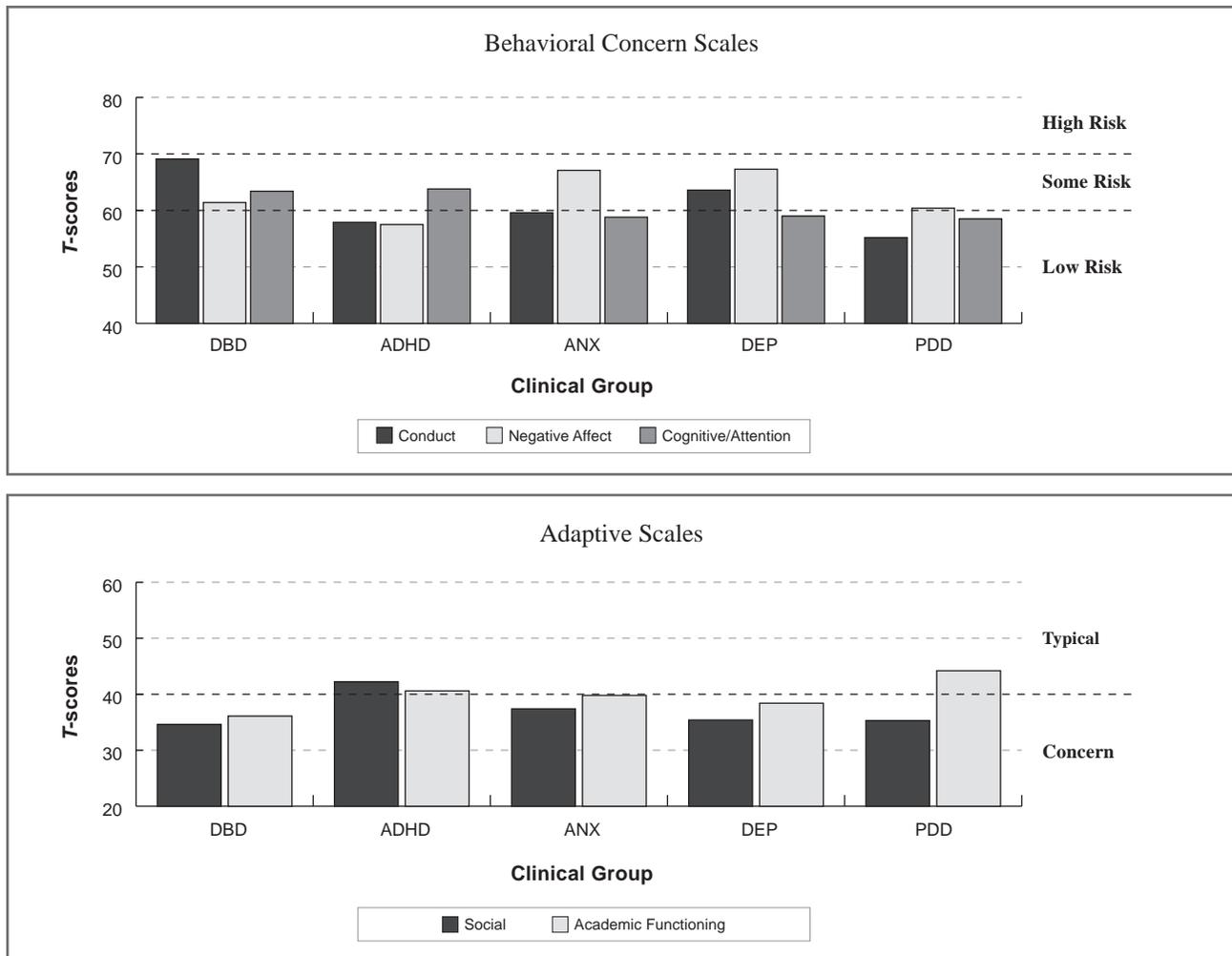
Classification Accuracy Statistic	Full Range of Scores	BIMAS Cut-Scores
Overall Correct Classification	78.3%	78.6%
Sensitivity	80.1%	73.4%
Specificity	77.7%	80.3%
Positive Predictive Power	54.6%	55.4%
Negative Predictive Power	92.1%	90.1%

**Table 11.22. Descriptive Statistics (T-scores) of the Different Clinical Samples and Effect Sizes (Comparison to the Norm) across the BIMAS–P Standard Scales**

Type of Scale	Scale	N	DBD	ADHD	ANX	DEP	PDD
			70	117	67	73	86
Behavioral Concern Scales	Conduct	M	<b>69.1</b>	57.9	59.6	63.6	55.2
		SD	<b>9.5</b>	8.8	11.1	8.9	10.1
		Cohen's <i>d</i>	<b>1.9</b>	0.8	1.0	1.4	0.5
	Negative Affect	M	69.1	57.9	<b>59.6</b>	<b>63.6</b>	55.2
		SD	9.3	9.1	<b>10.9</b>	<b>9.2</b>	9.6
		Cohen's <i>d</i>	1.9	0.8	<b>1.0</b>	<b>1.4</b>	0.5
	Cognitive/Attention	M	63.4	<b>63.8</b>	58.8	59.0	58.5
		SD	8.4	<b>9.9</b>	9.9	7.7	10.5
		Cohen's <i>d</i>	1.4	<b>1.4</b>	0.9	0.9	0.8
Adaptive Scales	Social	M	34.6	42.2	37.4	35.4	<b>35.3</b>
		SD	8.2	8.3	8.4	8.5	<b>1.3</b>
		Cohen's <i>d</i>	-1.6	-0.8	-1.3	-1.5	<b>-1.5</b>
	Academic Functioning	M	36.1	40.6	39.8	38.4	44.2
		SD	7.3	7.4	7.8	6.8	7.5
		Cohen's <i>d</i>	-1.4	-1.0	-1.0	-1.2	-0.6

Note. Clinical Ms (SDs) compared to values from the normative sample ( $N = 1,400, M = 50, SD = 10$ ). Cohen's *d* values of  $|0.2|$  = small effect,  $|0.5|$  = medium effect, and  $|0.8|$  = large effect. **Bolded** cells indicate the BIMAS–P scale with the largest effect size for each clinical group. Positive *d* values indicate that the clinical sample score was higher than the normative sample score; negative *d* values indicate that the normative sample score was higher than the clinical sample score. DBD = Disruptive Behavior Disorders, ADHD = Attention Deficit/Hyperactivity Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorders.

**Figure 11.2. Mean T-scores by Clinical Group: BIMAS–P Standard**



Note. DBD = Disruptive Behavior Disorders, ADHD = Attention Deficit/Hyperactivity Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorders.

## The BIMAS as a Screening Tool: Self-Report Ratings

Ratings on the BIMAS–SR form were completed by 350 youth with a reported clinical diagnosis (see Tables 11.11 and 11.12 for sample descriptions). The following sections present results regarding the ability of the BIMAS–SR ratings to differentiate clinical from non-clinical youth, and to differentiate between clinical groups.

### Differences Between Clinical and Non-Clinical Groups

Self-ratings from youth with a clinical diagnosis, like those of both teachers and parents, resulted in higher scores than did ratings from youth in the normative sample (see Table 11.23; ratings were approximately 0.75 to 1 *SD* above the normative *M* of 50 on the Behavioral Concern scales and approximately 0.75 to 1 *SD* below the normative mean on the Adaptive scales). Furthermore, an examination of the Cohen's *d* ratios revealed large effect sizes for all scales except for the Conduct scale, where the effect size was moderate. (In calculating effect sizes, the clinical *M*s and *SD*s presented in Table 11.23 were compared to the full normative sample, where  $N = 703$ ,  $M = 50$  and  $SD = 10$ .) All effect sizes were in the expected direction; that is, positive values on the Behavioral Concern scale indicate that the clinical sample means were *greater* than the normative means, while negative values on the Adaptive scales indicate that the clinical sample means were *lower* than the normative means.

As with the teacher and parent form analyses, two DFAs were performed using the five BIMAS–SR scale scores as predictors of group membership into the normative sample or the clinical sample using either the: 1) full-range of BIMAS scores, or 2) BIMAS cut-scores (i.e., a score greater than 60 for the Behavioral Concerns scales and/or a score of less than 40 for the Adaptive Skills scales). The performance of the clinical samples in each case was statistically different than that of the comparison group of students in the normative sample, as indicated by the Wilks' Lambda values (Full Range of Scores, Wilks' Lambda = .79,  $\chi^2 [5] = 245.2$ ,  $p < .001$ ; BIMAS Cut-Scores, Wilks' Lambda = .77,  $\chi^2 [5] = 270.4$ ,  $p < .001$ ). The standardized discriminant function coefficients indicate how heavily each BIMAS–SR variable contributed to the discrimination between the two groups. The Negative Affect scale was the most significantly contributing variable for the DFA that utilized the full range of scores, whereas the Social scale was the most significantly contributing variable for the cut-scores DFA (see Table 11.24). The overall correct classification rate was 71.5% when the full-range of scores were employed, and was 71.8% when the cut-scores were employed (the classification tables can be found in Table 11.25; see Table 11.26 for the full list of classification accuracy results).

## Differences Between Clinical Groups

In addition to predicting overall clinical versus non-clinical group membership, Table 11.27 and Figure 11.3 provide some evidence that youth self-report ratings on the BIMAS–SR can discriminate youth with various specific clinical disorders from non-clinical youth. An examination of the means for each clinical group generally revealed expected scale-level differences and moderate to large effect sizes. (In calculating Cohen's *d*, the clinical *M*s and *SD*s presented in Table 11.27 were compared to the full normative sample, where  $N = 703$ ,  $M = 50$  and  $SD = 10$ .) Mean scores were 1 to 1.5 *SD*s above the normative mean (normative  $M = 50$ ,  $SD = 10$ ) for the most discriminating scale for each clinical group. Specifically, Conduct was the most discriminating scale for the DBD group with a large effect size ( $d = 1.5$ ), Negative Affect was the most discriminating scale for both the Anxiety and Depression groups ( $d = 1.1$  and  $1.5$ , respectively) with large effect sizes, and the Social scale was the most discriminating for the PDD group with a large effect size ( $d = -1.4$ ). The ADHD group was the one group where the group mean fell just short of being 1 *SD* above the normative mean; however, a large effect size was still found ( $d = 0.8$ ).

**Table 11.23. Clinical Sample Descriptive Statistics (T-scores) and Effect Sizes (Comparison to the Norm) Across the BIMAS–SR Standard Scales**

BIMAS–SR Standard Scales	Clinical Sample			Cohen's <i>d</i>
	<i>N</i>	<i>M</i>	<i>SD</i>	
Conduct	350	57.3	9.7	0.7
Negative Affect	350	59.2	9.7	0.9
Cognitive/Attention	350	57.3	8.2	0.8
Social	350	41.4	9.7	-0.9
Academic Functioning	350	42.3	8.3	-0.8

*Note.* Clinical *M*s (*SD*s) compared to values from the normative sample ( $N = 703$ ,  $M = 50$ ,  $SD = 10$ ). Cohen's *d* values of  $|0.2|$  = small effect,  $|0.5|$  = medium effect, and  $|0.8|$  = large effect. Positive *d* values indicate that the clinical sample score was higher than the normative sample score; negative *d* values indicate that the normative sample score was higher than the clinical sample.

**Table 11.24. Standardized Discriminant Function Coefficients of the BIMAS–SR Standard Scales**

DFA	Full Range of Scores	BIMAS Cut-Scores
Conduct	-.145	.037
Negative Affect	.475	.402
Cognitive/Attention	.329	.435
Social	-.379	.595
Academic Functioning	-.251	-.055

**Table 11.25. Group Classification as Predicted by Scores on the BIMAS–SR Standard Scales**

		Predicted Group Membership					
		Full Range of Scores			BIMAS Cut-Scores		
		Normative	Clinical	Total	Normative	Clinical	Total
Actual Group Membership	Normative	<b>483</b>	216	699	<b>519</b>	181	700
	Clinical	83	<b>267</b>	350	115	<b>235</b>	350
	Total	566	483	1,049	634	416	1,050

Note. **Bolded** values indicate correctly classified cases.

**Table 11.26. Classification Accuracy of the BIMAS–SR Standard Scales**

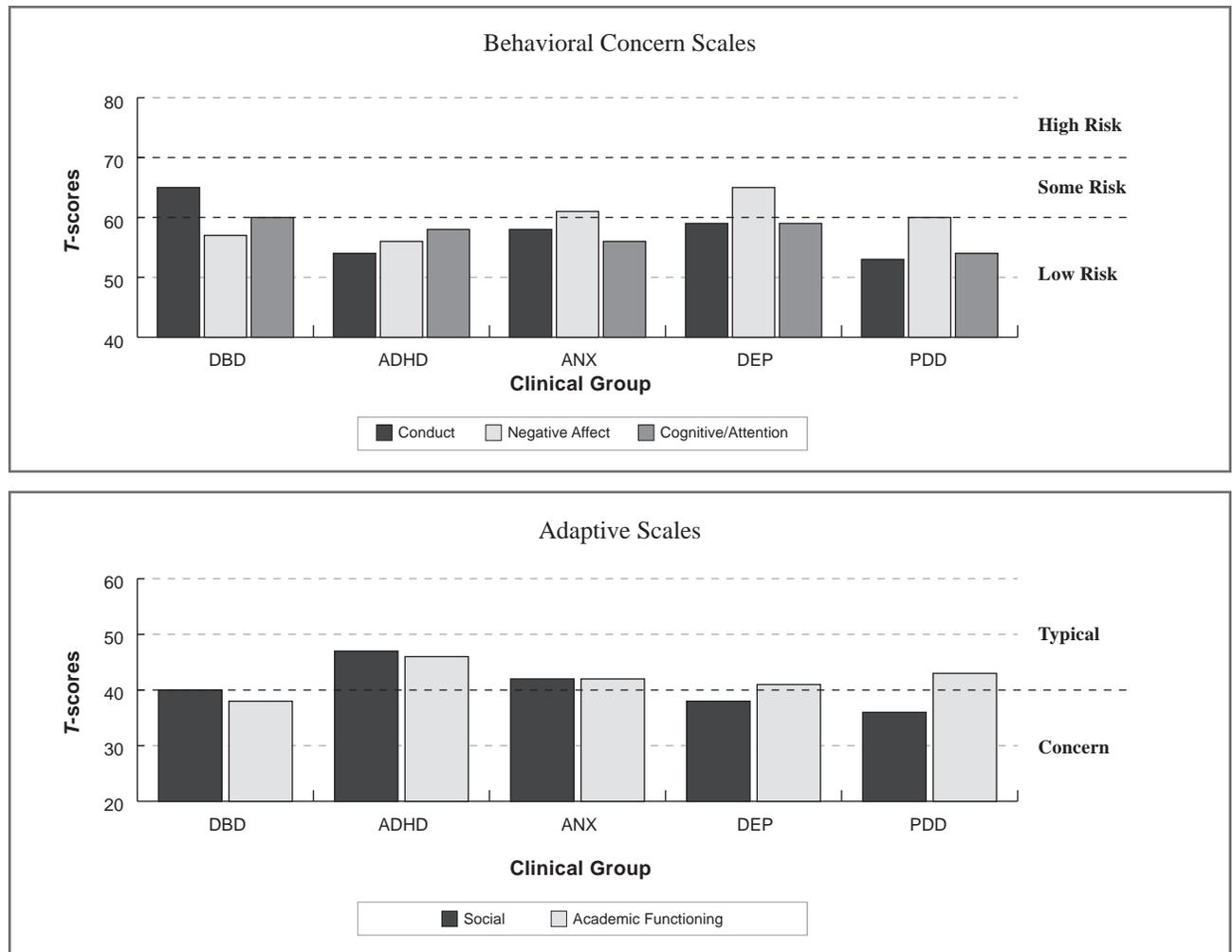
Classification Accuracy Statistic	Full Range of Scores	BIMAS Cut-Scores
Overall Correct Classification	71.5%	71.8%
Sensitivity	76.3%	67.1%
Specificity	69.1%	74.1%
Positive Predictive Power	55.3%	56.5%
Negative Predictive Power	85.3%	81.9%

**Table 11.27. Descriptive Statistics (T-scores) of the Different Clinical Samples and Effect Sizes (Comparison to the Norm) Across the BIMAS–SR Standard Scales**

Type of Scale	Scale	N	DBD	ADHD	ANX	DEP	PDD
			65	89	56	62	65
Behavioral Concern Scales	Conduct	M	<b>64.8</b>	53.7	57.7	59.2	53.2
		SD	<b>8.3</b>	7.8	0.4	8.8	8.7
		Cohen's <i>d</i>	<b>1.5</b>	0.4	0.8	0.9	0.3
	Negative Affect	M	57.0	55.7	<b>60.8</b>	<b>65.2</b>	60.0
		SD	8.3	8.8	<b>8.1</b>	<b>8.3</b>	0.5
		Cohen's <i>d</i>	0.7	0.6	<b>1.1</b>	<b>1.5</b>	1.0
	Cognitive/Attention	M	59.5	<b>57.6</b>	56.3	58.8	54.4
		SD	6.4	<b>7.4</b>	8.6	7.9	0.3
		Cohen's <i>d</i>	1.0	<b>0.8</b>	0.6	0.9	0.5
Adaptive Scales	Social	M	39.7	47.3	41.8	38.3	<b>36.4</b>
		SD	8.2	1.1	7.9	5.8	<b>9.1</b>
		Cohen's <i>d</i>	-1.0	-0.3	-0.8	-1.2	<b>-1.4</b>
	Academic Functioning	M	38.2	45.7	41.9	40.8	42.9
		SD	7.8	7.3	9.6	6.7	8.0
		Cohen's <i>d</i>	-1.2	-0.4	-0.8	-0.9	-0.7

Note. Clinical *M*s (*SD*s) compared to values from the normative sample ( $N = 1,400$ ,  $M = 50$ ,  $SD = 10$ ). Cohen's *d* values of  $|0.2|$  = small effect,  $|0.5|$  = medium effect, and  $|0.8|$  = large effect. **Bolded** cells indicate the BIMAS–P scale with the largest effect size for each clinical group. Positive *d* values indicate that the clinical sample score was higher than the normative sample score; negative *d* values indicate that the normative sample score was higher than the clinical sample score. DBD = Disruptive Behavior Disorders, ADHD = Attention Deficit/Hyperactivity Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorders.

Figure 11.3. Mean T-scores by Clinical Group: BIMAS–SR Standard



Note. DBD = Disruptive Behavior Disorders, ADHD = Attention Deficit/Hyperactivity Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorders.

## Validity of the BIMAS as a Progress Monitoring Tool

The BIMAS was built to serve the purpose of progress monitoring. That is, initial conceptualization and early BIMAS research efforts (as described in chapter 8, *Development*) had the goal of developing an instrument that would help mental health agencies and practitioners evaluate the effectiveness of services and interventions. Early empirical studies (e.g., Meier 1997, 1998, 2000, 2004; Meier, McDougal, & Bardos, 2008; Weinstock & Meier, 2003) examining the Intervention Items Selection Rules (IISRs; see chapter 8 for a detailed discussion on the IISRs) offer some support for the use of BIMAS as a progress monitoring tool. The following section reviews some of these studies.

## Change-Sensitivity of Early Versions of the BIMAS

As briefly described in chapter 8, *Development*, Weinstock and Meier (2003) compared the ability of two item selection methods, principal component analysis (PCA) and IISRs, to capture client changes on a self-report checklist which consisted of early BIMAS items. Participants were 615 university counseling center clients who received treatment at a large university counseling center over a period of 8 school years and they completed a 56-item self-report checklist at both intake and termination. The authors reported that 211 clients (34%) were male and 404 (66%) were female; 84% of the sample identified themselves as White, 6% were Asian, 6% were African-American, 3.5% were Hispanic, and 5% were Native American. Most clients were single (77%). Students ranged in age from 17 to 54 ( $M = 24.48$ ,  $SD = 5.69$ ), and the mean number of sessions attended was 10.84 ( $SD = 7.15$ ).

It was found that the use of IISRs and PCA item selection methods resulted in scales with differing psychometric properties. While the IISRs and PCA methods produced scales with high and roughly equivalent reliability estimates ( $\alpha = .84$  and  $.88$  at intake and termination on the 8-item PCA Scale;  $\alpha = .90$  and  $.93$  for the 25-item IISRs Scale at intake and termination, respectively). Effect size (ES) for the IISRs scale considerably exceeded those from the PCA scale (ES =  $.47$  for the PCA Scale; ES =  $.74$  for the IISRs Scale)<sup>2</sup>. These findings supported the hypothesis that trait-based, individual difference methods of item evaluation and selection (i.e., PCA) may not be optimal for measuring change-sensitive constructs. Moreover, the 25-item IISRs scale was roughly half the length of the original 56-item checklist, an important consideration in outcome assessment where the scale's length can influence both data quality and the respondent's willingness to complete the form.

In another study by Meier (2004), the change-sensitivity of items and subscales (i.e., the items' and subscales' effectiveness in progress monitoring) in an early version of the BIMAS called the Teacher/Counselor Elementary Student Outcomes Form (TCESOF) was evaluated using two samples of elementary school students receiving counseling. The TCESOF is a 22-item brief outcome measure completed by a teacher, counselor, or other school personnel familiar with a child of elementary school age who is completing such psychosocial interventions as counseling, psychotherapy, or family therapy. Teachers completed the TCESOF for 183 students at intake, and 130 TCESOF ratings were available at follow-up; a total of 88 matches were found between intake and follow-up forms. Demographic information available at intake for 153 students indicated that most were in Grades 1 (12%), 2 (17%), 3 (21%), 4 (15%), 5 (15%), and 6 (11%). Eighty-seven percent of the sample were White, 4% were African-American, and 4% Biracial; 65% were male and 35% were female. Students attended schools in about a half a dozen school districts in suburban and rural areas in upstate New York. Teachers were instructed to rate how often each of the behaviors, events, or situations occurred during the past month on the following scale: Never, Rarely, Sometimes, and Often. Never was explicitly defined as 0 instances of the behavior during the past month; Rarely as 1 time; Sometimes as 2–15 times; and Often as 16–30 times.

Overall, scales constructed based on IISRs evidenced greater sensitivity to change than the original full set of items. These change-sensitive scales included the *Positive Change* scale (i.e., items that evidenced an improvement over time; 5 items) and the *Negative Change* scale (i.e., items that evidenced a worsening of scores from intake to follow-up; 2 items). These change-sensitive scales showed

adequate internal consistency assessed by Cronbach's alpha ( $\alpha$  ranged from  $.37$  [on the 2-item *Negative Change* scale] to  $.79$ ). Moreover, larger ES was also evidenced on change-sensitive scales. ES was found to be  $.37$  and  $-.21$  on the *Positive Change* and *Negative Change* scales, respectively, compared to: (a) ES =  $.21$  for the original item set with 22 items; (b) ES =  $.16$  for a shorter version of the original item set (18 items) constructed for general screening purposes, called the *Revised-Original* scale; (c) ES =  $.25$  on the *Strength* scale which consists of 11 items strictly assessing positive behaviors (i.e., not constructed based on IISRs procedures), and (d) ES =  $.12$  on the *Distress/Problems* scale which comprised 11 negatively worded items (also not based on IISRs).

Other studies have also replicated the above results using samples from a broader range of psychosocial interventions. Meier, McDougal, and Bardos (2008), for example, examined the psychometric properties of a preliminary version of the BIMAS employed with parents of elementary school-aged children. Youth in the study received services from community mental agencies that provided individual, family, and group interventions to children, adolescents, and families in urban, suburban, and rural communities. Parents completed the outcome measure for 896 elementary school-aged children (66% identified as male and 34% as female). Sixty percent of the sample were White, 20% African-American, 8% Hispanic, 4% Biracial, 2% Native American, 3% Asian, and 3% Other. Mothers (67%) were the primary providers of data, and parents reported their child's school grade as 12% in Kindergarten, 14% in Grade 1, 13% in Grade 2, 14% in Grade 3, 19% in Grade 4, 15% in Grade 5, 11% in Grade 6, and remaining 2% in Grades 7-10.

Results indicated that scales formed with change-sensitive items evidenced adequate reliability estimates and larger effect sizes than scales that comprised the original item pool or a shorter list of items comprising the Total Screening scale (developed to be a stable indicator of mental health problems). Reliability estimates were adequate for the 32-item Original scale ( $\alpha = .88$  at both intake and follow-up), the 14-item Total Screening scale ( $\alpha = .80$  and  $.84$  at intake and follow-up, respectively), and for the group of Intervention Sensitive scales ( $\alpha$  ranged from  $.74$  to  $.83$ ). Reliability values indicate that all scales have sufficient internal consistency for subsequent use and interpretation. Because participant heterogeneity can decrease the power to detect treatment effects (Lipsey, 1998), scale results were computed for the largest homogeneous subgroups in the dataset (i.e., for the total sample, boys and girls, and Caucasian boys); the Original and Screening scales have lower ESs (ES =  $.15$  and  $.21$  for the two scales, respectively) than Intervention-Sensitive scales (ES ranged from  $.25$  to  $.43$ ).

<sup>2</sup> ES was computed using the most commonly employed method for calculating ES with multi-item scales: the mean of the post-test (at termination) was subtracted from the mean of the pre-test (at intake), and then divided by the pooled standard deviation.

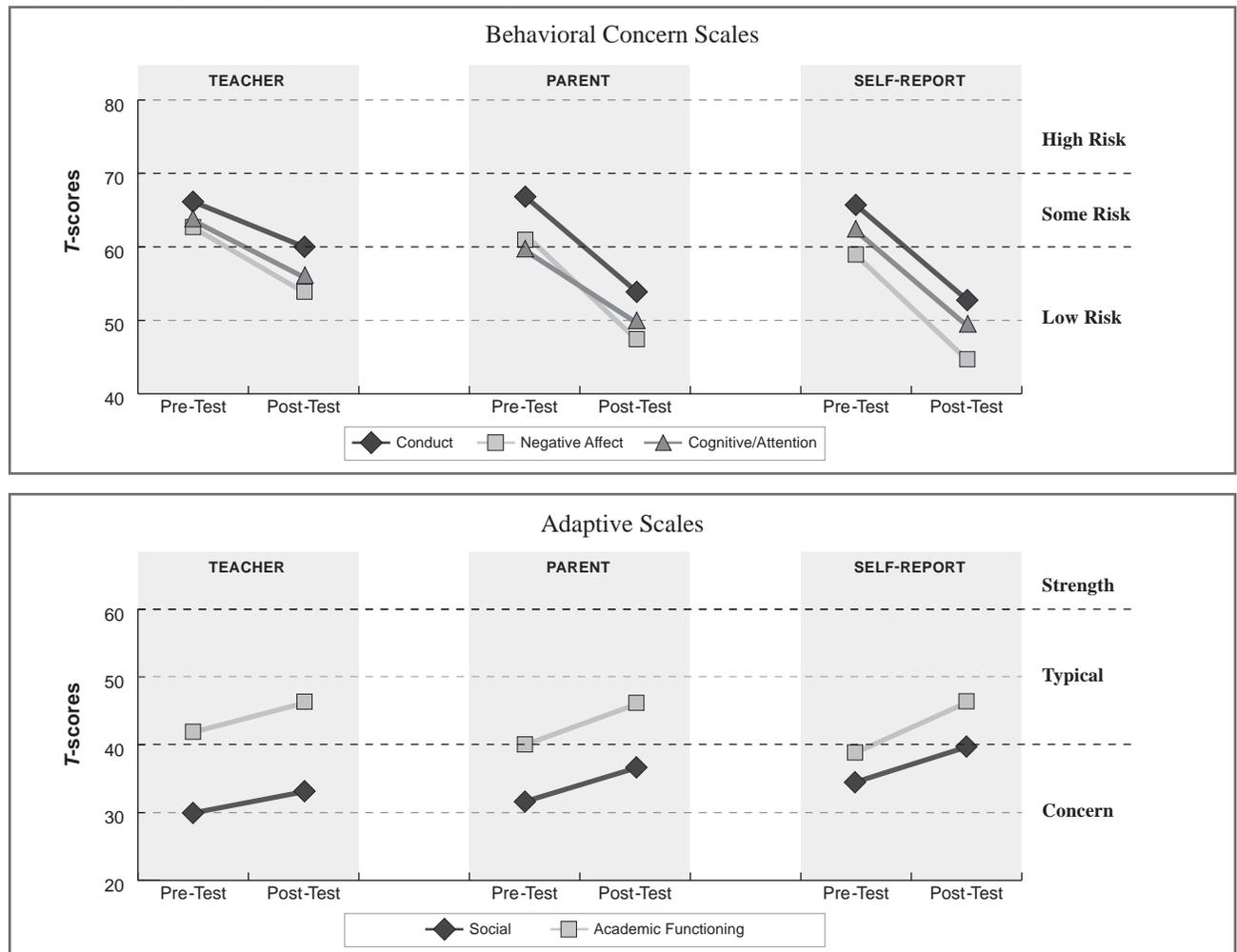
Thus, the above studies as well as other studies by Meier (1997, 1998, 2000) consistently provided empirical evidence that items (and their resulting aggregate scales) differ in their change-sensitivity compared to measures constructed using traditional, trait-based methods. Scales constructed using IISRs, like the BIMAS, produce briefer scales that are more sensitive to the effects produced by psychosocial interventions.

## Change-Sensitivity of the BIMAS

During the standardization of the BIMAS, a validity study was conducted to monitor the effectiveness of using the BIMAS as a progress monitoring tool. Specifically, a few sites that participated in the data collection also agreed to offer post-treatment data on some of the students who were receiving mental health services. In one such study, the BIMAS was used to monitor the progress of 46 students who participated in an anger management intervention. The sample consisted of 32 males and 14 females. There were 30 African American, 2 Hispanic, and 14 White students. The

students' ages ranged from 12 to 18 years old. The BIMAS was used to collect data prior to, and after, the completion of the intervention. The students' behaviors were rated by their teachers, parents, and by the youth themselves. Figure 11.4 and Tables 11.28 to 11.30 present the pre- and post-intervention data for the anger management group. Results from this study provide evidence that the BIMAS is sensitive to change, as it was effectively able to capture significant decreases in the Behavioral Concern scales as well as significant increases in the Adaptive Skills scales. Effect sizes (presented as Cohen's  $d$  ratios) revealed a large effect (i.e., all effect sizes  $\geq |0.8|$ ) across all rater-types and all scales, with the exception of the Social scale on the BIMAS-T where a moderate effect was found ( $d = -0.7$ ). All effect sizes were in the expected direction, indicating an improvement in functioning: positive effect sizes were found on the Behavioral Concern scales (i.e., scores decreased from pre-test to post-test), and negative effect sizes were found on the Adaptive scales (i.e., scores increased from pre-test to post-test).

Figure 11.4. Anger Management Treatment Group: Pre- to Post-Treatment BIMAS Standard T-scores



**Table 11.28. Pre- to Post-Performance of an Anger Management Treatment Group: BIMAS–T Standard T-Scores**

BIMAS–T Scale		Pre-Test	Post-Test	<i>t</i>	Cohen's <i>d</i>
Conduct	<i>M</i>	65.9	59.3	9.2	1.5
	<i>SD</i>	4.8	3.7		
Negative Affect	<i>M</i>	63.0	53.9	6.6	1.0
	<i>SD</i>	10.7	7.7		
Cognitive/Attention	<i>M</i>	63.3	55.3	7.3	1.2
	<i>SD</i>	6.6	6.9		
Social	<i>M</i>	30.0	34.4	-3.4	-0.7
	<i>SD</i>	5.5	7.2		
Academic Functioning	<i>M</i>	41.9	45.7	-5.2	-0.8
	<i>SD</i>	4.9	4.1		

Note. *N* = 46. All *ts* significant, *p* < .01. Cohen's *d* values of |0.2| = small effect, |0.5| = medium effect, and |0.8| = large effect. Positive *d* values indicate that the Pre-Test score was higher than the Post-Test score; negative *d* values indicate that the Pre-Test score was lower than the post-test score.

**Table 11.29. Pre- to Post-Performance of an Anger Management Treatment Group: BIMAS–P Standard T-Scores**

BIMAS–P Scale		Pre-Test	Post-Test	<i>t</i>	Cohen's <i>d</i>
Conduct	<i>M</i>	66.6	53.5	12.7	2.6
	<i>SD</i>	5.8	4.3		
Negative Affect	<i>M</i>	60.8	47.1	10.4	1.7
	<i>SD</i>	9.5	6.9		
Cognitive/Attention	<i>M</i>	59.4	49.5	10.3	2.0
	<i>SD</i>	5.4	4.6		
Social	<i>M</i>	31.7	37.5	-4.7	-1.0
	<i>SD</i>	4.9	6.9		
Academic Functioning	<i>M</i>	40.0	45.7	-7.3	-1.3
	<i>SD</i>	4.4	4.1		

Note. *N* = 46. All *ts* significant, *p* < .001. Cohen's *d* values of |0.2| = small effect, |0.5| = medium effect, and |0.8| = large effect. Positive *d* values indicate that the Pre-Test score was higher than the Post-Test score; negative *d* values indicate that the Pre-Test score was lower than the Post-Test score.

**Table 11.30. Pre- to Post-Performance of an Anger Management Treatment Group: BIMAS–SR Standard T-Scores**

BIMAS–SR Scale		Pre-Test	Post-Test	<i>t</i>	Cohen's <i>d</i>
Conduct	<i>M</i>	65.5	52.2	13.8	2.8
	<i>SD</i>	5.4	3.8		
Negative Affect	<i>M</i>	59.2	44.6	11.5	1.8
	<i>SD</i>	9.8	6.5		
Cognitive/Attention	<i>M</i>	62.7	49.6	12.9	2.4
	<i>SD</i>	6.6	4.2		
Social	<i>M</i>	35.1	39.5	-4.5	-0.8
	<i>SD</i>	6.2	4.8		
Academic Functioning	<i>M</i>	38.9	46.2	-10.1	-1.8
	<i>SD</i>	5.0	3.0		

Note. *N* = 46. All *ts* significant at *p* < .001. Cohen's *d* values of |0.2| = small effect, |0.5| = medium effect, and |0.8| = large effect. Positive *d* values indicate that the Pre-Test score was higher than the Post-Test score; negative *d* values indicate that the Pre-Test score was lower than the Post-Test score.

# 12 Concluding Comments

In recent years, as the Response to Intervention (RTI) movement in relation to Curriculum-Based Measurements (CBM) testing has been gaining momentum across many schools and districts, educators are beginning to realize that CBM benchmark test scores are only one piece of the puzzle in improving students' overall performance. More and more emphasis has been placed on the early identification and regular monitoring of behavioral/emotional/social problems in the schools. After all, if students are not behaving in the classroom, it is doubtful that much learning can be done, no matter how many CBM benchmark tests are given. Often times, students' academic scores naturally improve without any intervention in literacy or mathematics as soon as behavioral problems (e.g., lack of concentration, deviance, fidgeting) are addressed.

The RTI movement, in relation to curriculum performance, is a relatively recent movement. The application of the RTI framework in the student behavioral progress monitoring is an even newer concept. Despite its infancy, a recent online Google™ search for behavioral RTI generated over one million hits devoted to this topic, attesting to the widespread interest for schools and districts to proactively screen and monitor students, as well as the need to systematically organize all these behavioral data. It was at this perfect time that the Behavior Intervention Monitoring Assessment System (BIMAS™) was developed to provide school districts with the tool required to perform precisely these tasks.

As this *BIMAS Technical Manual* describes, the BIMAS was based on years of empirical research aimed at identifying items that are sensitive to change in response to intervention. This development process began even before the introduction of the RTI movement. The fact that BIMAS items were developed from the ground-up with the IISRs (Meier, 1997, 1998, 2000, 2004) means that the items are particularly suited for monitoring progress/change in the state of a youth.

This technical manual serves not only as a comprehensive reference to the BIMAS, but also as a general resource that provides insight into conducting behavioral, emotional, and social screening across the 3-Tiers in Kindergarten through Grade 12 and the use of the resultant data to assist in the design of district-/school-/grade-/class-wide or individual student-level interventions. The manual's coverage spans a presentation of the tool's background goal and rationale, a description of the measure, the BIMAS development process, as well as complete instructions regarding use, administration, scoring, and interpretation of the BIMAS. The importance of linking assessment to interventions is reflected in the manual by providing step-by-step guidelines on how the BIMAS results can be used to guide the planning and monitoring of interventions at the group level as well as the individual student level. As illustrated in the case study and throughout this manual, the BIMAS is suitable for use as both a screening and progress monitoring tool.

The BIMAS has been the subject of extensive statistical analysis, as evidenced in the appendices and the chapters on its reliability and validity. The BIMAS was normed using a sample of 3,500 respondents that was representative of the U.S. population.

The BIMAS is a useful tool in the screening process for identifying youth who may be at-risk for developing conduct, negative affective, and cognitive/attentional behavioral concerns. Furthermore, it allows for the identification of strengths in youth, as it contains a strength-based component for assessing social and academic functioning. The multi-informant forms help to elicit information from different raters, to shape treatment plans, and to measure treatment outcomes with the highly customizable Flex forms. The BIMAS must, however, be used as part of a dynamic evaluation process which synthesizes several types of information and careful clinical judgment. Decisions regarding diagnosis or high-stake placements should not be made solely on the basis of BIMAS results.

As research and experience continue to increase our collective understanding of issues relating to behavioral RTI, it is hoped that research with the BIMAS will play an integral role in shaping the future of RTI. The continuation of the highly beneficial iterative process is encouraged. We hope that the BIMAS will be the impetus for further thought, refinements, and improvements regarding the behavioral, emotional, and social issues that affect school-aged children and, likewise, that the ever-changing conceptualization of school-wide screening and progress monitoring will continue to lead to evolution in the BIMAS assessment.

Since the BIMAS is a web-based assessment, updates will be available to BIMAS users. In addition, BIMAS users can consult the MHS web site ([www.mhs.com](http://www.mhs.com)) for the latest BIMAS advances. Interested readers are invited and encouraged to send any comments or research results to the author or the publisher through the publisher's mailing address, below. Feedback will help in the continuing refinement of the BIMAS, with the ultimate goal being an instrument that is maximally reliable and valid, as well as useful to assessors.

Email may be directed to the MHS Research and Development Department at [r&d@mhs.com](mailto:r&d@mhs.com). Comments or research summaries/inquiries may be mailed to:

**In the U.S.A.**

Research and Development  
Multi-Health Systems Inc.  
P.O. Box 950, North Tonawanda, NY 14120-0950

**In Canada**

Research and Development  
Multi-Health Systems Inc.  
3770 Victoria Park Avenue, Toronto, ON M2H 3M6

# References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4<sup>th</sup> ed., text revision). Washington, DC: American Psychiatric Association.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- Anastasi, A., & Urbina S. (1997). *Psychological testing* (7<sup>th</sup> ed.). Upper Saddle River, NJ: Prentice Hall.
- Bardos, A. N. (2004). *Basic Achievement Skills Inventory (BASI)-Comprehensive Version administration manual*. San Antonio, TX: Pearson.
- Barrett, L. B. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality & Social Psychology Review, 10*, 20–46.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*, 588–606.
- Botcheva, L., White, C. R., & Huffman, L. C. (2002). Learning culture and outcomes measurement practices in community agencies. *American Journal of Evaluation, 23*, 421–434.
- Brandenburg, N. A., Friedman, R. M., & Silver, S. E. (1990). The epidemiology of childhood psychiatric disorders: Prevalence findings from recent studies. *Journal of the American Academy of Child and Adolescent Psychiatry, 29*, 76–83.
- Brems, C. (1993). *A comprehensive guide to child psychotherapy*. Boston: Allyn & Bacon.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models*. Newbury Park, CA: Sage.
- Brown, R. E., & Reed, C. S. (2002). An integral approach to evaluating outcome evaluation training. *American Journal of Evaluation, 23*, 1–17.
- Burlingame, G. M., Seaman, S., Johnson, J. E. Whipple, J., Richardson, E., Rees, F., Earnshaw, D., Spencer, R., Payne, M., & O'Neil, B. (2006). Sensitivity to change of the Brief Psychiatric Rating Scale—Extended (BPRS-E): An item and subscale analysis. *Psychological Services, 3*, 77–87.
- Cattell, R. B., & Burdsal, C. A. (1975). The radial parcel double factoring design: A solution to the item-vs.-parcel controversy. *Multivariate Behavioral Research, 10*, 165–179.
- Clement, P. W. (1999). *Outcomes and incomes: How to evaluate, improve, and market your psychotherapy practice by measuring outcomes*. New York: Guilford Press.
- Cohen, J. (1988). *Statistical power analysis in the behavioral sciences* (2<sup>nd</sup> ed.). Hillsdale, NJ: Erlbaum.
- Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology, 55*, 584–594.
- Collet, B., Ohan, J., & Myers, K. (2003). Ten-year review of rating scales. V: Scales assessing attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 42*, 1015–1037.
- Collins, L. M. (1991). Measurement in longitudinal research. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 137–148). Washington, DC: American Psychological Association.
- Conners, C. K. (2008). *Conners Comprehensive Behavior Rating Scales manual*. Toronto, Canada: Multi-Health Systems.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.
- Cronbach, L. J. (1951). *Coefficient alpha and the internal structure of tests*. *Psychometrika, 16*, 297–335.
- Cross, K. P., & Angelo, T. A. (1988). *Classroom assessment techniques*. Ann Arbor, MI: National Center for Research to Improve Postsecondary Teaching and Learning.
- Danziger, K. (1990). *Constructing the subject: Historical origins of psychological research*. New York: Cambridge University Press.
- Davis, S., & Meier, S. (2001). *The elements of managed care: A guide for helping professionals*. Pacific Grove, CA: Brooks/Cole.

- Dawis, R. (1992). Scale construction. In A. E. Kazdin (Ed.), *Methodological issues & strategies in clinical research* (pp. 193–213). Washington, DC: American Psychological Association.
- Dillman, D. A. (1978). *Mail and telephone surveys: The total design method*. New York: Wiley.
- Doss, L., & Reichle, J. (1991). Replacing excess behavior with an initial communicative repertoire. In J. Reichle, J. York, & J. Sigafoos (Eds.), *Implementing augmentative and alternative communication: Strategies for learners with severe disabilities* (pp. 215–237). Baltimore: Brooks.
- Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, *37*, 1097–1126.
- Epstein, S. (1980). The stability of behavior II. Implications for psychological research. *American Psychologist*, *35*, 790–806.
- Erbes, C., Polusny, M., Billig, J., Mylan, M., McGuire, K., Isenhardt, C., & Olson, D. (2004). Developing and applying a systematic process for evaluation of clinical outcome assessment instruments. *Psychological Services*, *1*, 31–39.
- Fiona, J. (2005, December). Response to Intervention in the Individuals with Disabilities Education Act (IDEA), 2004. Retrieved from International Reading Association, [http://www.reading.org/downloads/resources/IDEA\\_RTI\\_report.pdf](http://www.reading.org/downloads/resources/IDEA_RTI_report.pdf)
- Flesch, R. (1948). *A new readability yardstick*. *Journal of Applied Psychology*, *32*, 221–233.
- Froyd, J. E., Lambert, M. J., & Froyd, J. D. (1996). A review of practices of psychotherapy outcome measurement. *Journal of Mental Health*, *5*, 11–15.
- Gerrity, K. M., Jones, F. A., & Self, P. A. (1983). Developmental psychology for the clinical child psychologist. In C. E. Walker & M. Roberts (Eds.), *Handbook of Clinical Child Psychology*. New York: Wiley.
- Gibbs, D., Napp, D., Jolly, D., Westover, B., & Uhl, G. (2002). Increasing evaluation capacity within community-based HIV prevention programs. *Evaluation and Program Planning*, *25*, 216–269.
- Gresham, F. M. (2008). Best practices in diagnosis in a multitier problem-solving approach. In A. Thomas & J. Grimes (Eds.) *Best practices in school psychology V* (pp. 281–294). Bethesda, MD: National Association of School Psychologists.
- Guyatt, G., Walter, S., & Norman, G. (1987). Measuring change over time: Assessing the usefulness of evaluative instruments. *Journal of Chronic Disease*, *40*, 171–178.
- Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5<sup>th</sup> ed., pp. 84–135). New York: Wiley.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.
- Hughey, J., & Burdsal, C. (1982). 16PF-E structure using radial parcels versus items. *Journal of General Psychology*, *107*, 107–119.
- Individuals with Disabilities Education Improvement Act of 2004 (IDEA), Pub. L. No. 108–446, 118 Stat. 2647 (2004). [Amending 20 U.S.C. 1400 et seq.].
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.
- John, O. P., & Benet-Martínez, V. (2000). Measurement, scale construction, and reliability. In H. T. Reis and C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 339–369). New York, NY: Cambridge University Press.
- Kazdin, A. E. (1993). Treatment of conduct disorder: Progress and directions in psychotherapy research. *Development and Psychopathology*, *5*, 277–310.
- Kazdin, A. E. (2000). *Psychotherapy for children and adolescents: directions for research and practice*. New York: Oxford Press.
- Kessel, J. B., & Zimmerman, M. (1993). Reporting errors in studies of the diagnostic performance of self-administered questionnaires: Extent of the problem, recommendations for standardized presentation of results, and implications for the peer review process. *Psychological Assessment*, *5*, 395–399.
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel* (Research Branch Rep. No. 8–75). Memphis, TN: Naval Air Station.
- Kingery, P. M., & Walker, H. M. (2002). What we know about school safety. In M. Shinn, H. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches*. Bethesda, MD: National Association of School Psychologists.
- Klotz, M. B., & Canter, A., (2007). Response to Intervention: A primer for parents. National Association of School Psychologist (NASP) position paper. Retrieved from <http://www.nasponline.org/resources/handouts/rtiprimer.pdf>
- Lambert, M. J. (1994). Use of psychological tests for outcome assessment. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (pp. 75–97). Hillsdale, NJ: Lawrence Erlbaum.

- Lambert, M. J., Hatch, D. R., Kingston, M. D., & Edwards, B. C. (1986). Zung, Beck, and Hamilton Rating Scales as measures of treatment outcome: A meta-analytic comparison. *Journal of Consulting and Clinical Psychology, 54*, 54–59.
- LeBuffe, P. A., & Naglieri, J. A. (2003). *Devereux Early Childhood Assessment Clinical Form (DECA-C)*. Lewisville, NC: Kaplan Press.
- Lerew, C. D. (2004). The use of a cognitive strategy as an academic and behavioral intervention for children with attention-deficit/hyperactivity disorder. ProQuest Information & Learning). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 64* (Electronic; Print)
- Lipsey, M. W. (1983). A scheme for assessing measurement sensitivity in program evaluation and other applied research. *Psychological Bulletin, 94*, 152–165.
- Lipsey, M. W. (1998). Design sensitivity: Statistical power for applied experimental research. In L. Bickman & D. J. Rog (Eds.), *Handbook of applied social research methods* (pp. 39–68). Thousand Oaks, CA: Sage.
- March, J. S., Herbert W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391–410.
- Marston, D. (2005). Tiers of intervention in responsiveness to intervention: Prevention outcomes and learning disabilities identification patterns. *Journal of Learning Disabilities, 38*, 539–544.
- McCurdy, B. L., Mannella, M. C., & Eldridge, N. (2003). Positive behavior support in urban schools: Can we prevent the escalation of antisocial behavior? *Journal of Positive Behavior Interventions, 5*, 158–170.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.
- McDougal, J. L., Chafouleas, S. M., & Waterman, B. (2006). *A practitioner's guide to functional behavioral assessment and intervention in schools*. Champaign, IL: Research Press.
- McDougal, J. L., Graney, S. B., Wright, J. A., & Ardoin, S. P. (2009). *RTI in practice: A practical guide to implementing effective evidence-based interventions in your school*. Toronto, Canada: Wiley & Sons.
- Meier, S. T. (1994). *The chronic crisis in psychological measurement and assessment*. New York: Academic Press.
- Meier, S. T. (1997). Nomothetic item selection rules for tests of psychological interventions. *Psychotherapy Research, 7*, 419–427.
- Meier, S. T. (1998). Evaluating change-based item selection rules. *Measurement and evaluation in counseling and development, 31*, 15–27.
- Meier, S. T. (2000). Treatment sensitivity of the PE Form of the Social Skills Rating Scales: Implications for test construction procedures. *Measurement and Evaluation in Counseling and Development, 33*, 144–156.
- Meier, S. T. (2004). Improving design sensitivity through intervention-sensitive measures. *American Journal of Evaluation, 25*, 321–334.
- Meier, S. T. (2008). *Measuring change in counseling and psychotherapy*. New York: Guilford.
- Meier, S. T., McDougal, J., & Bardos, A. (2008). Development of a change-sensitive outcome measure for children receiving counseling. *Canadian Journal of School Psychology, 23*, 148–160.
- Meier, S. T., & Letsch, E. (2000). Data collection issues in an urban community mental health center: What is necessary and sufficient information for outcome assessment? *Professional Psychology: Research and Practice, 31*, 409–411.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3<sup>rd</sup> ed., pp. 13–103). Washington, DC: American Council on Education and National Council on Measurement in Education.
- Metzler, C. W., Biglan, A., Rusby, J. C., & Sprague, J. R. (2001). Evaluation of a comprehensive behavior management program to improve school-wide positive behavior support. *Education and Treatment of Children, 24*, 448–479.
- Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The Partners for Change Outcome Management System. *Journal of Clinical Psychology, 61*, 199–208.
- Morrison, G. M., & Skiba, R. (2001). Predicting violence from school misbehavior: *Promises and perils. Psychology in the Schools, 2*, 173–184.
- Nakasato, J. (2000). Data-based decision making in Hawaii's behavior support effort. *Journal of Positive Behavior Interventions, 2*, 247–251.
- Naglieri, J. A., & Bardos, A. N. (1997). General Ability Measure for Adults. *Applied Psychology, 79*, 493–505.
- Naglieri, J. A., & Chambers, K. M. (2009). Psychometric issues and current scales for assessing Autism Spectrum Disorders. In S. Goldstein, J. A. Naglieri, & S. Ozonoff (Eds.), *Assessment of Autism Spectrum Disorders* (pp. 55–90). New York, NY: Guilford.
- Naglieri, J. A., LeBuffe, P. A., & Pfeiffer, S. I. (1994). *Devereux Scales of Mental Disorders*. San Antonio, TX: The Psychological Corporation.
- Naglieri, J. A., McNeish, T. J., & Bardos, A. N. (1991). *Draw A Person: Screening Procedure for Emotional Disturbance examiner's manual*. Austin, TX: PRO-ED Inc.

- Nelson, J. R., Martella, R. M., & Marchand-Martella, N. (2002). Maximizing student learning: The effects of a comprehensive school-based program for preventing problem behaviors. *Journal of Emotional and Behavioral Disorders, 10*, 136–148.
- Nikitina, A. (2006). SMART Goal Setting: A Surefire Way To Achieve Your Goals. Retrieved from <http://www.goal-setting-guide.com/smart-goals.html>
- Reichle, J., McEvoy, M., Davis, C., Rogers, E., Feeley, K., Johnston, S., et al. (1996). Coordinating preservice and in-service training of early interventionists to serve preschoolers who engage in challenging behavior. In L. K. Kogel, R. L. Koegel, & G. Dunlap (Eds.), *Positive behavioral support: Including people with difficult behavior in the community* (pp. 227-265). Baltimore: Brookes.
- Reid, J. B., & Patterson, G. R. (1991). Early prevention and intervention with conduct problems: A social interactional model for the integration of research and practice. In G. Stoner, M. R. Shinn, & H. M. Walker (Eds.), *Interventions for achievement and behavior problems* (pp. 715–740). Silver Spring, MD: National Association of School Psychologists.
- Spiegel, S. (1987). The interpersonal world of the infant. *Contemporary Psychoanalysis, 23*, 6–17.
- StatSoft. (2001). Statistica (Version 6) [Computer software]. Tulsa, OK: Author.
- Stiffman, A. R., Orme, J. G., Evans, D. A., Feldman, R. A., & Keeney, P. A. (1984). A brief measure of children's behavior problems: The Behavior Rating Index for Children. *Measurement and Evaluation in Counseling and Development, 16*, 83–90.
- Strupp, H. H., Horowitz, L. M. and Lambert, M. J. (Eds.). (1997). *Measuring patient changes in mood, anxiety, and personality disorders: Toward a core battery*. Washington, DC: American Psychological Association.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5<sup>th</sup> ed.). Boston: Pearson/Allyn & Bacon.
- Thayer, C. E., & Fine, A. H. (2001). Evaluation and outcome measurement in the non-profit sector: Stakeholder participation. *Evaluation & Program Planning, 24*, 103–108.
- Thorndike, R. M. (1982). *Data collection and analysis*. New York: Gardner.
- Tilly III, W. D. (2008). The evolution of school psychology to science based practice: Problem solving and the three-tiered model. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V* (pp. 17–36). Bethesda, MD: National Association of School Psychologists.
- Tobin, T., & Sugai, G. (1999). Using sixth grade school records to predict violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders, 7*, 40–53.
- Tryon, W. W. (1991). *Activity measurement in psychology and medicine*. New York: Plenum.
- Tufte, E. (1983). *The visual display of quantitative information*. Connecticut, USA: Graphic Press.
- U.S. Department of Education, Office of Special Education Programs. (2010). National Technical Assistance Center on Positive Behavioral Interventions and Supports. Retrieved from [www.phis.org](http://www.phis.org).
- Urbina, K. (2004). *Essentials of psychological testing*. New York: Wiley.
- Vermeersch, D. A., Whipple, J. L., Lambert, M. J., Hawkins, E. J., Burchfield, C. M., & Okiishi, J. C. (2004). Outcome questionnaire: Is it sensitive to changes in counseling center clients? *Journal of Counseling Psychology, 51*, 38–49.
- Walker, H. M., & Severson, H. (1992). *Systematic screening of behavior disorders (SSBD): A multiple gating procedure*. Longmont, CO: Sopris West.
- Walker, H. M., & Shinn, M. R. (2002). Structuring school-based interventions to achieve integrated primary, secondary and tertiary prevention goals for safe and effective schools. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems: Vol. II. Preventive and remedial approaches* (pp. 1–26). Bethesda, MD: National Association of School Psychologists.
- Weinstock, M., & Meier, S. T. (2003). A comparison of two item selection methodologies for measuring change in university counseling center clients. *Measurement and Evaluation in Counseling & Development, 36*, 66–75.
- Weisz, J. R., Huey, S. J., & Weersing, V. R. (1998). Psychotherapy outcome research with children and adolescents: The state of the art. In T. H. Ollendick & R. J. Prinz (Eds.), *Advances in Clinical Child Psychology* (Vol. 20, pp. 49–91). New York: Plenum.
- Weisz, J. R., Weiss, B., & Donenberg, G. R. (1992). The lab versus the clinic: Effects of child and adolescent psychotherapy. *American Psychologist, 47*, 1578–1585. New York: Wiley.
- World Health Organization. (2004). *International Statistical Classification of Diseases and Health Related Problems, 10th revision* (2<sup>nd</sup> ed.; ICD-10). Geneva: World Health Organization.
- Wright, J. A., & Dusek, J. B. (1998). Compiling base rates for disruptive behaviors from student discipline referral data. *School Psychology Review, 27*, 147–183.
- Zachary, R. A., & Gorsuch, R. L. (1985). Continuous norming: Implications for the WAIS-R. *Journal of Clinical Psychology, 41*, 86–94.

# Appendix A

## BIMAS Standard Items by Scale

This appendix provides a list of items by scale for all BIMAS™ Standard rater forms. The items by scale are identical across teacher, parent, and self-report forms (except that items on the self-report form are written in the first person voice). Clinician form items are listed separately (the item “attended his/her scheduled therapy appointments” appears only on the clinician-completed form, and the Academic Functioning scale [with the exception of the item “followed directions”] is excluded from this form). Reverse-scored items are noted with an (R).

### BIMAS–Teacher Standard & BIMAS–Parent Standard

#### BIMAS Behavioral Concern Scales

##### Conduct (9 items)

Item #	Item
2	appeared angry.
9	engaged in risk-taking behavior.
13	fought with others (verbally, physically, or both).
17	lied or cheated.
21	lost his/her temper when upset.
25	was aggressive (threatened or bullied others).
29	was suspected of using alcohol and/or drugs.
31	was sent to an authority for discipline.
32	was suspected of smoking or chewing tobacco.

##### Negative Affect (7 items)

Item #	Item
5	appeared sleepy or tired.
8	appeared depressed.
12	acted sad or withdrawn.
16	was easily embarrassed or felt ashamed.
20	appeared anxious (worried or nervous).
24	expressed thoughts of hurting himself/herself.
27	was emotional or upset.

##### Cognitive/Attention (7 items)

Item #	Item
3	had trouble paying attention.
6	was impulsive.
10	had problems staying on task.
14	acted without thinking.
18	had trouble remembering.
22	had trouble with organizing and planning.
28	fidgeted.

### BIMAS Adaptive Scales

##### Social (6 items)

Item #	Item
1	shared what he/she was thinking about.
7	spoke clearly with others.
11	maintained friendships.
15	appeared comfortable when relating to others.
19	was generally friendly with others.
23	worked out problems with others.

##### Academic Functioning (5 items)

Item #	Item
4	followed directions.
26 (R)	received failing grades at school.
30	worked up to his/her academic potential.
33	was prepared for class.
34 (R)	was absent from school.

**BIMAS–Self-Report Standard****BIMAS Behavioral Concern Scales****Conduct (9 items)**

Item #	Item
2	felt angry.
9	did something risky.
13	fought with others (verbally, physically, or both).
17	lied or cheated.
21	lost my temper when I was upset.
25	threatened or bullied others.
29	used alcohol and/or drugs.
31	was sent to an authority for discipline.
32	smoked or chewed tobacco.

**Negative Affect (7 items)**

Item #	Item
5	felt sleepy or tired.
8	was depressed.
12	was sad or withdrawn.
16	was easily embarrassed or felt ashamed.
20	was anxious (worried or nervous).
24	had thoughts of hurting myself.
27	felt emotional or upset.

**Cognitive/Attention (7 items)**

Item #	Item
3	had trouble paying attention.
6	was impulsive.
10	had problems staying on task.
14	acted without thinking.
18	had trouble remembering things.
22	had trouble with organizing and planning.
28	fidgeted.

**BIMAS Adaptive Scales****Social (6 items)**

Item #	Item
1	shared my thoughts with others.
7	communicated clearly with others.
11	maintained friendships.
15	felt relaxed interacting with others.
19	was friendly with others.
23	worked out problems with others.

**Academic Functioning (5 items)**

Item #	Item
4	followed directions.
26 (R)	received failing grades at school.
30	tried my hardest when it came to schoolwork.
33	went prepared for class.
34 (R)	was absent from school.

**BIMAS–Clinician Standard****BIMAS Behavioral Concern Scales****Conduct (9 items)**

Item #	Item
2	appeared angry.
9	engaged in risk-taking behavior.
13	fought with others (verbally, physically, or both).
17	lied or cheated.
21	lost his/her temper when upset.
25	was aggressive (threatened or bullied others).
28	was suspected of using alcohol and/or drugs.
29	was sent to an authority for discipline.
30	was suspected of smoking or chewing tobacco.

**Negative Affect (7 items)**

Item #	Item
5	appeared sleepy or tired.
8	appeared depressed.
12	acted sad or withdrawn.
16	was easily embarrassed or felt ashamed.
20	appeared anxious (worried or nervous).
24	expressed thoughts of hurting himself/herself.
26	was emotional or upset.

**Cognitive/Attention (7 items)**

Item #	Item
3	had trouble paying attention.
6	was impulsive.
10	had problems staying on task.
14	acted without thinking.
18	had trouble remembering.
22	had trouble with organizing and planning.
27	fidgeted.

**BIMAS Adaptive Scale & Clinician Form Adaptive Items****Social (6 items)**

Item #	Item
1	shared what he/she was thinking about.
7	spoke clearly with others.
11	maintained friendships.
15	appeared comfortable when relating to others.
19	was generally friendly with others.
23	worked out problems with others.

**Clinician Form Adaptive Items**

Item #	Item
4	followed directions.
31	attended his/her scheduled therapy appointments.

# Appendix B

## BIMAS Flex Items by Scale

This appendix provides a list of BIMAS™ Flex items for each of the Standard form anchor items. It also provides the valence for each Flex item (i.e., whether a Flex item was positively or negatively worded). The Flex items are identical across teacher, parent, self-report, and clinician forms (except that items on the self-report form are written in the first person voice). The clinician form has parallel Flex items across the Conduct, Negative Affect, Cognitive/Attention, and Social Scales. One Standard form item (“attended his/her scheduled therapy appointments”) and its corresponding list of Flex items appear only on the clinician-completed forms. Three Standard anchor items on the Academic Functioning scale from the other rater forms are also excluded from the Clinician form.

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Conduct Items				BIMAS Conduct Items
<b>2</b>	<b>appeared angry.</b>	<b>appeared angry.</b>	<b>felt angry.</b>	<b>2</b>	<b>appeared angry.</b>
Negative	was angry at peers.	was angry at peers.	was angry with my classmates/friends.	Negative	was angry at peers.
Negative	was angry at teachers.	was angry at teachers.	was angry with my teachers.	Negative	was angry at teachers.
Negative	was angry at parents.	was angry at parents.	was angry with my parents.	Negative	was angry at parents.
Negative	acted sulky or moody.	acted sulky or moody.	felt sulky or moody.	Negative	acted sulky or moody.
Negative	yelled at others.	yelled at others.	yelled at others.	Negative	yelled at others.
Negative	got frustrated easily.	got frustrated easily.	felt frustrated.	Negative	got frustrated easily.
Negative	got angry when feelings were hurt.	got angry when feelings were hurt.	got angry when someone hurt my feelings.	Negative	got angry when feelings were hurt.
Negative	destroyed things when angry.	destroyed things when angry.	broke things when I got angry.	Negative	destroyed things when angry.
Positive	was content.	was content.	felt content.	Positive	was content.
Positive	used a strategy to control emotions.	used a strategy to control emotions.	used a strategy to control my emotions.	Positive	used a strategy to control emotions.
Positive	showed frustration in a positive way.	showed frustration in a positive way.	showed frustration in a positive way.	Positive	showed frustration in a positive way.
Positive	controlled anger.	controlled anger.	controlled my anger.	Positive	controlled anger.
<b>9</b>	<b>engaged in risk-taking behavior.</b>	<b>engaged in risk-taking behavior.</b>	<b>did something risky.</b>	<b>9</b>	<b>engaged in risk-taking behavior.</b>
Negative	took part in risk-taking behaviors when upset.	took part in risk-taking behaviors when upset.	did something risky when I was upset.	Negative	took part in risk-taking behaviors when upset.
Negative	took part in risk-taking behaviors when anxious.	took part in risk-taking behaviors when anxious.	did something risky when I felt anxious.	Negative	took part in risk-taking behaviors when anxious.
Negative	broke the law.	broke the law.	broke the law.	Negative	broke the law.
Negative	spent time with friends who broke the law.	spent time with friends who broke the law.	spent time with friends who broke the law.	Negative	spent time with friends who broke the law.
Negative	set fires.	set fires.	set fires.	Negative	set fires.
Negative	played with matches.	played with matches.	played with matches.	Negative	played with matches.
Negative	took part in risky sports.	took part in risky sports.	took part in risky sports.	Negative	took part in risky sports.
Negative	used illegal drugs.	used illegal drugs.	used illegal drugs.	Negative	used illegal drugs.
Negative	drank alcohol.	drank alcohol.	drank alcohol.	Negative	drank alcohol.
Negative	used prescription drugs that he/she was not supposed to use.	used prescription drugs that he/she was not supposed to use.	used drugs that I'm not supposed to use.	Negative	used prescription drugs that he/she was not supposed to use.
Positive	made responsible choices.	made responsible choices.	made responsible choices.	Positive	made responsible choices.

*Continued...*

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Conduct Items			BIMAS Conduct Items	
Positive	stopped and thought before acting.	stopped and thought before acting.	stopped and thought before acting.	Positive	stopped and thought before acting.
Positive	abstained from alcohol.	abstained from alcohol.	stayed away from alcohol.	Positive	abstained from alcohol.
Positive	abstained from drugs.	abstained from drugs.	stayed away from drugs.	Positive	abstained from drugs.
Positive	abstained from using prescription drugs that he/she was not supposed to.	abstained from using prescription drugs that he/she was not supposed to.	stayed away from using drugs I'm not supposed to.	Positive	abstained from using prescription drugs that he/she was not supposed to.
Positive	used an appropriate strategy to calm down when upset.	used an appropriate strategy to calm down when upset.	used a strategy to calm myself down when upset.	Positive	used an appropriate strategy to calm down when upset.
Positive	used an appropriate strategy to calm down when anxious.	used an appropriate strategy to calm down when anxious.	used a strategy to calm myself down when I felt anxious.	Positive	used an appropriate strategy to calm down when anxious.
Positive	spent time with positive friends.	spent time with positive friends.	spent time with positive friends.	Positive	spent time with positive friends.
13	<b>fought with others (verbally, physically, or both).</b>	<b>fought with others (verbally, physically, or both).</b>	<b>fought with others (verbally, physically, or both).</b>	13	<b>fought with others (verbally, physically, or both).</b>
Negative	argued with peers.	argued with peers.	argued with my classmates/friends.	Negative	argued with peers.
Negative	argued with teachers.	argued with teachers.	argued with my teachers.	Negative	argued with teachers.
Negative	argued with parents.	argued with parents.	argued with my parents.	Negative	argued with parents.
Negative	argued with siblings.	argued with siblings.	argued with my siblings.	Negative	argued with siblings.
Negative	talked back to parents.	talked back to parents.	talked back to my parents.	Negative	talked back to parents.
Negative	talked back to teachers.	talked back to teachers.	talked back to my teachers.	Negative	talked back to teachers.
Negative	physically hurt peers.	physically hurt peers.	physically fought with my classmates/friends.	Negative	physically hurt peers.
Negative	physically hurt parents.	physically hurt parents.	physically fought with my parents.	Negative	physically hurt parents.
Negative	physically hurt teachers.	physically hurt teachers.	physically fought with my teachers.	Negative	physically hurt teachers.
Negative	physically hurt siblings.	physically hurt siblings.	physically fought with my siblings.	Negative	physically hurt siblings.
Negative	threatened peers.	threatened peers.	threatened my classmates/friends.	Negative	threatened peers.
Negative	threatened teachers.	threatened teachers.	threatened my teachers.	Negative	threatened teachers.
Negative	threatened parents.	threatened parents.	threatened my parents.	Negative	threatened parents.
Negative	threatened siblings.	threatened siblings.	threatened my siblings.	Negative	threatened siblings.
Negative	yelled at peers.	yelled at peers.	yelled at my classmates/friends.	Negative	yelled at peers.
Negative	yelled at teachers.	yelled at teachers.	yelled at my teachers.	Negative	yelled at teachers.
Negative	yelled at parents.	yelled at parents.	yelled at my parents.	Negative	yelled at parents.
Negative	yelled at siblings.	yelled at siblings.	yelled at my siblings.	Negative	yelled at siblings.
Negative	picked fights with others.	picked fights with others.	picked fights with others.	Negative	picked fights with others.
Negative	arranged for a peer to fight with someone.	arranged for a peer to fight with someone.	got classmates or friends to fight with each other.	Negative	arranged for a peer to fight with someone.
Negative	made others do what he/she wanted.	made others do what he/she wanted.	made others do what I wanted.	Negative	made others do what he/she wanted.
Negative	bullied others.	bullied others.	bullied others.	Negative	bullied others.
Positive	showed regret after a fight.	showed regret after a fight.	regretted fighting.	Positive	showed regret after a fight.
Positive	kept hands and feet to himself/herself.	kept hands and feet to himself/herself.	kept my hands and feet to myself.	Positive	kept hands and feet to himself/herself.
Positive	was respectful to adults.	was respectful to adults.	was respectful to adults.	Positive	was respectful to adults.
Positive	walked away from a fight.	walked away from a fight.	walked away from a fight.	Positive	walked away from a fight.
Positive	prevented a fight.	prevented a fight.	prevented a fight.	Positive	prevented a fight.
Positive	stopped an argument.	stopped an argument.	stopped an argument.	Positive	stopped an argument.
Positive	managed frustration well.	managed frustration well.	managed my frustration.	Positive	managed frustration well.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Conduct Items			BIMAS Conduct Items	
<b>Positive</b>	avoided an argument.	avoided an argument.	avoided an argument.	<b>Positive</b>	avoided an argument.
<b>17</b>	<b>lied or cheated.</b>	<b>lied or cheated.</b>	<b>lied or cheated.</b>	<b>17</b>	<b>lied or cheated.</b>
<b>Negative</b>	lied to parents.	lied to parents.	lied to my parents.	<b>Negative</b>	lied to parents.
<b>Negative</b>	lied to teachers.	lied to teachers.	lied to my teachers.	<b>Negative</b>	lied to teachers.
<b>Negative</b>	lied to peers.	lied to peers.	lied to my classmates/ friends.	<b>Negative</b>	lied to peers.
<b>Negative</b>	denied doing something that others had seen him/ her do.	denied doing something that others had seen him/ her do.	denied doing things that others had seen me do.	<b>Negative</b>	denied doing something that others had seen him/ her do.
<b>Negative</b>	cheated in a game.	cheated in a game.	cheated in a game.	<b>Negative</b>	cheated in a game.
<b>Negative</b>	cheated on a test.	cheated on a test.	cheated on a test.	<b>Negative</b>	cheated in a test.
<b>Negative</b>	cheated on homework.	cheated on homework.	cheated on homework.	<b>Negative</b>	cheated on homework.
<b>Negative</b>	blamed others for his/her own actions.	blamed others for his/her own actions.	blamed others for my actions.	<b>Negative</b>	blamed others for his/her own actions.
<b>Positive</b>	took responsibility for his/her actions.	took responsibility for his/her actions.	took responsibility for my actions.	<b>Positive</b>	took responsibility for his/her actions.
<b>Positive</b>	played by the rules.	played by the rules.	played by the rules.	<b>Positive</b>	played by the rules.
<b>Positive</b>	was honest.	was honest.	was honest.	<b>Positive</b>	was honest.
<b>Positive</b>	was fair with others.	was fair with others.	was fair with others.	<b>Positive</b>	was fair with others.
<b>21</b>	<b>lost his/her temper when upset.</b>	<b>lost his/her temper when upset.</b>	<b>lost my temper when I was upset.</b>	<b>21</b>	<b>lost his/her temper when upset.</b>
<b>Negative</b>	lost his/her temper when challenged.	lost his/her temper when challenged.	lost my temper when I was challenged.	<b>Negative</b>	lost his/her temper when challenged.
<b>Negative</b>	overreacted to mild conflicts with peers.	overreacted to mild conflicts with peers.	overreacted to mild conflicts with classmates/ friends.	<b>Negative</b>	overreacted to mild conflicts with peers.
<b>Negative</b>	was difficult to calm down when upset.	was difficult to calm down when upset.	had a hard time calming down after I was upset.	<b>Negative</b>	was difficult to calm down when upset.
<b>Negative</b>	swore or used bad language when upset.	swore or used bad language when upset.	swore or used bad language when I was upset.	<b>Negative</b>	swore or used bad language when upset.
<b>Negative</b>	had a tantrum.	had a tantrum.	had a tantrum or blow-up.	<b>Negative</b>	had a tantrum.
<b>Negative</b>	shouted or yelled at others in anger.	shouted or yelled at others in anger.	shouted or yelled at others in anger.	<b>Negative</b>	shouted or yelled at others in anger.
<b>Negative</b>	hit or kicked things when upset.	hit or kicked things when upset.	hit or kicked things when I was upset.	<b>Negative</b>	hit or kicked things when upset.
<b>Positive</b>	was patient.	was patient.	was patient.	<b>Positive</b>	was patient.
<b>Positive</b>	managed frustration well.	managed frustration well.	managed my frustration.	<b>Positive</b>	managed frustration well.
<b>Positive</b>	reacted calmly to small conflicts.	reacted calmly to small conflicts.	reacted calmly to small conflicts.	<b>Positive</b>	reacted calmly to small conflicts.
<b>Positive</b>	used a strategy to calm down when upset.	used a strategy to calm down when upset.	did something to calm myself down when I was upset.	<b>Positive</b>	used a strategy to calm down when upset.
<b>Positive</b>	walked away from conflict.	walked away from conflict.	walked away from conflict.	<b>Positive</b>	walked away from conflict.
<b>Positive</b>	was able to control his/her emotions.	was able to control his/her emotions.	was able to control my emotions.	<b>Positive</b>	was able to control his/her emotions.
<b>25</b>	<b>was aggressive (threatened or bullied others).</b>	<b>was aggressive (threatened or bullied others).</b>	<b>threatened or bullied others.</b>	<b>25</b>	<b>was aggressive (threatened or bullied others).</b>
<b>Negative</b>	physically hurt peers.	physically hurt peers.	physically hurt my classmates/friends.	<b>Negative</b>	physically hurt peers.
<b>Negative</b>	physically hurt parents.	physically hurt parents.	physically hurt my parents.	<b>Negative</b>	physically hurt parents.
<b>Negative</b>	physically hurt teachers.	physically hurt teachers.	physically hurt my teachers.	<b>Negative</b>	physically hurt teachers.
<b>Negative</b>	threatened peers.	threatened peers.	threatened my classmates/ friends.	<b>Negative</b>	threatened peers.
<b>Negative</b>	threatened teachers.	threatened teachers.	threatened my teachers.	<b>Negative</b>	threatened teachers.
<b>Negative</b>	threatened parents.	threatened parents.	threatened my parents.	<b>Negative</b>	threatened parents.
<b>Negative</b>	intimidated a teacher.	intimidated a teacher.	intimidated a teacher.	<b>Negative</b>	intimidated a teacher.

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Conduct Items			BIMAS Conduct Items	
Negative	intimidated a parent.	intimidated a parent.	intimidated my parent.	Negative	intimidated a parent.
Negative	intimidated a peer.	intimidated a peer.	intimidated a classmate/ friend.	Negative	intimidated a peer.
Negative	bullied others.	bullied others.	bullied others.	Negative	bullied others.
Negative	picked on other children.	picked on other children.	picked on other children.	Negative	picked on other children.
Negative	got others to gang up on peers.	got others to gang up on peers.	got others to gang up on classmates/friends.	Negative	got others to gang up on peers.
Negative	got others to do violent things.	got others to do violent things.	got others to do violent things.	Negative	got others to do violent things.
Negative	was aggressive without being provoked.	was aggressive without being provoked.	threatened or bullied others without being provoked.	Negative	was aggressive without being provoked.
Negative	was aggressive when provoked.	was aggressive when provoked.	threatened or bullied others when provoked.	Negative	was aggressive when provoked.
Negative	spread rumors about others.	spread rumors about others.	spread rumors about others.	Negative	spread rumors about others.
Negative	left others out of group activities on purpose.	left others out of group activities on purpose.	left others out of group activities on purpose.	Negative	left others out of group activities on purpose.
Negative	said mean things to others.	said mean things to others.	said mean things to others.	Negative	said mean things to others.
Negative	started a fight.	started a fight.	started a fight.	Negative	started a fight.
Negative	bit or spit.	bit or spit.	bit or spit.	Negative	bit or spit.
Negative	annoyed others on purpose.	annoyed others on purpose.	annoyed others on purpose.	Negative	annoyed others on purpose.
Negative	was cruel to animals.	was cruel to animals.	was cruel to animals.	Negative	was cruel to animals.
Negative	destroyed things that belong to others.	destroyed things that belong to others.	destroyed things that belong to others.	Negative	destroyed things that belong to others.
Negative	swore or used bad language.	swore or used bad language.	swore or used bad language.	Negative	swore or used bad language.
Positive	treated others' property with respect.	treated others' property with respect.	treated others' things with respect.	Positive	treated others' property with respect.
Positive	got along well with others.	got along well with others.	got along well with others.	Positive	got along well with others.
Positive	showed regret after a fight.	showed regret after a fight.	regretted fighting.	Positive	showed regret after a fight.
Positive	kept hands and feet to himself/herself.	kept hands and feet to himself/herself.	kept my hands and feet to myself.	Positive	kept hands and feet to himself/herself.
Positive	avoided verbal arguments.	avoided verbal arguments.	avoided arguments.	Positive	avoided verbal arguments.
Positive	refrained from making threats.	refrained from making threats.	avoided making threats.	Positive	refrained from making threats.
Positive	refrained from bullying others.	refrained from bullying others.	avoided bullying others.	Positive	refrained from bullying others.
Positive	refrained from saying things about others that were not true.	refrained from saying things about others that were not true.	avoided saying things about others that were not true.	Positive	refrained from saying things about others that were not true.
Positive	was respectful of others' feelings.	was respectful of others' feelings.	was respectful of others' feelings.	Positive	was respectful of others' feelings.
29	<b>was suspected of using alcohol and/or drugs.</b>	<b>was suspected of using alcohol and/or drugs.</b>	<b>used alcohol and/or drugs.</b>	28	<b>was suspected of using alcohol and/or drugs.</b>
Negative	was suspected of drinking alcohol.	was suspected of drinking alcohol.	drank alcohol.	Negative	was suspected of drinking alcohol.
Negative	was suspected of using drugs.	was suspected of using drugs.	used drugs that I'm not supposed to.	Negative	was suspected of using drugs.
Negative	was caught with alcohol.	was caught with alcohol.	got caught with alcohol.	Negative	was caught with alcohol.
Negative	was caught with drugs.	was caught with drugs.	got caught with drugs.	Negative	was caught with drugs.
Negative	spent time with peers who use drugs.	spent time with peers who use drugs.	spent time with friends who use drugs.	Negative	spent time with peers who use drugs.
Negative	spent time with peers who drink alcohol.	spent time with peers who drink alcohol.	spent time with friends who drink alcohol.	Negative	spent time with peers who drink alcohol.
Negative	was under the influence of alcohol.	was under the influence of alcohol.	drank alcohol.	Negative	was under the influence of alcohol.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Conduct Items			BIMAS Conduct Items	
Negative	was under the influence of drugs.	was under the influence of drugs.	used drugs that I'm not supposed to.	Negative	was under the influence of drugs.
Negative	spoke about drinking alcohol.	spoke about drinking alcohol.	talked about drinking alcohol.	Negative	spoke about drinking alcohol.
Negative	spoke about using drugs.	spoke about using drugs.	talked about using drugs.	Negative	spoke about using drugs.
Positive	abstained from alcohol.	abstained from alcohol.	stayed away from alcohol.	Positive	abstained from alcohol.
Positive	abstained from drugs.	abstained from drugs.	stayed away from drugs.	Positive	abstained from drugs.
Positive	refrained from talking about alcohol.	refrained from talking about alcohol.	refrained from talking about alcohol.	Positive	refrained from talking about alcohol.
Positive	refrained from talking about drugs.	refrained from talking about drugs.	refrained from talking about drugs.	Positive	refrained from talking about drugs.
Positive	was sober and coherent.	was sober and coherent.	was sober.	Positive	was sober and coherent.
<b>31</b>	<b>was sent to an authority for discipline.</b>	<b>was sent to an authority for discipline.</b>	<b>was sent to an authority for discipline.</b>	<b>29</b>	<b>was sent to an authority for discipline.</b>
Negative	used bad language toward a teacher.	used bad language toward a teacher.	used bad language toward a teacher.	Negative	used bad language toward a teacher.
Negative	used bad language toward a parent.	used bad language toward a parent.	used bad language toward my parent.	Negative	used bad language toward a parent.
Negative	used bad language toward a peer.	used bad language toward a peer.	used bad language toward a classmate/friend.	Negative	used bad language toward a peer.
Negative	ruined or damaged property.	ruined or damaged property.	caused damage to property.	Negative	ruined or damaged property.
Negative	set fire(s).	set fire(s).	set fire(s).	Negative	set fire(s).
Negative	assaulted a teacher.	assaulted a teacher.	assaulted a teacher.	Negative	assaulted a teacher.
Negative	assaulted a peer.	assaulted a peer.	assaulted a friend/classmate.	Negative	assaulted a peer.
Negative	assaulted a parent.	assaulted a parent.	assaulted my parent.	Negative	assaulted a parent.
Positive	was respectful to teacher(s).	was respectful to teacher(s).	was respectful to teacher(s).	Positive	was respectful to teacher(s).
Positive	was respectful to parent(s).	was respectful to parent(s).	was respectful to my parent(s).	Positive	was respectful to parent(s).
Positive	was respectful to peers.	was respectful to peers.	was respectful to my classmates and friends.	Positive	was respectful to peers.
Positive	was respectful of others' things.	was respectful of others' things.	was respectful of others' things.	Positive	was respectful of others' things.
Positive	used acceptable language at school.	used acceptable language at school.	used acceptable language at school.	Positive	used acceptable language at school.
Positive	followed rules.	followed rules.	followed rules.	Positive	followed rules.
Positive	followed adult requests without arguments.	followed adult requests without arguments.	did what an adult asked me to do without arguing.	Positive	followed adult requests without arguments.
<b>32</b>	<b>was suspected of smoking or chewing tobacco.</b>	<b>was suspected of smoking or chewing tobacco.</b>	<b>smoked or chewed tobacco.</b>	<b>30</b>	<b>was suspected of smoking or chewing tobacco.</b>
Negative	was suspected of smoking tobacco.	was suspected of smoking tobacco.	smoked tobacco.	Negative	was suspected of smoking tobacco.
Negative	was observed using tobacco.	was observed using tobacco.	was seen smoking tobacco.	Negative	was observed using tobacco.
Negative	spoke about using tobacco.	spoke about using tobacco.	talked about using tobacco.	Negative	spoke about using tobacco.
Negative	spent time with peers who use tobacco.	spent time with peers who use tobacco.	spent time with friends who use tobacco.	Negative	spent time with peers who use tobacco.
Positive	abstained from smoking tobacco.	abstained from smoking tobacco.	stayed away from smoking tobacco.	Positive	abstained from smoking tobacco.
Positive	refrained from talking about tobacco.	refrained from talking about tobacco.	refrained from talking about tobacco.	Positive	refrained from talking about tobacco.

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Negative Affect Items				BIMAS Negative Affect Items
<b>5</b>	<b>appeared sleepy or tired.</b>	<b>appeared sleepy or tired.</b>	<b>felt sleepy or tired.</b>	<b>5</b>	<b>appeared sleepy or tired.</b>
Negative	had trouble sleeping.	had trouble sleeping.	had a hard time sleeping.	Negative	had trouble sleeping.
Negative	slept too much.	slept too much.	felt I slept more than I should have.	Negative	slept too much.
Negative	was tired.	was tired.	felt tired.	Negative	was tired.
Negative	had low energy.	had low energy.	felt sluggish.	Negative	had low energy.
Negative	fell asleep in class.	fell asleep in class.	fell asleep in class.	Negative	fell asleep in class.
Negative	worried so much that he/she had trouble sleeping.	worried so much that he/she had trouble sleeping.	worried so much that I had trouble sleeping.	Negative	worried so much that he/she had trouble sleeping.
Positive	was well rested.	was well rested.	felt well rested.	Positive	was well rested.
Positive	was alert.	was alert.	felt alert.	Positive	was alert.
Positive	was energetic.	was energetic.	felt energetic.	Positive	was energetic.
Positive	was active.	was active.	was active.	Positive	was active.
Positive	slept well.	slept well.	slept well.	Positive	slept well.
<b>8</b>	<b>appeared depressed.</b>	<b>appeared depressed.</b>	<b>was depressed.</b>	<b>8</b>	<b>appeared depressed.</b>
Negative	was pessimistic.	was pessimistic.	felt that things won't work out for me.	Negative	was pessimistic.
Negative	had a negative self-image.	had a negative self-image.	felt bad about myself.	Negative	had a negative self-image.
Negative	withdrew from peers.	withdrew from peers.	wanted to be on my own rather than with my classmates/friends.	Negative	withdrew from peers.
Negative	was uninterested in doing things he/she usually enjoys.	was uninterested in doing things he/she usually enjoys.	didn't feel like doing the things I usually enjoy.	Negative	was uninterested in doing things he/she usually enjoys.
Negative	felt guilty.	felt guilty.	felt guilty.	Negative	felt guilty.
Negative	was tearful.	was tearful.	felt like crying.	Negative	was tearful.
Negative	felt worthless.	felt worthless.	felt worthless.	Negative	felt worthless.
Negative	was irritable.	was irritable.	felt irritable.	Negative	was irritable.
Negative	had low energy.	had low energy.	had low energy.	Negative	had low energy.
Positive	was optimistic.	was optimistic.	was optimistic.	Positive	was optimistic.
Positive	was confident.	was confident.	felt confident.	Positive	was confident.
Positive	took part in group activities.	took part in group activities.	took part in group activities.	Positive	took part in group activities.
Positive	enjoyed taking part in activities.	enjoyed taking part in activities.	enjoyed taking part in activities.	Positive	enjoyed taking part in activities.
Positive	was happy and content.	was happy and content.	felt happy and content.	Positive	was happy and content.
<b>12</b>	<b>acted sad or withdrawn.</b>	<b>acted sad or withdrawn.</b>	<b>was sad or withdrawn.</b>	<b>12</b>	<b>acted sad or withdrawn.</b>
Negative	kept to himself/herself.	kept to himself/herself.	kept to myself.	Negative	kept to himself/herself.
Negative	was sad.	was sad.	felt sad.	Negative	was sad.
Negative	cried.	cried.	cried.	Negative	cried.
Negative	was tearful.	was tearful.	was tearful.	Negative	was tearful.
Negative	withdrew from peers.	withdrew from peers.	wanted to be on my own rather than with my classmates/friends.	Negative	withdrew from peers.
Negative	was uninterested in doing things he/she usually enjoys.	was uninterested in doing things he/she usually enjoys.	didn't feel like doing the things I usually enjoy.	Negative	was uninterested in doing things he/she usually enjoys.
Negative	had trouble concentrating.	had trouble concentrating.	had trouble concentrating.	Negative	had trouble concentrating.
Negative	felt overly guilty.	felt overly guilty.	felt overly guilty.	Negative	felt overly guilty.
Negative	was cranky or irritable.	was cranky or irritable.	felt cranky or irritable.	Negative	was cranky or irritable.
Negative	had to push himself/herself to do schoolwork.	had to push himself/herself to do schoolwork.	had to push myself to do my schoolwork.	Negative	had to push himself/herself to do schoolwork.
Negative	avoided talking about feelings.	avoided talking about feelings.	avoided talking about my feelings.	Negative	avoided talking about feelings.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Negative Affect Items				BIMAS Negative Affect Items
Positive	was happy and content.	was happy and content.	felt happy and content.	Positive	was happy and content.
Positive	took part in group activities.	took part in group activities.	took part in group activities.	Positive	took part in group activities.
Positive	enjoyed activities.	enjoyed activities.	enjoyed activities.	Positive	enjoyed activities.
Positive	spent time with friends.	spent time with friends.	spent time with my friends.	Positive	spent time with friends.
Positive	was able to concentrate on tasks.	was able to concentrate on tasks.	was able to concentrate on tasks.	Positive	was able to concentrate on tasks.
Positive	enjoyed being with other people.	enjoyed being with other people.	enjoyed being with other people.	Positive	enjoyed being with other people.
Positive	talked about feelings.	talked about feelings.	talked about my feelings.	Positive	talked about feelings.
16	<b>was easily embarrassed or felt ashamed.</b>	<b>was easily embarrassed or felt ashamed.</b>	<b>was easily embarrassed or felt ashamed.</b>	16	<b>was easily embarrassed or felt ashamed.</b>
Negative	was reluctant to try new things.	was reluctant to try new things.	was reluctant to try new things.	Negative	was reluctant to try new things.
Negative	was reluctant to meet new people.	was reluctant to meet new people.	was reluctant to meet new people.	Negative	was reluctant to meet new people.
Negative	was easily hurt by criticism.	was easily hurt by criticism.	felt hurt by criticism.	Negative	was easily hurt by criticism.
Negative	was upset if someone criticized his/her work.	was upset if someone criticized his/her work.	felt upset if someone criticized my work.	Negative	was upset if someone criticized his/her work.
Negative	was awkward in social settings.	was awkward in social settings.	felt awkward with other people.	Negative	was awkward in social settings.
Negative	was easily embarrassed.	was easily embarrassed.	was easily embarrassed.	Negative	was easily embarrassed.
Negative	feared being embarrassed in front of others.	feared being embarrassed in front of others.	was afraid to be embarrassed in front of others.	Negative	feared being embarrassed in front of others.
Negative	was nervous doing things with others.	was nervous doing things with others.	felt nervous doing things with others.	Negative	was nervous doing things with others.
Negative	did not take part in group discussions.	did not take part in group discussions.	had trouble taking part in group discussions.	Negative	did not take part in group discussions.
Negative	was reluctant to ask for help when needed.	was reluctant to ask for help when needed.	was reluctant to ask for help when needed.	Negative	was reluctant to ask for help when needed.
Positive	was confident.	was confident.	felt confident.	Positive	was confident.
Positive	tried new activities.	tried new activities.	tried new activities.	Positive	tried new activities.
Positive	was relaxed and at ease around others.	was relaxed and at ease around others.	felt relaxed around others.	Positive	was relaxed and at ease around others.
Positive	took criticism well.	took criticism well.	felt okay about my work being criticized.	Positive	took criticism well.
Positive	was comfortable doing things in front of others.	was comfortable doing things in front of others.	felt comfortable doing things in front of others.	Positive	was comfortable doing things in front of others.
Positive	interacted with others.	interacted with others.	interacted with others.	Positive	interacted with others.
Positive	took part in group discussions.	took part in group discussions.	took part in group discussions.	Positive	took part in group discussions.
Positive	asked for help with schoolwork when needed.	asked for help with schoolwork when needed.	asked for help with my schoolwork.	Positive	asked for help with schoolwork when needed.
Positive	asked for help with homework when needed.	asked for help with homework when needed.	asked for help with my homework.	Positive	asked for help with homework when needed.
20	<b>appeared anxious (worried or nervous).</b>	<b>appeared anxious (worried or nervous).</b>	<b>was anxious (worried or nervous).</b>	20	<b>appeared anxious (worried or nervous).</b>
Negative	had trouble controlling worries.	had trouble controlling worries.	had trouble controlling my worries.	Negative	had trouble controlling worries.
Negative	was anxious about meeting new people.	was anxious about meeting new people.	was anxious about meeting new people.	Negative	was anxious about meeting new people.
Negative	was worried about getting a bad grade.	was worried about getting a bad grade.	worried about getting a bad grade.	Negative	was worried about getting a bad grade.
Negative	was worried about his/her family.	was worried about his/her family.	worried about my family.	Negative	was worried about his/her family.

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Negative Affect Items				BIMAS Negative Affect Items
Negative	was very worried about a family member being harmed.	was very worried about a family member being harmed.	worried a lot about a family member getting hurt.	Negative	was very worried about a family member being harmed.
Negative	worried about small things.	worried about small things.	worried about small things.	Negative	worried about small things.
Negative	was uneasy being away from home.	was uneasy being away from home.	felt uneasy being away from home.	Negative	was uneasy being away from home.
Negative	refused to go to school due to separation fears.	refused to go to school due to separation fears.	refused to go to school because I was worried about being away from my family.	Negative	refused to go to school due to separation fears.
Negative	was worried about doing things in front of people.	was worried about doing things in front of people.	worried about doing things in front of people.	Negative	was worried about doing things in front of people.
Negative	was worried about things before they happened.	was worried about things before they happened.	worried about things before they happened.	Negative	was worried about things before they happened.
Negative	was anxious about taking part in group activities.	was anxious about taking part in group activities.	was anxious about doing things with others.	Negative	was anxious about taking part in group activities.
Negative	was anxious in social settings.	was anxious in social settings.	felt anxious when I was with other people.	Negative	was anxious in social settings.
Negative	was agitated, “on-edge,” or jittery.	was agitated, “on-edge,” or jittery.	felt nervous or jumpy.	Negative	was agitated, “on-edge,” or jittery.
Negative	was irritable.	was irritable.	felt irritable.	Negative	was irritable.
Negative	was easily annoyed by others.	was easily annoyed by others.	felt easily annoyed by others.	Negative	was easily annoyed by others.
Negative	worried so much that he/she had trouble sleeping.	worried so much that he/she had trouble sleeping.	worried so much that I had trouble sleeping.	Negative	worried so much that he/she had trouble sleeping.
Positive	was relaxed.	was relaxed.	felt relaxed.	Positive	was relaxed.
Positive	was relaxed in social settings.	was relaxed in social settings.	felt relaxed when I was with other people.	Positive	was relaxed in social settings.
Positive	was confident about school performance.	was confident about school performance.	felt confident about my grades.	Positive	was confident about school performance.
Positive	was comfortable about being away from home.	was comfortable about being away from home.	felt comfortable being away from home.	Positive	was comfortable about being away from home.
Positive	felt secure about the well-being of his/her family members.	felt secure about the well-being of his/her family members.	felt that my family members were safe.	Positive	felt secure about the well-being of his/her family members.
Positive	went to school without fear or distress.	went to school without fear or distress.	went to school without fear or distress.	Positive	went to school without fear or distress.
Positive	was relaxed with his/her family.	was relaxed with his/her family.	felt relaxed with my family.	Positive	was relaxed with his/her family.
Positive	did not worry about small things.	did not worry about small things.	avoided worrying about small things.	Positive	did not worry about small things.
Positive	used a strategy to control worries.	used a strategy to control worries.	used a strategy to control my worries.	Positive	used a strategy to control worries.
Positive	was at ease taking part in small group activities.	was at ease taking part in small group activities.	felt at ease taking part in small group activities.	Positive	was at ease taking part in small group activities.
Positive	interacted with new people.	interacted with new people.	interacted with new people.	Positive	interacted with new people.
24	<b>expressed thoughts of hurting himself/herself.</b>	<b>expressed thoughts of hurting himself/herself.</b>	<b>had thoughts of hurting myself.</b>	24	<b>expressed thoughts of hurting himself/herself.</b>
Negative	expressed suicidal thoughts.	expressed suicidal thoughts.	had thoughts of ending my life.	Negative	expressed suicidal thoughts.
Negative	was preoccupied with themes of death.	was preoccupied with themes of death.	thought about things related to death.	Negative	was preoccupied with themes of death.
Negative	listened to music with violent themes.	listened to music with violent themes.	listened to music with violent themes.	Negative	listened to music with violent themes.
Negative	drew pictures of death or despair.	drew pictures of death or despair.	drew pictures of death or despair.	Negative	drew pictures of death or despair.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Negative Affect Items				BIMAS Negative Affect Items
Negative	dwelled on negative thoughts.	dwelled on negative thoughts.	had negative thoughts.	Negative	dwelled on negative thoughts.
Negative	expressed thoughts of self-harm.	expressed thoughts of self-harm.	had thoughts of hurting myself.	Negative	expressed thoughts of self-harm.
Negative	hurt himself/herself on purpose.	hurt himself/herself on purpose.	hurt myself on purpose.	Negative	hurt himself/herself on purpose.
Negative	cut himself/herself on purpose.	cut himself/herself on purpose.	cut myself on purpose.	Negative	cut himself/herself on purpose.
Negative	was withdrawn from friends and family.	was withdrawn from friends and family.	withdrew from my friends and my family.	Negative	was withdrawn from friends and family.
Negative	took part in risk-taking behavior when upset.	took part in risk-taking behavior when upset.	took part in risky activity when upset.	Negative	took part in risk-taking behavior when upset.
Negative	expressed thoughts of helplessness.	expressed thoughts of helplessness.	felt helpless.	Negative	expressed thoughts of helplessness.
Negative	expressed thoughts of hopelessness.	expressed thoughts of hopelessness.	felt hopeless.	Negative	expressed thoughts of hopelessness.
Negative	expressed thoughts of worthlessness.	expressed thoughts of worthlessness.	felt worthless.	Negative	expressed thoughts of worthlessness.
Positive	expressed positive thoughts.	expressed positive thoughts.	had positive thoughts.	Positive	expressed positive thoughts.
Positive	talked about plans for the future.	talked about plans for the future.	made plans for the future.	Positive	talked about plans for the future.
Positive	expressed positive thoughts about the future.	expressed positive thoughts about the future.	had positive thoughts about the future.	Positive	expressed positive thoughts about the future.
Positive	took part in positive activities.	took part in positive activities.	took part in positive activities.	Positive	took part in positive activities.
Positive	set positive goals.	set positive goals.	set positive goals for myself.	Positive	set positive goals.
<b>27</b>	<b>was emotional or upset.</b>	<b>was emotional or upset.</b>	<b>felt emotional or upset.</b>	<b>26</b>	<b>was emotional or upset.</b>
Negative	was overly sensitive to criticism.	was overly sensitive to criticism.	was upset by criticism.	Negative	was overly sensitive to criticism.
Negative	overreacted to a minor issue.	overreacted to a minor issue.	overreacted to a small problem.	Negative	overreacted to a minor issue.
Negative	cried easily.	cried easily.	cried easily.	Negative	cried easily.
Negative	changed his/her mood quickly.	changed his/her mood quickly.	was moody.	Negative	changed his/her mood quickly.
Negative	became emotional when challenged.	became emotional when challenged.	became emotional when someone did not agree with me.	Negative	became emotional when challenged.
Negative	became emotional when anxious.	became emotional when anxious.	became emotional when I felt nervous about something.	Negative	became emotional when anxious.
Negative	got upset when challenged.	got upset when challenged.	got upset when someone did not agree with me.	Negative	got upset when challenged.
Negative	got upset if things did not go according to plan.	got upset if things did not go according to plan.	got upset if things did not go according to my plan.	Negative	got upset if things did not go according to plan.
Negative	overreacted to a minor disagreement with peers.	overreacted to a minor disagreement with peers.	overreacted to a minor disagreement with my classmates/friends.	Negative	overreacted to a minor disagreement with peers.
Positive	used a strategy to control emotions.	used a strategy to control emotions.	used a strategy to control my emotions.	Positive	used a strategy to control emotions.
Positive	controlled his/her emotions.	controlled his/her emotions.	controlled my emotions.	Positive	controlled his/her emotions.
Positive	solved minor problems with peers without help from adults.	solved minor problems with peers without help from adults.	solved minor problems with my classmates/friends without help from adults.	Positive	solved minor problems with peers without help from adults.
Positive	adapted well to stress.	adapted well to stress.	handled stress well.	Positive	adapted well to stress.
Positive	responded well to criticism.	responded well to criticism.	took criticism well.	Positive	responded well to criticism.

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Cognitive/Attention Items				BIMAS Cognitive/Attention Items
<b>3</b>	<b>had trouble paying attention.</b>	<b>had trouble paying attention.</b>	<b>had trouble paying attention.</b>	<b>3</b>	<b>had trouble paying attention.</b>
Negative	was easily distracted by noises.	was easily distracted by noises.	was easily distracted by noises.	Negative	was easily distracted by noises.
Negative	had trouble completing tasks.	had trouble completing tasks.	had trouble completing tasks.	Negative	had trouble completing tasks.
Negative	had a hard time following directions.	had a hard time following directions.	had a hard time following directions.	Negative	had a hard time following directions.
Negative	had to be reminded to complete schoolwork.	had to be reminded to complete schoolwork.	had to be reminded to complete schoolwork.	Negative	had to be reminded to complete schoolwork.
Negative	had to be reminded to complete chores.	had to be reminded to complete chores.	had to be reminded to complete chores.	Negative	had to be reminded to complete chores.
Negative	was careless while doing schoolwork.	was careless while doing schoolwork.	was careless while doing schoolwork.	Negative	was careless while doing schoolwork.
Negative	was careless while doing chores.	was careless while doing chores.	was careless while doing chores.	Negative	was careless while doing chores.
Negative	made careless mistakes.	made careless mistakes.	made careless mistakes.	Negative	made careless mistakes.
Negative	did not pay attention to details.	did not pay attention to details.	had trouble paying attention to details.	Negative	did not pay attention to details.
Negative	did not seem to listen to what was being said to him/her.	did not seem to listen to what was being said to him/her.	had trouble listening to what was being said to me.	Negative	did not seem to listen to what was being said to him/her.
Negative	forgot to hand in completed homework.	forgot to hand in completed homework.	forgot to hand in completed homework.	Negative	forgot to hand in completed homework.
Negative	forgot the required materials for homework.	forgot the required materials for homework.	forgot the things I needed for my homework.	Negative	forgot the required materials for homework.
Negative	forgot the required materials for schoolwork.	forgot the required materials for schoolwork.	forgot the things I needed for my schoolwork.	Negative	forgot the required materials for schoolwork.
Negative	had trouble concentrating.	had trouble concentrating.	had trouble concentrating.	Negative	had trouble concentrating.
Positive	paid attention to details.	paid attention to details.	paid attention to details.	Positive	paid attention to details.
Positive	was attentive.	was attentive.	was attentive.	Positive	was attentive.
Positive	stayed focused despite distracting sounds.	stayed focused despite distracting sounds.	stayed focused despite distracting sounds.	Positive	stayed focused despite distracting sounds.
Positive	followed directions.	followed directions.	followed directions.	Positive	followed directions.
Positive	stayed on task.	stayed on task.	stayed on task.	Positive	stayed on task.
Positive	completed tasks or chores without reminders.	completed tasks or chores without reminders.	completed tasks or chores without reminders.	Positive	completed tasks or chores without reminders.
Positive	finished tasks or chores neatly.	finished tasks or chores neatly.	finished tasks or chores neatly.	Positive	finished tasks or chores neatly.
Positive	handed in completed homework on time.	handed in completed homework on time.	handed in completed homework on time.	Positive	handed in completed homework on time.
Positive	was prepared with the required materials for homework or schoolwork.	was prepared with the required materials for homework or schoolwork.	was prepared with all the things I need to do my homework or schoolwork.	Positive	was prepared with the required materials for homework or schoolwork.
Positive	was focused on teacher instructions.	was focused on teacher instructions.	was focused on teacher instructions.	Positive	was focused on teacher instructions.
Positive	was focused on independent schoolwork.	was focused on independent schoolwork.	was focused on independent schoolwork.	Positive	was focused on independent schoolwork.
Positive	was focused during group activities.	was focused during group activities.	was focused during group activities.	Positive	was focused during group activities.
Positive	listened to what was being said to him/her.	listened to what was being said to him/her.	listened to what was being said to me.	Positive	listened to what was being said to him/her.
<b>6</b>	<b>was impulsive.</b>	<b>was impulsive.</b>	<b>was impulsive.</b>	<b>6</b>	<b>was impulsive.</b>
Negative	interrupted a teacher while he/she was speaking.	interrupted a teacher while he/she was speaking.	interrupted my teacher while he/she was speaking.	Negative	interrupted a teacher while he/she was speaking.
Negative	interrupted a peer while he/she was speaking.	interrupted a peer while he/she was speaking.	interrupted a classmate/friend while he/she was speaking.	Negative	interrupted a peer while he/she was speaking.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Cognitive/Attention Items				BIMAS Cognitive/Attention Items
Negative	interrupted a parent while he/she was speaking.	interrupted a parent while he/she was speaking.	interrupted my parent while he/she was speaking.	Negative	interrupted a parent while he/she was speaking.
Negative	buted into others' conversations.	buted into others' conversations.	buted into others' conversations.	Negative	buted into others' conversations.
Negative	buted into others' games.	buted into others' games.	buted into others' games.	Negative	buted into others' games.
Negative	had trouble waiting for his/her turn to speak.	had trouble waiting for his/her turn to speak.	had trouble waiting for my turn to speak.	Negative	had trouble waiting for his/her turn to speak.
Negative	had trouble waiting for his/her turn in a game.	had trouble waiting for his/her turn in a game.	had trouble waiting for my turn in a game.	Negative	had trouble waiting for his/her turn in a game.
Negative	blurted out an answer before the question was finished.	blurted out an answer before the question was finished.	blurted out an answer before the question was finished.	Negative	blurted out an answer before the question was finished.
Negative	did not think about the consequences of his/her actions.	did not think about the consequences of his/her actions.	didn't think about the consequences of my actions.	Negative	did not think about the consequences of his/her actions.
Negative	called out in the classroom.	called out in the classroom.	called out in the classroom.	Negative	called out in the classroom.
Negative	got overexcited.	got overexcited.	got overexcited.	Negative	got overexcited.
Positive	waited his/her turn before speaking.	waited his/her turn before speaking.	waited my turn before speaking.	Positive	waited his/her turn before speaking.
Positive	waited for his/her turn in a game.	waited for his/her turn in a game.	waited for my turn in a game.	Positive	waited for his/her turn in a game.
Positive	waited for a question to be finished before answering the question.	waited for a question to be finished before answering the question.	waited for a question to be finished before answering the question.	Positive	waited for a question to be finished before answering the question.
Positive	waited for instructions to be finished before asking a question.	waited for instructions to be finished before asking a question.	waited for instructions to be finished before asking a question.	Positive	waited for instructions to be finished before asking a question.
Positive	thought before acting.	thought before acting.	thought before acting.	Positive	thought before acting.
Positive	waited to be called on by the teacher before talking.	waited to be called on by the teacher before talking.	waited to be called on by my teacher before talking.	Positive	waited to be called on by the teacher before talking.
Positive	completed schoolwork with care.	completed schoolwork with care.	completed my schoolwork with care.	Positive	completed schoolwork with care.
Positive	made a plan for a task or activity.	made a plan for a task or activity.	made a plan for a task or activity.	Positive	made a plan for a task or activity.
10	<b>had problems staying on task.</b>	<b>had problems staying on task.</b>	<b>had problems staying on task.</b>	10	<b>had problems staying on task.</b>
Negative	gave up easily on difficult tasks.	gave up easily on difficult tasks.	gave up easily on difficult tasks.	Negative	gave up easily on difficult tasks.
Negative	gave up easily on boring tasks.	gave up easily on boring tasks.	gave up easily on boring tasks.	Negative	gave up easily on boring tasks.
Negative	had trouble staying focused on schoolwork.	had trouble staying focused on schoolwork.	had trouble staying focused on my schoolwork.	Negative	had trouble staying focused on schoolwork.
Negative	had trouble staying focused on homework.	had trouble staying focused on homework.	had trouble staying focused on my homework.	Negative	had trouble staying focused on homework.
Negative	had trouble staying focused on one activity at a time.	had trouble staying focused on one activity at a time.	had a hard time focusing on one activity at a time.	Negative	had trouble staying focused on one activity at a time.
Negative	had trouble staying focused on directions.	had trouble staying focused on directions.	had trouble staying focused on directions.	Negative	had trouble staying focused on directions.
Negative	had trouble staying focused on classroom lessons.	had trouble staying focused on classroom lessons.	had trouble staying focused on classroom lessons.	Negative	had trouble staying focused on classroom lessons.
Negative	avoided doing challenging schoolwork.	avoided doing challenging schoolwork.	avoided doing schoolwork that was hard.	Negative	avoided doing challenging schoolwork.
Negative	needed reminders of academic instructions.	needed reminders of academic instructions.	needed reminders of how to do my schoolwork.	Negative	needed reminders of academic instructions.
Negative	had trouble following directions even when he/she seemed to understand them.	had trouble following directions even when he/she seemed to understand them.	had a hard time following directions.	Negative	had trouble following directions even when he/she seemed to understand them.

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Cognitive/Attention Items				BIMAS Cognitive/Attention Items
Negative	had trouble completing schoolwork even when he/she seemed to understand what to do.	had trouble completing schoolwork even when he/she seemed to understand what to do.	had trouble completing schoolwork.	Negative	had trouble completing schoolwork even when he/she seemed to understand what to do.
Positive	stayed focused on schoolwork.	stayed focused on schoolwork.	stayed focused on schoolwork.	Positive	stayed focused on schoolwork.
Positive	stayed focused on homework.	stayed focused on homework.	stayed focused on homework.	Positive	stayed focused on homework.
Positive	stayed focused on activities.	stayed focused on activities.	stayed focused on my activities.	Positive	stayed focused on activities.
Positive	persisted with difficult tasks.	persisted with difficult tasks.	kept trying with difficult tasks.	Positive	persisted with difficult tasks.
Positive	persisted with boring tasks.	persisted with boring tasks.	persisted with boring tasks.	Positive	persisted with boring tasks.
Positive	followed directions.	followed directions.	followed directions.	Positive	followed directions.
Positive	completed schoolwork without reminders.	completed schoolwork without reminders.	did my schoolwork without reminders.	Positive	completed schoolwork without reminders.
Positive	completed homework without reminders.	completed homework without reminders.	did my homework without reminders.	Positive	completed homework without reminders.
14	<b>acted without thinking.</b>	<b>acted without thinking.</b>	<b>acted without thinking.</b>	14	<b>acted without thinking.</b>
Negative	answered before a question was fully asked.	answered before a question was fully asked.	answered before a question was fully asked.	Negative	answered before a question was fully asked.
Negative	interrupted a teacher while he/she was speaking.	interrupted a teacher while he/she was speaking.	interrupted a teacher while he/she was speaking.	Negative	interrupted a teacher while he/she was speaking.
Negative	interrupted a peer while he/she was speaking.	interrupted a peer while he/she was speaking.	interrupted a classmate/friend while he/she was speaking.	Negative	interrupted a peer while he/she was speaking.
Negative	interrupted a parent while he/she was speaking.	interrupted a parent while he/she was speaking.	interrupted my parent while he/she was speaking.	Negative	interrupted a parent while he/she was speaking.
Negative	butted into others' conversations.	butted into others' conversations.	butted into others' conversations.	Negative	butted into others' conversations.
Negative	butted into others' games.	butted into others' games.	butted into others' games.	Negative	butted into others' games.
Negative	had trouble waiting for his/her turn to speak in class.	had trouble waiting for his/her turn to speak in class.	had trouble waiting for my turn to speak in class.	Negative	had trouble waiting for his/her turn to speak in class.
Negative	had trouble waiting for his/her turn in a game.	had trouble waiting for his/her turn in a game.	had trouble waiting for my turn in a game.	Negative	had trouble waiting for his/her turn in a game.
Negative	blurted out an answer before the question was finished.	blurted out an answer before the question was finished.	blurted out an answer before the question was finished.	Negative	blurted out an answer before the question was finished.
Negative	did not think about the consequences of his/her actions.	did not think about the consequences of his/her actions.	didn't think about the consequences of my actions.	Negative	did not think about the consequences of his/her actions.
Negative	called out in the classroom.	called out in the classroom.	called out in the classroom.	Negative	called out in the classroom.
Negative	got overexcited.	got overexcited.	got overexcited.	Negative	got overexcited.
Positive	waited his/her turn before speaking.	waited his/her turn before speaking.	waited my turn before speaking.	Positive	waited his/her turn before speaking.
Positive	waited for his/her turn in a game.	waited for his/her turn in a game.	waited my turn in a game.	Positive	waited for his/her turn in a game.
Positive	waited for a question to be finished before answering the question.	waited for a question to be finished before answering the question.	waited for a question to be finished before answering the question.	Positive	waited for a question to be finished before answering the question.
Positive	waited for the teacher to finish giving instructions before asking a question.	waited for the teacher to finish giving instructions before asking a question.	waited for my teacher to finish giving instructions before asking a question.	Positive	waited for the teacher to finish giving instructions before asking a question.
Positive	thought before acting.	thought before acting.	thought before acting.	Positive	thought before acting.
Positive	raised his/her hand before speaking.	raised his/her hand before speaking.	raised my hand before speaking.	Positive	raised his/her hand before speaking.
Positive	completed schoolwork with care.	completed schoolwork with care.	completed schoolwork with care.	Positive	completed schoolwork with care.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Cognitive/Attention Items				BIMAS Cognitive/Attention Items
Positive	made a plan for a task or activity.	made a plan for a task or activity.	made a plan for a task or activity.	Positive	made a plan for a task or activity.
18	had trouble remembering things.	had trouble remembering things.	had trouble remembering things.	18	had trouble remembering things.
Negative	had trouble remembering directions.	had trouble remembering directions.	had trouble remembering directions.	Negative	had trouble remembering directions.
Negative	forgot to do homework.	forgot to do homework.	forgot to do my homework.	Negative	forgot to do homework.
Negative	forgot to take the required materials for schoolwork.	forgot to bring the required materials for schoolwork.	forgot to bring the things I needed to do my schoolwork.	Negative	forgot to bring the required materials for schoolwork.
Negative	forgot to bring materials needed for homework.	forgot to take materials needed for homework.	forgot to bring the things I needed to do my homework.	Negative	forgot to bring materials needed for homework.
Negative	forgot things already learned.	forgot things already learned.	forgot things I had already learned.	Negative	forgot things already learned.
Negative	forgot to turn in completed work.	forgot to turn in completed work.	forgot to turn in my completed work.	Negative	forgot to turn in completed work.
Negative	lost schoolwork.	lost schoolwork.	lost my schoolwork.	Negative	lost schoolwork.
Negative	lost materials required to do schoolwork.	lost materials required to do schoolwork.	lost materials I needed to do my schoolwork.	Negative	lost materials required to do schoolwork.
Negative	lost materials required to do homework.	lost materials required to do homework.	lost the things I needed to do my homework.	Negative	lost materials required to do homework.
Negative	lost own belongings (e.g., books, toys).	lost own belongings (e.g., books, toys).	lost my belongings (e.g., books, toys).	Negative	lost own belongings (e.g., books, toys).
Negative	needed reminders to do schoolwork.	needed reminders to do schoolwork.	needed reminders to do my schoolwork.	Negative	needed reminders to do schoolwork.
Negative	needed reminders to do homework.	needed reminders to do homework.	needed reminders to do my homework.	Negative	needed reminders to do homework.
Negative	needed reminders to complete chores.	needed reminders to complete chores.	needed reminders to complete my chores.	Negative	needed reminders to complete chores.
Negative	needed reminders to complete tasks.	needed reminders to complete tasks.	needed reminders to complete tasks.	Negative	needed reminders to complete tasks.
Positive	followed one-step directions without reminders.	followed one-step directions without reminders.	followed one-step directions without reminders.	Positive	followed one-step directions without reminders.
Positive	followed two-step directions without reminders.	followed two-step directions without reminders.	followed two-step directions without reminders.	Positive	followed two-step directions without reminders.
Positive	followed three-step directions without reminders.	followed three-step directions without reminders.	followed three-step directions without reminders.	Positive	followed three-step directions without reminders.
Positive	followed directions of four or more steps without reminders.	followed directions of four or more steps without reminders.	followed directions of four or more steps without reminders.	Positive	followed directions of four or more steps without reminders.
Positive	completed homework without reminders.	completed homework without reminders.	completed my homework without reminders.	Positive	completed homework without reminders.
Positive	completed schoolwork without reminders.	completed schoolwork without reminders.	completed my schoolwork without reminders.	Positive	completed schoolwork without reminders.
Positive	completed chores without reminders.	completed chores without reminders.	completed my chores without reminders.	Positive	completed chores without reminders.
Positive	completed activities without reminders.	completed activities without reminders.	completed activities without reminders.	Positive	completed activities without reminders.
Positive	turned in completed work on time.	turned in completed work on time.	turned in completed work on time.	Positive	turned in completed work on time.
Positive	went to class prepared with all the required materials.	came to class prepared with all the required materials.	came to class prepared with all the required materials.	Positive	came to class prepared with all the required materials.
Positive	kept track of own belongings (e.g., books, toys).	kept track of own belongings (e.g., books, toys).	kept track of my own belongings (e.g., books, toys).	Positive	kept track of own belongings (e.g., books, toys).

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Cognitive/Attention Items				BIMAS Cognitive/Attention Items
Positive	brought home the required materials for homework.	took home the required materials for homework.	brought home the things I needed to do my homework.	Positive	brought home the required materials for homework.
22	<b>had trouble with organizing and planning.</b>	<b>had trouble with organizing and planning.</b>	<b>had trouble with organizing and planning.</b>	22	<b>had trouble with organizing and planning.</b>
Negative	forgot to take the required materials for schoolwork.	forgot to bring the required materials for schoolwork.	forgot to bring the things I needed to do my schoolwork.	Negative	forgot to bring the required materials for schoolwork.
Negative	forgot to bring materials needed for homework.	forgot to take materials needed for homework.	forgot to bring the things I needed to do my homework.	Negative	forgot to bring materials needed for homework.
Negative	was late for school.	was late for school.	was late for school.	Negative	was late for school.
Negative	arrived late to class.	arrived late to class.	arrived late to class.	Negative	arrived late to class.
Negative	had a messy desk at school.	had a messy desk at school.	had a messy desk at school.	Negative	had a messy desk at school.
Negative	lost schoolwork.	lost schoolwork.	lost my schoolwork.	Negative	lost schoolwork.
Negative	lost materials required for schoolwork.	lost materials required for schoolwork.	lost things I needed for schoolwork.	Negative	lost materials required for schoolwork.
Negative	lost materials required to do homework.	lost materials required to do homework.	lost things I needed for homework.	Negative	lost materials required to do homework.
Negative	lost own belongings (e.g., books, toys).	lost own belongings (e.g., books, toys).	lost my belongings (e.g., books, toys).	Negative	lost own belongings (e.g., books, toys).
Negative	had trouble completing tasks on time.	had trouble completing tasks on time.	had trouble completing tasks on time.	Negative	had trouble completing tasks on time.
Negative	completed a task at the last minute.	completed a task at the last minute.	completed a task at the last minute.	Negative	completed a task at the last minute.
Negative	needed help from an adult to plan and organize tasks.	needed help from an adult to plan and organize tasks.	needed help from an adult to plan and organize tasks.	Negative	needed help from an adult to plan and organize tasks.
Negative	needed help to get started on a task.	needed help to get started on a task.	needed help to get started on a task.	Negative	needed help to get started on a task.
Negative	had trouble with multiple-step problem solving.	had trouble with multiple-step problem solving.	had trouble solving problems with multiple steps.	Negative	had trouble with multiple-step problem solving.
Negative	forgot to turn in completed work.	forgot to turn in completed work.	forgot to turn in my completed work.	Negative	forgot to turn in completed work.
Negative	had trouble finishing what he/she started.	had trouble finishing what he/she started.	had trouble finishing what I started.	Negative	had trouble finishing what he/she started.
Positive	turned in completed work on time.	turned in completed work on time.	turned in completed work on time.	Positive	turned in completed work on time.
Positive	went to class prepared with all the required materials.	came to class prepared with all the required materials.	came to class prepared with all the materials I needed.	Positive	came to class prepared with all the required materials.
Positive	kept track of own belongings (e.g., books, toys).	kept track of own belongings (e.g., books, toys).	kept track of my own belongings (e.g., books, toys).	Positive	kept track of own belongings (e.g., books, toys).
Positive	brought home the required materials to do homework.	took home the required materials to do homework.	brought home the things I needed to do my homework.	Positive	brought home the required materials to do homework.
Positive	made a plan before beginning a task.	made a plan before beginning a task.	made a plan before beginning a task.	Positive	made a plan before beginning a task.
Positive	met required deadlines.	met required deadlines.	met required deadlines.	Positive	met required deadlines.
Positive	completed tasks by himself/herself.	completed tasks by himself/herself.	completed tasks by myself.	Positive	completed tasks by himself/herself.
Positive	kept his/her things organized.	kept his/her things organized.	kept my things organized.	Positive	kept his/her things organized.
Positive	was on time.	was on time.	was on time.	Positive	was on time.
28	<b>fidged.</b>	<b>fidged.</b>	<b>fidged.</b>	27	<b>fidged.</b>
Negative	squirmed in his/her seat/ chair.	squirmed in his/her seat/ chair.	squirmed in my seat/ chair.	Negative	squirmed in his/her seat/ chair.
Negative	fidged with his/her hands or feet.	fidged with his/her hands or feet.	fidged with my hands or feet.	Negative	fidged with his/her hands or feet.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Cognitive/Attention Items				BIMAS Cognitive/Attention Items
Negative	made distracting noises.	made distracting noises.	made distracting noises.	Negative	made distracting noises.
Negative	left his/her seat without permission.	left his/her seat without permission.	left my seat without permission.	Negative	left his/her seat without permission.
Negative	played with objects during class.	played with objects during class.	played with objects during class.	Negative	played with objects during class.
Negative	made distracting noises with objects (e.g., by tapping a pencil).	made distracting noises with objects (e.g., by tapping a pencil).	made distracting noises with objects (e.g., by tapping a pencil).	Negative	made distracting noises with objects (e.g., by tapping a pencil).
Negative	had trouble working quietly.	had trouble working quietly.	had trouble working quietly.	Negative	had trouble working quietly.
Negative	had trouble playing quietly.	had trouble playing quietly.	had trouble playing quietly.	Negative	had trouble playing quietly.
Negative	was unable to stay still.	was unable to stay still.	was unable to stay still.	Negative	was unable to stay still.
Positive	used a strategy to fidget quietly without distracting others.	used a strategy to fidget quietly without distracting others.	found a way to fidget quietly without distracting others.	Positive	used a strategy to fidget quietly without distracting others.
Positive	stayed in his/her seat.	stayed in his/her seat.	stayed in my seat.	Positive	stayed in his/her seat.
Positive	worked quietly.	worked quietly.	worked quietly.	Positive	worked quietly.
Positive	played quietly.	played quietly.	played quietly.	Positive	played quietly.
Positive	sat in his/her seat without fidgeting and/or squirming.	sat in his/her seat without fidgeting and/or squirming.	sat in my seat without fidgeting and/or squirming.	Positive	sat in his/her seat without fidgeting and/or squirming.
	BIMAS Social Items				BIMAS Social Items
1	shared what he/she was thinking about.	shared what he/she was thinking about.	shared my thoughts with others.	1	shared what he/she was thinking about.
Negative	was unwilling to share thoughts with his/her teacher.	was unwilling to share thoughts with his/her teacher.	was unwilling to share my thoughts with my teacher.	Negative	was unwilling to share thoughts with his/her teacher.
Negative	was unwilling to share thoughts with his/her family.	was unwilling to share thoughts with his/her family.	was unwilling to share my thoughts with my family.	Negative	was unwilling to share thoughts with his/her family.
Negative	was unwilling to share thoughts with his/her friends.	was unwilling to share thoughts with his/her friends.	was unwilling to share my thoughts with my friends.	Negative	was unwilling to share thoughts with his/her friends.
Negative	was unwilling to share thoughts with his/her class.	was unwilling to share thoughts with his/her class.	was unwilling to share my thoughts with my class.	Negative	was unwilling to share thoughts with his/her class.
Negative	was anxious about sharing thoughts with the teacher.	was anxious about sharing thoughts with the teacher.	was anxious about sharing my thoughts with my teacher.	Negative	was anxious about sharing thoughts with the teacher.
Negative	was anxious about sharing thoughts with his/her family.	was anxious about sharing thoughts with his/her family.	was anxious about sharing my thoughts with my family.	Negative	was anxious about sharing thoughts with his/her family.
Negative	was anxious about sharing thoughts with his/her friends.	was anxious about sharing thoughts with his/her friends.	was anxious about sharing my thoughts with my friends.	Negative	was anxious about sharing thoughts with his/her friends.
Negative	did not expand on thoughts when talking with others.	did not expand on thoughts when talking with others.	had trouble expanding my thoughts when talking with others.	Negative	did not expand on thoughts when talking with others.
Negative	did not ask questions when he/she did not understand.	did not ask questions when he/she did not understand.	didn't ask questions to clarify when I didn't understand something.	Negative	did not ask questions when he/she did not understand.
Negative	was overly defensive when asked questions.	was overly defensive when asked questions.	felt overly defensive when someone asked me questions.	Negative	was overly defensive when asked questions.
Negative	showed less emotion than was appropriate for the situation.	showed less emotion than was appropriate for the situation.	showed less emotion than was appropriate for the situation.	Negative	showed less emotion than was appropriate for the situation.
Negative	was in his/her own world.	was in his/her own world.	was in my own world.	Negative	was in his/her own world.
Positive	was engaged in a conversation.	was engaged in a conversation.	was engaged in a conversation.	Positive	was engaged in a conversation.

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Social Items			BIMAS Social Items	
Positive	contributed to a group discussion.	contributed to a group discussion.	contributed to a group discussion.	Positive	contributed to a group discussion.
Positive	expanded on ideas when talking.	expanded on ideas when talking.	expanded on my ideas when I was talking.	Positive	expanded on ideas when talking.
Positive	asked questions when he/she did not understand something.	asked questions when he/she did not understand something.	asked questions when I did not understand something.	Positive	asked questions when he/she did not understand something.
Positive	shared thoughts and/or feelings with his/her teachers.	shared thoughts and/or feelings with his/her teachers.	shared my thoughts and/or feelings with my teachers.	Positive	shared thoughts and/or feelings with his/her teachers.
Positive	shared thoughts and/or feelings with his/her family.	shared thoughts and/or feelings with his/her family.	shared my thoughts and/or feelings with my family.	Positive	shared thoughts and/or feelings with his/her family.
Positive	shared thoughts and/or feelings with his/her class.	shared thoughts and/or feelings with his/her class.	shared my thoughts and/or feelings with my class.	Positive	shared thoughts and/or feelings with his/her class.
Positive	shared thoughts and/or feelings with his/her friends.	shared thoughts and/or feelings with his/her friends.	shared my thoughts and/or feelings with my friends.	Positive	shared thoughts and/or feelings with his/her friends.
Positive	was comfortable when answering questions.	was comfortable when answering questions.	felt comfortable when I was answering questions.	Positive	was comfortable when answering questions.
<b>7</b>	<b>spoke clearly with others.</b>	<b>spoke clearly with others.</b>	<b>communicated clearly with others.</b>	<b>7</b>	<b>spoke clearly with others.</b>
Negative	had a hard time pronouncing words correctly.	had a hard time pronouncing words correctly.	had a hard time pronouncing words correctly.	Negative	had a hard time pronouncing words correctly.
Negative	had trouble finding the right words.	had trouble finding the right words.	had trouble finding the right words.	Negative	had trouble finding the right words.
Negative	was hard to understand when he/she spoke.	was hard to understand when he/she spoke.	felt I was hard to understand when I spoke.	Negative	was hard to understand when he/she spoke.
Negative	had trouble expressing a thought or idea.	had trouble expressing a thought or idea.	had trouble expressing a thought or idea.	Negative	had trouble expressing a thought or idea.
Negative	was unable to expand on a thought or idea.	was unable to expand on a thought or idea.	was unable to expand on a thought or idea.	Negative	was unable to expand on a thought or idea.
Negative	had trouble making a point.	had trouble making a point.	had trouble making a point.	Negative	had trouble making a point.
Negative	stuttered.	stuttered.	stuttered.	Negative	stuttered.
Negative	was scattered when he/she spoke.	was scattered when he/she spoke.	had a hard time organizing my thoughts when I spoke.	Negative	was scattered when he/she spoke.
Negative	had trouble combining words into sentences.	had trouble combining words into sentences.	had trouble combining words into sentences.	Negative	had trouble combining words into sentences.
Negative	had trouble starting a conversation.	had trouble starting a conversation.	had trouble starting a conversation.	Negative	had trouble starting a conversation.
Negative	had trouble keeping a conversation going.	had trouble keeping a conversation going.	had trouble keeping a conversation going.	Negative	had trouble keeping a conversation going.
Positive	spoke clearly.	spoke clearly.	spoke clearly.	Positive	spoke clearly.
Positive	expressed thoughts and ideas clearly.	expressed thoughts and ideas clearly.	expressed my thoughts and ideas clearly.	Positive	expressed thoughts and ideas clearly.
Positive	was easily understood when speaking.	was easily understood when speaking.	felt I was easily understood when I spoke.	Positive	was easily understood when speaking.
Positive	spoke fluently.	spoke fluently.	spoke fluently.	Positive	spoke fluently.
Positive	pronounced words correctly.	pronounced words correctly.	pronounced words correctly.	Positive	pronounced words correctly.
Positive	started a conversation.	started a conversation.	started a conversation.	Positive	started a conversation.
Positive	kept a conversation going.	kept a conversation going.	kept a conversation going.	Positive	kept a conversation going.
<b>11</b>	<b>maintained friendships.</b>	<b>maintained friendships.</b>	<b>maintained friendships.</b>	<b>11</b>	<b>maintained friendships.</b>
Negative	fought with friends.	fought with friends.	fought with friends.	Negative	fought with friends.
Negative	was ignored by friends.	was ignored by friends.	was ignored by my friends.	Negative	was ignored by friends.
Negative	left phone calls/texts from friends unreturned.	left phone calls/texts from friends unreturned.	didn't return my friends' phone calls/texts.	Negative	left phone calls/texts from friends unreturned.
Negative	was bossy toward friends.	was bossy toward friends.	was bossy toward my friends.	Negative	was bossy toward friends.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Social Items				BIMAS Social Items
Negative	was the last one to be picked for a team or a game.	was the last one to be picked for a team or a game.	was the last one to be picked for a team or a game.	Negative	was the last one to be picked for a team or a game.
Negative	was rude to friends.	was rude to friends.	was rude to friends.	Negative	was rude to friends.
Negative	spread rumors about friends.	spread rumors about friends.	spread rumors about my friends.	Negative	spread rumors about friends.
Negative	got others to gang up on friends.	got others to gang up on friends.	got others to gang up on my friends.	Negative	got others to gang up on friends.
Negative	left friends out of group activities on purpose.	left friends out of group activities on purpose.	left friends out of group activities on purpose.	Negative	left friends out of group activities on purpose.
Positive	got invited to play with friends.	got invited to play with friends.	got invited to play with friends.	Positive	got invited to play with friends.
Positive	interacted well with friends.	interacted well with friends.	had a good time with friends.	Positive	interacted well with friends.
Positive	got invited to go out with friends.	got invited to go out with friends.	got invited to go out with friends.	Positive	got invited to go out with friends.
Positive	kept in touch with friends.	kept in touch with friends.	kept in touch with my friends.	Positive	kept in touch with friends.
Positive	was kind toward friends.	was kind toward friends.	was kind toward my friends.	Positive	was kind toward friends.
Positive	was involved in activities with friends.	was involved in activities with friends.	was involved in activities with my friends.	Positive	was involved in activities with friends.
Positive	kept confidences.	kept confidences.	kept a secret.	Positive	kept confidences.
15	<b>appeared comfortable when relating to others.</b>	<b>appeared comfortable when relating to others.</b>	<b>felt relaxed interacting with others.</b>	15	<b>appeared comfortable when relating to others.</b>
Negative	preferred being alone to being with other people.	preferred being alone to being with other people.	preferred being alone rather than being with other people.	Negative	preferred being alone to being with other people.
Negative	avoided friends.	avoided friends.	avoided my friends.	Negative	avoided friends.
Negative	avoided talking with teachers.	avoided talking with teachers.	avoided talking with teachers.	Negative	avoided talking with teachers.
Negative	was overly defensive when asked questions.	was overly defensive when asked questions.	felt overly defensive when someone asked me questions.	Negative	was overly defensive when asked questions.
Negative	was awkward during social situations.	was awkward during social situations.	felt awkward with other people.	Negative	was awkward during social situations.
Negative	had trouble starting a conversation.	had trouble starting a conversation.	had trouble starting a conversation.	Negative	had trouble starting a conversation.
Negative	had trouble keeping a conversation going.	had trouble keeping a conversation going.	had trouble keeping a conversation going.	Negative	had trouble keeping a conversation going.
Negative	was nervous talking to peers.	was nervous talking to peers.	was nervous talking to my classmates/friends.	Negative	was nervous talking to peers.
Negative	was nervous talking in class.	was nervous talking in class.	was nervous talking in class.	Negative	was nervous talking in class.
Negative	was nervous talking to teachers.	was nervous talking to teachers.	was nervous talking to teachers.	Negative	was nervous talking to teachers.
Positive	took part in class discussions.	took part in class discussions.	took part in class discussions.	Positive	took part in class discussions.
Positive	enjoyed group activities.	enjoyed group activities.	enjoyed doing things in a group.	Positive	enjoyed group activities.
Positive	took part in a conversation with parents.	took part in a conversation with parents.	took part in a conversation with my parents.	Positive	took part in a conversation with parents.
Positive	took part in a conversation with peers.	took part in a conversation with peers.	took part in a conversation with my classmates/friends.	Positive	took part in a conversation with peers.
Positive	took part in a conversation with adults.	took part in a conversation with adults.	took part in a conversation with adults.	Positive	took part in a conversation with adults.
Positive	started a conversation.	started a conversation.	started a conversation.	Positive	started a conversation.
Positive	kept a conversation going.	kept a conversation going.	kept a conversation going.	Positive	kept a conversation going.
Positive	was outgoing.	was outgoing.	was outgoing.	Positive	was outgoing.

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Social Items				BIMAS Social Items
19	<b>was generally friendly with others.</b>	<b>was generally friendly with others.</b>	<b>was friendly with others.</b>	19	<b>was generally friendly with others.</b>
Negative	was rude to friends.	was rude to friends.	was rude to my friends.	Negative	was rude to friends.
Negative	was rude to parents.	was rude to parents.	was rude to my parents.	Negative	was rude to parents.
Negative	was rude to teachers.	was rude to teachers.	was rude to my teachers.	Negative	was rude to teachers.
Negative	bullied others.	bullied others.	bullied others.	Negative	bullied others.
Negative	teased others.	teased others.	teased others.	Negative	teased others.
Negative	said mean things to others.	said mean things to others.	said mean things to others.	Negative	said mean things to others.
Negative	said things that were not true of others.	said things that were not true of others.	said things that were not true of others.	Negative	said things that were not true of others.
Negative	blamed others for his/her misbehavior.	blamed others for his/her misbehavior.	blamed others for the things I did wrong.	Negative	blamed others for his/her misbehavior.
Negative	argued with parents.	argued with parents.	argued with my parents.	Negative	argued with parents.
Negative	argued with teachers.	argued with teachers.	argued with my teachers.	Negative	argued with teachers.
Negative	argued with peers.	argued with peers.	argued with my classmates/friends.	Negative	argued with peers.
Negative	argued with siblings.	argued with siblings.	argued with my siblings.	Negative	argued with siblings.
Negative	annoyed parents on purpose.	annoyed parents on purpose.	annoyed my parents on purpose.	Negative	annoyed parents on purpose.
Negative	annoyed teachers on purpose.	annoyed teachers on purpose.	annoyed my teachers on purpose.	Negative	annoyed teachers on purpose.
Negative	annoyed peers on purpose.	annoyed peers on purpose.	annoyed my classmates/friends on purpose.	Negative	annoyed peers on purpose.
Negative	annoyed siblings on purpose.	annoyed siblings on purpose.	annoyed my siblings on purpose.	Negative	annoyed siblings on purpose.
Negative	was left out of group work by peers.	was left out of group work by peers.	was left out of group work by my classmates.	Negative	was left out of group work by peers.
Negative	was left out of group activities by peers.	was left out of group activities by peers.	was left out of group activities by my friends.	Negative	was left out of group activities by peers.
Negative	was left out of games by peers.	was left out of games by peers.	was left out of games by my classmates/friends.	Negative	was left out of games by peers.
Negative	preferred being alone to being with other people.	preferred being alone to being with other people.	preferred being alone to being with other people.	Negative	preferred being alone to being with other people.
Positive	was polite to teachers.	was polite to teachers.	was polite to my teachers.	Positive	was polite to teachers.
Positive	was polite to parents.	was polite to parents.	was polite to my parents.	Positive	was polite to parents.
Positive	was polite to peers.	was polite to peers.	was polite to my classmates/friends.	Positive	was polite to peers.
Positive	was polite to siblings.	was polite to siblings.	was polite to my siblings.	Positive	was polite to siblings.
Positive	took part in class discussions.	took part in class discussions.	took part in class discussions.	Positive	took part in class discussions.
Positive	was interested in what other people were doing.	was interested in what other people were doing.	was interested in what other people were doing.	Positive	was interested in what other people were doing.
Positive	was included by peers in group work.	was included by peers in group work.	was included by classmates in group work.	Positive	was included by peers in group work.
Positive	was included by peers in games.	was included by peers in games.	was included by my friends in games.	Positive	was included by peers in games.
Positive	helped others.	helped others.	helped others.	Positive	helped others.
Positive	played well with peers.	played well with peers.	played well with my classmates/friends.	Positive	played well with peers.
Positive	played well with siblings.	played well with siblings.	played well with my siblings.	Positive	played well with siblings.
Positive	offered help if someone was hurt.	offered help if someone was hurt.	offered help if someone was hurt.	Positive	offered help if someone was hurt.
Positive	offered sympathy if someone was upset.	offered sympathy if someone was upset.	was kind to someone if he/she was upset.	Positive	offered sympathy if someone was upset.
Positive	took responsibility for his/her own mistakes and/or misbehavior.	took responsibility for his/her own mistakes and/or misbehavior.	took responsibility for the things I did.	Positive	took responsibility for his/her own mistakes and/or misbehavior.
Positive	had fun with friends.	had fun with friends.	had fun with friends.	Positive	had fun with friends.
Positive	smiled at friends.	smiled at friends.	smiled at my friends.	Positive	smiled at friends.
Positive	was considerate.	was considerate.	was considerate.	Positive	was considerate.

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Social Items				BIMAS Social Items
23	<b>worked out problems with others.</b>	<b>worked out problems with others.</b>	<b>worked out problems with others.</b>	23	<b>worked out problems with others.</b>
Negative	tried to get even with someone.	tried to get even with someone.	tried to get even with someone.	Negative	tried to get even with someone.
Negative	kept to himself/herself when something was wrong.	kept to himself/herself when something was wrong.	kept to myself when something was wrong.	Negative	kept to himself/herself when something was wrong.
Negative	needed help from an adult to resolve problems with peers.	needed help from an adult to resolve problems with peers.	needed help from an adult to resolve problems with classmates/friends.	Negative	needed help from an adult to resolve problems with peers.
Negative	made small conflicts with others worse.	made small conflicts with others worse.	made small conflicts with others worse.	Negative	made small conflicts with others worse.
Negative	didn't care about others' feelings.	didn't care about others' feelings.	didn't care about others' feelings.	Negative	didn't care about others' feelings.
Negative	had trouble controlling anger.	had trouble controlling anger.	had trouble controlling my anger.	Negative	had trouble controlling anger.
Negative	said mean things to others.	said mean things to others.	said mean things to others.	Negative	said mean things to others.
Negative	was easily annoyed by others.	was easily annoyed by others.	was easily annoyed by others.	Negative	was easily annoyed by others.
Negative	threatened a peer after a disagreement.	threatened a peer after a disagreement.	threatened a classmate/friend after a disagreement.	Negative	threatened a peer after a disagreement.
Negative	threatened a teacher after a disagreement.	threatened a teacher after a disagreement.	threatened a teacher after a disagreement.	Negative	threatened a teacher after a disagreement.
Negative	threatened a parent after a disagreement.	threatened a parent after a disagreement.	threatened my parent after a disagreement.	Negative	threatened a parent after a disagreement.
Negative	argued with teachers.	argued with teachers.	argued with teachers.	Negative	argued with teachers.
Negative	argued with parents.	argued with parents.	argued with parents.	Negative	argued with parents.
Negative	argued with peers.	argued with peers.	argued with classmates/friends.	Negative	argued with peers.
Negative	argued with siblings.	argued with siblings.	argued with my siblings.	Negative	argued with siblings.
Positive	was forgiving to others.	was forgiving to others.	was forgiving to others.	Positive	was forgiving to others.
Positive	mediated a conflict.	mediated a conflict.	helped to stop an argument.	Positive	mediated a conflict.
Positive	asked an adult for help to mediate a conflict.	asked an adult for help to mediate a conflict.	asked an adult for help to stop an argument.	Positive	asked an adult for help to mediate a conflict.
Positive	worked out problems with peers by himself/herself.	worked out problems with peers by himself/herself.	worked out problems with my classmates/friends by myself.	Positive	worked out problems with peers by himself/herself.
Positive	worked out problems with siblings by himself/herself.	worked out problems with siblings by himself/herself.	worked out problems with my siblings by myself.	Positive	worked out problems with siblings by himself/herself.
Positive	stayed calm while working out problems with others.	stayed calm while working out problems with others.	stayed calm while working out problems with others.	Positive	stayed calm while working out problems with others.
Positive	used a strategy to resolve a conflict.	used a strategy to resolve a conflict.	used a strategy to resolve an argument.	Positive	used a strategy to resolve a conflict.
	<b>BIMAS Academic Functioning Items</b>				<b>BIMAS Clinician Form Adaptive Items</b>
4	<b>followed directions.</b>	<b>followed directions.</b>	<b>followed directions.</b>	4	<b>followed directions.</b>
Negative	had trouble following verbal directions.	had trouble following verbal directions.	had trouble following verbal directions.	Negative	had trouble following verbal directions.
Negative	had trouble following written instructions.	had trouble following written instructions.	had trouble following written instructions.	Negative	had trouble following written instructions.
Negative	had trouble following multi-step directions.	had trouble following multi-step directions.	had trouble following directions with many steps.	Negative	had trouble following multi-step directions.
Negative	responded poorly to redirection.	responded poorly to redirection.	had trouble with being redirected to what I was supposed to be doing.	Negative	responded poorly to redirection.

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Academic Functioning Items				BIMAS Clinician Form Adaptive Items
Negative	needed extra explanation of instructions.	needed extra explanation of instructions.	needed an extra explanation of instructions.	Negative	needed extra explanation of instructions.
Negative	did not ask questions when he/she did not understand directions.	did not ask questions when he/she did not understand directions.	didn't ask questions when I did not understand directions.	Negative	did not ask questions when he/she did not understand directions.
Positive	followed one-step directions.	followed one-step directions.	followed one-step directions.	Positive	followed one-step directions.
Positive	followed two-step directions.	followed two-step directions.	followed two-step directions.	Positive	followed two-step directions.
Positive	followed three-step directions.	followed three-step directions.	followed three-step directions.	Positive	followed three-step directions.
Positive	followed directions of four or more steps.	followed directions of four or more steps.	followed directions of four or more steps.	Positive	followed directions of four or more steps.
Positive	followed verbal directions.	followed verbal directions.	followed verbal directions.	Positive	followed verbal directions.
Positive	followed written directions.	followed written directions.	followed written directions.	Positive	followed written directions.
Positive	asked questions when he/she did not understand directions.	asked questions when he/she did not understand directions.	asked questions when I did not understand directions.	Positive	asked questions when he/she did not understand directions.
Positive	remembered directions without reminders.	remembered directions without reminders.	remembered directions without reminders.	Positive	remembered directions without reminders.
26-R	<b>received failing grades at school.</b>	<b>received failing grades at school.</b>	<b>received failing grades at school.</b>		N/A
Negative	had trouble with math.	had trouble with math.	had trouble with math.		
Negative	had trouble with reading.	had trouble with reading.	had trouble with reading.		
Negative	had trouble with writing.	had trouble with writing.	had trouble with writing.		
Negative	did not complete homework.	did not complete homework.	didn't complete my homework.		
Negative	left homework incomplete.	left homework incomplete.	left my homework incomplete.		
Negative	did poorly on exams and/or quizzes.	did poorly on exams and/or quizzes.	did poorly on exams and/or quizzes.		
Negative	did not complete work at school.	did not complete work at school.	left schoolwork undone.		
Negative	did not prepare for tests and/or quizzes.	did not prepare for tests and/or quizzes.	didn't prepare for tests and/or quizzes.		
Negative	had poor attendance.	had poor attendance.	had poor attendance.		
Negative	showed poor motivation at school.	showed poor motivation at school.	had trouble getting motivated at school.		
Negative	misbehaved in class.	misbehaved in class.	misbehaved in class.		
Negative	was anxious about exams and/or quizzes, even though he/she was prepared.	was anxious about exams and/or quizzes, even though he/she was prepared.	was anxious about exams and/or quizzes, even though I was prepared.		
Negative	had trouble with math, even when he/she tried hard.	had trouble with math, even when he/she tried hard.	had trouble with math, even when I tried hard.		
Negative	had trouble with reading, even when he/she tried hard.	had trouble with reading, even when he/she tried hard.	had trouble with reading, even when I tried hard.		
Positive	prepared for a test or a quiz.	prepared for a test or a quiz.	prepared for a test or a quiz.		
Positive	was prepared for test/quizzes.	was prepared for tests/quizzes.	was prepared for tests/quizzes.		
Positive	did well on homework assignments.	did well on homework assignments.	did well on homework assignments.		
Positive	did well on exams and quizzes.	did well on exams and quizzes.	did well on exams and quizzes.		
Positive	had good attendance.	had good attendance.	had good attendance.		
Positive	completed assigned work at school.	completed assigned work at school.	completed assigned work at school.		

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Academic Functioning Items				
Positive	completed homework.	completed homework.	completed homework.		
Positive	completed homework before doing other activities.	completed homework before doing other activities.	completed my homework before doing other activities.		
Positive	worked hard at something that was hard for him/her.	worked hard at something that was hard for him/her.	worked hard at something that was hard for me.		
Positive	behaved properly in class.	behaved properly in class.	behaved properly in class.		
Positive	worked hard at school.	worked hard at school.	worked hard at school.		
Positive	took notes during class.	took notes during class.	took notes during class.		
Positive	asked for help when he/she needed help with schoolwork.	asked for help when he/she needed help with schoolwork.	asked for help when I needed help with schoolwork.		
Positive	asked for help when he/she needed help with homework.	asked for help when he/she needed help with homework.	asked for help when I needed help with homework.		
Positive	sought feedback to improve schoolwork.	sought feedback to improve schoolwork.	asked for feedback to improve my schoolwork.		
Positive	showed initiative to learn more.	showed initiative to learn more.	showed initiative to learn more.		
<b>30</b>	<b>worked up to his/her academic potential.</b>	<b>worked up to his/her academic potential.</b>	<b>tried my hardest when it came to schoolwork.</b>		<b>N/A</b>
Negative	did not complete homework.	did not complete homework.	didn't complete my homework.		
Negative	left homework incomplete.	left homework incomplete.	left my homework incomplete.		
Negative	did poorly on exams and/or quizzes.	did poorly on exams and/or quizzes.	did poorly on exams and/or quizzes.		
Negative	did not complete work at school.	did not complete work at school.	didn't complete work at school.		
Negative	had poor attendance.	had poor attendance.	had poor attendance.		
Negative	showed poor motivation at school.	showed poor motivation at school.	had trouble getting motivated at school.		
Negative	misbehaved in class.	misbehaved in class.	misbehaved in class.		
Negative	had trouble with math, even when he/she tried hard.	had trouble with math, even when he/she tried hard.	had trouble with math, even when I tried hard.		
Negative	had trouble with reading, even when he/she tried hard.	had trouble with reading, even when he/she tried hard.	had trouble with reading, even when I tried hard.		
Negative	was anxious about exams and/or quizzes, even though he/she was prepared.	was anxious about exams and/or quizzes, even though he/she was prepared.	was anxious about exams and/or quizzes, even though I was prepared.		
Positive	did well on homework assignments.	did well on homework assignments.	did well on homework assignments.		
Positive	did well on exams and quizzes.	did well on exams and quizzes.	did well on exams and quizzes.		
Positive	had good attendance.	had good attendance.	had good attendance.		
Positive	completed assigned work at school.	completed assigned work at school.	completed assigned work at school.		
Positive	completed homework.	completed homework.	completed homework.		
Positive	behaved properly in class.	behaved properly in class.	behaved properly in class.		
Positive	worked hard at school.	worked hard at school.	worked hard at school.		
Positive	completed schoolwork without help.	completed schoolwork without help.	completed schoolwork without help.		
Positive	took notes during class.	took notes during class.	took notes during class.		
Positive	asked for help when he/she needed help with schoolwork.	asked for help when he/she needed help with schoolwork.	asked for help when I needed help with schoolwork.		

Continued...

Behavior Intervention Monitoring Assessment System (BIMAS™)

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Academic Functioning Items				BIMAS Clinician Form Adaptive Items
<b>Positive</b>	asked for help when he/she needed help with homework.	asked for help when he/she needed help with homework.	asked for help when I needed help with homework.		
	<b>N/A</b>	<b>N/A</b>	<b>N/A</b>	<b>31</b>	<b>attended his/her scheduled therapy appointments.</b>
				<b>Negative</b>	was absent for his/her appointment due to illness.
				<b>Negative</b>	did not report to his/her appointment.
				<b>Negative</b>	wanted to go home after arriving at his/her appointment.
				<b>Negative</b>	complained about somatic symptoms during his/her appointment.
				<b>Negative</b>	requested to leave his/her appointment.
				<b>Negative</b>	skipped his/her appointment.
				<b>Negative</b>	reported late for his/her appointment.
				<b>Negative</b>	did not show up for his/her appointment and did not call to cancel.
				<b>Negative</b>	canceled his/her appointment at the last minute.
				<b>Positive</b>	had good attendance.
				<b>Positive</b>	reported to his/her appointment.
				<b>Positive</b>	stayed until the session was complete.
				<b>Positive</b>	worked well with his/her clinician.
				<b>Positive</b>	arrived at his/her session on time.
				<b>Positive</b>	made scheduled appointments.
<b>33</b>	<b>was prepared for class.</b>	<b>was prepared for class.</b>	<b>went prepared to class.</b>		<b>N/A</b>
<b>Negative</b>	was unprepared for a test or quiz.	was unprepared for a test or quiz.	was unprepared for a test or quiz.		
<b>Negative</b>	was unprepared for PE/gym class.	was unprepared for PE/gym class.	was unprepared for PE/gym class.		
<b>Negative</b>	forgot to take the required materials for schoolwork.	forgot to bring the required materials for schoolwork.	forgot to bring the things I needed to do my schoolwork.		
<b>Negative</b>	forgot to bring materials needed for homework.	forgot to take materials needed for homework.	forgot to bring the things I needed to do my homework.		
<b>Negative</b>	forgot to turn in completed work.	forgot to turn in completed work.	forgot to turn in my completed work.		
<b>Negative</b>	didn't hand in homework on time.	didn't hand in homework on time.	didn't hand in my homework on time.		
<b>Negative</b>	lost schoolwork.	lost schoolwork.	lost my schoolwork.		
<b>Negative</b>	lost materials required for schoolwork.	lost materials required for schoolwork.	lost things I needed to do my schoolwork.		
<b>Negative</b>	lost materials required to do homework.	lost materials required to do homework.	lost things I needed to do my homework.		
<b>Negative</b>	lost own belongings (e.g., books, toys).	lost own belongings (e.g., books, toys).	lost my belongings (e.g., books, toys).		

Continued...

Standard Form Item Number/Flex Valence	Teacher	Parent	Self-Report	Standard Form Item Number/Flex Valence	Clinician
	BIMAS Academic Functioning Items				
Negative	arrived late to class.	arrived late to class.	arrived late to class.		
Positive	turned in completed work on time.	turned in completed work on time.	turned in completed work on time.		
Positive	went to class prepared with all the required materials.	came to class prepared with all the required materials.	came to class prepared with all the materials I need.		
Positive	kept track of own belongings (e.g., books, toys).	kept track of own belongings (e.g., books, toys).	kept track of my own belongings (e.g., books, toys).		
Positive	brought home the required materials for homework.	took home the required materials for homework.	brought home the things I needed to do my homework.		
Positive	went to class on time.	came to class on time.	came to class on time.		
Positive	prepared for a test or quiz.	prepared for a test or quiz.	prepared for a test or quiz.		
Positive	was prepared for PE/gym class.	was prepared for PE/gym class.	was prepared for PE/gym class.		
34_R	<b>was absent from school.</b>	<b>was absent from school.</b>	<b>was absent from school.</b>		N/A
Negative	was absent due to illness.	was absent due to illness.	was absent due to illness.		
Negative	skipped class(es).	skipped class(es).	skipped class(es).		
Negative	arrived late to class.	arrived late to class.	arrived late to class.		
Negative	complained about feeling sick.	complained about feeling sick.	complained about feeling sick.		
Negative	reported late to school.	reported late to school.	arrived late to school.		
Negative	wanted to go home after arriving at school.	wanted to go home after arriving at school.	wanted to go home after arriving at school.		
Negative	was too tired to go to school.	was too tired to go to school.	was too tired to go to school.		
Negative	was too sick to go to school.	was too sick to go to school.	was too sick to go to school.		
Negative	missed the bus.	missed the bus.	missed my bus.		
Negative	refused to go to school.	refused to go to school.	refused to go to school.		
Positive	arrived at school on time.	arrived at school on time.	arrived at school on time.		
Positive	attended classes.	attended classes.	attended classes.		
Positive	stayed in school all day.	stayed in school all day.	stayed in school all day.		

# Appendix C

## Reliable Change Index Values

This appendix provides the Reliable Change Index (RCI) values for the BIMAS™ Standard that are built into the *Standard Individual Progress Report: Significant Change Over Time* (Report ID: 15) for assessing statistically significant change over time (see the *Scores for Progress and Outcome Monitoring* section in chapter 5, *Understanding and Interpreting BIMAS Scores*, for more information on the use of RCI in interpretation).

The values needed to establish statistical significance, when comparing Time 1 to Time 2 scores, are provided in Table C.1 for the BIMAS Standard. For a Time 1 to Time 2 difference to be statistically significant, the absolute difference between the two *T*-scores on a scale must be equal to or greater than the value provided in the table.

**Table C.1. RCI Values for the BIMAS Standard**

Type of Scale	Scale	Behavioral Concern Scales		
		BIMAS–T	BIMAS–P	BIMAS–SR
Behavioral Concern Scales	Conduct	8	11	11
	Negative Affect	9	8	9
	Cognitive/Attention	7	10	10
Adaptive Scales	Social	7	5	8
	Academic Functioning	8	11	9

# Appendix D

## Values for Rater Comparisons

This appendix provides the minimum differences required for the scores between two raters to be considered statistically significant for the BIMAS™ Standard. The values are built into the *Standard Individual Comparison Between Raters* (Report ID: 8) for assessing differences in *T*-scores between raters (see *Statistically Significant Differences Between Raters* section in chapter 5, *Understanding and Interpreting BIMAS Scores*, for more information on differences between raters' scores).

The minimum values needed to establish statistical significance, when comparing scores between any given pair of raters, are provided in Tables D.1 to D.2. For the difference between any pair of raters to be statistically significant, the absolute difference between the two *T*-scores on a scale must be equal to or greater than the value provided in the relevant table.

**Table D.1. Minimum Values for Comparisons Between Raters on the BIMAS Standard Using Combined Gender Norms (Default Scoring Option)**

Type of Scale	Scale	Age Group	Required Difference Between Rater Pairs				
			T1 vs. T2	P1 vs. P2	T vs. P	T vs. S	P vs. S
Behavioral Concern Scales	Conduct	5–6	7	9	8	n/a	n/a
		7–9	8	8	8	n/a	n/a
		10–11	6	9	8	n/a	n/a
		12–13	5	7	6	7	7
		14–16	6	7	7	7	7
	17–18	6	9	8	7	9	
	Negative Affect	5–6	9	12	11	n/a	n/a
		7–9	10	11	10	n/a	n/a
		10–11	8	10	9	n/a	n/a
		12–13	9	9	9	9	9
		14–16	10	9	9	10	9
	17–18	9	10	10	9	10	
	Cognitive/Attention	5–6	7	8	8	n/a	n/a
		7–9	9	9	9	n/a	n/a
		10–11	7	8	7	n/a	n/a
12–13		8	6	7	8	8	
14–16		8	8	8	8	8	
17–18	9	8	9	9	8		
Adaptive Scales	Social	5–6	12	11	11	n/a	n/a
		7–9	10	11	10	n/a	n/a
		10–11	10	11	11	n/a	n/a
		12–13	9	8	8	10	9
		14–16	9	9	9	10	10
	17–18	9	9	9	9	9	
	Academic Functioning	5–6	13	15	14	n/a	n/a
		7–9	12	13	13	n/a	n/a
		10–11	10	12	11	n/a	n/a
		12–13	9	9	9	10	10
		14–16	11	8	9	11	10
		17–18	9	11	10	10	10

*Note.* T = Teacher; P = Parent; S = Self. n/a = Not applicable (because Self-Reports are not available for youth under the age of 12 years).

Table D.2. Minimum Values for Comparisons Between Raters on the BIMAS Standard Using Gender-Specific Norms

Type of Scale	Scale	Gender	Age Group	Required Difference Between Rater Pairs				
				T1 vs. T2	P1 vs. P2	T vs. P	T vs. S	P vs. S
Behavioral Concern Scales	Conduct	Male	5-6	7	10	8	n/a	n/a
			7-9	7	9	8	n/a	n/a
			10-11	7	8	8	n/a	n/a
			12-13	5	7	7	7	8
			14-16	6	7	6	7	8
			17-18	6	9	8	7	9
		Female	5-6	6	9	8	n/a	n/a
			7-9	10	8	9	n/a	n/a
			10-11	6	10	8	n/a	n/a
			12-13	5	6	6	6	7
			14-16	7	7	7	7	7
			17-18	7	10	8	7	9
	Negative Affect	Male	5-6	10	14	12	n/a	n/a
			7-9	10	10	10	n/a	n/a
			10-11	9	9	9	n/a	n/a
			12-13	8	9	8	9	9
			14-16	10	9	9	10	9
			17-18	9	11	10	9	10
		Female	5-6	7	10	9	n/a	n/a
			7-9	11	11	11	n/a	n/a
			10-11	8	12	10	n/a	n/a
			12-13	10	11	10	10	10
			14-16	11	9	10	10	9
			17-18	10	10	10	9	10
	Cognitive/Attention	Male	5-6	7	8	7	n/a	n/a
			7-9	8	10	9	n/a	n/a
			10-11	8	8	8	n/a	n/a
			12-13	8	7	7	9	9
			14-16	8	7	7	8	8
			17-18	10	10	10	9	9
Female		5-6	8	9	8	n/a	n/a	
		7-9	9	9	9	n/a	n/a	
		10-11	6	8	7	n/a	n/a	
		12-13	7	6	6	8	7	
		14-16	8	9	9	8	9	
		17-18	8	8	8	8	8	
Adaptive Scales	Social	Male	5-6	10	10	10	n/a	n/a
			7-9	10	12	11	n/a	n/a
			10-11	10	12	11	n/a	n/a
			12-13	9	9	9	10	10
			14-16	9	11	10	10	11
			17-18	10	8	9	9	9
		Female	5-6	14	12	13	n/a	n/a
			7-9	10	10	10	n/a	n/a
			10-11	9	10	9	n/a	n/a
			12-13	9	7	8	9	8
			14-16	10	7	9	10	9
			17-18	8	9	9	10	10
	Academic Functioning	Male	5-6	15	15	15	n/a	n/a
			7-9	12	13	13	n/a	n/a
			10-11	12	14	13	n/a	n/a
			12-13	8	10	9	10	11
			14-16	12	9	11	12	10
			17-18	9	12	11	11	12
		Female	5-6	11	15	13	n/a	n/a
			7-9	12	13	13	n/a	n/a
			10-11	8	9	9	n/a	n/a
			12-13	9	10	9	10	11
			14-16	9	9	9	11	11
			17-18	10	10	10	9	9

Note. T = Teacher; P = Parent; S = Self. n/a = Not applicable (because Self-Reports are not available for youth under the age of 12 years).

# Appendix E.1 Risk Level Pyramids

## Risk Level Pyramids BIMAS™—Teacher Standard

Springfield School District  
2010–2011

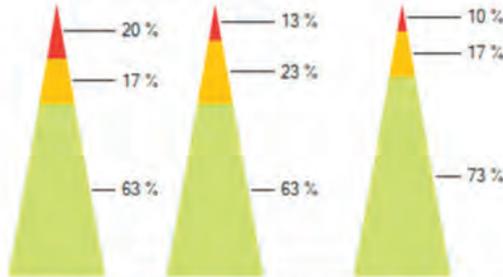
Universal Assessment: 1

Schools Selected: Lincoln Middle School  
MacDonald Elementary  
Rouge Elementary  
Grades Selected: K, 1, 2, 3, 4, 5, 6, 7, 8

**Total For District**  
500 Students

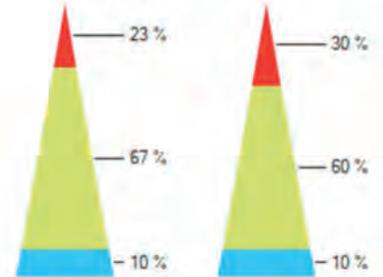
### Behavioral Concern Scales

Percent of Students



### Adaptive Scales

Percent of Students

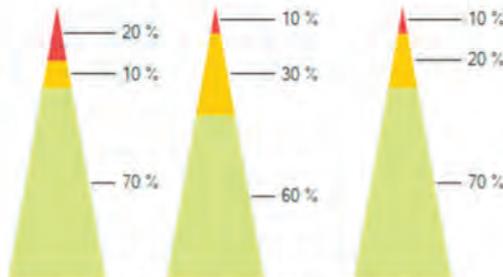


Levels Of Risk	Conduct	Negative Affect	Cognitive/ Attention	Levels Of Functioning	Social	Academic Functioning
High Risk	100 (20%)	65 (13%)	50 (10%)	Concern	115 (23%)	150 (30%)
Some Risk	85 (17%)	115 (23%)	85 (17%)	Typical	335 (67%)	300 (60%)
Low Risk	315 (63%)	315 (63%)	365 (73%)	Strength	50 (10%)	50 (10%)
<b>Total</b>	500 (100%)	500 (100%)	500 (100%)	<b>Total</b>	500 (100%)	500 (100%)

**Lincoln Middle School**  
200 Students

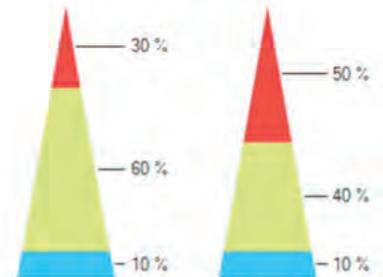
### Behavioral Concern Scales

Percent of Students



### Adaptive Scales

Percent of Students

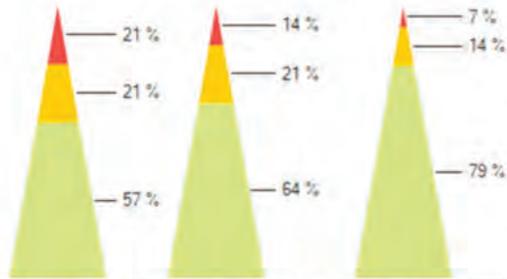


Levels Of Risk	Conduct	Negative Affect	Cognitive/ Attention	Levels Of Functioning	Social	Academic Functioning
High Risk	40 (20%)	20 (10%)	20 (10%)	Concern	60 (30%)	100 (50%)
Some Risk	20 (10%)	60 (30%)	40 (20%)	Typical	120 (60%)	80 (40%)
Low Risk	140 (70%)	120 (60%)	140 (70%)	Strength	20 (10%)	20 (10%)
<b>Total</b>	200 (100%)	200 (100%)	200 (100%)	<b>Total</b>	200 (100%)	200 (100%)

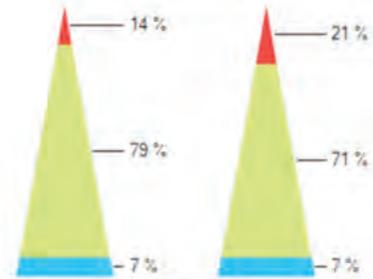
Note. All names are fictional.

**MacDonald Elementary**  
200 Students

**Behavioral Concern Scales**  
Percent of Students



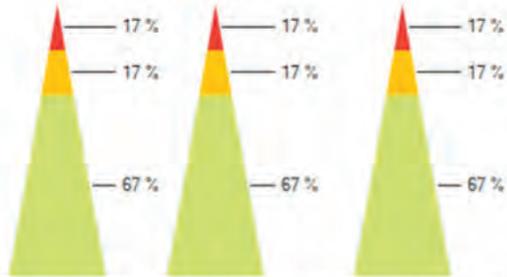
**Adaptive Scales**  
Percent of Students



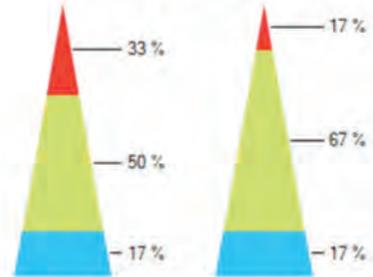
Levels Of Risk	Conduct	Negative Affect	Cognitive/ Attention	Levels Of Functioning	Social	Academic Functioning
High Risk	42 (21%)	28 (14%)	14 (7%)	Concern	28 (14%)	42 (21%)
Some Risk	42 (21%)	42 (21%)	28 (14%)	Typical	158 (79%)	142 (71%)
Low Risk	114 (57%)	128 (64%)	158 (79%)	Strength	14 (7%)	14 (7%)
<b>Total</b>	200 (100%)	200 (100%)	200 (100%)	<b>Total</b>	200 (100%)	200 (100%)

**Rouge Elementary**  
100 Students

**Behavioral Concern Scales**  
Percent of Students



**Adaptive Scales**  
Percent of Students



Levels Of Risk	Conduct	Negative Affect	Cognitive/ Attention	Levels Of Functioning	Social	Academic Functioning
High Risk	17 (17%)	17 (17%)	17 (17%)	Concern	33 (33%)	17 (17%)
Some Risk	17 (17%)	17 (17%)	17 (17%)	Typical	50 (50%)	67 (67%)
Low Risk	67 (67%)	67 (67%)	67 (67%)	Strength	17 (17%)	17 (17%)
<b>Total</b>	100 (100%)	100 (100%)	100 (100%)	<b>Total</b>	100 (100%)	100 (100%)

Note. All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



# Appendix E.2a Student List by Risk Level

(Assessment Report Version)

## Student List by Risk Level BIMAS™-Teacher Standard

Springfield School District  
MacDonald Elementary  
2010-2011

Universal Assessment(s): 1  
Grade(s): 4  
Class(es): All  
Filtered by Demographic Variable: None

Click on student name to view the student's BIMAS Standard Individual Assessment Report for the selected Universal Assessment.

### Conduct Scale

In Prog Monitor = In Progress Monitoring at time of assessment

● = Yes ○ = No

High Risk Group (T-score = 70+) in Universal Assessment 1		
Universal Assessment 1		
Student Name	T-score	In Prog Monitor
<a href="#">Farrah, Colleen</a>	83	●
<a href="#">Klumming, Heather</a>	81	●
<a href="#">Nathanson, Andy</a>	71	●

Some Risk Group (T-score = 60-69) in Universal Assessment 1		
Universal Assessment 1		
Student Name	T-score	In Prog Monitor
<a href="#">Deppe, Joey</a>	63	○
<a href="#">Lola, Jojo</a>	69	●
<a href="#">Taylor, Dudley</a>	63	○

Low Risk Group (T-score < 60) in Universal Assessment 1		
Universal Assessment 1		
Student Name	T-score	In Prog Monitor
<a href="#">Black, Juliana</a>	57	○
<a href="#">Chan, Christina</a>	55	○
<a href="#">Fernandez, Jose</a>	57	○
<a href="#">Fong, Gary</a>	46	○
<a href="#">Platt, Ken</a>	57	○
<a href="#">Pratt, Beverly</a>	55	○
<a href="#">Sakamoto, Catherine</a>	57	○
<a href="#">Smith, Janice</a>	55	○

Sort by Student Last Name  Sort by Student First Name

Note. All names are fictional.



# Appendix E.2b Student List by Risk Level

(Progress Report Version)

## Student List by Risk Level BIMAS™-Teacher Standard

Springfield School District  
MacDonald Elementary  
2010-2011

Universal Assessment(s): 1-2  
Grade(s): 4  
Class(es): All

Filtered by Demographic Variable: None

Click on student name to view the student's BIMAS Standard Individual Progress Report for the selected Universal Assessments.

### Conduct Scale

In Prog Monitor = In Progress Monitoring at time of assessment      N/A = No Score Available      • = Yes ○ = No

High Risk Group (T-score = 70+) in Universal Assessment 1				
Student Name	UA 1		UA 2	
	T-score	In Prog Monitor	T-score	In Prog Monitor
<a href="#">Farrah, Colleen</a>	83	•	65	•
<a href="#">Klumming, Heather</a>	81	•	62	•
<a href="#">Nathanson, Andy</a>	71	•	65	•

Some Risk Group (T-score = 60-69) in Universal Assessment 1				
Student Name	UA 1		UA 2	
	T-score	In Prog Monitor	T-score	In Prog Monitor
<a href="#">Deppe, Joey</a>	63	○	55	•
<a href="#">Lola, Jolo</a>	69	•	46	•
<a href="#">Taylor, Dudley</a>	63	○	61	•

Low Risk Group (T-score < 60) in Universal Assessment 1				
Student Name	UA 1		UA 2	
	T-score	In Prog Monitor	T-score	In Prog Monitor
<a href="#">Black, Juliana</a>	57	○	57	○
<a href="#">Chan, Christina</a>	55	○	53	○
<a href="#">Fernandez, Jose</a>	57	○	53	○
<a href="#">Fong, Gary</a>	46	○	59	○
<a href="#">Platt, Ken</a>	57	○	55	○
<a href="#">Pratt, Beverly</a>	55	○	51	○
<a href="#">Sakamoto, Catherine</a>	57	○	53	○
<a href="#">Smith, Janice</a>	55	○	53	○

Sort by Student Last Name    Sort by Student First Name    Sort by Student Score from UA 1

Note: All names are fictional.

# Appendix E.3

## Class/Group Student Scores

### Class/Group Student Scores BIMAS™-Teacher Standard

Springfield School District  
MacDonald Elementary  
2010–2011

Grade: 4  
Class: 4A  
Universal Assessment: 1

Click on student name to view the student's BIMAS Standard Individual Assessment Report for the selected Universal Assessment.

Student Name	Behavioral Concern Scales Higher T-scores indicate MORE concerns.			Adaptive Scales Higher T-scores indicate LESS concerns.	
	Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
<a href="#">Black, Juliana</a>	57	54	52	47	54
<a href="#">Chan, Christina</a>	55	56	54	53	66
<a href="#">Deppe, Joey</a>	63	60	75	73	54
<a href="#">Farrah, Colleen</a>	83	81	59	39	31
<a href="#">Fernandez, Jose</a>	57	50	52	47	51
<a href="#">Fong, Gary</a>	46	56	55	47	54
<a href="#">Klumming, Heatbet</a>	81	68	66	39	34
<a href="#">Lofa, Jojo</a>	69	54	52	47	54
<a href="#">Nathanson, Andy</a>	71	81	56	49	34
<a href="#">Platt, Ken</a>	57	54	54	47	47
<a href="#">Pratt, Beverly</a>	55	50	51	53	51
<a href="#">Sakamoto, Catherine</a>	57	50	56	47	47
<a href="#">Smith, Janice</a>	55	56	51	49	47
<a href="#">Taylor, Dudley</a>	63	60	66	47	46
<b>Total in High Risk</b>	3 21 %	2 14 %	1 7 %		
<b>Total in Some Risk</b>	3 21 %	3 21 %	2 14 %		
<b>Total in Low Risk</b>	8 57 %	9 64 %	11 79 %		
<b>Total in Concern</b>				2 14 %	3 21 %
<b>Total in Typical</b>				11 79 %	10 71 %
<b>Total in Strength</b>				1 7 %	1 7 %

Sort by Student Last Name  Sort by Student First Name

Note. All names are fictional.

# Appendix E.4

## Standard Individual Assessment Report

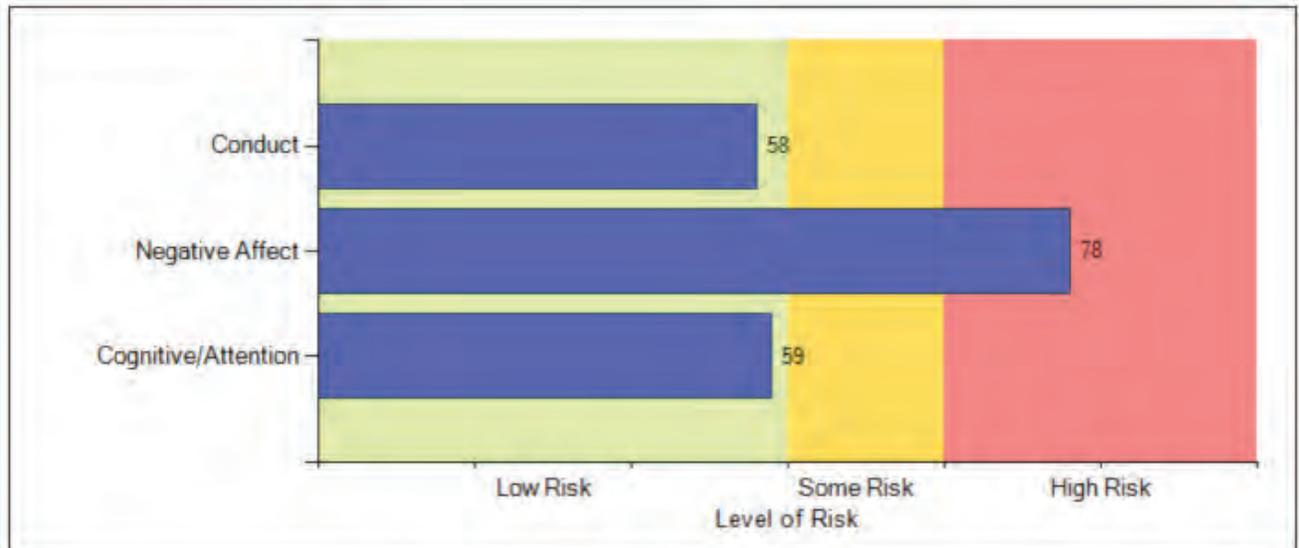
### Standard Individual Assessment Report BIMAS™-Teacher Standard

Springfield School District  
Lincoln Middle School

Student: John L Oakes  
Gender: Male  
Age: 12  
DOB: 06/13/1990  
Grade: 6  
Class: English 1200  
Rater: T. Mr. Steve Murr  
Date of Assessment: 12/9/2010  
Universal Assessment: 1

#### Behavioral Concern Scales: T-scores

Higher scores indicate MORE concerns.



The following table summarizes the BIMAS-Teacher Standard results for John L Oakes on the Behavioral Concern Scales. Please refer to the *BIMAS Technical Manual* for more information on the interpretation of these results.

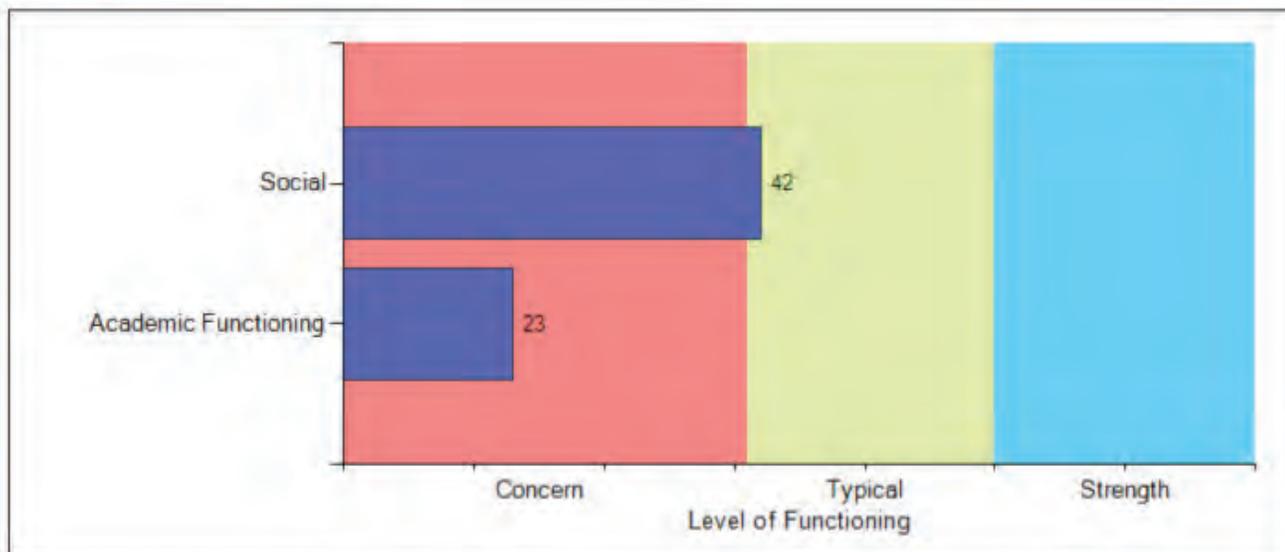
Scale	T-score	Percentile	90% CI	Scale Descriptor
Conduct	58	92	54-62	Low Risk
Negative Affect	78	99	72-84	High Risk
Cognitive/Attention	59	82	54-64	Low Risk

Note: CI = Confidence Interval.

Note. All names are fictional.

**Adaptive Scales: T-scores**

Higher scores indicate FEWER concerns.



The following table summarizes the BIMAS–Teacher Standard results for John L Oakes on the Adaptive Scales. Please refer to the *BIMAS Technical Manual* for more information on the interpretation of these results.

Scale	T-score	Percentile	90% CI	Scale Descriptor
Social	42	21	36-48	Typical
Academic Functioning	23	1	17-29	Concern

Note: CI = Confidence Interval.

Note. All names are fictional.

**Behavioral Concern Scales: Item Scores by Scale**

The following section provides the item-level analyses for each of the Behavioral Concern Scales. Items identified as Concern/Mild Concern should be further explored and considered when designing and monitoring individual intervention programs.

The following response key applies to the table in this section.

**Item Score:**

- 0 = **Never** (Observed 0 times or not observed)
- 1 = **Rarely** (Observed 1 - 2 times or to a minimum extent)
- 2 = **Sometimes** (Observed 3 - 4 times or to a moderate extent)
- 3 = **Often** (Observed 5 - 6 times or to a significant extent)
- 4 = **Very Often** (Observed 7 or more times or to an extreme extent)

<b>Behavioral Concern Scales</b> Higher scores indicate MORE concerns				
<b>Conduct</b>				
T-score	Percentile	90% CI	Scale Descriptor	
58	92	54-62	Low Risk	
Item			Item Score	Item Descriptor
21	lost temper		2	Concern
2	angry		1	Mild Concern
13	fought (verbally/physically/both)		1	Mild Concern
9	risky behavior		0	No Concern
17	lied/cheated		0	No Concern
25	aggressive		0	No Concern
29	alcohol and/or drug use		0	No Concern
31	sent to disciplinary authority		0	No Concern
32	tobacco use		0	No Concern
<b>Negative Affect</b>				
T-score	Percentile	90% CI	Scale Descriptor	
78	99	72-84	High Risk	
Item			Item Score	Item Descriptor
12	sad/withdrawn		4	Concern
24	thoughts of hurting self		3	Concern
8	depressed		2	Concern
5	sleepy/tired		2	Mild Concern
27	emotional or easily upset		2	Mild Concern
20	anxious		1	No Concern
16	embarrassed/ashamed		0	No Concern
<b>Cognitive/Attention</b>				
T-score	Percentile	90% CI	Scale Descriptor	
59	82	54-64	Low Risk	
Item			Item Score	Item Descriptor
10	problems staying on task		3	Concern
22	trouble with organization and planning		2	Mild Concern
28	fidged		2	Mild Concern
3	trouble paying attention		1	No Concern
18	trouble remembering		1	No Concern
6	impulsive		0	No Concern
14	acted without thinking		0	No Concern

Note: CI = Confidence Interval.

Note. All names are fictional.

**Adaptive Scales: Item Scores by Scale**

The following section provides the item-level analyses for each of the Adaptive Scales. Items identified as Concern/Mild Concern should be further explored and considered when designing and monitoring individual intervention programs. In addition, an examination of items identified as Positive may also aid in intervention program design.

The following response key applies to the table in this section.

(R) = This item has been reverse scored.

Item Score:

- 0 = Never (Observed 0 times or not observed)
- 1 = Rarely (Observed 1 - 2 times or to a minimum extent)
- 2 = Sometimes (Observed 3 - 4 times or to a moderate extent)
- 3 = Often (Observed 5 - 6 times or to a significant extent)
- 4 = Very Often (Observed 7 or more times or to an extreme extent)

Adaptive Scales				
Higher scores indicate FEWER concerns				
Social				
T-score	Percentile	90% CI	Scale Descriptor	
42	21	36-48	Typical	
Item			Item Score	Item Descriptor
15	relating with others		2	Concern
1	shared thoughts		2	Mild Concern
11	maintained friendships		3	Fair
19	friendly		3	Fair
23	worked out problems with others		3	Fair
7	spoke clearly with others		4	Fair
Academic Functioning				
T-score	Percentile	90% CI	Scale Descriptor	
23	1	17-29	Concern	
Item			Item Score	Item Descriptor
33	prepared for class		0	Concern
26	failing grades (R)		1	Concern
30	worked up to academic potential		1	Concern
34	absent from school (R)		1	Concern
4	followed directions		2	Mild Concern

Note: (R) = Reverse scored Item. CI = Confidence Interval.

**Summary of Results**

The following section summarizes the BIMAS–Teacher Standard results for John L Oakes. Scores reported in this section include the obtained T-score along with the 90% confidence interval (CI: i.e., there is a 90% probability that the true T-score falls within this range) and percentile ranking (i.e., percentage of students of similar age in the general population obtaining a score at or below the one obtained by John), as well as the risk level of the score. Higher scores on the Behavioral Concern Scales indicate a greater number of problematic behaviors, while higher scores on the Adaptive Scales indicates a greater number of positive behaviors.

**Behavioral Concern Scales**

Ratings of John L Oakes on the **Conduct scale** yielded a T-score of 58 (90% CI = 54-62), which fell in the **Low Risk** range and was ranked at the **92<sup>nd</sup> percentile**. Ratings on this scale indicate the youth’s risk level in manifesting externalizing problems related to anger management, aggression, substance use, and/or deviance.

Ratings of John L Oakes on the **Negative Affect scale** yielded a T-score of 78 (90% CI = 72-84), which fell in the **High Risk** range and was ranked at the **99<sup>th</sup> percentile**. Ratings on this scale indicate the youth’s risk level in manifesting internalizing problems related to anxiety and/or depression.

Ratings of John L Oakes on the **Cognitive/Attention scale** yielded a T-score of 59 (90% CI = 54-64), which fell in the **Low Risk** range and was ranked at the **82<sup>nd</sup> percentile**. Ratings on this scale indicate the youth’s risk level in manifesting problems related to inattention, impulsivity, planning/organizational skills, and/or memory.

**Adaptive Scales**

Ratings of John L Oakes on the **Social scale** yielded a T-score of 42 (90% CI = 36-48), which fell in the **Typical** range and was ranked at the **21<sup>st</sup> percentile**. Ratings on this scale indicate the youth’s willingness and capacity to successfully engage in behaviors that develop and maintain relationships with others, such as sharing thoughts and being friendly.

Ratings of John L Oakes on the **Academic Functioning scale** yielded a T-score of 23 (90% CI = 17-29), which fell in the **Concern** range and was ranked at the **1<sup>st</sup> percentile**. Ratings on this scale indicate the youth’s willingness and capacity to perform well in an academic setting, demonstrated by behaviors such as following directions, maintaining good attendance, and/or being prepared for classes.

Note. All names are fictional.



# Appendix E.5

## Demographics Breakdown

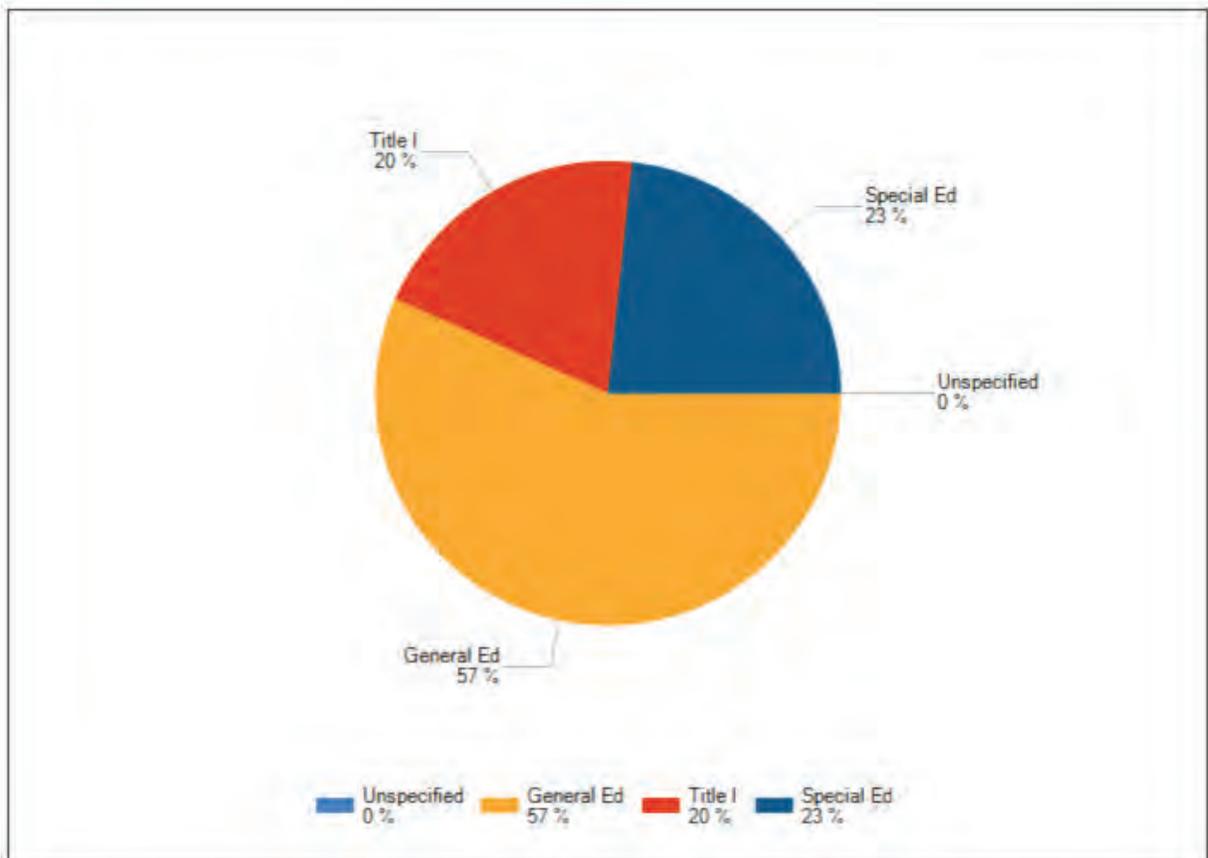
Pie Chart Version (Default)

### Service Code Breakdown BIMAS™-Teacher Standard

Springfield School District  
2010-2011

Universal Assessment:1

Schools Selected: Lincoln Middle School  
MacDonald Elementary  
Rouge Elementary  
Grades Selected: K, 1, 2, 3, 4, 5, 6, 7, 8



Service Code	Number of Students	Percent of Students
Unspecified	0	0 %
General Ed	285	57 %
Title I	100	20 %
Special Ed	115	23 %
<b>Total</b>	<b>500</b>	<b>100%</b>

Display in Histogram

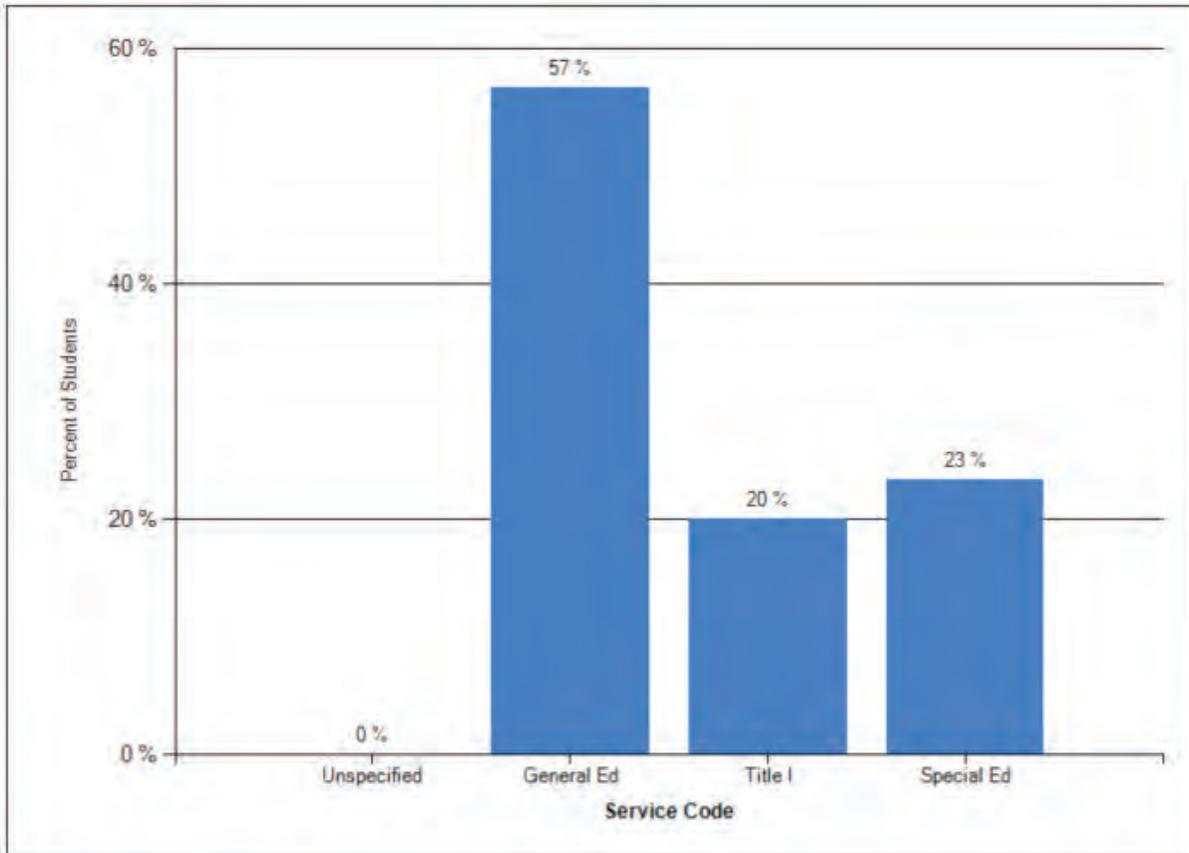
### Histogram Version

#### Service Code Breakdown BIMAS™-Teacher Standard

Springfield School District  
2010-2011

Universal Assessment:1

Schools Selected: Lincoln Middle School  
MacDonald Elementary  
Rouge Elementary  
Grades Selected: K, 1, 2, 3, 4, 5, 6, 7, 8



Display Y-Axis in **Percent** of Students

Display Y-Axis in **Number** of Students

Service Code	Number of Students	Percent of Students
Unspecified	0	0 %
General Ed	285	57 %
Title I	100	20 %
Special Ed	115	23 %
<b>Total</b>	<b>500</b>	<b>100%</b>

Display in Pie Chart

Note. All names are fictional.

# Appendix E.6

## Risk Level by Demographics

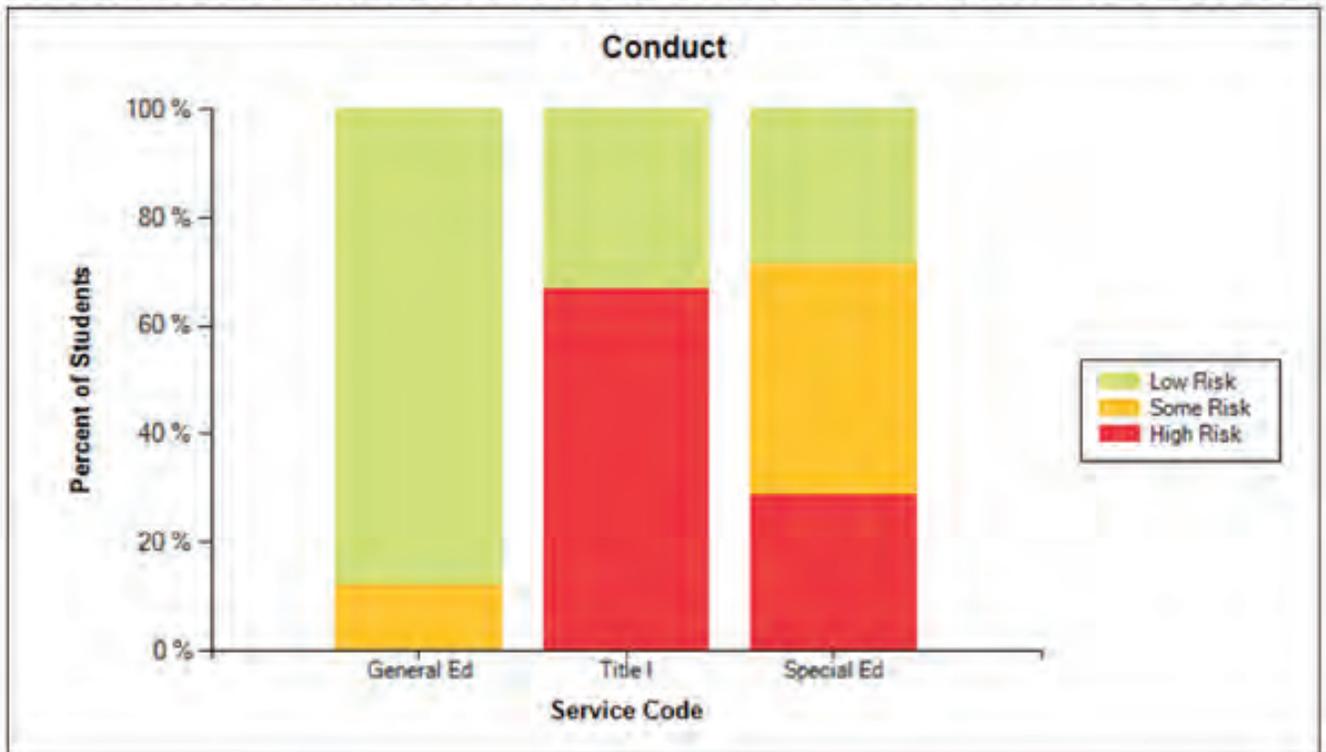
### Risk Level by Service Code BIMAS™-Teacher Standard

Springfield School District  
2010-2011

Universal Assessment:1

Schools Selected: Lincoln Middle School  
MacDonald Elementary  
Rouge Elementary  
Grades Selected: K, 1, 2, 3, 4, 5, 6, 7, 8

Click on any of the columns in the graph to obtain a Student List by Risk Level report filtered by that specific demographic variable category.



Select graphical display:

- Show High Risk Only
- Show High & Some Risk Only
- Show All Levels of Risk



Service Code	T-score			Low Risk Total	T-score		Some Risk Total	T-score			High Risk Total	Total
	< 50	50-54	55-59		60-64	65-69		70-74	75-79	80-85		
General Ed	18 %	12 %	59 %	88 %	6 %	6 %	12 %	0 %	0 %	0 %	0 %	100 %
Title I	17 %	0 %	17 %	33 %	0 %	0 %	0 %	33 %	0 %	33 %	67 %	100 %
Special Ed	29 %	0 %	0 %	29 %	14 %	29 %	43 %	0 %	0 %	29 %	29 %	100 %

DISPLAY TABLE BY NUMBER OF STUDENTS

Note. All names are fictional.

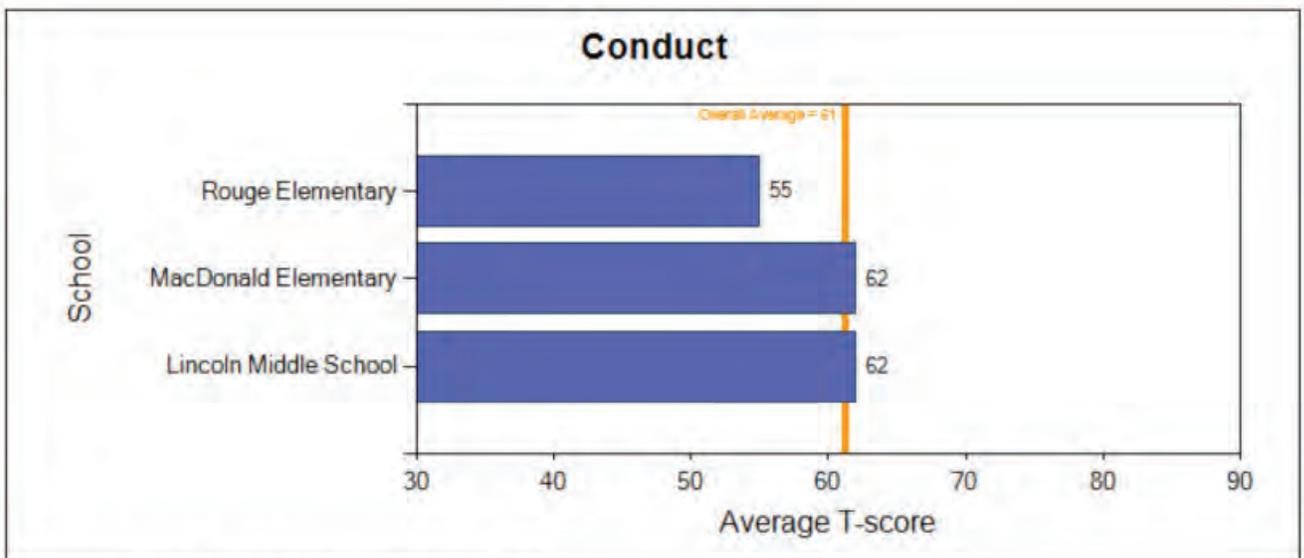
# Appendix E.7

## Average Score Comparison Report

### Average Score Comparison Report BIMAS™-Teacher Standard

Springfield School District  
2010-2011

Universal Assessment: 1  
 Schools Selected: Lincoln Middle School  
 MacDonald Elementary  
 Rouge Elementary  
 Grade: K, 1, 2, 3, 4, 5, 6, 7, 8  
 Service Code Selected: General Ed, Title I, Special Ed



Service Code Selected: General Ed, Title I, Special Ed

School	Universal Assessment 1	
	Average T-score	Average Level of Risk
Rouge Elementary	55	Low Risk
MacDonald Elementary	62	Some Risk
Lincoln Middle School	62	Some Risk
<b>Overall Average</b>	<b>61</b>	<b>Some Risk</b>

Note: Overall average denotes the mean score of all the students in the current selection rather than an arithmetic mean of each school's average T-score.

Note. All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



# Appendix E.8

## Standard Individual Comparison Between Raters

### Standard Individual Comparison Between Raters BIMAS™ Standard

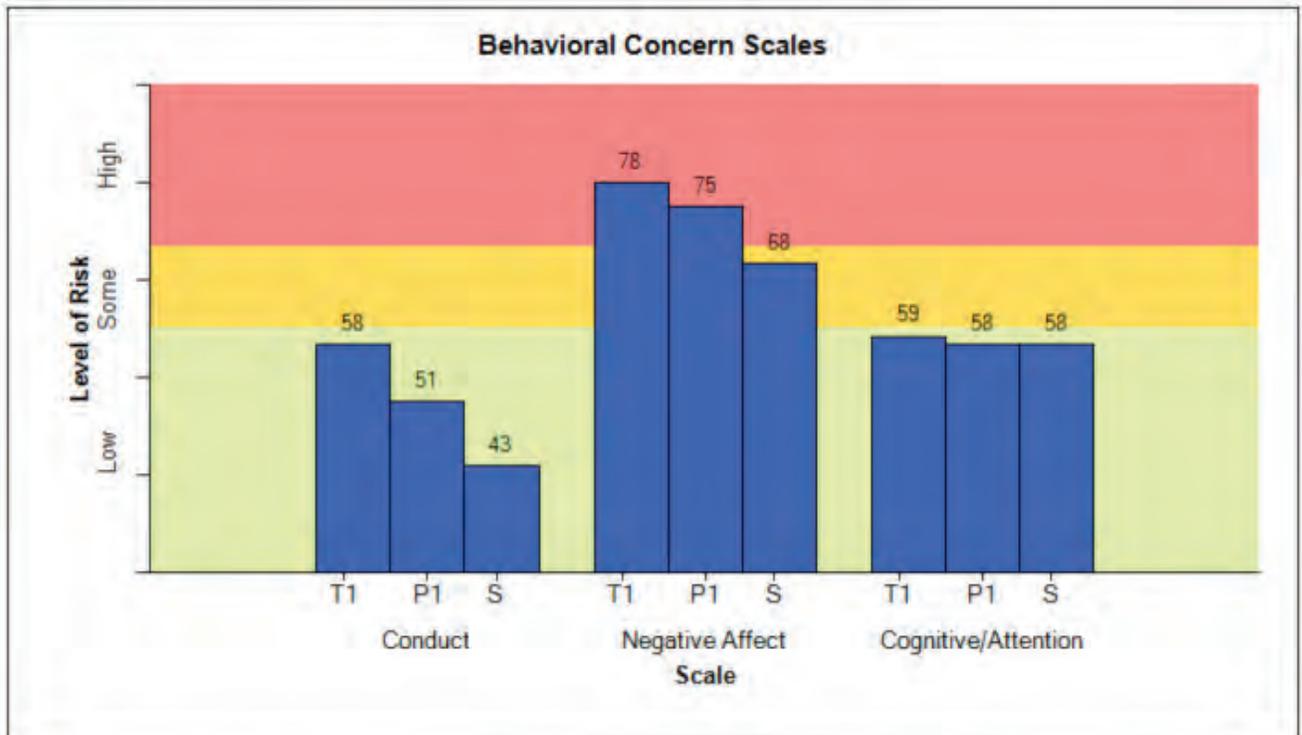
Springfield School District  
Lincoln Middle School

Student: John L. Oakes  
Gender: Male  
Age: 12  
DOB: 6/13/1990  
Grade: 6  
Class: English 1200

Rater	Teacher 1	Parent 1	Self
Rater Name	Mr. Steve Murr	Peggy Oakes	John Oakes
Date of Assessment	12/09/2010	12/10/2010	12/10/2010

### Behavioral Concern Scales: T-scores Comparison Between Raters

Higher scores indicate MORE concerns



Note. All names are fictional.

## Behavior Intervention Monitoring Assessment System (BIMAS™)

The following table displays T-scores, Confidence Intervals, and Percentiles for each Behavioral Concern scale, as well as any statistically significant (90% level of significance, adjusted for multiple comparisons), differences in T-scores between pairs of raters. If a pair of ratings is not noted in the "Statistically Significant Differences Between Raters" column, then the difference between those two raters did not reach statistical significance.

**Note:** CI = Confidence Interval, P = Parent, T = Teacher, and S = Self.

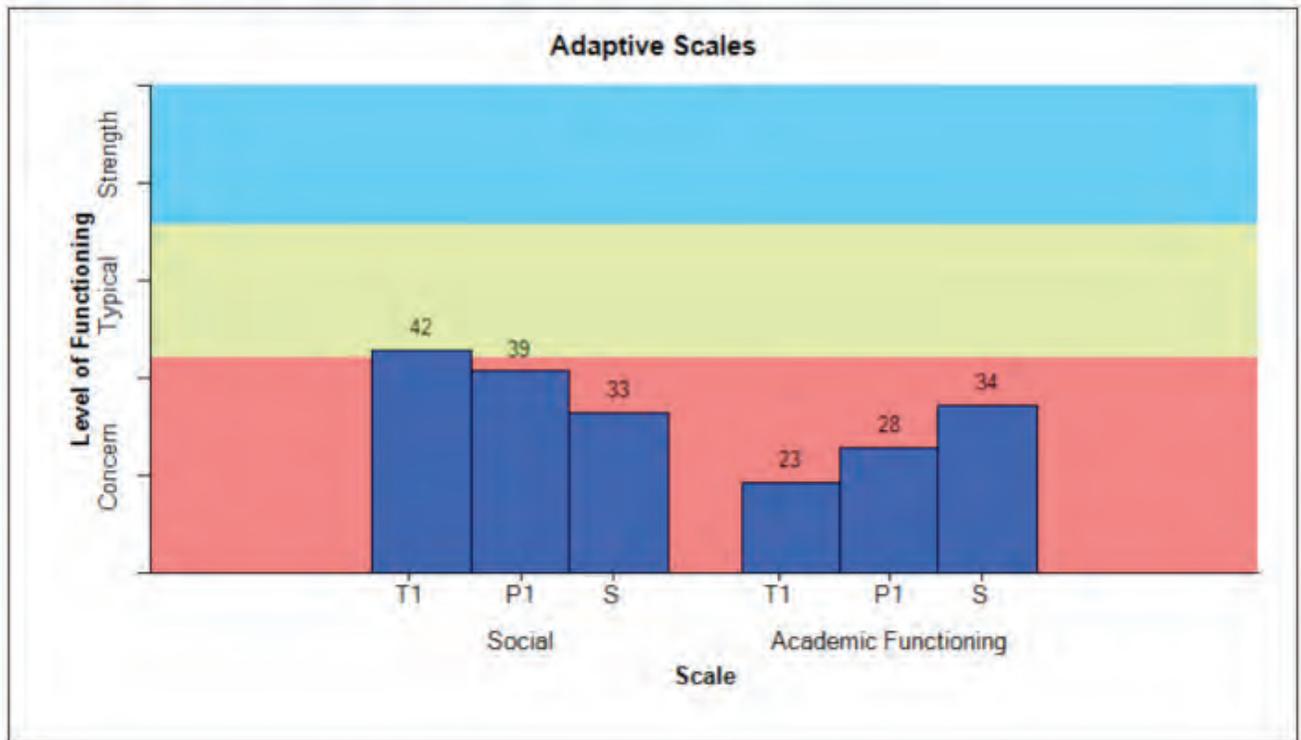
Click on any of the Raters in the table below to view the corresponding BIMAS Standard Individual Assessment Report by the rater.

Scale		Rater			Statistically Significant Differences Between Raters
		T1	P1	S	
Conduct	T-score	58	51	43	T1 > P1, S; P1 > S;
	90% CI	54-62	46-56	38-48	
	Percentile	92	55	25	
	Level of Risk	Low Risk	Low Risk	Low Risk	
Negative Affect	T-score	78	75	68	T1 > S;
	90% CI	72-84	67-83	62-74	
	Percentile	99	99	96	
	Level of Risk	High Risk	High Risk	Some Risk	
Cognitive/Attention	T-score	59	58	58	
	90% CI	54-64	52-64	51-65	
	Percentile	82	79	80	
	Level of Risk	Low Risk	Low Risk	Low Risk	

*Note.* All names are fictional.

### Adaptive Scales: T-scores Comparison Between Raters

Higher scores indicate FEWER concerns



The following table displays T-scores, Confidence Intervals, and Percentiles for each Adaptive Scale, as well as any statistically significant (90% level of significance, adjusted for multiple comparisons), differences in T-scores between pairs of raters. If a pair of ratings is not noted in the "Statistically Significant Differences Between Raters" column, then the difference between those two raters did not reach statistical significance.

**Note:** CI = Confidence Interval, P = Parent, T = Teacher, and S = Self.

Click on any of the Raters in the table below to view the corresponding BIMAS Standard Individual Assessment Report by the rater.

Scale		Rater			Statistically Significant Differences Between Raters
		T1	P1	S	
Social	T-score	42	39	33	
	90% CI	36-48	32-46	25-41	
	Percentile	21	13	4	
	Level of Functioning	Typical	Concern	Concern	
Academic Functioning	T-score	23	28	34	S > T1;
	90% CI	17-29	18-38	26-42	
	Percentile	1	2	5	
	Level of Functioning	Concern	Concern	Concern	

*Note.* All names are fictional.

**Behavioral Concern Scales: Item Scores Comparison Between Raters by Scale**

Higher scores indicate MORE concerns

The following table provides an item-level analysis for each of the Behavioral Concern Scales. Items identified as Concern/Mild Concern by any of the raters should be further explored and considered when designing and monitoring individual intervention programs.

P = Parent, T = Teacher, and S = Self.

The following response key applies to the table in this section.

Item Score:

- 0 = Never (Observed 0 times or not observed)
- 1 = Rarely (Observed 1 - 2 times or to a minimum extent)
- 2 = Sometimes (Observed 3 - 4 times or to a moderate extent)
- 3 = Often (Observed 5 - 6 times or to a significant extent)
- 4 = Very Often (Observed 7 or more times or to an extreme extent)
- ? = Omitted Item

- = Concern
- = Mild Concern
- = No Concern

<b>Behavioral Concern Scales</b>				
Higher scores indicate MORE concerns				
Item No.	Item Stem	Rater's Item Raw Score		
		T1	P1	S
<b>Conduct</b>				
2	angry	1 Mild Concern	1 No Concern	0 No Concern
9	risky behavior(s)	0 No Concern	0 No Concern	0 No Concern
13	fought (verbally/physically/both)	1 Mild Concern	0 No Concern	0 No Concern
17	lied/cheated	0 No Concern	0 No Concern	0 No Concern
21	lost temper	2 Concern	1 No Concern	1 No Concern
25	physical/verbal aggressiveness	0 No Concern	0 No Concern	0 No Concern
29	alcohol and/or drug use	0 No Concern	0 No Concern	0 No Concern
31	referred to disciplinary authority	0 No Concern	0 No Concern	0 No Concern
32	tobacco use	0 No Concern	0 No Concern	0 No Concern
<b>Scale Score</b>		T = 58	T = 51	T = 43
<b>Scale Descriptor</b>		Low Risk	Low Risk	Low Risk

Note. All names are fictional.

Appendix E.8: Standard Individual Comparison Between Raters

Item No.	Item Stem	Rater's Item Raw Score		
		T1	P1	S
<b>Negative Affect</b>				
5	sleepy/tired	2 Mild Concern	4 Concern	3 Mild Concern
8	depressed	2 Concern	3 Concern	3 Concern
12	sad/withdrawn	4 Concern	3 Concern	2 Mild Concern
16	embarrassed/ashamed	0 No Concern	1 No Concern	2 Mild Concern
20	anxious/worried/nervous	1 No Concern	1 No Concern	1 No Concern
24	thoughts of hurting self	3 Concern	3 Concern	3 Concern
27	emotional/easily upset	2 Mild Concern	3 Concern	2 Mild Concern
Scale Score Scale Descriptor		T = 78 High Risk	T = 75 High Risk	T = 68 Some Risk
Item No.	Item Stem	Rater's Item Raw Score		
		T1	P1	S
<b>Cognitive/Attention</b>				
3	trouble paying attention	1 No Concern	2 No Concern	3 Concern
6	impulsive	0 No Concern	1 No Concern	0 No Concern
10	problems staying on task	3 Concern	2 No Concern	2 No Concern
14	acted without thinking	0 No Concern	1 No Concern	2 No Concern
18	trouble remembering	1 No Concern	1 No Concern	2 Mild Concern
22	organization & planning problems	2 Mild Concern	2 Mild Concern	2 Mild Concern
28	fidgeded	2 Mild Concern	2 Mild Concern	1 No Concern
Scale Score Scale Descriptor		T = 59 Low Risk	T = 58 Low Risk	T = 58 Low Risk

Note. All names are fictional.

**Adaptive Scales: Item Scores Comparison Between Raters by Scale**

Higher scores indicate FEWER concerns

The following table provides the item-level analysis for each of the Adaptive Scales. Items identified as Concern/Mild Concern by any of the raters should be further explored and considered when designing and monitoring individual intervention programs. In addition, an examination of items identified as Positive may also aid in intervention program design.

P = Parent, T = Teacher, and S = Self.

The following response key applies to the table in this section.

(R) = This item has been reverse scored.

**Item Score:**

0 = Never (Observed 0 times or not observed)

1 = Rarely (Observed 1 - 2 times or to a minimum extent)

2 = Sometimes (Observed 3 - 4 times or to a moderate extent)

3 = Often (Observed 5 - 6 times or to a significant extent)

4 = Very Often (Observed 7 or more times or to an extreme extent)

? = Omitted Item

= Concern

= Mild Concern

= Fair

= Positive

<b>Adaptive Scales</b>				
Higher scores indicate FEWER concerns				
Item No.	Item Stem	Rater's Item Raw Score		
		T1	P1	S
<b>Social</b>				
1	shared thoughts	2 Mild Concern	1 Concern	1 Concern
7	communicated clearly	4 Fair	4 Positive	3 Fair
11	maintained friendships	3 Fair	3 Fair	2 Concern
15	comfortable with social interactions	2 Concern	2 Mild Concern	0 Concern
19	friendly	3 Fair	3 Fair	2 Mild Concern
23	worked out problems with others	3 Fair	1 Concern	2 Fair
Scale Score Scale Descriptor		T = 42 Typical	T = 39 Concern	T = 33 Concern
Item No.	Item Stem	Rater's Item Raw Score		
		T1	P1	S
<b>Academic Functioning</b>				
4	followed directions	2 Mild Concern	2 Mild Concern	3 Fair
26	failing grades(R)	3 Concern	3 Concern	3 Concern
30	worked up to academic potential	1 Concern	1 Concern	2 Mild Concern
33	prepared for classes	0 Concern	2 Mild Concern	2 Mild Concern
34	absent from school(R)	3 Concern	2 Concern	2 Mild Concern
Scale Score Scale Descriptor		T = 23 Concern	T = 28 Concern	T = 34 Concern

(R) = Reverse scored item.

Note. All names are fictional.

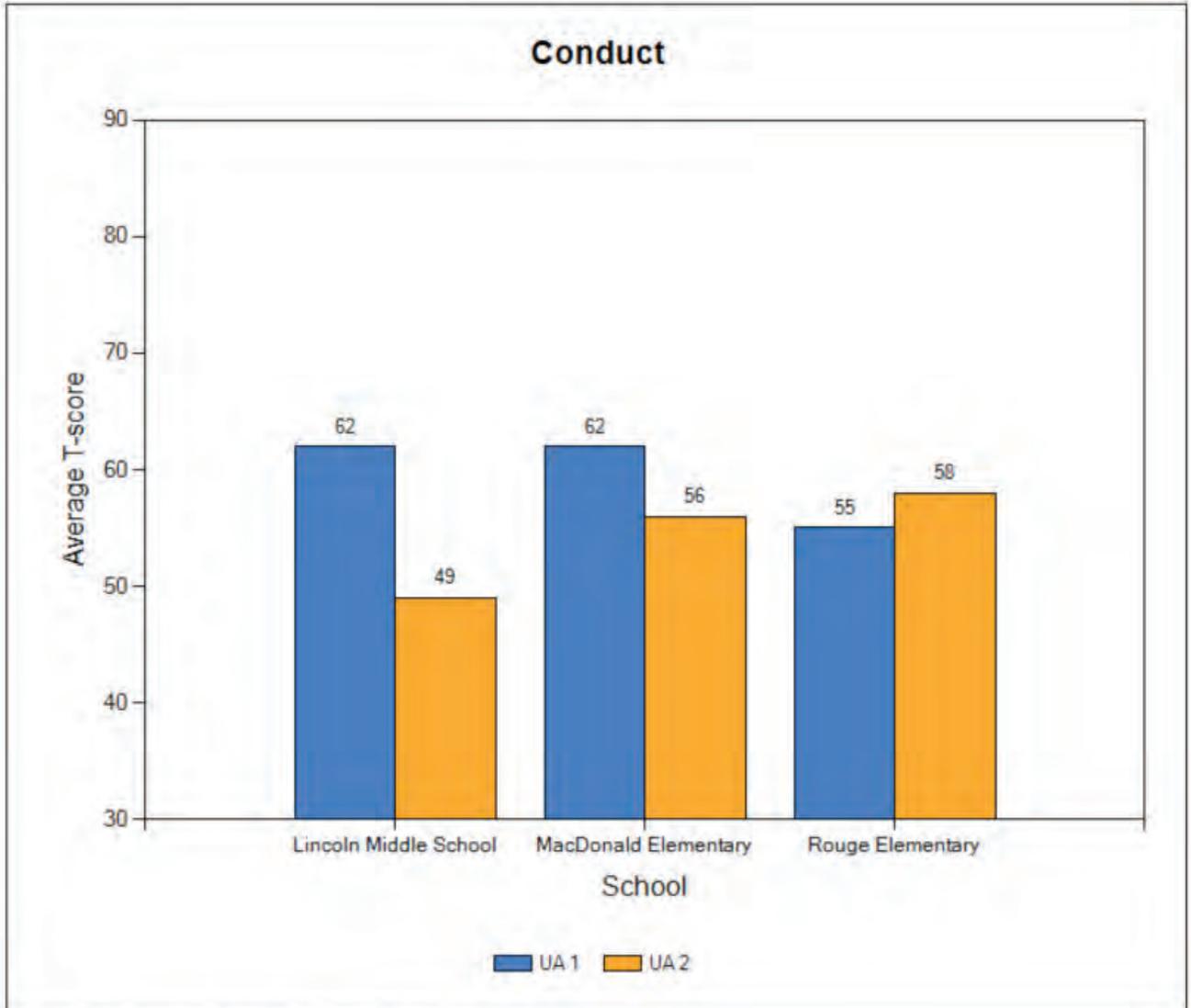
# Appendix E.9

## Average Score Comparison: Progress Report

### Average Score Comparison: Progress Report BIMAS™-Teacher Standard

Springfield School District  
2010

Universal Assessment(s): 1, 2  
 School(s) Selected: Lincoln Middle School  
 MacDonald Elementary  
 Rouge Elementary  
 Grade(s) Selected: All  
 Service Code Selected: General Ed, Title I, Special Ed



Service Code Selected: General Ed, Title I, Special Ed

Behavior Intervention Monitoring Assessment System (BIMAS™)

School	Universal Assessment 1		Universal Assessment 2	
	Average T-score	Average Level of Risk	Average T-score	Average Level of Risk
Lincoln Middle School	62	Some Risk	49	Low Risk
MacDonald Elementary	62	Some Risk	56	Low Risk
Rouge Elementary	55	Low Risk	58	Low Risk
<b>Overall Average</b>	<b>61</b>	<b>Some Risk</b>	<b>55</b>	<b>Low Risk</b>

**Note:** Overall average denotes the mean score of all the students in the current selection rather than an arithmetic mean of each school's average T-score.

*Note.* All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



# Appendix E.10

## Class/Group

### Student Scores: Progress Report

#### Class/Group Student Scores: Progress Report BIMAS™-Teacher Standard

Springfield School District  
MacDonald Elementary  
2010–2011

Grade: 4  
Class: 4A  
Universal Assessments: 1,2

Student Name	Conduct Scale	
	Higher T-scores indicate MORE concerns.	
	UA1	UA2
<a href="#">Black, Juliana</a>	57	57
<a href="#">Chan, Christina</a>	55	53
<a href="#">Deppe, Joey</a>	63	55
<a href="#">Farrah, Colleen</a>	83	65
<a href="#">Fernandez, Jose</a>	57	53
<a href="#">Fong, Gary</a>	46	59
<a href="#">Klumming, Heather</a>	81	62
<a href="#">Lola, Jojo</a>	69	46
<a href="#">Nathanson, Andy</a>	71	65
<a href="#">Platt, Ken</a>	57	55
<a href="#">Pratt, Beverly</a>	55	51
<a href="#">Sakamoto, Catherine</a>	57	53
<a href="#">Smith, Janice</a>	55	53
<a href="#">Taylor, Dudley</a>	63	61
<b>Total In High Risk</b>	<b>3</b> 21 %	<b>0</b> 0 %
<b>Total In Some Risk</b>	<b>3</b> 21 %	<b>4</b> 29 %
<b>Total In Low Risk</b>	<b>8</b> 57 %	<b>10</b> 71 %

Note. All names are fictional.

Behavior Intervention Monitoring Assessment System (BIMAS™)

Student Name	Negative Affect Scale	
	Higher T-scores indicate MORE concerns.	
	UA1	UA2
<a href="#">Black, Juliana</a>	54	56
<a href="#">Chan, Christina</a>	56	54
<a href="#">Deppe, Joey</a>	68	63
<a href="#">Farrah, Colleen</a>	81	81
<a href="#">Fernandez, Jose</a>	50	47
<a href="#">Fong, Gary</a>	56	56
<a href="#">Klumming, Heather</a>	68	63
<a href="#">Lola, Jojo</a>	54	47
<a href="#">Nathanson, Andy</a>	81	66
<a href="#">Platt, Ken</a>	54	56
<a href="#">Pratt, Beverly</a>	50	47
<a href="#">Sakamoto, Catherine</a>	50	54
<a href="#">Smith, Janice</a>	56	54
<a href="#">Taylor, Dudley</a>	68	56
<b>Total In High Risk</b>	<b>2</b> 14 %	<b>1</b> 7 %
<b>Total In Some Risk</b>	<b>3</b> 21 %	<b>3</b> 21 %
<b>Total In Low Risk</b>	<b>9</b> 64 %	<b>10</b> 71 %

Note. All names are fictional.

Student Name	Cognitive/Attention Scale	
	Higher T-scores indicate MORE concerns.	
	UA1	UA2
<a href="#">Black, Juliana</a>	52	51
<a href="#">Chan, Christina</a>	54	52
<a href="#">Deppe, Joey</a>	75	66
<a href="#">Farrah, Colleen</a>	59	55
<a href="#">Fernandez, Jose</a>	52	50
<a href="#">Fong, Gary</a>	55	54
<a href="#">Klumming, Heather</a>	66	65
<a href="#">Lola, Jojo</a>	52	47
<a href="#">Nathanson, Andy</a>	56	55
<a href="#">Platt, Ken</a>	54	52
<a href="#">Pratt, Beverly</a>	51	50
<a href="#">Sakamoto, Catherine</a>	56	55
<a href="#">Smith, Janice</a>	51	50
<a href="#">Taylor, Dudley</a>	66	54
<b>Total In High Risk</b>	<b>1</b> 7 %	<b>0</b> 0 %
<b>Total In Some Risk</b>	<b>2</b> 14 %	<b>2</b> 14 %
<b>Total In Low Risk</b>	<b>11</b> 79 %	<b>12</b> 86 %

Note. All names are fictional.

Behavior Intervention Monitoring Assessment System (BIMAS™)

Student Name	Social Scale Higher T-scores indicate FEWER concerns	
	UA1	UA2
<a href="#">Black, Juliana</a>	47	53
<a href="#">Chan, Christina</a>	53	57
<a href="#">Deppe, Joey</a>	73	65
<a href="#">Farrah, Colleen</a>	39	43
<a href="#">Fernandez, Jose</a>	47	51
<a href="#">Fong, Gary</a>	47	43
<a href="#">Klumming, Heather</a>	39	53
<a href="#">Lola, Jojo</a>	47	55
<a href="#">Nathanson, Andy</a>	49	51
<a href="#">Platt, Ken</a>	47	65
<a href="#">Pratt, Beverly</a>	53	65
<a href="#">Sakamoto, Catherine</a>	47	49
<a href="#">Smith, Janice</a>	49	51
<a href="#">Taylor, Dudley</a>	47	45
<b>Total In Concern</b>	<b>2</b> 14 %	<b>0</b> 0 %
<b>Total In Typical</b>	<b>11</b> 79 %	<b>11</b> 79 %
<b>Total In Strength</b>	<b>1</b> 7 %	<b>3</b> 21 %

Note. All names are fictional.

Student Name	Academic Functioning Scale Higher T-scores indicate FEWER concerns	
	UA1	UA2
<a href="#">Black, Juliana</a>	54	51
<a href="#">Chan, Christina</a>	66	66
<a href="#">Deppe, Joey</a>	54	43
<a href="#">Farrah, Colleen</a>	31	43
<a href="#">Fernandez, Jose</a>	51	54
<a href="#">Fong, Gary</a>	54	49
<a href="#">Klumming, Heather</a>	34	43
<a href="#">Lola, Jojo</a>	54	66
<a href="#">Nathanson, Andy</a>	34	43
<a href="#">Platt, Ken</a>	47	51
<a href="#">Pratt, Beverly</a>	51	49
<a href="#">Sakamoto, Catherine</a>	47	47
<a href="#">Smith, Janice</a>	47	51
<a href="#">Taylor, Dudley</a>	45	45
<b>Total In Concern</b>	<b>3</b> 21 %	<b>0</b> 0 %
<b>Total In Typical</b>	<b>10</b> 71 %	<b>12</b> 86 %
<b>Total In Strength</b>	<b>1</b> 7 %	<b>2</b> 14 %

Sort by Student Last Name  Sort by Student First Name

Note. All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



# Appendix E.11

## Risk Level by Demographics: Progress Report

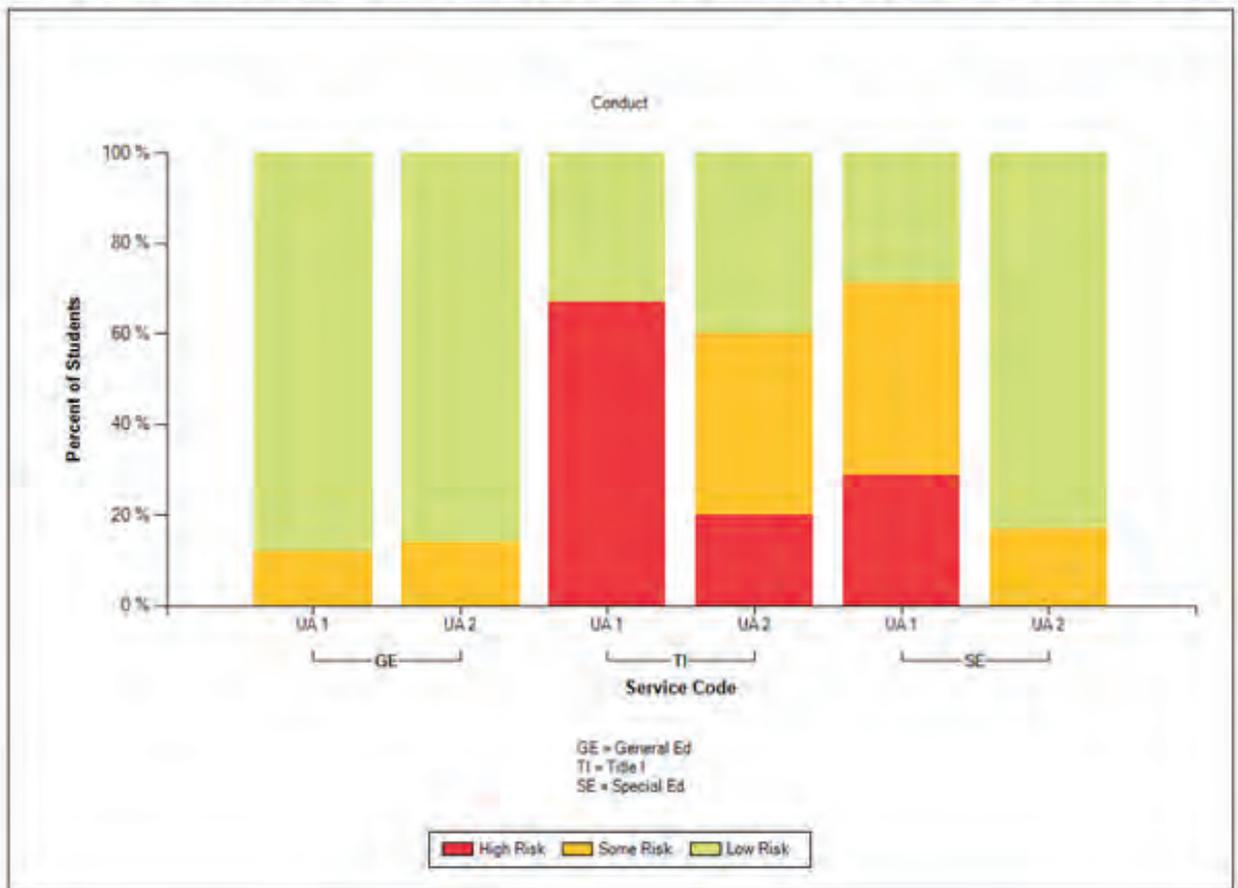
### Risk Level by Service Code: Progress Report BIMAS™-Teacher Standard

Springfield School District  
2010

Universal Assessment: 1, 2

Schools Selected: Lincoln Middle School  
MacDonald Elementary  
Rouge Elementary  
Grades Selected: K, 1, 2, 3, 4, 5, 6, 7, 8

Click on any of the columns in the graph to obtain a Student List by Risk Level report filtered by that specific demographic variable category.



Select graphical display:

- Show High Risk Only
- Show Some Risk Only
- Show Low Risk Only
- Show All Levels of Risk



Note. All names are fictional.

Behavior Intervention Monitoring Assessment System (BIMAS™)

Service Code	UA	T-score			Lbvr Risk Total	T-score		Some Risk Total	T-score			High Risk Total	Total
		< 50	50-54	55-59		60-64	65-69		70-74	75-79	80-85		
General Ed	UA 1	18 %	12 %	59 %	88 %	6 %	6 %	12 %	0 %	0 %	0 %	0 %	100 %
	UA 2	29 %	36 %	21 %	86 %	14 %	0 %	14 %	0 %	0 %	0 %	0 %	100 %
Title I	UA 1	17 %	0 %	17 %	33 %	0 %	0 %	0 %	33 %	0 %	33 %	67 %	100 %
	UA 2	20 %	20 %	0 %	40 %	20 %	20 %	40 %	0 %	20 %	0 %	20 %	100 %
Special Ed	UA 1	29 %	0 %	0 %	29 %	14 %	29 %	43 %	0 %	0 %	29 %	29 %	100 %
	UA 2	17 %	0 %	67 %	83 %	0 %	17 %	17 %	0 %	0 %	0 %	0 %	100 %

DISPLAY TABLE BY NUMBER OF STUDENTS

Note. All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



# Appendix E.12

## Flex Individual Progress Report

**Flex Individual Progress Report**  
**BIMAS™-Teacher Flex**  
Springfield School District  
Lincoln Middle School

Student: John L. Oakes  
Gender: Male  
Age: 12  
DOB: 06/13/1990  
Grade: 6  
Class: English 1200  
Rater: T. Mr. Steve Murr  
Dates of Flex Entry: 12/17/2010 through 2/11/2011

### Flex Item Scores: Scale-Level Comparisons Across Administrations

The following section provides the interpretive guidelines on the Flex items across all scales.

The following response key applies to the tables and graphs for all Flex items.

#### Item Score:

0 = **Never** (Observed 0 times or not observed)

1 = **Rarely** (Observed 1 - 2 times or to a minimum extent)

2 = **Sometimes** (Observed 3 - 4 times or to a moderate extent)

3 = **Often** (Observed 5 - 6 times or to a significant extent)

4 = **Very Often** (Observed 7 or more times or to an extreme extent)

? = **Omitted Item**

N/A = Item was not available on Flex forms on the selected date(s) of Flex entry

#### Item Descriptor Legend:

##### Negative (-) Items:

Higher item scores indicate MORE concerns

 = Concern

 = Mild Concern

 = No Concern

##### Positive (+) Items:

Higher item scores indicate FEWER concerns

 = Concern

 = Mild Concern

 = Fair

 = Positive

*Note.* All names are fictional.

### A) Behavioral Concern Scales

The following table provides the progress on Flex items for the specified dates of entries on selected Behavioral Concern Scale(s).

Negative Affect Scale									
Standard Form Anchor Item:									
8. appeared depressed									
Flex Item:									
had a negative self-image (-)									
Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	3	3	2	2	1	1	0	1	0
Flex Item:									
took part in group activities (+)									
Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	1	2	3	3	3	3	3	3	3
Standard Form Anchor Item:									
24. expressed thoughts of hurting himself/herself									
Flex Item:									
expressed thoughts of self-harm (-)									
Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	3	2	1	1	1	0	0	0	0

### B) Adaptive Scales

The following table provides the progress on Flex items for the specified dates of entries on selected Adaptive Scale(s).

Academic Functioning Scale									
Standard Form Anchor Item:									
33. was prepared for class									
Flex Item:									
turned in completed work on time (+)									
Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	2	2	3	3	3	3	3	3	4
Standard Form Anchor Item:									
34. was absent from school(R)									
Flex Item:									
attended classes (+)									
Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	2	3	3	3	4	4	4	4	4

Note. All names are fictional.

**Flex Item Scores: Graphs & Comments**

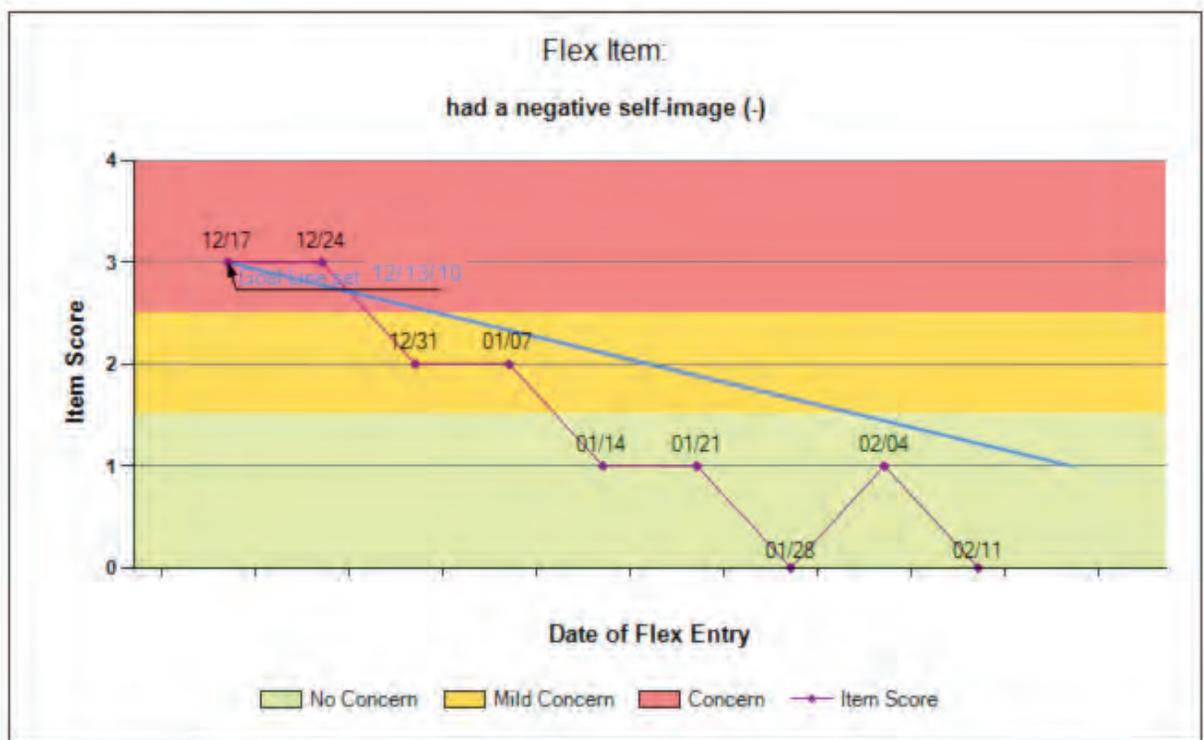
**A) Behavioral Concern Scales**

The following table provides the progress on Flex items for the specified dates of entries on selected Behavioral Concern Scale(s).

Negative Affect Scale									
Standard Form Anchor Item:									
8. appeared depressed									
Flex Item:									
had a negative self-image (-)									
Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	3	3	2	2	1	1	0	1	0

Comments:

- 12/17/10: John does not seem to focus more on his weaknesses than strengths.
- 12/24/10: He seemed very concerned about other finding him "odd."
- 12/31/10: N/A
- 01/07/11: He is making good progress.
- 01/14/11: N/A
- 01/21/11: N/A
- 01/28/11: He seems to exuberate confidence this week.
- 02/04/11: N/A
- 02/11/11: N/A



Note. All names are fictional.

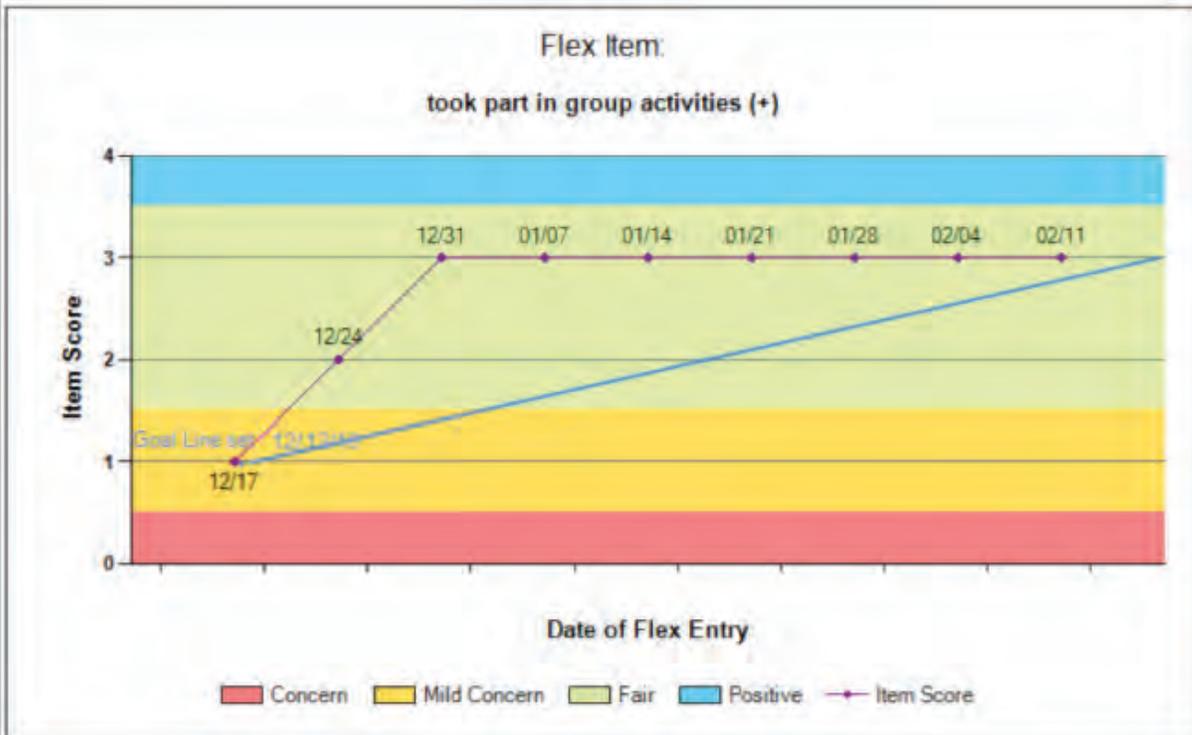
Flex Item:

took part in group activities (+)

Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	1	2	3	3	3	3	3	3	3

Comments:

- 12/17/10: Encouraged John to participate in the group treasure hunt game on Wed. Still see reluctance.
- 12/24/10: N/A
- 12/31/10: N/A
- 01/07/11: He seems more willing to open up this week.
- 01/14/11: N/A
- 01/21/11: N/A
- 01/28/11: N/A
- 02/04/11: N/A
- 02/11/11: N/A



Note. All names are fictional.

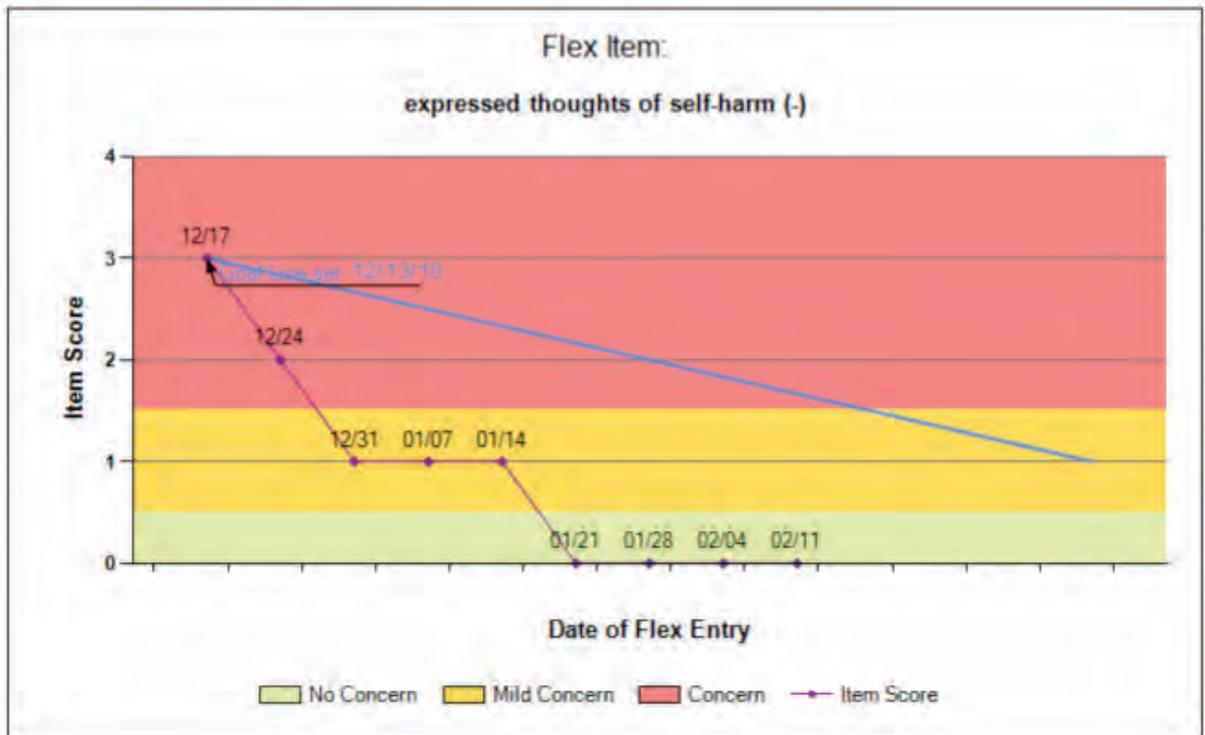
**Standard Form Anchor Item:**  
**24. expressed thoughts of hurting himself/herself**

Flex Item:  
**expressed thoughts of self-harm (-)**

Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	3	2	1	1	1	0	0	0	0

Comments:

- 12/17/10: John related to me that he sometimes thinks about harming himself.
- 12/24/10: N/A
- 12/31/10: N/A
- 01/07/11: He rarely talks about harming himself anymore.
- 01/14/11: N/A
- 01/21/11: N/A
- 01/28/11: N/A
- 02/04/11: N/A
- 02/11/11: N/A



Note. All names are fictional.

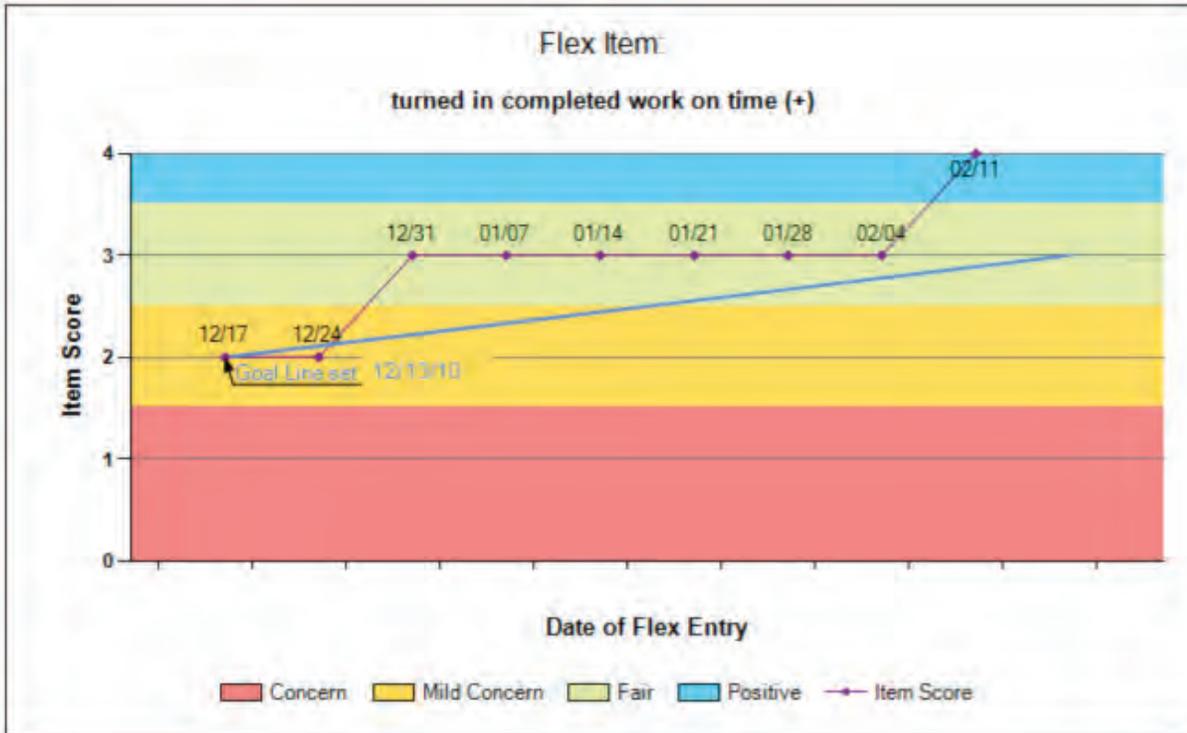
**B) Adaptive Scales**

The following table provides the progress on Flex items for the specified dates of entries on selected Adaptive Scale(s).

Academic Functioning Scale									
Standard Form Anchor Item:									
33. was prepared for class									
Flex Item:									
turned in completed work on time (+)									
Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	2	2	3	3	3	3	3	3	4

Comments:

- 12/17/10: I was able to see more effort this week.
- 12/24/10: N/A
- 12/31/10: N/A
- 01/07/11: He was on time in turing his work most of the time this week.
- 01/14/11: N/A
- 01/21/11: N/A
- 01/28/11: N/A
- 02/04/11: N/A
- 02/11/11: N/A



Note. All names are fictional.

Standard Form Anchor Item:  
 34. was absent from school(R)

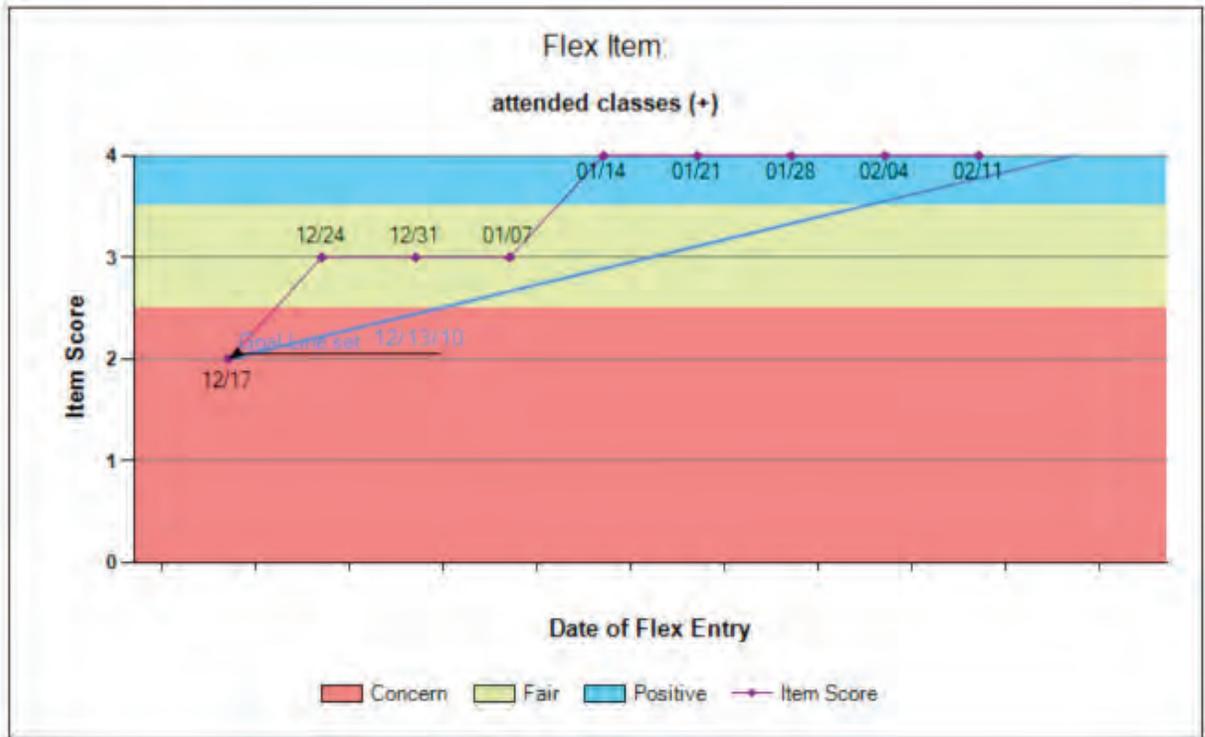
Flex Item:

attended classes (+)

Date	12/17	12/24	12/31	01/07	01/14	01/21	01/28	02/04	02/11
Item Score	2	3	3	3	4	4	4	4	4

Comments:

N/A



Note. All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



# Appendix E.13

## Flex Individual Comparison Between Raters

### Flex Individual Comparison Between Raters

#### BIMAS™ Flex

Springfield School District  
Lincoln Middle School

Student: John L Oakes  
Gender: Male  
Age: 12  
DOB: 06/13/1990  
Grade: 6  
Class: English 1200

Rater	Teacher 1	Parent 1	Self
Rater Name	Mr. Steve Murr	Peggy Oakes	John Oakes
Date of Assessment	12/17/2010	12/17/2010	12/17/2010

#### Flex Item Scores: Scale-Level Comparisons Between Raters

The following section provides the interpretive guidelines on the Flex items across all scales.

The following response key applies to the tables and graphs for all Flex items.

##### Item Score:

0 = **Never** (Observed 0 times or not observed)

1 = **Rarely** (Observed 1 - 2 times or to a minimum extent)

2 = **Sometimes** (Observed 3 - 4 times or to a moderate extent)

3 = **Often** (Observed 5 - 6 times or to a significant extent)

4 = **Very Often** (Observed 7 or more times or to an extreme extent)

? = **Omitted Item**

N/A = Item was not available on Flex forms on the selected date(s) of Flex entry

##### Item Descriptor Legend:

###### Negative (-) Items:

Higher item scores indicate MORE concerns

 = Concern

 = Mild Concern

 = No Concern

###### Positive (+) Items:

Higher item scores indicate FEWER concerns

 = Concern

 = Mild Concern

 = Fair

 = Positive

*Note.* All names are fictional.

### A) Behavioral Concern Scales

The following table provides the comparison for selected raters on Flex items for the specified date of entry on Behavioral Concern Scale(s).

Negative Affect			
Standard Form Anchor Item: 8. appeared depressed.			
Flex Item: had a negative self-image (-)			
Rater	Teacher 1	Parent 1	Self
Item Score	3 Concern	3 Concern	4 Concern
Standard Form Anchor Item: 8. appeared depressed.			
Flex Item: took part in group activities (+)			
Rater	Teacher 1	Parent 1	Self
Item Score	1 Mild Concern	1 Mild Concern	0 Concern
Standard Form Anchor Item: 24. expressed thoughts of hurting himself/herself.			
Flex Item: expressed thoughts of self-harm (-)			
Rater	Teacher 1	Parent 1	Self
Item Score	3 Concern	3 Concern	4 Concern

### B) Adaptive Scales

The following table provides the comparison for selected raters on Flex items for the specified date of entry on Adaptive Scale(s).

Academic Functioning			
Standard Form Anchor Item: 33. was prepared for class.			
Flex Item: turned in completed work on time (+)			
Rater	Teacher 1	Parent 1	Self
Item Score	2 Mild Concern	2 Mild Concern	0 Concern
Standard Form Anchor Item: 34. was absent from school.(R)			
Flex Item: attended classes (+)			
Rater	Teacher 1	Parent 1	Self
Item Score	2 Concern	3 Fair	3 Fair

Note. All names are fictional.

**Flex Item Scores: Graphs & Comments**

**A) Behavioral Concern Scales**

The following section provides the comparison for selected raters on Flex items for the specified date(s) of entry on Behavioral Concern Scale(s) in a graph.

Note: P = Parent, T = Teacher, S = Self



Note. All names are fictional.

Standard Form Anchor Item:

8. appeared depressed.

Flex Item:

took part in group activities (+)

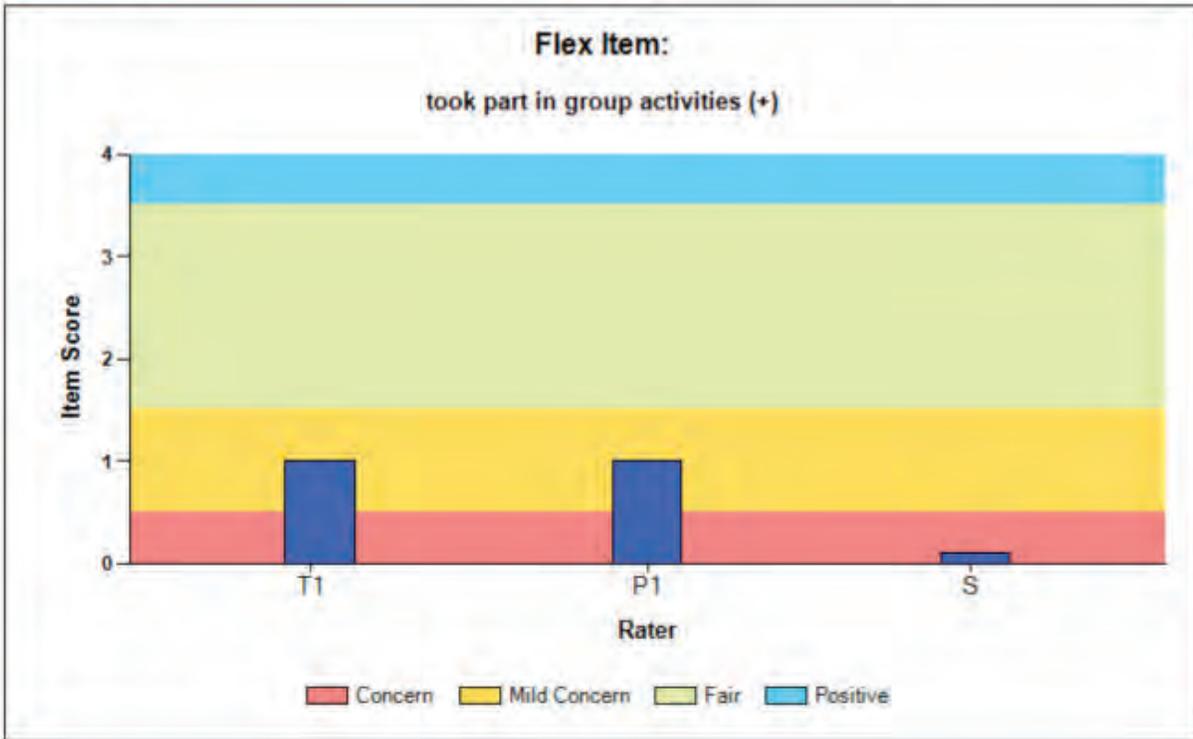
Rater	Teacher 1	Parent 1	Self
Item Score	1 Mild Concern	1 Mild Concern	0 Concern

Comments:

Teacher 1: Encouraged John to participate in the group treasure hunt game on Wed. Still see reluctance.

Parent 1: N/A.

Self: N/A



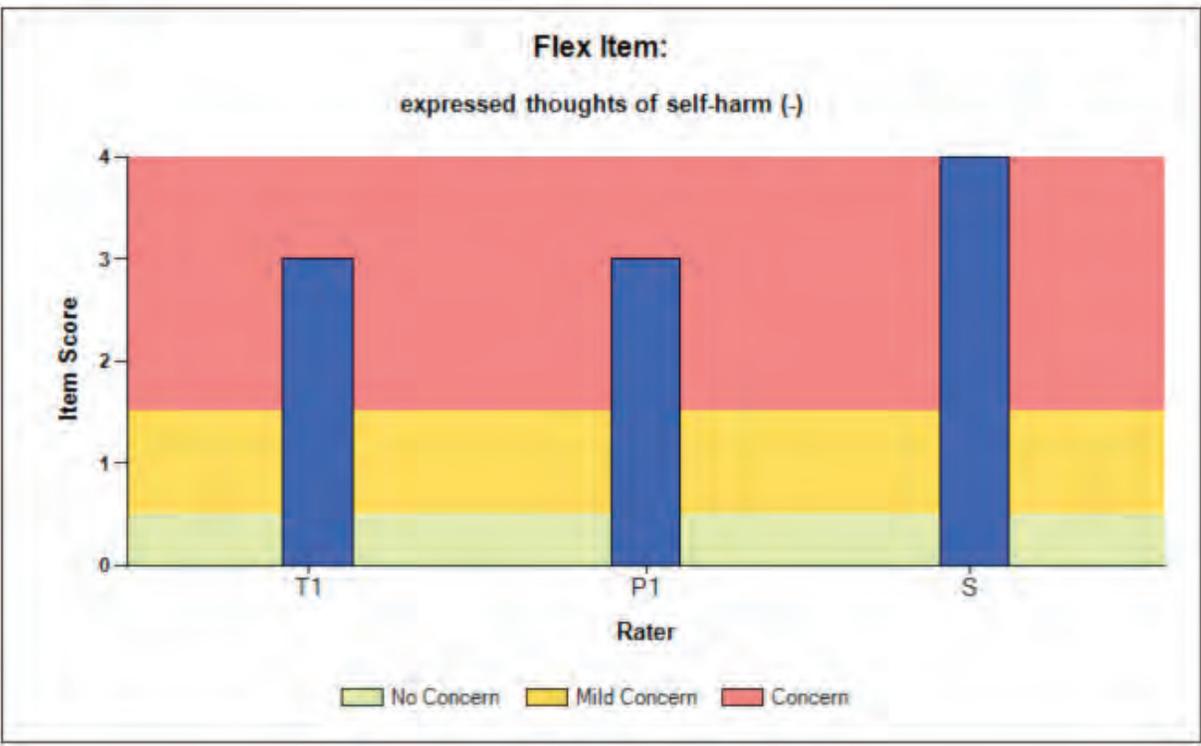
Note. All names are fictional.

**Standard Form Anchor Item:**  
**24. expressed thoughts of hurting himself/herself.**

Flex Item:  
**expressed thoughts of self-harm (-)**

Rater	Teacher 1	Parent 1	Self
Item Score	3 Concern	3 Concern	4 Concern

**Comments:**  
 Teacher 1: John related to me that he sometimes thinks about harming himself.  
 Parent 1: N/A  
 Self: N/A

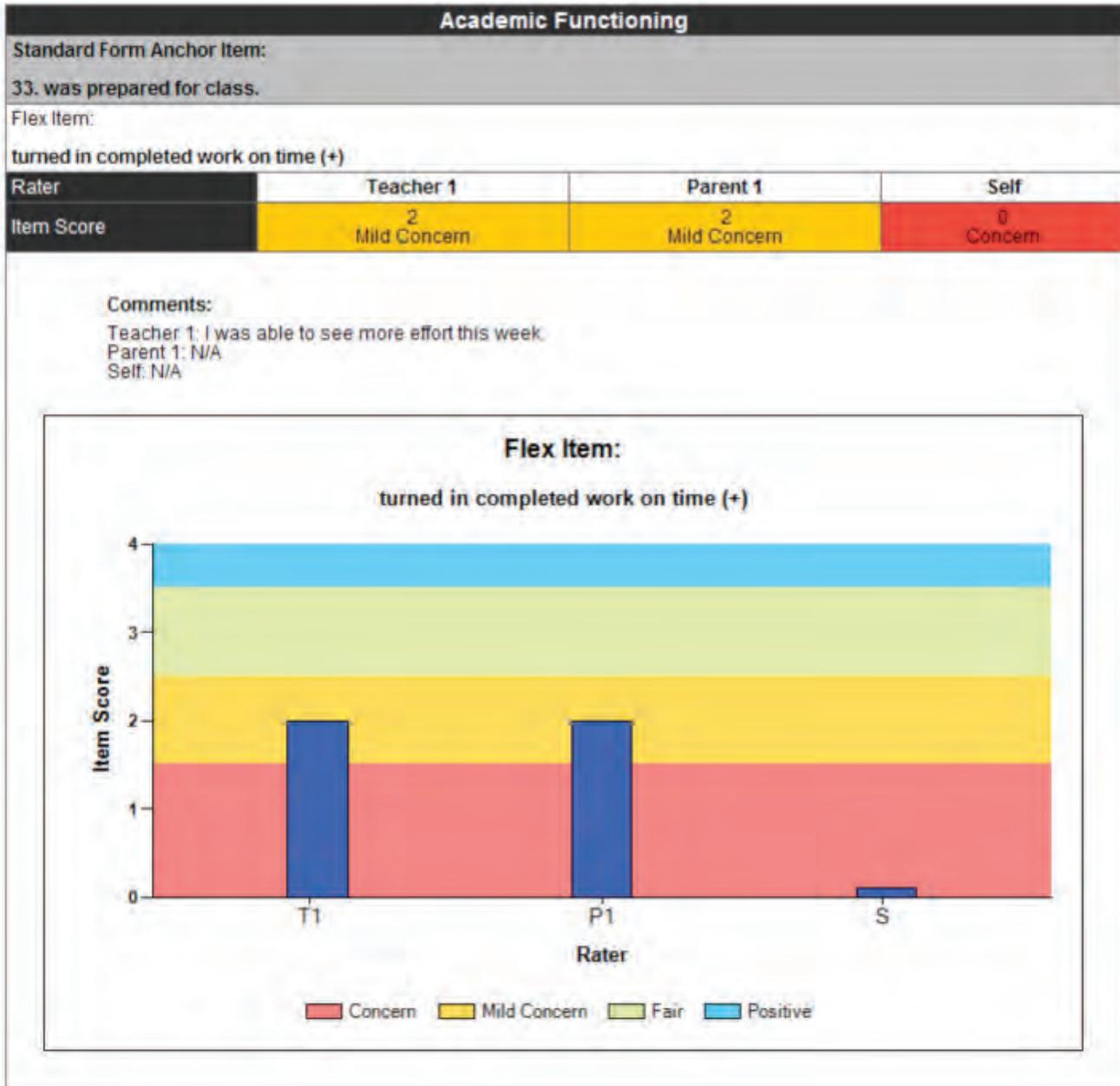


*Note.* All names are fictional.

**B) Adaptive Scales**

The following section provides the comparison for selected raters on Flex items for the specified date of entry on Adaptive Scale(s) in a graph.

Note: P = Parent; T = Teacher; S = Self



Note. All names are fictional.



Note. All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



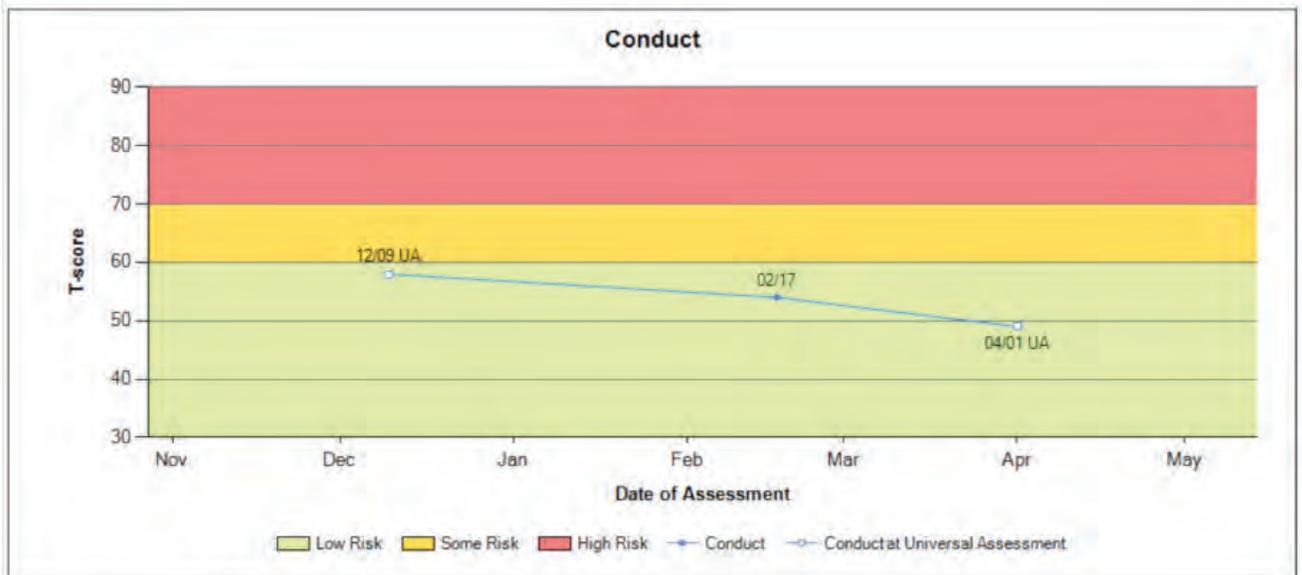
# Appendix E.14

## Standard Individual Progress Report

### Standard Individual Progress Report BIMAS™-Teacher Standard

Springfield School District  
Lincoln Middle School

Student: John L. Oakes  
 Gender: Male  
 Age: 12  
 DOB: 06/13/1990  
 Grade: 6  
 Class: English 1200  
 Rater: T. Mr. Steve Murr  
 Dates of Assessment: 12/9/2010 through 4/1/2011



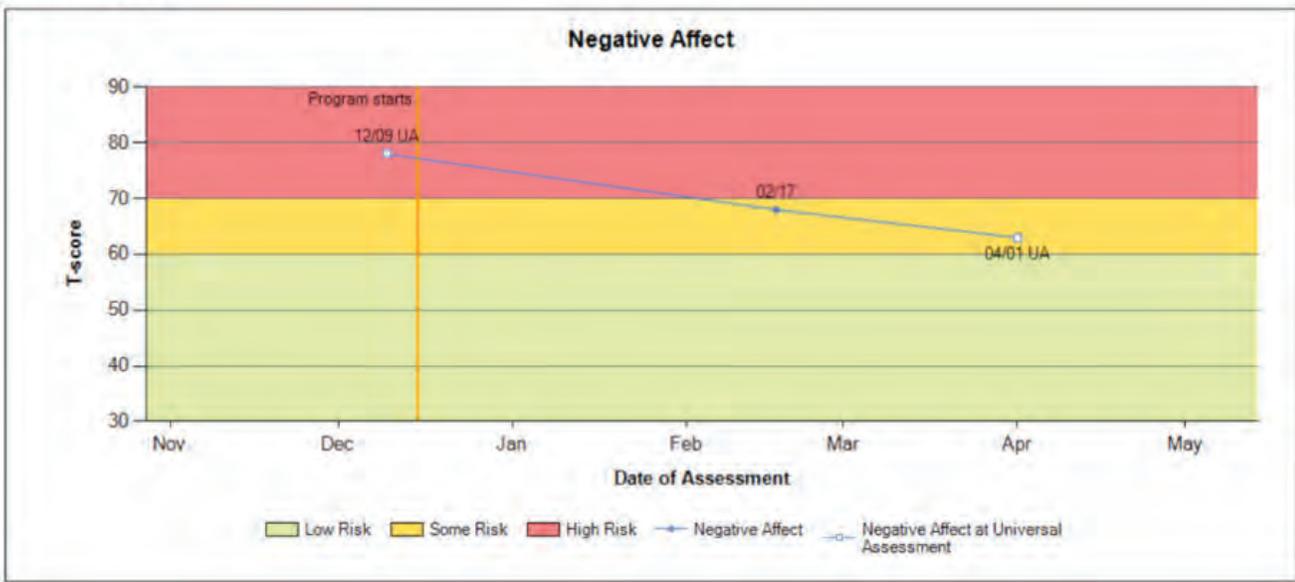
Date of Assessment	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Conduct T-score	58	54	49

UA = Date of Universal Assessment

#### Intervention Notes

[Edit Intervention Dates/Notes](#)

*Note.* All names are fictional.



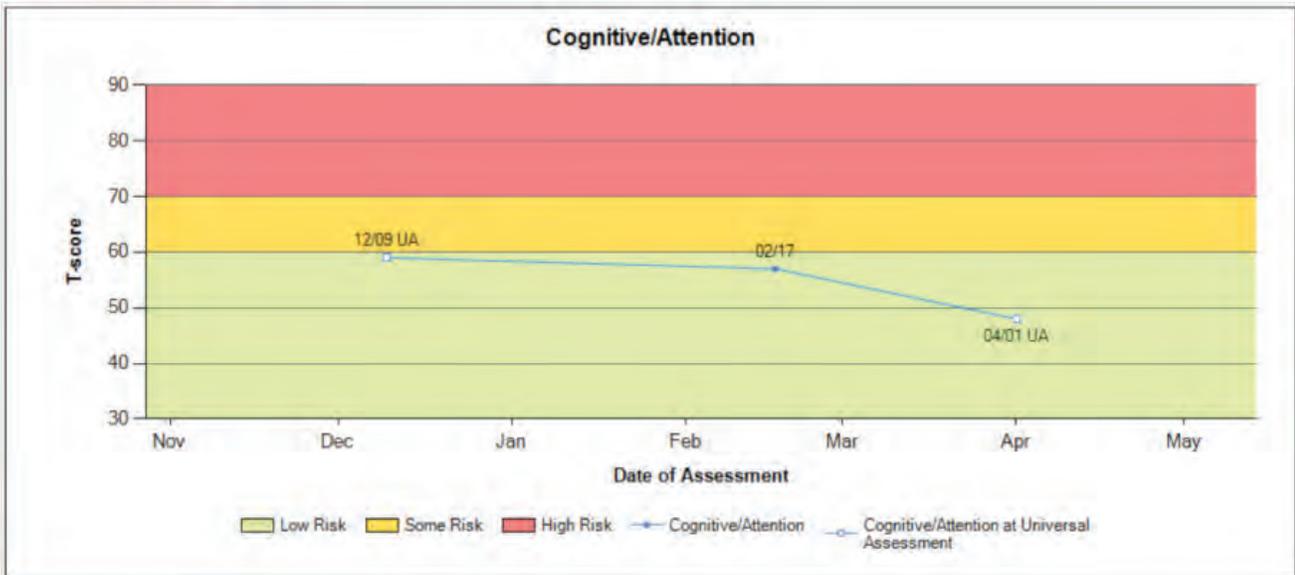
Date of Assessment	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Negative Affect T-score	78	68	63

UA = Date of Universal Assessment

**Intervention Notes**

12/15/2010 Program starts  
Weekly support group begins

[Edit Intervention Dates/Notes](#)



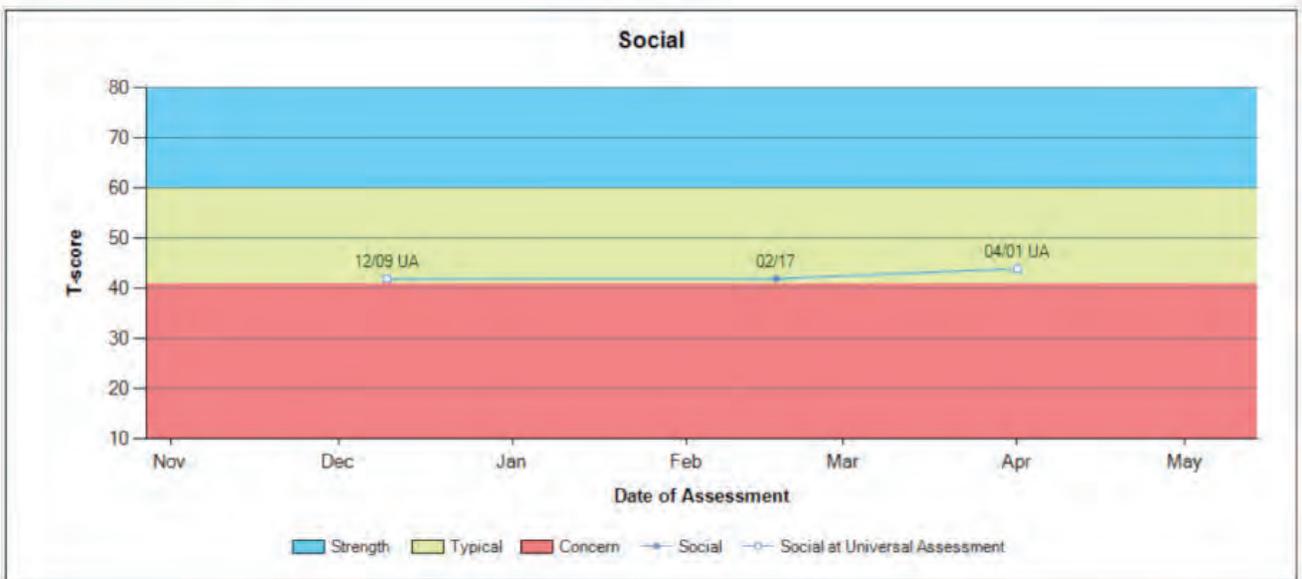
Date of Assessment	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Cognitive/Attention T-score	59	57	48

UA = Date of Universal Assessment

**Intervention Notes**

[Edit Intervention Dates/Notes](#)

Note. All names are fictional.

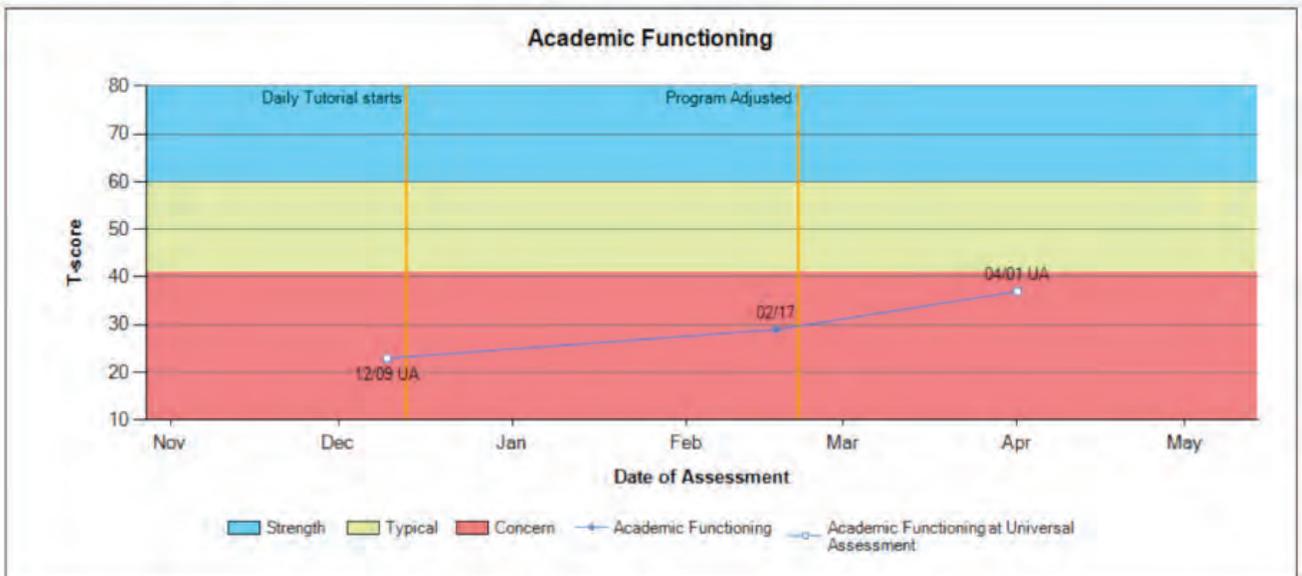


Date of Assessment	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Social T-score	42	42	44

UA = Date of Universal Assessment

**Intervention Notes**

[Edit Intervention Dates/Notes](#)



Date of Assessment	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Academic Functioning T-score	23	29	37

UA = Date of Universal Assessment

**Intervention Notes**

- 12/13/2010 **Daily Tutorial starts**  
Remedial class to address difficulties in Math and Reading.
- 2/21/2011 **Program Adjusted**  
Weekly intensive training on Math to address slower progress in scores

[Edit Intervention Dates/Notes](#)

**Related Reports:**

To determine if any of the changes between assessments in this report were statistically significant:

[Standard Individual Progress Report: Significant Change Over Time](#)

To monitor progress at the item level:

[Standard Individual Progress Report: Item Analysis](#)

*Note.* All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



# Appendix E.15a

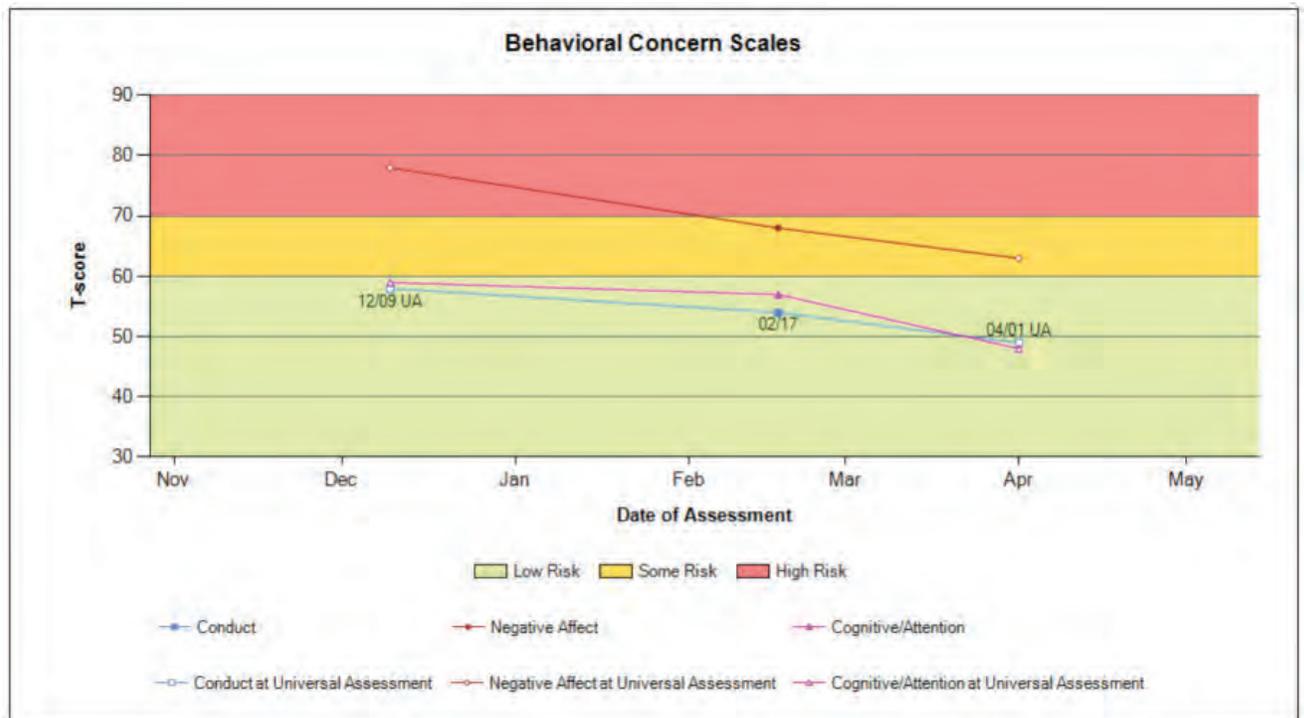
## Standard Individual Progress Report: Significant Change Over Time

(Statistically Significant Change [Based on T-Scores] Version)

### Standard Individual Progress Report: Significant Change over Time BIMAS™-Teacher Standard

Springfield School District  
Lincoln Middle School

Student: John L. Oakes  
Gender: Male  
Age: 12  
DOB: 06/13/1990  
Grade: 6  
Class: English 1200  
Rater: T. Mr. Steve Murr  
Dates of Assessment: 12/9/2010 through 4/1/2011



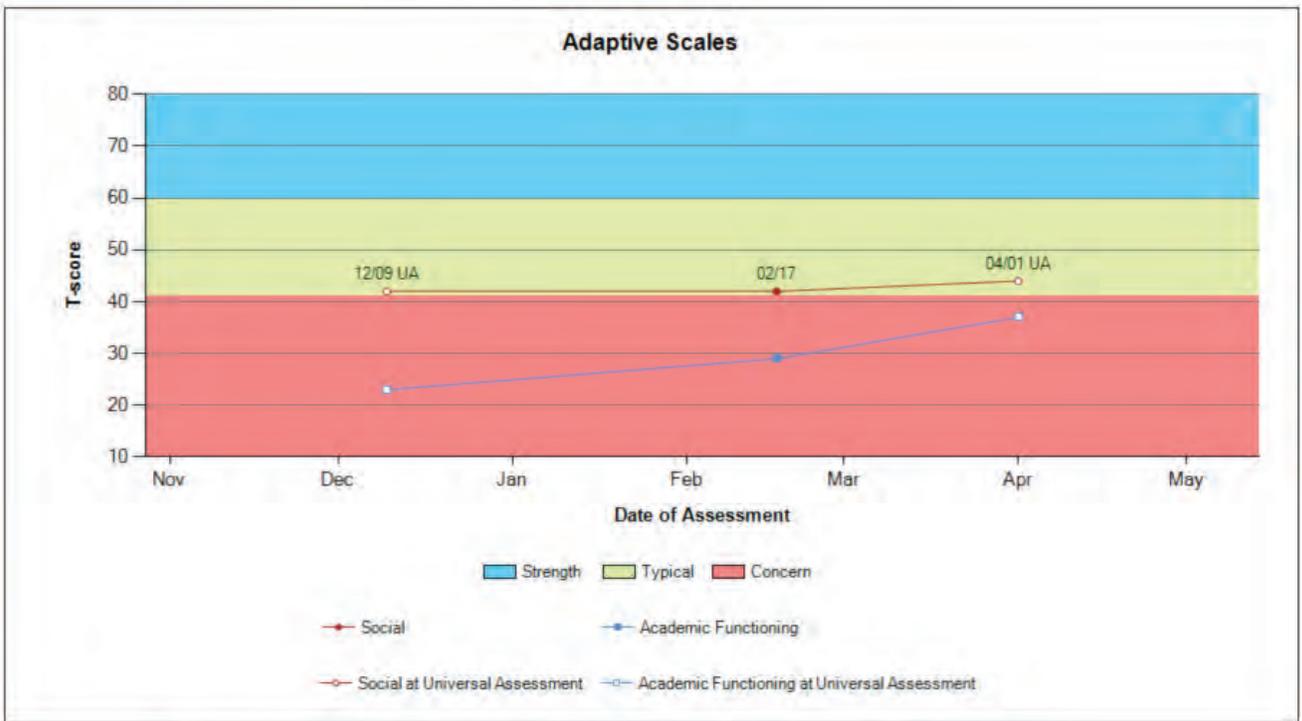
Note. All names are fictional.

Behavior Intervention Monitoring Assessment System (BIMAS™)

Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
<b>Conduct</b>	Raw Score	4	2	0
	T-score	58	54	49
	90% CI	54-62	50-58	45-53
	Percentile	92	88	45
	Level of Risk	Low	Low	Low
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
<b>Negative Affect</b>	Raw Score	14	10	8
	T-score	78	68	63
	90% CI	72-84	62-74	57-69
	Percentile	99	96	89
	Level of Risk	High	Some	Some
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
<b>Cognitive/Attention</b>	Raw Score	9	8	3
	T-score	59	57	48
	90% CI	54-64	52-62	43-53
	Percentile	82	75	42
	Level of Risk	Low	Low	Low

<sup>UA</sup> = Date of Universal Assessment

Note. All names are fictional.



Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Social	Raw Score	17	17	18
	T-score	42	42	44
	90% CI	36-48	36-48	38-50
	Percentile	21	21	27
	Level of Functioning	Typical	Typical	Typical
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Academic Functioning	Raw Score	5	10	13
	T-score	23	29	37
	90% CI	17-29	23-35	31-43
	Percentile	1	2	10
	Level of Functioning	Concern	Concern	Concern

<sup>UA</sup> = Date of Universal Assessment

### Assess Change Over Time

Display Statistically Significant Change (Based on T-scores)

Display Effect Size (Based on Raw Scores)

Note. All names are fictional.

### Behavioral Concern Scales-Change Over Time: Statistically Significant Change (based on T-scores)

Please refer to the *BIMAS Technical Manual* for more information on the interpretation of statistically significant change.

Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Conduct	T-score	58	54	49	Improved
	Level of Risk	Low	Low	Low	
	Significant Change ( $p < .10$ )	12/09 <sup>UA</sup> -02/17 (No Change) 02/17-04/01 <sup>UA</sup> (No Change)			
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Negative Affect	T-score	78	68	63	Improved
	Level of Risk	High	Some	Some	
	Significant Change ( $p < .10$ )	12/09 <sup>UA</sup> -02/17 (Improved) 02/17-04/01 <sup>UA</sup> (No Change)			
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Cognitive/Attention	T-score	59	57	48	Improved
	Level of Risk	Low	Low	Low	
	Significant Change ( $p < .10$ )	12/09 <sup>UA</sup> -02/17 (No Change) 02/17-04/01 <sup>UA</sup> (Improved)			

<sup>UA</sup> = Date of Universal Assessment

### Behavioral Concern Scales: Interpretation Guidelines for Statistically Significant Change

**Improved:** Behavioral concerns have significantly decreased (i.e., scores have improved) between administrations.

**No Change:** The amount of change in T-scores between administrations did not reach statistical significance.

**Worse:** Behavioral concerns have become significantly more pronounced (i.e., scores have increased) between administrations.

### Adaptive Scales-Change Over Time: Statistically Significant Change (based on T-scores)

Please refer to the *BIMAS Technical Manual* for more information on the interpretation of statistically significant change.

Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Social	T-score	42	42	44	No Change
	Level of Functioning	Typical	Typical	Typical	
	Significant Change ( $p < .10$ )	12/09 <sup>UA</sup> -02/17 (No Change) 02/17-04/01 <sup>UA</sup> (No Change)			
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Academic Functioning	T-score	23	29	37	Improved
	Level of Functioning	Concern	Concern	Concern	
	Significant Change ( $p < .10$ )	12/09 <sup>UA</sup> -02/17 (No Change) 02/17-04/01 <sup>UA</sup> (Improved)			

<sup>UA</sup> = Date of Universal Assessment

### Adaptive Scales: Interpretation Guidelines for Statistically Significant Change

**Improved:** Concerns regarding adaptive behaviors have significantly decreased (i.e., scores have improved) between administrations.

**No Change:** The amount of change in T-scores between administrations did not reach statistical significance.

**Worse:** Concerns regarding adaptive behaviors have become significantly more pronounced (i.e., scores have decreased) between administrations.

#### Related Reports:

To review Intervention Dates/Notes relating to specific scale(s) for the selected assessment dates:

[Standard Individual Progress Report](#)

To monitor progress at the item level:

[Standard Individual Progress: Item Analysis](#)

Note. All names are fictional.

# Appendix E.15b

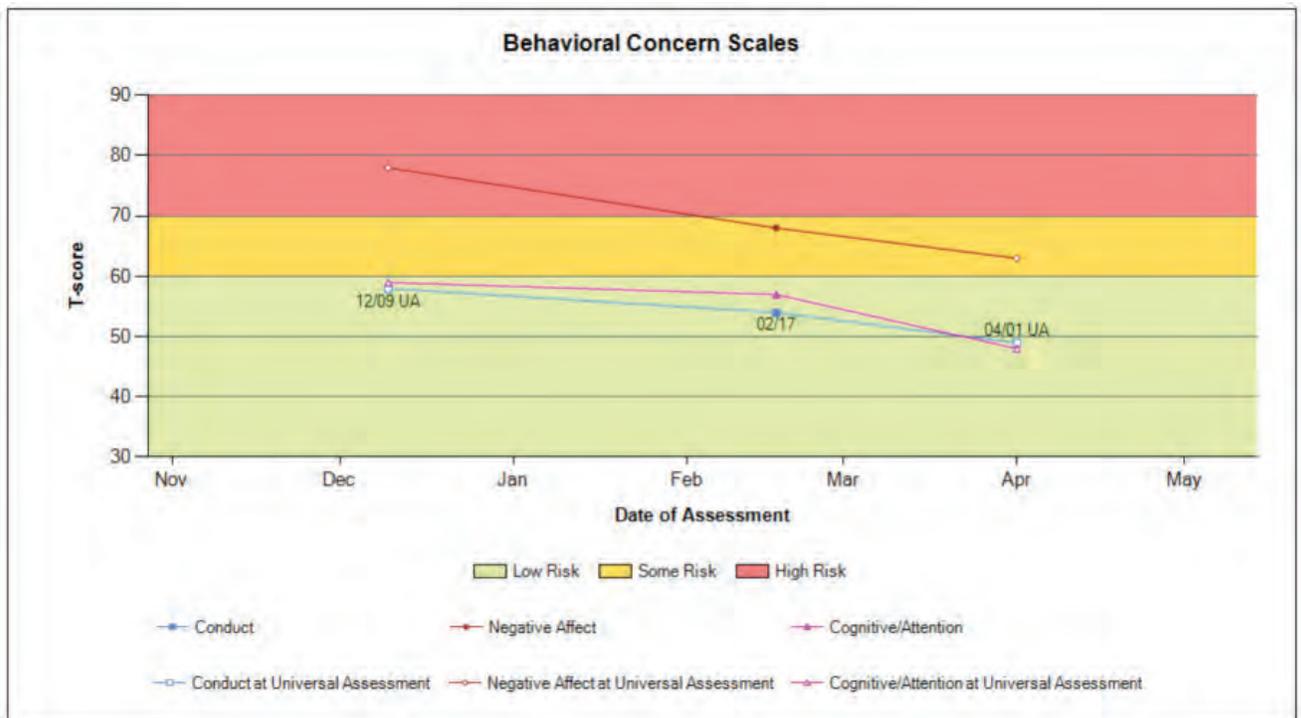
## Standard Individual Progress Report: Significant Change Over Time

(Effect Size [Based on raw scores] Version)

### Standard Individual Progress Report: Significant Change over Time BIMAS™-Teacher Standard

Springfield School District  
Lincoln Middle School

Student: John L. Oakes  
 Gender: Male  
 Age: 12  
 DOB: 06/13/1990  
 Grade: 6  
 Class: English 1200  
 Rater: T. Mr. Steve Murr  
 Dates of Assessment: 12/9/2010 through 4/1/2011



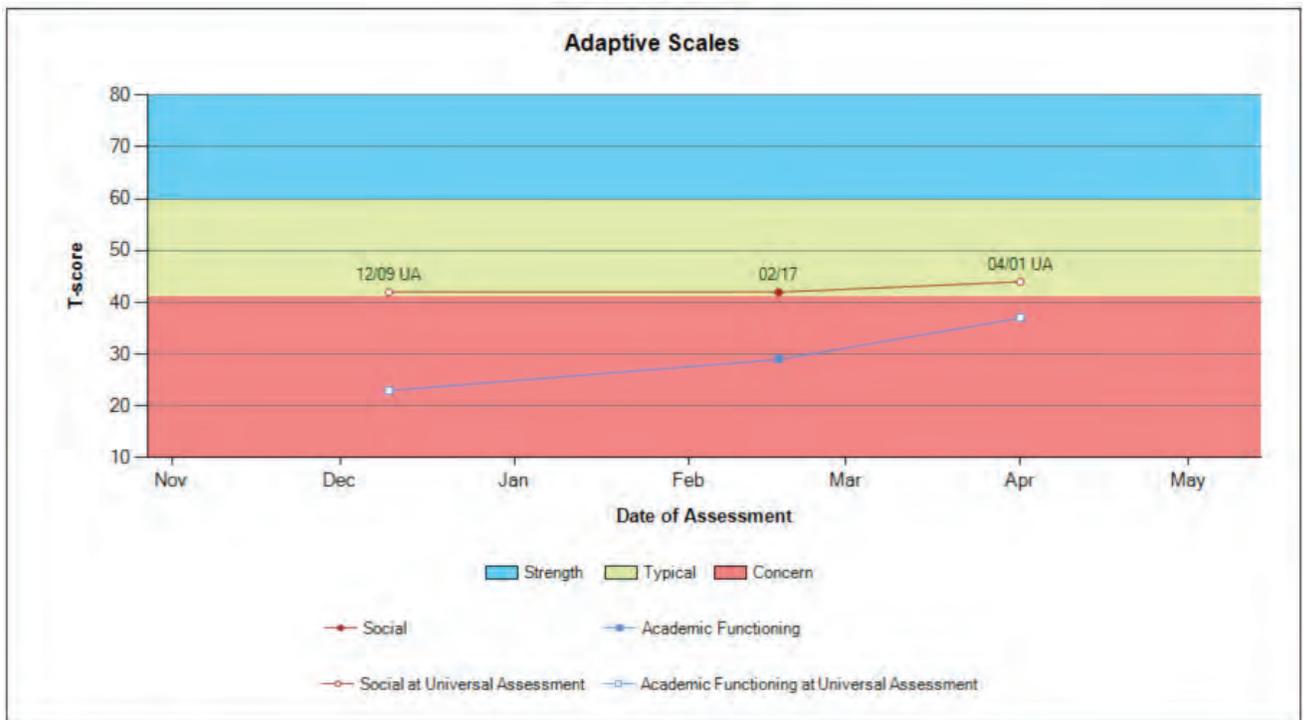
Note. All names are fictional.

Behavior Intervention Monitoring Assessment System (BIMAS™)

Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
<b>Conduct</b>	Raw Score	4	2	0
	T-score	58	54	49
	90% CI	54-62	50-58	45-53
	Percentile	92	88	45
	Level of Risk	Low	Low	Low
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
<b>Negative Affect</b>	Raw Score	14	10	8
	T-score	78	68	63
	90% CI	72-84	62-74	57-69
	Percentile	99	96	89
	Level of Risk	High	Some	Some
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
<b>Cognitive/Attention</b>	Raw Score	9	8	3
	T-score	59	57	48
	90% CI	54-64	52-62	43-53
	Percentile	82	75	42
	Level of Risk	Low	Low	Low

<sup>UA</sup> = Date of Universal Assessment

Note. All names are fictional.



Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Social	Raw Score	17	17	18
	T-score	42	42	44
	90% CI	36-48	36-48	38-50
	Percentile	21	21	27
	Level of Functioning	Typical	Typical	Typical
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>
Academic Functioning	Raw Score	5	10	13
	T-score	23	29	37
	90% CI	17-29	23-35	31-43
	Percentile	1	2	10
	Level of Functioning	Concern	Concern	Concern

<sup>UA</sup> = Date of Universal Assessment

Note. All names are fictional.

### Assess Change Over Time

Display Statistically Significant Change (Based on T-scores)

Display Effect Size (Based on Raw Scores)

### Behavioral Concern Scales-Change Over Time: Effect Size (based on Raw Scores)

Please refer to the *BIMAS Technical Manual* for more information on the interpretation of effect size.

Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Conduct	Raw Score	4	2	0	-0.61 (Improved)
	Level of Risk	Low	Low	Low	
	Effect Size	12/09 <sup>UA</sup> -02/17: -0.31(No Change) 02/17-04/01 <sup>UA</sup> : -0.50(No Change)			
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Negative Affect	Raw Score	14	10	8	-0.66 (Improved)
	Level of Risk	High	Some	Some	
	Effect Size	12/09 <sup>UA</sup> -02/17: -0.44(No Change) 02/17-04/01 <sup>UA</sup> : -0.29(No Change)			
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Cognitive/Attention	Raw Score	9	8	3	-0.77 (Improved)
	Level of Risk	Low	Low	Low	
	Effect Size	12/09 <sup>UA</sup> -02/17: -0.13(No Change) 02/17-04/01 <sup>UA</sup> : -0.79(Improved)			

<sup>UA</sup> = Date of Universal Assessment

### Behavioral Concern Scales: Interpretation Guidelines for Effect Size

**Much Improved** (ES ≤ -1.50): Behavioral concerns have much decreased (i.e., scores have improved to a large extent) between administrations.

**Improved** (ES = -0.51 to -1.49): Behavioral concerns have decreased (i.e., scores have improved to some extent) between administrations.

**No Change** (ES = -0.50 to + 0.50): The amount of change in raw scores between administrations did not produce a statistically significant effect size.

**Worse** (ES = 0.51 to 1.49): Behavioral concerns have become a little more pronounced (i.e., scores have increased to some extent) between administrations.

**Much Worse** (ES ≥ 1.50): Behavioral concerns have become much more pronounced (i.e., scores have increased to a large extent) between administrations.

### Adaptive Scales-Change Over Time: Effect Size (based on Raw Scores)

Please refer to the *BIMAS Technical Manual* for more information on the interpretation of effect size.

Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Social	Raw Score	17	17	18	0.22 (No Change)
	Level of Functioning	Typical	Typical	Typical	
	Effect Size	12/09 <sup>UA</sup> -02/17: 0.00(No Change) 02/17-04/01 <sup>UA</sup> : 0.22(No Change)			
Date of Assessment		12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	Overall (12/09-04/01)
Academic Functioning	Raw Score	5	10	13	2.26 (Much Improved)
	Level of Functioning	Concern	Concern	Concern	
	Effect Size	12/09 <sup>UA</sup> -02/17: 1.41(Improved) 02/17-04/01 <sup>UA</sup> : 0.73(Improved)			

<sup>UA</sup> = Date of Universal Assessment

### Adaptive Scales: Interpretation Guidelines for Effect Size

**Much Improved** (ES ≥ 1.50): Concerns regarding adaptive behaviors have much decreased (i.e., scores have improved to a large extent) between administrations.

**Improved** (ES = 0.51 to 1.49): Concerns regarding adaptive behaviors have decreased (i.e., scores have improved to some extent) between administrations.

**No Change** (ES = -0.50 to + 0.50): The amount of change in raw scores between administrations did not produce a statistically significant effect size.

**Worse** (ES = -0.51 to -1.49): Concerns regarding adaptive behaviors have become a little more pronounced (i.e., scores have decreased to some extent) between administrations.

**Much Worse** (ES ≤ -1.50): Concerns regarding adaptive behaviors have become much more pronounced (i.e., scores have decreased to a large extent) between administrations.

### Related Reports:

To review Intervention Dates/Notes relating to specific scale(s) for the selected assessment dates:

[Standard Individual Progress Report](#)

To monitor progress at the item level:

[Standard Individual Progress - Item Analysis](#)

Note. All names are fictional.



# Appendix E.16

## Standard Individual Progress Report: Item Analysis

### Standard Individual Progress Report: Item Analysis BIMAS™-Teacher Standard

Springfield School District  
Lincoln Middle School

Student: John L. Oakes  
Gender: Male  
Age: 12  
DOB: 06/13/1990  
Grade: 6  
Class: English 1200  
Rater: T. Mr. Steve Murr  
Dates of Assessment: 12/9/2010 through 4/1/2011

### Item Scores: Scale-Level Comparisons Across Administrations

#### A) Behavioral Concern Scales

The following table provides the item-level progress on the specified dates of assessments for each of the Behavioral Concern Scales. Items identified as Concern/Mild Concern should be further explored and considered when designing and monitoring individual intervention programs.

The following response key applies to the table in this section.

Item Score:

- 0 = **Never** (Observed 0 times or not observed)
- 1 = **Rarely** (Observed 1 - 2 times or to a minimum extent)
- 2 = **Sometimes** (Observed 3 - 4 times or to a moderate extent)
- 3 = **Often** (Observed 5 - 6 times or to a significant extent)
- 4 = **Very Often** (Observed 7 or more times or to an extreme extent)
- ? = **Omitted Item**

Item Descriptor Legend:

-  = Concern
-  = Mild Concern
-  = No Concern

*Note.* All names are fictional.

Behavioral Concern Scales				<a href="#">Graph All 34 items</a>	
Higher scores indicate MORE concerns					
<b>Conduct</b>				<input type="checkbox"/> Graph all Conduct items	
Item No.	Date of Assessment				
	Item Score				
	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>		
2	angry	1	1	0	<input checked="" type="checkbox"/> Graph this item
9	risky behavior	0	0	0	<input type="checkbox"/> Graph this item
13	fought (verbally/physically/both)	1	0	0	<input type="checkbox"/> Graph this item
17	lied/cheated	0	0	0	<input type="checkbox"/> Graph this item
21	lost temper	2	1	0	<input checked="" type="checkbox"/> Graph this item
25	aggressive	0	0	0	<input type="checkbox"/> Graph this item
29	alcohol and/or drug use	0	0	0	<input type="checkbox"/> Graph this item
31	sent to disciplinary authority	0	0	0	<input type="checkbox"/> Graph this item
32	tobacco use	0	0	0	<input type="checkbox"/> Graph this item
<b>Negative Affect</b>				<input type="checkbox"/> Graph all Negative Affect items	
Item No.	Date of Assessment				
	Item Score				
	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>		
5	sleepy/tired	2	1	2	<input checked="" type="checkbox"/> Graph this item
8	depressed	2	2	2	<input type="checkbox"/> Graph this item
12	sad/withdrawn	4	3	2	<input checked="" type="checkbox"/> Graph this item
16	embarrassed/ashamed	0	0	2	<input type="checkbox"/> Graph this item
20	anxious	1	1	0	<input type="checkbox"/> Graph this item
24	thoughts of hurting self	3	2	0	<input checked="" type="checkbox"/> Graph this item
27	emotional or upset	2	1	0	<input type="checkbox"/> Graph this item

Cognitive/Attention				<input type="checkbox"/> Graph all Cognitive/Attention items
Item No.	Date of Assessment			
	Item Score			
	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	
3	trouble paying attention			<input type="checkbox"/> Graph this item
	1	1	0	
6	impulsive			<input type="checkbox"/> Graph this item
	0	0	0	
10	problems staying on task			<input checked="" type="checkbox"/> Graph this item
	3	2	3	
14	acted without thinking			<input type="checkbox"/> Graph this item
	0	0	0	
18	trouble remembering			<input type="checkbox"/> Graph this item
	1	1	0	
22	trouble with organizing and planning			<input type="checkbox"/> Graph this item
	2	2	0	
28	fidgeted			<input type="checkbox"/> Graph this item
	2	2	0	

## B) Adaptive Scales

The following table provides the item-level progress on the specified dates of assessments for each of the Adaptive Scales. Items identified as Concern/Mild Concern should be further explored and considered when designing and monitoring individual intervention programs. In addition, an examination of items identified as Positive may also aid in intervention program design.

The following response key applies to the table in this section.

(R) = This item has been reverse scored.

Item Score:

0 = **Never** (Observed 0 times or not observed)

1 = **Rarely** (Observed 1 - 2 times or to a minimum extent)

2 = **Sometimes** (Observed 3 - 4 times or to a moderate extent)

3 = **Often** (Observed 5 - 6 times or to a significant extent)

4 = **Very Often** (Observed 7 or more times or to an extreme extent)

? = **Omitted Item**

Item Descriptor Legend:

= Concern

= Mild Concern

= Fair

= Positive

Note. All names are fictional.

<b>Adaptive Scales</b>				
Higher scores indicate FEWER concerns				
<b>Social</b>			<input type="checkbox"/> Graph all Social items	
Item No.	Date of Assessment			
	Item Score			
	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	
1	shared thoughts			<input type="checkbox"/> Graph this item
	2	2	3	
7	spoke clearly with others			<input type="checkbox"/> Graph this item
	4	4	3	
11	maintained friendships			<input type="checkbox"/> Graph this item
	3	3	3	
15	comfortable relating with others			<input type="checkbox"/> Graph this item
	2	2	3	
19	friendly			<input type="checkbox"/> Graph this item
	3	3	3	
23	worked out problems with others			<input type="checkbox"/> Graph this item
	3	3	3	
<b>Academic Functioning</b>			<input type="checkbox"/> Graph all Academic Functioning items	
Item No.	Date of Assessment			
	Item Score			
	12/09 <sup>UA</sup>	02/17	04/01 <sup>UA</sup>	
4	followed directions			<input checked="" type="checkbox"/> Graph this item
	2	2	4	
26	failing grades (R)			<input type="checkbox"/> Graph this item
	1	2	3	
30	worked up to academic potential			<input type="checkbox"/> Graph this item
	1	2	1	
33	prepared for class			<input type="checkbox"/> Graph this item
	0	2	2	
34	absent from school (R)			<input checked="" type="checkbox"/> Graph this item
	1	2	3	

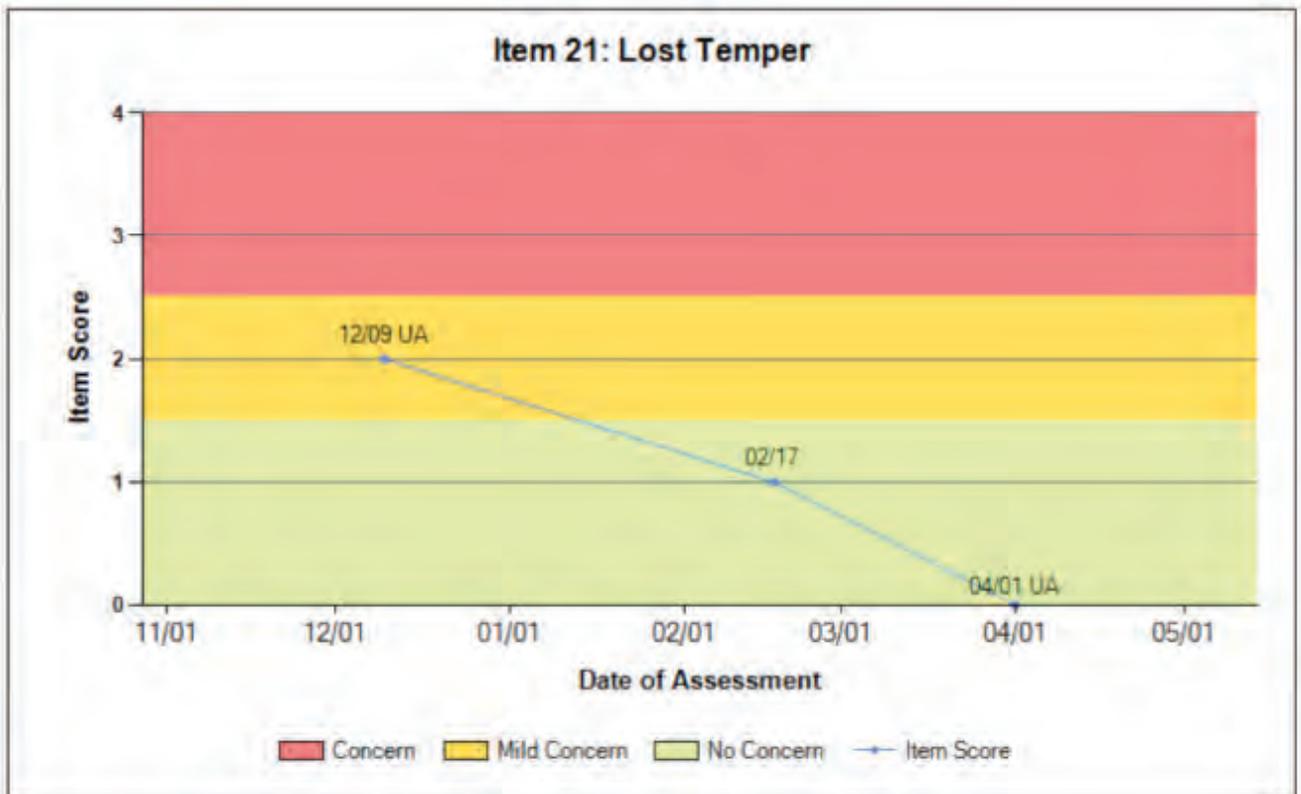
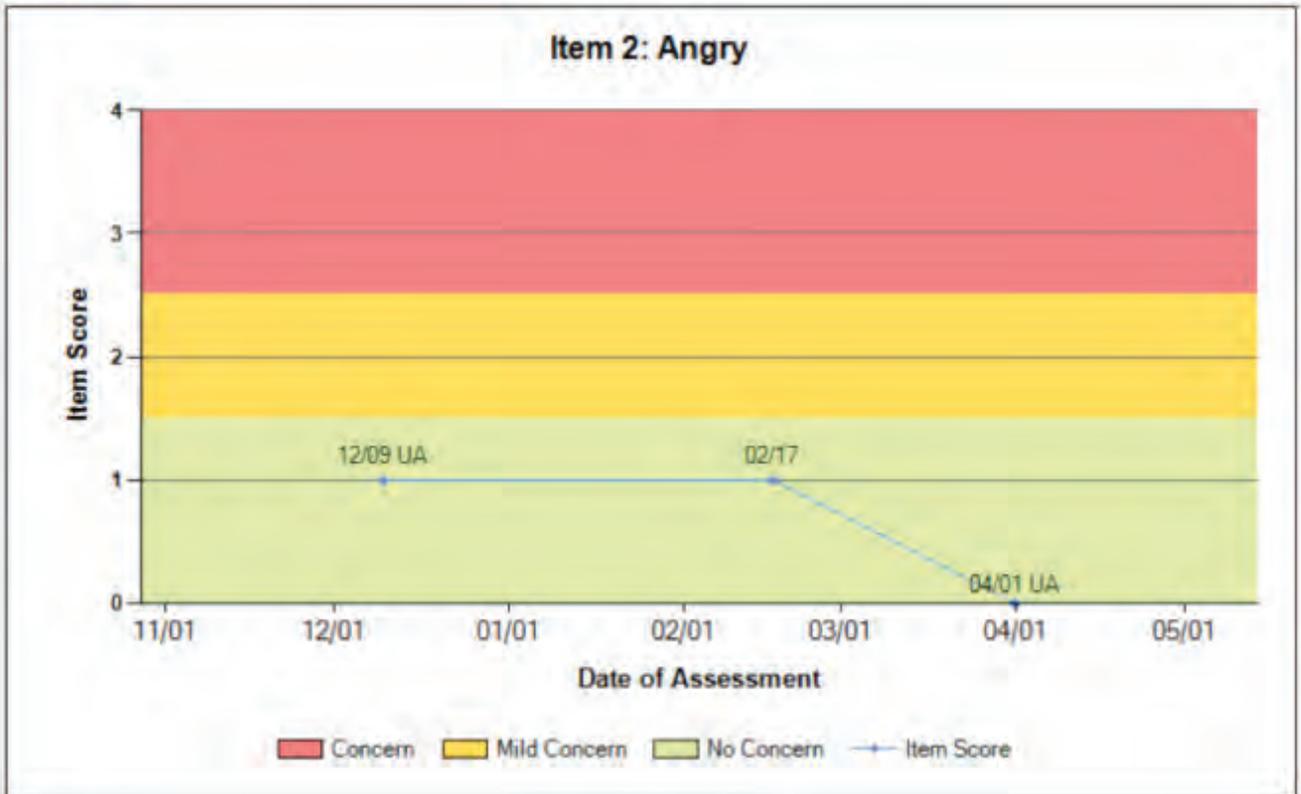
Note: (R) = Reverse scored item.  
 Note: All names are fictional.

**Item Scores: Graphs**

**A) Behavioral Concern Scales**

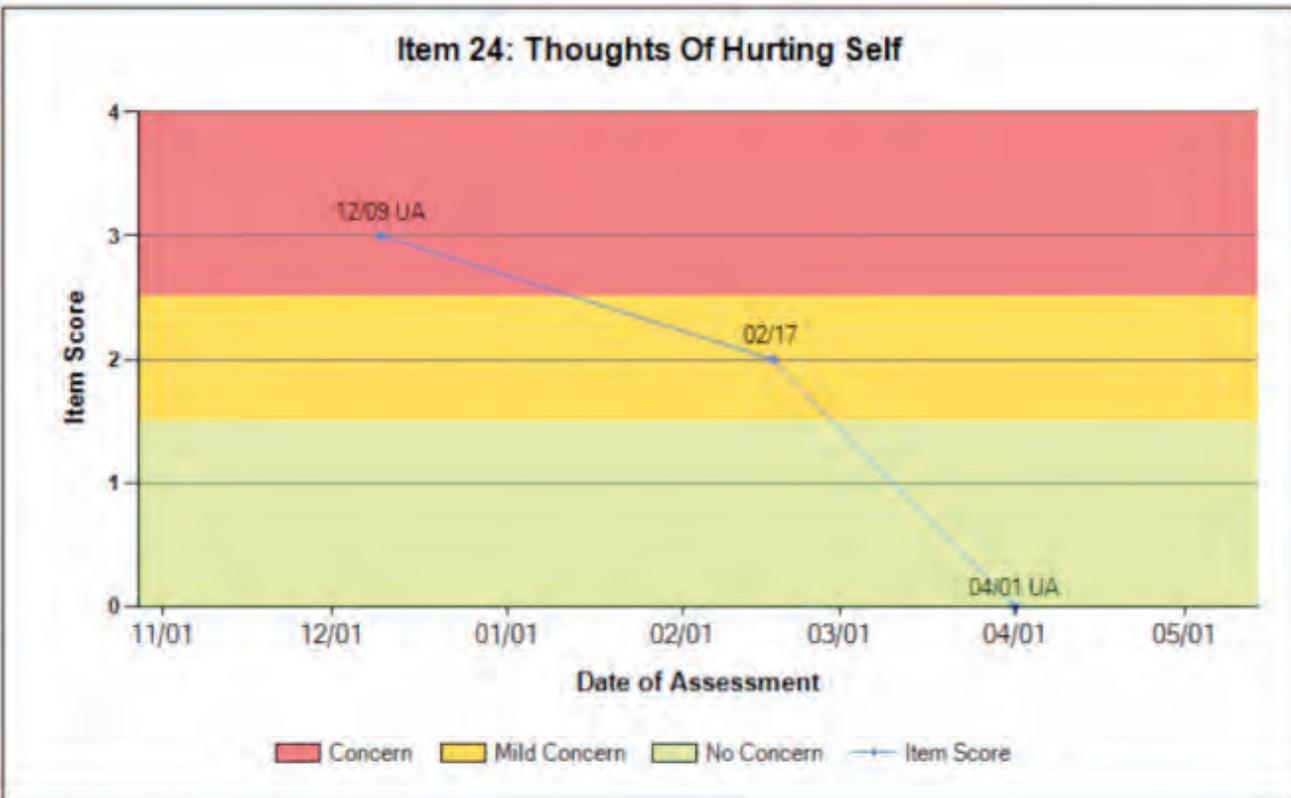
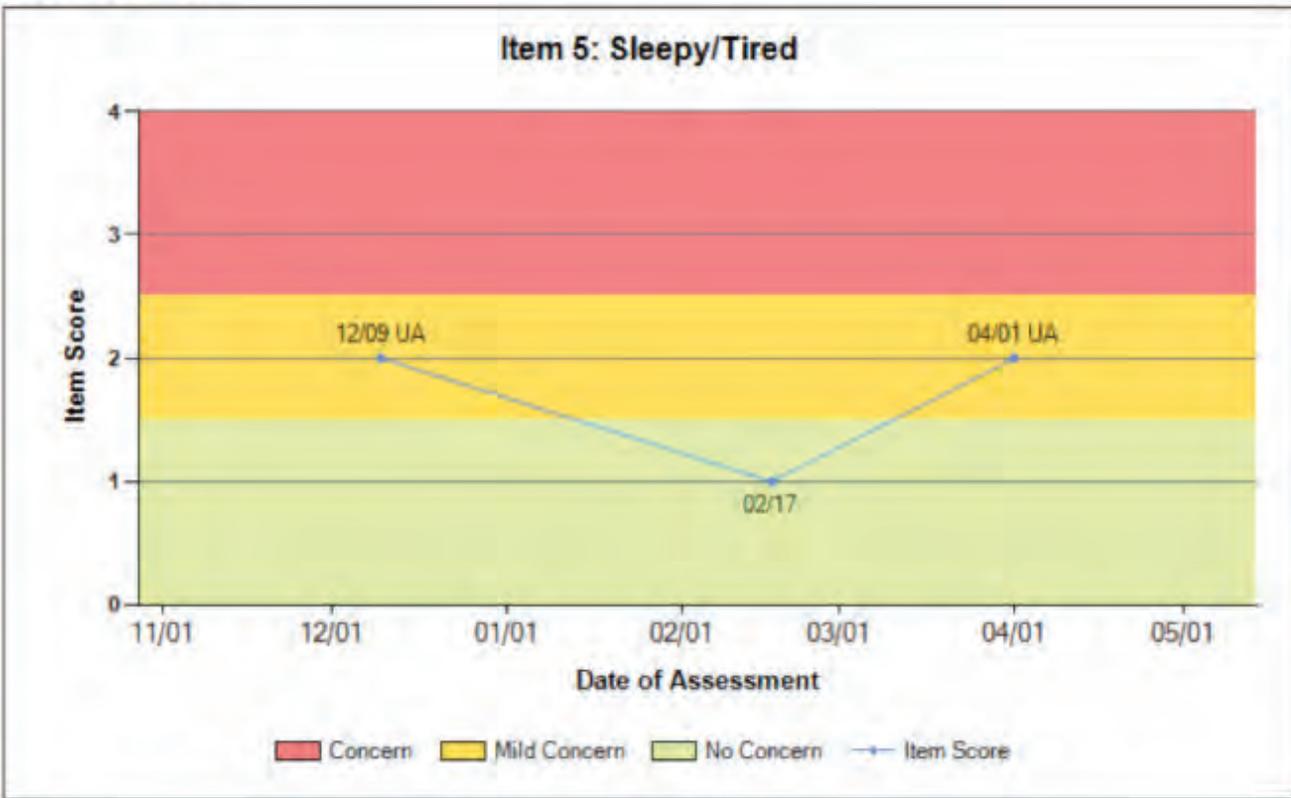
Higher scores indicate MORE concerns

**Conduct**



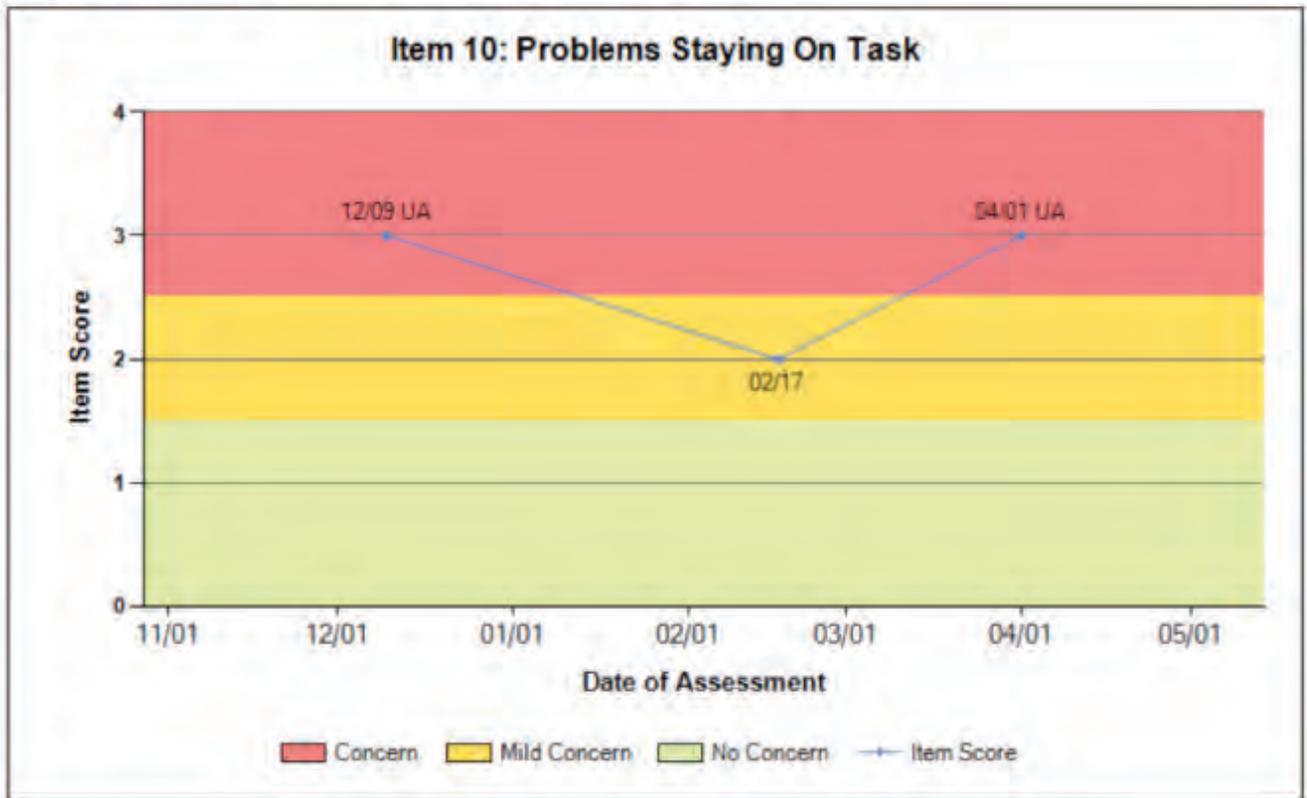
Note. All names are fictional.

Negative Affect



Note. All names are fictional.

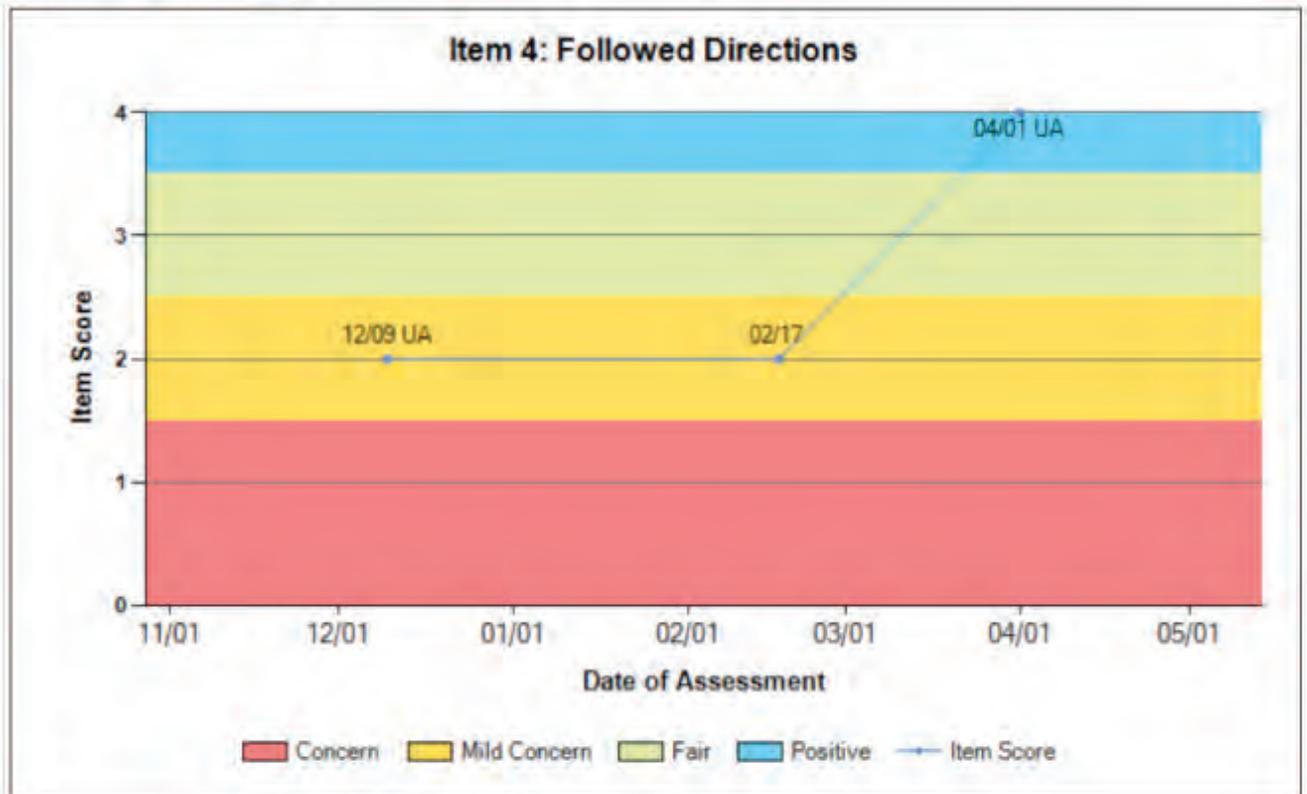
**Cognitive/Attention**



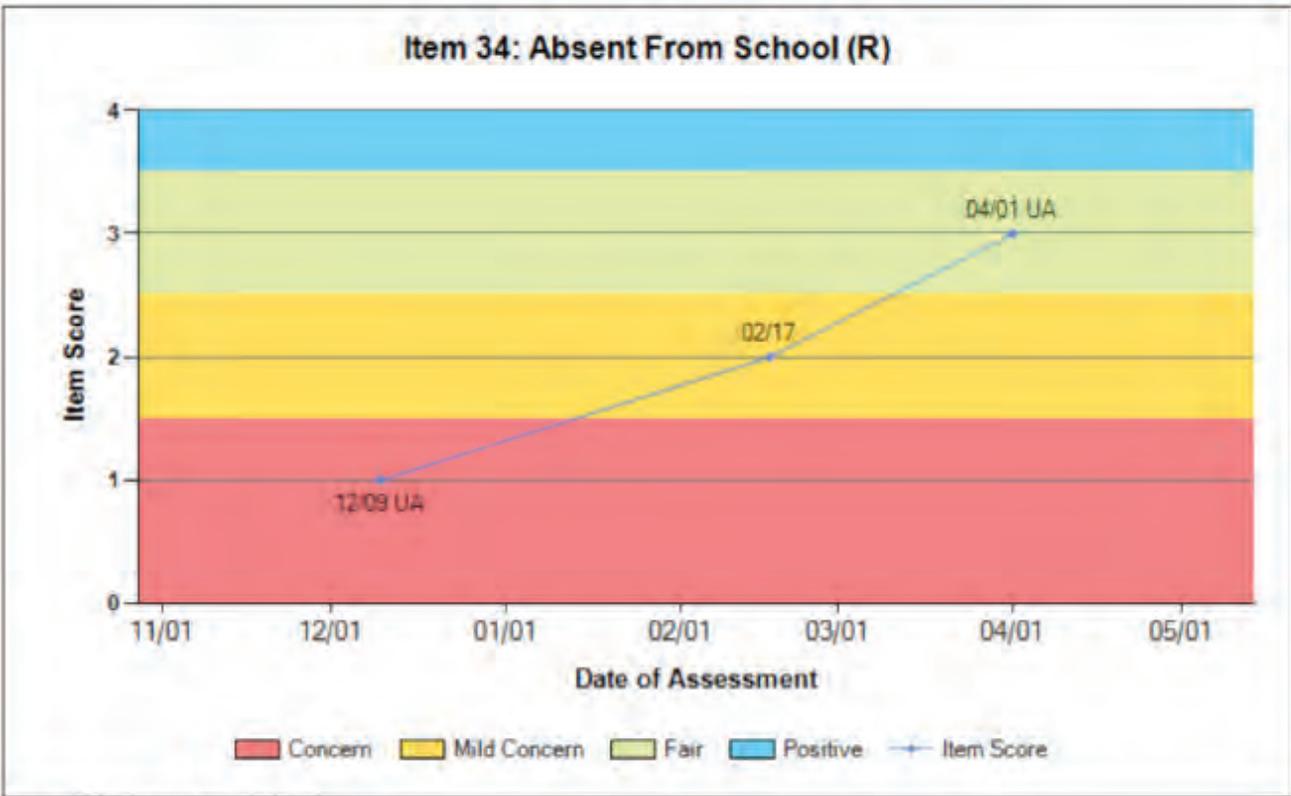
**B) Adaptive Scales**

Higher scores indicate FEWER concerns

**Academic Functioning**



*Note.* All names are fictional.



Note: (R) = Reverse scored Item.

**Related Reports:**

To monitor progress at the scale level:

[Standard Individual Progress Report](#)

Note. All names are fictional.



Copyright © 2011 Multi-Health Systems Inc. All rights reserved.



# Appendix F

## Generalizability of the BIMAS Standard Across Race/Ethnic Groups

Looking at scores across race and ethnic groups provides a way to examine the generalizability of the BIMAS™: the more similar the results, the more utility the BIMAS has for use with diverse populations. Analyses were conducted to determine if similar scores would be obtained across the race/ethnic groups from the normative sample (see chapter 9, *Standardization*, for a description of the normative sample). Multivariate analyses of covariance (MANCOVAs) were conducted to analyze the effects of race/ethnicity on scores from the BIMAS–Teacher Standard (BIMAS–T Standard), BIMAS–Parent Standard (BIMAS–P Standard), and BIMAS–Self-Report Standard (BIMAS–SR Standard). The African American, Hispanic, and White groups were the only groups with large enough sample sizes to be included in the analyses. The demographic characteristics of the samples (e.g., age and gender of the rated youth) were controlled for by including these variables as covariates. Because differences in the composition of the groups (e.g., covariates) were statistically controlled in these analyses, adjusted means are presented. In addition to significance levels ( $p < .05$ ), an estimate of effect size (Partial  $\eta^2$ ) is provided for every effect. The following planned comparisons were

designed before the analyses were conducted in order to assess pair-wise differences (given that the omnibus  $F$ -test was significant):

1. White versus African American
2. White versus Hispanic

Results at the multivariate level indicated significant, but small, effects of race/ethnicity for all rater types (see Table F.1). Univariate results are presented in Tables F.2 to F.4, revealing several significant effects (which is not surprising given the large sample size), although all effect sizes are small. An examination of the means reveals that ratings of White children were very close to the normative mean of 50, while ratings of African American and Hispanic children fell either slightly above, or slightly below this normative mean (scores never deviated more than 3.5  $T$ -scores from the mean). The results of these analyses indicate that although some isolated differences were statistically significant, on the whole, there were very few meaningful differences between scores on the BIMAS across the races/ethnicities as demonstrated by the small effect sizes.

**Table F.1. Race/Ethnicity Effects: Multivariate Results**

Rater	Wilks' Lambda	$F$	$df$	$p$	Partial $\eta^2$
Teacher	.93	8.35	10, 2174	< .001	.037
Parent	.98	2.33	10, 2578	.010	.009
Self-Report	.90	15.21	10, 1212	< .001	.050

*Note.* Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14.

**Table F.2. Differences Between Race/Ethnic Groups: BIMAS–T Standard**

Scale		Race/Ethnicity			$F$ (2, 1293)	$p$	Partial $\eta^2$	Planned Comparisons
		AA ( $n = 250$ )	HI ( $n = 169$ )	WH ( $n = 677$ )				
Conduct	$M$	52.5	50.4	51.3	5.05	.007	.009	AA > WH HI = WH
	$SD$	6.5	7.3	7.5				
Negative Affect	$M$	50.0	51.4	50.1	1.96	.141	.004	n/a
	$SD$	7.6	8.5	8.8				
Cognitive/ Attention	$M$	51.8	52.4	49.0	14.98	< .001	.027	AA > WH HI > WH
	$SD$	8.7	9.8	10.1				
Social	$M$	49.6	48.3	50.9	5.37	.005	.010	AA = WH WH > HI
	$SD$	9.3	10.5	10.8				
Academic	$M$	48.0	48.6	50.6	8.51	< .001	.015	WH > AA WH > HI
	$SD$	8.9	10.1	10.3				

*Note.* AA = African American, HI = Hispanic, WH = White, n/a = not applicable. Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14. For the planned comparisons, the '>' symbol indicates a statistically significant difference ( $p < .05$ ), and the '=' symbol indicates no significant difference.

**Table F.3. Differences Between Race/Ethnic Groups: BIMAS–P Standard**

Scale		Race/Ethnicity			<i>F</i> (2, 1293)	<i>p</i>	Partial $\eta^2$	Planned Comparisons
		AA ( <i>n</i> = 236)	HI ( <i>n</i> = 203)	WH ( <i>n</i> = 859)				
Conduct	<i>M</i>	51.6	50.3	50.7	1.47	.230	.002	n/a
	<i>SD</i>	6.2	7.0	5.4				
Negative Affect	<i>M</i>	51.2	50.3	50.1	1.44	.238	.002	n/a
	<i>SD</i>	6.5	7.3	5.7				
Cognitive/ Attention	<i>M</i>	52.0	50.7	50.1	3.73	.024	.006	AA > WH HI = WH
	<i>SD</i>	6.9	7.8	6.0				
Social	<i>M</i>	48.7	49.7	49.8	1.34	.261	.002	n/a
	<i>SD</i>	7.0	7.8	6.1				
Academic	<i>M</i>	47.2	48.7	49.9	8.41	< .001	.013	WH > AA HI = WH
	<i>SD</i>	6.5	7.3	5.7				

Note. AA = African American, HI = Hispanic, WH = White, n/a = not applicable. Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14. For the planned comparisons, the '>' symbol indicates a statistically significant difference ( $p < .05$ ), and the '=' symbol indicates no significant.

**Table F.4. Differences Between Race/Ethnic Groups: BIMAS–SR Standard**

Scale		Race/Ethnicity			<i>F</i> (2, 610)	<i>p</i>	Partial $\eta^2$	Planned Comparisons
		AA ( <i>n</i> = 127)	HI ( <i>n</i> = 136)	WH ( <i>n</i> = 352)				
Conduct	<i>M</i>	52.5	47.7	49.6	9.12	< .001	.029	AA > WH WH > HI
	<i>SD</i>	9.2	9.2	9.2				
Negative Affect	<i>M</i>	50.6	46.7	49.9	6.76	.001	.022	AA = WH WH > HI
	<i>SD</i>	9.7	9.8	9.8				
Cognitive/ Attention	<i>M</i>	51.2	46.5	50.5	9.70	< .001	.031	AA = WH WH > HI
	<i>SD</i>	9.9	9.9	9.9				
Social	<i>M</i>	49.3	52.9	50.4	4.65	.010	.015	AA = WH HI > WH
	<i>SD</i>	9.9	9.9	9.9				
Academic	<i>M</i>	47.3	50.8	51.1	9.02	< .001	.029	WH > AA HI = WH
	<i>SD</i>	8.9	8.9	8.9				

Note. AA = African American, HI = Hispanic, WH = White. Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14. For the planned comparisons, the '>' symbol indicates a statistically significant difference ( $p < .05$ ), and the '=' symbol indicates no significant difference.

# Appendix G

## Age and Gender Effects

This appendix provides the results of the age and gender effects analyses that were conducted in order to determine the normative groups (see chapter 9, *Standardization*, for details). Multivariate Analyses of Variance (MANOVAs) were employed to examine the relationships between gender and age, and BIMAS™ Standard raw scale scores. Results at the multivariate level revealed significant main effects for Age and Gender, as well as for the Age × Gender interaction for all three raters (see Table G.1). At the univariate level, Age was found to significantly affect all scales (see Table G.2), while Gender significantly affected the majority of the scales (see Table G.3). These main effects were qualified by several significant Age × Gender interactions (see Table G.4).

**Table G.1. Age and Gender Effects: Multivariate Results**

Rater	Independent Variable	Wilks' Lambda	F	df	p	Partial $\eta^2$
Teacher	Age	.71	16.81	25, 4381.2	< .001	.066
	Gender	.86	39.49	5, 1179	< .001	.143
	Age × Gender	.86	7.38	25, 4381.3	< .001	.030
Parent	Age	.85	9.38	25, 5113.1	< .001	.033
	Gender	.95	15.21	5, 1376	< .001	.052
	Age × Gender	.93	3.84	25, 5113.1	< .001	.014
Self-Report	Age	.72	16.81	25, 4381.3	< .001	.066
	Gender	.92	11.81	5, 658	< .001	.082
	Age × Gender	.86	7.38	25, 4381.3	< .001	.030

*Note.* Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14.

**Table G.2. Age Effects: Univariate Results**

Rater	Scale	F	df	p	Partial $\eta^2$
Teacher	Conduct	21.47	5, 1183	< .001	.083
	Negative Affect	4.82	5, 1183	< .001	.020
	Cognitive/Attention	14.16	5, 1183	< .001	.056
	Social	16.66	5, 1183	< .001	.066
	Academic Functioning	23.77	5, 1183	< .001	.091
Parent	Conduct	17.05	5, 1380	< .001	.058
	Negative Affect	9.68	5, 1380	< .001	.034
	Cognitive/Attention	11.89	5, 1380	< .001	.041
	Social	6.90	5, 1380	< .001	.024
	Academic Functioning	9.06	5, 1380	< .001	.032
Self-Report	Conduct	4.37	2, 662	.013	.013
	Negative Affect	8.16	2, 662	< .001	.024
	Cognitive/Attention	10.15	2, 662	< .001	.030
	Social	5.30	2, 662	.005	.016
	Academic Functioning	4.71	2, 662	.009	.014

*Note.* Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14.

**Table G.3. Gender Effects: Univariate Results**

Rater	Scale	<i>F</i>	<i>df</i>	<i>p</i>	Partial $\eta^2$
Teacher	Conduct	1.78	1, 1183	.189	.001
	Negative Affect	7.17	1, 1183	.008	.006
	Cognitive/Attention	83.33	1, 1183	< .001	.066
	Social	0.57	1, 1183	.450	.000
	Academic Functioning	52.11	1, 1183	< .001	.042
Parent	Conduct	15.51	1, 1380	< .001	.011
	Negative Affect	1.31	1, 1380	.253	.001
	Cognitive/Attention	49.30	1, 1380	< .001	.034
	Social	21.00	1, 1380	< .001	.016
	Academic Functioning	8.77	1, 1380	< .001	.014
Self-Report	Conduct	11.73	1, 662	.001	.017
	Negative Affect	1.07	1, 662	.301	.002
	Cognitive/Attention	1.15	1, 662	.284	.020
	Social	8.07	1, 662	.005	.012
	Academic Functioning	30.81	1, 662	< .001	.044

Note. Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14.

**Table G.4. Age  $\times$  Gender Effects: Univariate Results**

Rater	Scale	<i>F</i>	<i>df</i>	<i>p</i>	Partial $\eta^2$
Teacher	Conduct	5.75	5, 1183	< .001	.024
	Negative Affect	4.29	5, 1183	.001	.018
	Cognitive/Attention	1.63	5, 1183	.148	.007
	Social	7.29	5, 1183	< .001	.030
	Academic Functioning	4.94	5, 1183	< .001	.020
Parent	Conduct	3.14	5, 1380	.008	.011
	Negative Affect	5.28	5, 1380	< .001	.019
	Cognitive/Attention	9.77	5, 1380	< .001	.034
	Social	3.82	5, 1380	.002	.014
	Academic Functioning	6.78	5, 1380	< .001	.024
Self-Report	Conduct	1.34	2, 662	.264	.004
	Negative Affect	2.31	2, 662	.100	.007
	Cognitive/Attention	0.44	2, 662	.647	.001
	Social	11.27	2, 662	< .001	.033
	Academic Functioning	0.01	2, 662	.950	.000

Note. Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14.

# Appendix H

## Raw Score Standard Error of Measurement Values

This appendix provides the standard error of measurement (*SEM*) values for raw scores on the BIMAS™ Standard for all three rater types: BIMAS–Teacher Standard (BIMAS–T Standard), BIMAS–Parent Standard (BIMAS–P Standard), and BIMAS–Self-Report Standard (BIMAS–SR Standard). For most purposes, *SEM* values for *T*-scores (see *Standard Error of Measurement* in chapter 10, *Reliability*) are sufficient; however, raw score values are provided for research purposes (see Tables H.1 to H.3).

**Table H.1. Standard Error of Measurement: BIMAS–T Standard Raw Scores**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
Total Sample		1.21	1.26	1.63	1.52	1.52
Combined Gender	5–6	1.25	1.60	1.87	1.64	1.63
	7–9	1.60	1.57	2.02	1.64	1.64
	10–11	1.18	1.34	1.64	1.79	1.56
	12–13	0.50	1.02	1.18	1.00	1.01
	14–16	1.05	1.07	1.28	1.30	1.55
	17–18	1.65	1.27	1.73	1.27	1.75
Male	5–6	1.20	1.73	1.85	1.48	1.79
	7–9	1.58	1.46	1.97	1.54	1.69
	10–11	1.15	1.15	1.76	1.68	1.79
	12–13	0.55	0.74	1.29	1.01	1.03
	14–16	1.06	0.69	1.44	1.43	1.67
	17–18	1.31	1.31	1.94	1.24	1.85
Female	5–6	1.35	1.41	1.95	1.98	1.37
	7–9	1.79	1.70	2.13	1.75	1.55
	10–11	1.24	1.39	1.52	1.84	1.28
	12–13	0.40	1.35	0.96	1.01	0.97
	14–16	1.00	1.38	1.01	1.17	1.31
	17–18	1.91	1.17	1.33	1.23	1.51

**Table H.2. Standard Error of Measurement: BIMAS–P Standard Raw Scores**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
Total Sample		1.72	1.57	1.80	1.56	1.59
Combined Gender	5–6	1.87	1.75	1.71	1.60	1.66
	7–9	1.73	1.63	1.71	1.62	1.64
	10–11	1.69	1.44	1.82	1.93	1.69
	12–13	1.59	1.49	1.55	1.33	1.46
	14–16	1.50	1.47	1.97	1.64	1.66
	17–18	2.21	1.76	2.07	1.29	1.47
Male	5–6	1.95	1.91	1.64	1.62	1.61
	7–9	1.81	1.49	1.75	1.86	1.58
	10–11	1.55	1.30	1.88	2.10	1.67
	12–13	1.85	1.61	1.82	1.46	1.57
	14–16	1.55	1.54	2.08	1.73	1.75
	17–18	2.21	1.46	1.85	1.24	1.65
Female	5–6	1.79	1.58	1.89	1.63	1.73
	7–9	1.74	1.73	1.73	1.44	1.73
	10–11	1.95	1.76	1.85	1.62	1.61
	12–13	1.29	1.33	1.16	1.20	1.29
	14–16	1.31	1.35	1.49	1.40	1.47
	17–18	2.04	1.96	2.08	1.30	1.26

**Table H.3. Standard Error of Measurement: BIMAS–SR Standard Raw Scores**

Gender and Age Group		Behavioral Concern Scales			Adaptive Scales	
		Conduct	Negative Affect	Cognitive/Attention	Social	Academic Functioning
Total Sample		2.09	1.87	2.01	1.82	1.82
Combined Gender	12–13	2.08	1.82	2.02	1.97	1.90
	14–16	1.99	1.81	1.98	1.81	1.65
	17–18	2.17	1.94	2.05	1.68	1.96
Male	12–13	2.05	1.57	2.03	1.85	1.92
	14–16	2.02	1.73	1.95	1.90	1.72
	17–18	2.27	2.10	2.15	1.79	1.99
Female	12–13	2.20	2.01	1.93	2.01	1.86
	14–16	1.91	1.90	1.99	1.67	1.49
	17–18	1.93	1.84	1.96	1.47	1.93

# Appendix I

## Group Membership MANOVAs

Multivariate analyses for variance (MANOVAs) were conducted to investigate whether scores on the BIMAS™ Standard could distinguish between clinical and non-clinical groups, as well as between various clinical groups. This appendix presents results for the BIMAS–Teacher Standard (BIMAS–T Standard), BIMAS–Parent Standard (BIMAS–P Standard), and BIMAS–Self-Report Standard (BIMAS–SR Standard).

*T*-scores from the clinical samples (see *The BIMAS Standard as a Screening Tool* section in chapter 11, *Validity*, for a sample description) were compared to each other and to those of the normative sample (see chapter 9, *Standardization*, for a sample description). The BIMAS Standard scale scores were the dependent variables (DV). The independent variable (IV) of interest was group membership. In addition to reporting significance levels, an estimate of effect size (partial  $\eta^2$ ) is provided for every effect. Values for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14.

The following planned comparisons were designed before the analyses were conducted in order to assess pair-wise differences:

1. target clinical group versus the normative sample
2. target clinical group versus the other clinical groups

The target clinical groups were defined as follows:

- Conduct: Disruptive Behavior Disorders (DBD; includes Conduct Disorder, Oppositional Defiant Disorder)
- Negative Affect: Anxiety (ANX; includes Generalized Anxiety Disorder, Obsessive Compulsive Disorder, and Social Phobia) and Depression (DEP; includes Major Depressive Disorder and Dysthymia)

- Cognitive/Attention: Attention-Deficit/Hyperactivity Disorder (ADHD)
- Social: Pervasive Developmental Disorders (PDD; includes Autism and Asperger’s Syndrome)
- Academic Functioning: all clinical groups

Results at the multivariate level indicated significant effects of group membership for all rater types (see Table I.1). Significant effects, with large effect sizes, were found for all scales at the univariate level (see Tables I.2 to I.4). Results of the planned comparisons revealed that for all rater types, the target clinical group scored significantly higher than the normative sample across Behavioral Concern scales (Conduct, Negative Affect, and Cognitive/Attention; higher scores indicate more behavioral concerns). Moreover, with only one exception, the target clinical group scored significantly higher than other clinical groups across the three Behavioral Concern scales for all rater types. On the Adaptive scales (Social and Academic Functioning), as expected, the target clinical group scored significantly lower than the normative sample for all rater types (lower scores indicate more concerns in adaptive functioning). In addition, other than one exception, the target clinical group also scored significantly lower than other clinical groups on the two Adaptive scales across all rater forms.

**Table I.1. Group Membership Effects: Multivariate Results**

Rater	Wilks' Lambda	F	df	p	Partial $\eta^2$
Teacher	.45	54.15	30, 7430	< .001	.146
Parent	.54	47.87	25, 6884.5	< .001	.115
Self-Report	.63	21.00	25, 3812.9	< .001	.090

*Note.* Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14.

**Table I.2. Differences Between Groups: BIMAS–T Standard**

Scale	Group Membership										F (6, 1861)	Partial $\eta^2$	Planned Comparisons
	Norm (N = 1399)	DBD (N = 122)	ADHD (N = 93)	ANX (N = 54)	DEP (N = 60)	PDD (N = 95)	LD (N = 45)						
Conduct	M	70.5	62.0	63.2	63.4	58.1	59.5						
	SD	9.0	9.5	12.0	10.0	9.8	11.3						
Behavioral Concern Scales	M	50.7	65.6	63.4	71.2	72.0	64.1						
	SD	8.8	10.2	8.7	10.4	8.9	10.2						
Cognitive/Attention	M	50.2	68.6	69.1	64.8	63.9	67.4						
	SD	11.1	9.7	8.2	11.7	9.0	6.8						
Social	M	48.7	35.3	39.5	34.4	36.5	33.4						
	SD	10.0	8.4	11.0	11.9	9.3	9.1						
Adaptive Scales	M	50.4	37.2	40.2	38.6	39.8	39.5						
	SD	10.1	8.4	9.3	11.9	9.9	10.3						

Note. All *F*s significant at  $p < .001$ . Norm = Normative Sample, DBD = Disruptive Behavioral Disorders, ADHD = Attention Deficit/Hyperactive Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorders, LD = Learning Disorders. Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14. For the planned comparisons, the '>' symbol indicates a statistically significant difference ( $p < .05$ ).

**Table I.3. Differences Between Groups: BIMAS–P Standard**

Scale	Group Membership							F (5, 1803)	Partial $\eta^2$	Planned Comparisons
	Norm (N = 1396)	DBD (N = 70)	ADHD (N = 117)	ANX (N = 67)	DEP (N = 73)	PDD (N = 86)				
Conduct	M	51.1	69.1	57.9	59.6	63.6	55.2			
	SD	8.9	9.5	8.8	11.1	8.9	10.1			
Behavioral Concern Scales	M	50.7	61.4	57.5	67.1	67.3	60.4			
	SD	9.0	9.3	9.1	10.9	9.2	9.5			
Cognitive/Attention	M	49.4	63.4	63.8	58.8	59.0	58.5			
	SD	9.5	8.4	9.9	9.9	7.7	10.5			
Social	M	49.8	34.6	42.2	37.4	35.4	35.3			
	SD	9.4	8.2	8.3	8.4	8.5	11.3			
Adaptive Scales	M	49.7	36.1	40.6	39.7	38.4	44.2			
	SD	9.0	7.2	7.4	7.8	6.8	7.5			

Note. All *F*s significant at  $p < .001$ . Norm = Normative Sample, DBD = Disruptive Behavioral Disorders, ADHD = Attention Deficit/Hyperactive Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorders. Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14. For the planned comparisons, the '>' symbol indicates a statistically significant difference ( $p < .05$ ), and the '=' symbol indicates no significant difference.

Table I.4. Differences Between Groups: BIMAS-SR Standard

Scale			Group Membership						F (5, 1030)	Partial $\eta^2$	Planned Comparisons
			Norm (N = 699)	DBD (N = 65)	ADHD (N = 89)	ANX (N = 56)	DEP (N = 62)	PDD (N = 65)			
Conduct	M		50.1	64.8	53.7	57.7	59.1	53.2	41.71	.168	DBD > Other Clinical DBD > Norm
	SD		9.6	8.3	7.8	10.4	8.7	8.7			
Behavioral Concern Scales	M		50.2	57.0	55.7	60.8	65.2	60.0	48.30	.190	ANX, DEP > Other Clinical ANX, DEP > Norm
	SD		10.2	8.3	8.8	8.1	8.3	10.5			
Cognitive/ Attention	M		49.4	59.5	57.6	56.3	58.8	54.4	34.21	.142	ADHD = Other Clinical ADHD > Norm
	SD		10.1	6.4	7.3	8.6	7.9	10.3			
Social	M		49.6	39.7	47.3	41.8	38.3	36.4	41.61	.168	Other Clinical > PDD Norm > PDD
	SD		10.8	8.2	11.1	7.9	5.7	9.1			
Adaptive Scales	M		49.6	38.2	45.7	41.9	40.8	42.9	35.26	.146	Norm > All Clinical
	SD		9.7	7.8	7.2	9.6	6.7	8.0			

Note. All *F*s significant at  $p < .001$ . Norm = Normative Sample, DBD = Disruptive Behavioral Disorders, ADHD = Attention Deficit/Hyperactive Disorder, ANX = Anxiety, DEP = Depression, PDD = Pervasive Developmental Disorders. Ranges for evaluating partial  $\eta^2$  are small = .01, medium = .06, large = .14. For the planned comparisons, the '>' symbol indicates a statistically significant difference ( $p < .05$ ), and the '=' symbol indicates no significance difference.

# Index

## A

- across-informant correlations. *See* consistency between raters
- Adaptive scales. *See* scales
- administration
  - age range 16
  - appropriate raters 16
  - completion time 16
  - considerations 4, 13–16
  - ethical and legal guidelines 18
  - group 17
  - options 13
  - procedure 18–20
  - reading levels 16
  - setting 17
  - time frame 17
  - verbal 17–18
- age by gender interaction 207–208
- age effects 73–74
- age range 16, 73
- assessing change over time 35–38, 43–44, 47–48, 52–54, 58–61
- Assessment reports. *See* reports

## B

- Bardos, Dr. Achilles xii
- Behavior Concern scales. *See* scales
- BIMAS Flex. *See* forms
- BIMAS Online 1
  - administration 2, 18–20
  - hardware and software requirements 21
  - scoring 2, 20–21
- BIMAS, overview
  - components 1–2
  - formats 2
  - main features 1
  - multi-informant form options 2
  - report types 2
  - uses/applicability 3
- BIMAS Standard. *See* forms

## C

- careless responding 4
- case study 49–61
  - Tier 1 (Universal Level) 49–54
  - Tiers 2 and 3 (Targeted and Intensive Levels) 54–61

- CBRS, relationship with BIMAS 81, 84–86
- CFA. *See* confirmatory factor analysis
- change-sensitivity 10–11, 64–65, 95–98
- choosing the appropriate form
  - BIMAS Flex 14–15
  - BIMAS Standard 13–14
  - overview 13
  - within the RTI Framework 15
- clinical assessment 3–4, 9–11
- clinical groups
  - data collection 66–67, 86–87
  - demographic characteristics 87
  - MANOVA 211–213
- clinically meaningful change 35–38, 47–48, 59–61
- clinician form 2, 106
  - development of 66, 67
- combined gender norms 29, 73
- Comparative reports. *See* reports
- comparing results between
  - groups 38–39, 43, 44, 51, 52–53
  - raters 38–39, 46, 47, 56–58, 59–60
- concerns in school-aged youth 7
- conducting research with the BIMAS 100
- confidence intervals 29, 31, 33–34, 77
- confidentiality 18, 19
- confirmatory factor analysis
  - method and results 67, 82
- consent, informed 18
- consistency between raters 80
- consistency, internal 75–76
- contact the publisher, MHS iv
- convergent validity. *See* relationship between the BIMAS and other measures
- copyright information iii, iv
  - legal guidelines 18
- correcting errors 20
- correlations
  - across-informant. *See* consistency between raters
  - between BIMAS and other measures. *See* CBRS, relationship with BIMAS
  - scale-level 82–84
- Cronbach's alpha 75, 76

## D

- data collection, method 66–67, 71
- debriefing respondents 18, 20
- demographic information
  - normative sample, collection of 71
  - on rater forms 19

Demographic reports. *See* reports  
 development of the BIMAS 63–69  
   BIMAS Flex 67–68  
   BIMAS Standard 64–67  
   goals 63–64  
   Item Descriptors for BIMAS Standard 68–69  
   preliminary research 64–67  
   preliminary structure and items 65–66  
   rationale 63  
 different raters 28, 38, 46, 47, 66, 80  
 disability  
   IDEA 2004 7, 63  
   rater with 17  
 disclaimer iv  
 discriminative validity. *See* validity: construct:  
   screening  
 double-entry, scoring option 21  
 DSM-IV-TR  
   data collection, clinical sample 71

## E

early identification and intervention  
   1, 3, 7–8, 63, 99  
 effects, demographic 205–206, 207–208  
 effect size  
   change-sensitivity, early versions of the BIMAS  
     95–97  
   change-sensitivity, the BIMAS 97–98  
   comparisons between normative and clinical  
     samples  
     parent form 90–92  
     self-report form 93–94  
     teacher form 88–90  
 interpreting change  
   31, 37–38, 48, 60–61, 193–196  
 elevated scores. *See* high scale scores, interpreting  
 environments, administration 17  
 errors, correcting 20  
 ethical guidelines 18  
 ethnicity/race. *See* race/ethnicity

## F

factor analysis  
   confirmatory 67, 82  
 features  
   online scoring 20–21  
   overview 1  
 feedback to the publisher xiii, 100  
 fixing errors 20  
 formats, available  
   administration and scoring 2, 13–21  
 forms  
   BIMAS Clinician. *See* clinician form  
   BIMAS Flex 2, 14–16, 47, 59–60, 107–129  
   development of 67–68

rater forms  
   BIMAS Parent. *See* parent form  
   BIMAS Self-Report. *See* self form  
   BIMAS Teacher. *See* teacher form  
 BIMAS Standard 1–2, 14, 105–106  
   development of 64–67  
 freedom to withdraw 18  
 frequency  
   item responses 14, 34, 66, 68–69

## G

gender effects 73–74, 207–208  
 gender norms 29, 68, 73  
 geographic region distributions  
   normative sample  
     parent form 73  
     self-report form 73  
     teacher form 72  
 goals for BIMAS development. *See* development of  
   the BIMAS  
 grade levels, recommended 16  
 group administration 17  
 Group Assessment (GA) 23–25, 44  
 group evaluation. *See* Response to Intervention  
   (RTI): Tier 1 (Universal Level)  
 group membership, effects  
   clinical sample  
     parent form 91–92, 211–212  
     self-report form 93–95, 211, 213  
     teacher form 88–90, 211–212  
 guidelines  
   administration 16–17  
   ethical and legal 18  
   interpretation 31–39, 41–48

## H

hardware requirements 21  
 high group average scores, interpreting 39, 43  
   in case study 51  
 high scale scores, interpreting 32–35  
   in case study 54–56

## I

ICD 71  
 IDEA 2004  
   and early intervention 7–8, 63  
 IEP. *See* Individualized Education Program (IEP)  
 IISRs. *See* Intervention Item Selection Rules  
   (IISRs)  
 improper responses 20  
 individual administration 17  
 individual evaluation 44–48, 54–61  
 Individualized Education Program (IEP) 3, 67  
 individual level reports. *See* reports  
 Individuals with Disabilities Education Improvement

Act. *See* IDEA 2004  
 information confidentiality 18, 19  
 information, demographic. *See* demographic information  
 informed consent 17, 18  
 intercorrelations, scale. *See* correlations:  
   scale-level  
 internal consistency 75–76  
 interpretation of BIMAS results across reports  
   caution 4, 44  
   discrepant results between raters. *See* comparing results between: raters  
   step-by-step method  
     application of report types, overview 42  
     group evaluation 41–44  
       in case study 49–54  
     individual evaluation 44–48  
       in case study 54–61  
 interpretation of BIMAS scores. *See* scores  
 inter-rater reliability 80  
 intervention  
   importance of early 7, 63  
   in case study  
     group 52  
     individual 58–61  
   planning of  
     group 43  
     individual 3, 47  
 Intervention Group 41, 50, 141, 161–165  
 Intervention Item Selection Rules (IISRs)  
   10–11, 64–65, 95–97  
 Item Descriptors 34–35, 68–69  
 item-level scores, interpreting 34–35, 45–46  
   in case study 55–56  
 item responses 34, 66  
 items  
   development of 64–66  
   omissions 20  
 items by scale  
   Flex items 107–129  
   Standard items 105–106

## L

legal guidelines 18  
 levels of functioning 33–35  
 levels of risk 32, 34–35  
 levels, reading 16  
 low scale scores, interpreting 32–33  
   in case study 55–56

## M

materials, test 18, 19–20  
 maximum item omissions 20  
 McDougal, Dr. James xii

meaningful change  
   group level, interpreting 44  
     in case study 52–54  
   individual level, interpreting 35–38, 47–48  
     in case study 59–61  
 means  
   clinical sample  
     parent form 91  
       across clinical groups 92, 212  
     self form 93  
       across clinical groups 94–95, 213  
     teacher form 89  
       across clinical groups 89–90, 212  
   normative sample 74  
     across race/ethnic groups 206  
 Meier, Dr. Scott xii  
 MHS. *See* Multi-Health Systems  
 missing responses  
   data collection during standardization 71  
   during administration 20  
 monitoring change over time  
   group level, interpreting 44  
     in case study 52–54  
   individual level, interpreting 35–38, 47–48  
     in case study 59–61  
 Multi-Health Systems  
   contact information iv, 100  
   copyright information iii, iv  
   website 100  
 multiple  
   raters 2, 38–39, 46  
   responses 20  
 multi-student level reports. *See* reports

## N

negative predictive power 87  
   parent form 91  
   self form 94  
   teacher form 89  
 normative sample  
   data collection 66  
   description  
     parent form 72–73  
     self form 73  
     teacher form 72  
 norming 73–74

## O

omitted responses 20  
 online  
   administration 2, 13, 16–20  
   requirements 21  
   scoring 2, 20–21  
 original item set 65–66  
 outcome assessment 3, 8–11

outcome monitoring  
scores for 35–38

overall correct classification rate 81, 87  
parent form 91  
self form 94  
teacher form 89

## P

paper-and-pencil forms  
administration 2, 13, 16–20  
scoring, including scanning and importing  
2, 13, 18–21

parceling items 82

parental education level 71–72

parent form 2, 14

development of 66–68

means and standard deviations 74

normative sample 72–73

parent ratings 2, 14, 71–73

PEL. *See* parental education level

percentiles

in norming 74

interpretation 33

positive predictive power 87

parent form 91

self form 94

teacher form 89

preliminary item development 65–66

primary diagnosis, clinical sample 86–87

principles of use 4

program evaluation 3

progress monitoring 3, 8

scores for 35–38

validity of the BIMAS in 95–98

Progress reports. *See* reports

prorating scores 20

psychosocial factors. *See* effects, demographic

publisher's address iii, 100

## Q

qualifications, scoring and interpreting 4

## R

race/ethnicity

categories 71

distributions 72–73

effects 205–206

rater comparisons, values for 133–134

raters

appropriate 16–17

rights 18

rationale for the BIMAS 63

raw scores 32

RCI. *See* Reliable Change Index

readability analyses 16

reading BIMAS items aloud for a rater. *See*  
administration: verbal

reading levels 16

re-administration 3, 13–16, 17

referrals 8

relationship between the BIMAS and other  
measures 84–86

reliability

definition 75

internal consistency 75–76

inter-rater 80

test-retest 78–79

Reliable Change Index

definition 36–37

effect size estimates, comparison with 38

in case study 60

interpreting change with 37

values 131

remote administration 17

reports

levels, overview 27

individual level 25–26

multi-student level 24–25

options 29

types

Assessment reports

about 2–3, 23

Class/Group Student Scores

23, 24, 27, 31, 41, 42, 50, 141

Risk Level Pyramids

24, 27, 31, 42, 49, 135–136

Standard Individual Assessment Report 23,

25, 27, 31–35, 42, 45, 55–56, 143–146

Student List by Risk Level

23, 24, 41, 42, 50, 137

Comparative reports

about 3, 28

Average Score Comparison Report

24, 28, 31, 39, 42, 44, 51, 151

Standard Individual Comparison Between

Raters 25, 27–28, 31, 38–39, 42, 46, 57–5

8, 153–158

Demographic reports

about 3, 28–29

Demographics Breakdown 24, 27, 28–29, 31

, 42, 43, 50–51, 147–148

Risk Level by Demographics 24, 27, 28–29,

31, 42, 43, 51, 149

Progress reports

about 3, 27–28

Average Score Comparison: Progress Report

25, 27, 42, 44, 52–53, 159–160

Class/Group Student Scores: Progress Report

42, 44, 161–165

Flex Individual Comparison Between Raters

26–28, 42, 47, 59, 177–183

- Flex Individual Progress Report
    - 26–28, 47, 59, 169–175
  - Risk Level by Demographics: Progress Report
    - 25, 27–29, 42, 44, 53, 167–168
  - Standard Individual Progress Report
    - 26, 27–28, 47, 60, 185–188
  - Standard Individual Progress Report: Item Analysis
    - 26–28, 47–48, 61, 197–204
  - Standard Individual Progress Report: Significant Change Over Time
    - 26–28, 47–48, 60–61, 189–192
  - Student List by Risk Level
    - 23–24, 27, 42, 44, 53–54, 139
  - research, use of the BIMAS in 100
  - response bias 4
  - response options 66
  - responses
    - correcting 19–20
    - omitted 20
  - Response to Intervention (RTI) 7–8, 15–16, 99
    - 3-Tier model 7–8
    - Tier 1 (Universal Level) 8, 49–54
    - Tier 2 and 3 (Targeted and Intense Levels) 8, 16, 54–61
  - reviewing treatment plans 3
  - RTI. *See* Response to Intervention (RTI)
- S**
- samples
    - clinical 86–87
    - demographic characteristics of 87
    - description of 86
    - normative 71–73
  - scales
    - Adaptive scales
      - about 2
      - development of 67, 69
      - interpreting scores
        - item level 35, 45–46, 56
        - scale level 33, 45, 55
    - items
      - Flex 121–129
      - Standard 105–106
    - scales
      - Academic Functioning 2, 33, 83, 105–106
      - Social 2, 33, 83, 105–106
  - Behavioral Concern scales
    - about 1
    - development of 67, 68–69
    - interpreting scores
      - item level 34–35, 45, 56
      - scale level 32, 45, 55
    - items
      - Flex 107–121
      - Standard 105–106
  - scales
    - Cognitive/Attention 1, 32, 83, 105–106
    - Conduct 1, 32, 83, 105–106
    - Negative Affect 1, 32, 83, 105–106
  - intercorrelations 82, 84
  - structure 81–83
  - scale scores, interpreting 32–34, 45
    - in case study 55
  - scorers, qualifications of 4
  - scores
    - comparison of
      - group averages 38–39, 43, 51
      - scores between raters 38, 46, 56–58
    - confidence intervals 33–34
    - item scores
      - interpreting 34–35, 45–46, 55–56
      - Item Descriptors, development of 68–69
    - number of students at each risk/functioning level
      - 41, 43, 44, 49–54
    - percentage of students at each risk/functioning level
      - 41, 43, 44, 49–54
    - progress and monitoring
      - effect size 37–38, 48, 60–61
      - Reliable Change Index (RCI) 36–37, 38, 48, 60–61
    - prorating 20
    - raw 32
    - scale scores
      - interpreting 32–34, 45, 55
      - Scale Descriptors, development of 32–33
    - Standardized 32, 73–74
    - T-scores 32–33
    - types of scores, overview 31
  - scoring
    - online 2, 13, 20–21
    - overview 13
  - screening 3, 14, 15, 41, 86–95
  - self form 2, 14
    - development 66–68
    - means and standard deviations 74
    - normative sample 73
  - self ratings 2, 14, 71, 73
  - service code
    - general education 28, 50–51, 53, 147–148, 149
    - special education 28, 50–51, 53, 147–148, 149
    - Title 1 28, 50–51, 53, 147–148, 149
  - settings, administration 17–18
  - significant change over time
    - 26, 31, 48, 60–61, 189–192
  - skipped responses 20
  - software. *See* BIMAS Online
  - standard error of measurement 77–78, 209–210
  - standard error of prediction 78–79
  - standardization process 71–74
    - age and gender effects 207–208
    - data collection 71

generalizability of the BIMAS Standard across  
 race/ethnic groups 205–206  
 geographic region distribution 72–73  
 normative samples 71–73  
 norming procedures 73–74  
 parental education level distribution 72  
 race/ethnicity distribution 72–73  
 standardization samples 71–73  
 standardized scores 32–35, 73–74  
 statistically significant change  
 26, 36–38, 48, 131, 189–192  
 in case study 60–61  
 structure  
 form 1–2, 13–16  
 development of 65–68  
 scale 1–2, 81–84  
 summary of results section in Standard Individual  
 Assessment Report 25, 29, 146

## T

teacher form 2, 14  
 development of 66–68  
 means and standard deviations 74  
 normative sample 72  
 teacher ratings 2, 14, 71–73  
 technical support 21  
 test materials 18, 19–20  
 test-retest reliability 75, 78–79  
 time frame, administration 17  
 time series graph 3, 27, 36, 47–48, 59–60  
 tools, online administration 21  
 translations 69  
 treatment  
 developing plans 3, 43, 47  
 evaluation of effectiveness  
 14–16, 35–38, 44, 47–48  
 in case study 52–54, 59–61  
 treatment response  
 monitoring. *See* treatment  
 T-score rights on BIMAS Online 29  
 T-scores  
 average T-score 39, 43, 44, 51, 52–53  
 definition 32  
 interpretation 32–33  
 non-linear, transformation of 74  
 report option 29

## U

Universal Assessment (UA) 23–25, 41, 44  
 use, principles of 4  
 users  
 appropriate 4  
 qualifications 4

## V

validity  
 construct  
 consistency of scores across rater types 80  
 generalizability across race/ethnic groups  
 205–206  
 progress monitoring 95–98  
 screening 86–95, 211–213  
 content 81–84  
 convergent 84–86  
 definition 81  
 visual displays 35–36

## W

website, MHS iv



## Behavior Intervention Monitoring Assessment System

James L. McDougal, Psy.D.,  
Achilles N. Bardos, Ph.D., &  
Scott T. Meier, Ph.D.

### TECHNICAL MANUAL

The Behavior Intervention Monitoring Assessment System (BIMAS™) is an empirically-based behavior rating and online data management system specifically developed for monitoring response to behavioral and psychosocial intervention in children and adolescents aged 5 to 18 years.

This multi-informant assessment, applicable across multiple settings, has forms for teachers, parents, youth, and clinicians. The BIMAS is a brief, repeatable, and change-sensitive measure that is useful for screening, progress monitoring, outcome assessment, and program evaluation within the Response to Intervention (RTI) framework.

The BIMAS Standard forms, featuring empirically-based change-sensitive items, are useful for universal screening of behavioral, social, emotional, and academic difficulties and strengths as well as assessing response to intervention.

The BIMAS Flex items can be used to develop individualized goals for intervention and to monitor progress on a more frequent basis.

Dynamic web-based Assessment, Progress, Comparative, and Demographic Reports are available from the BIMAS Online allowing for real-time exploration of assessment results at group and individual levels.

Thousands of ratings from across North America, years of intensive research, and sophisticated statistical analyses resulted in the BIMAS. This manual will tell you everything you need to know about the development, administration, scoring, and interpretation of the BIMAS, as well as its standardization, reliability, and validity.

#### USA

P.O. Box 950  
North Tonawanda, NY  
14120-0950  
Phone: 1.800.456.3003  
Fax: 1.888.540.4484

#### CANADA

3770 Victoria Park Ave.  
Toronto, ON M2H 3M6  
Phone: 1.800.268.6011  
Fax: 1.888.540.4484

#### INTERNATIONAL

Phone: +1.416.492.2627  
Fax: +1.416.492.3343

#### WEBSITE

[www.mhs.com](http://www.mhs.com)

#### EMAIL

[customerservice@mhs.com](mailto:customerservice@mhs.com)