



Ohio Mental Health Consumer Outcomes System

Report 12: *Reliable Change and Clinical Significance for Outcomes Instruments*

The purpose of this report is to explore the concepts of Reliable Change and Clinical Significance in assessing changes when comparing treatment outcomes between two time points for various assessment scales employed in the Ohio Mental Health Consumer Outcomes System. In this report, we compute the Reliable Change Index for Outcomes scales where previous research had not already identified values. We also explore the concept of Clinical Significance and report the cutoff scores established for the Ohio Scales Problem Severity and Functioning scales from previous publications. Aggregate reporting techniques using Reliable Change and Clinical Significance are presented.

SUMMARY

- The Reliable Change Index is a measure to assess the magnitude of change score necessary to be considered statistically reliable and not due to random measurement error. It is a function of the standard deviation and reliability coefficient of an instrument.
- Clinical Significance is a determination of a change in clinical status of moving from a dysfunctional range to a functional range while making a Reliable Improvement.
- In order to be considered as a reliable and clinically significant improvement, both criteria for Reliable Change and Clinical Significance should be met when comparing assessment scores between two time points.
- Ogles, Melendez, Davis, & Lunnen (1999) computed Reliable Change Indexes for Functioning and Problem Severity for the Ohio Scales, youth and parent versions respectively. Healy (2005) calculated the Reliable Change Index for the Community Functioning Scale in the Adult Provider Form A.
- A simple aggregate reporting technique based on the Reliable Change Index is to count the number and calculate the percent of cases that make reliable improvement, no reliable change, or reliable deterioration.
- A simple aggregate report using Clinical Significance would be to report the number, and calculate percent, of cases that are in functional and dysfunctional ranges, and report how many and what percentage of cases move from the dysfunctional range to functional range.
- The Reliable Change Index of various scales in the instruments employed in the Ohio Mental Health Consumer Outcomes System are presented (see page 6 for detail). In this report, the information needed for determining Clinical Significance is available only for the Ohio Scales Problem Severity and Functioning scales.

What is Reliable Change and Clinical Significance?

The use of pretest and posttest group scores and the calculation of statistical significance in traditional psychotherapy outcomes evaluation has been criticized as inadequate in providing information about treatment effects for specific individuals. At the same time, the use of statistical significance may have little clinical relevance for typical users of Outcomes data (Jacobson, Follette, & Revenstorf, 1984). Various research efforts in defining what constitutes improvement in psychotherapy research have been launched and most of these researchers focus on the concept of *Reliable and Clinically Significant Change* (some examples: Jacobson & Truax, 1991; Lunnen & Ogles, 1998; McGlinchey, Atkins, & Jacobson, 2002).

The broad concept of Reliable and Clinically Significant Change¹ can be addressed in two parts:

Reliable Change - Is the change of sufficient magnitude to be confident that the change is beyond that which could be attributed to measurement error?

Clinical Significance – How does the end state of the client compare with the scores observed in socially and clinically meaningful comparison groups?

Most of the literature explores these concepts together, as one builds upon the other. In this report, we will discuss the concept of Reliable Change for all instruments, but because we don't have measures for the Outcomes instruments from a non-clinical group of adults, we only discuss the concept of Clinically Significant Change as it relates to the Ohio Scales Functioning and Problem Severity scales. Further research effort would be needed to delineate the Clinically Significant Changes for the Adult Outcomes instruments.

Reliable Change (RC) is based upon the reliability or consistency of the measurement instrument, and is the difference on an assessment tool between two time points that would not be due to simple measurement error. A reliable change is usually assessed as the difference between pretest and posttest scores against a certain critical level. However, an assessment of Reliable Change is valuable at any point in treatment as a check of progress. The Reliable Change Index is evidence to show that changes between two time points (i.e., pretest to posttest, initial to follow up/termination) are not an artifact of measurement error.

Clinical Significance refers to a return to non-clinical functioning from a dysfunctional population (Jacobson & Truax, 1991). They suggested three methods to assess this concept.

- A. There is at least a 2 standard deviation (SD) distance from the mean of the original dysfunctional group.
- B. The subsequent result should fall within 2 standard deviations from the mean of the non-clinical population.
- C. There is a greater likelihood that the post treatment score falls closer to the mean of the non-clinical population than the mean of the clinical group. Jacobson, Follette, & Revenstorf (1984) proposed a statistical way to calculate the cutting score for the variable of interest.

Jacobson and Truax (1991) recommended using method B or C when norms for both clinical and non-clinical populations were available. In the case of the overlap between the distributions of the two populations, Jacobson and Truax (1991) suggested method C as a better choice. The information regarding the non-clinical population necessary to make a determination of Clinical Significance using methods B or C is only available for the Ohio Scales Functioning and Problem Severity scales. While method A could be used to determine Clinical Significance for the other scales, we do not know the extent of overlap between the functional and dysfunctional groups and

¹ Jacobson & Truax (1991) & Evans, Margison, & Barkham (1998) provided precise summaries of the calculation with illustrated examples in explaining the two concepts.

thus have no way to verify that two standard deviations from the mean of the dysfunctional group is actually close to the mean of the functional population.

In order to attribute an effect to treatment using these methods, both the criteria of *Reliable Change* and *Clinical Significance* must be met.

How to calculate a Reliable Change Index

The following calculation of the Reliable Change Index (RCI) was adapted from Jacobson & Truax (1991). The Reliable Change Index is a function of the standard deviation and reliability of the instrument:

$$RCI = \frac{x_2 - x_1}{SE_{diff}}$$

Where x_1 is the score at time 1 (usually the initial/pretest score) and x_2 is the time 2 score (follow up/posttest score) of the same subject. SE_{diff} is the standard error of the difference between the two test scores. It can be computed directly from the standard error of the measurement S_E :

$$SE_{diff} = \sqrt{2 (S_E)^2}$$

$$S_E = SD_1 \times \sqrt{1-r}$$

Where SD_1 is the standard deviation of the initial assessment and r is the reliability coefficient of the instrument².

When RCI is greater than 1.96, it is unlikely that the posttest score is due to random measurement error and the change is reliable ($p < .05$). The Reliable Change is easier to comprehend when presented in terms of a score range for the instrument. So working the other way round, the critical level for Reliable Change can be calculated using the RCI (which can be set at 1.96) multiplied by the SE_{diff} that will give a range that indicates the cut off level for reliable change between the two time points.

In this report, we used the first assessment of each unique individual in the Outcomes statewide database on July 10, 2006. We computed the Cronbach's Alpha and the standard deviation of each measure. Please refer to the appendix for detail.

The Reliable Change Index is a standardized measure based upon the normal curve. While this formulation is convenient when comparing across change in scales, it is not of much use in the field where a clinician might have to plug values into a formula in order to determine if a reliable change has been made. In order to make the concept of Reliable Change more readily available to users, we have calculated the minimum amount of change on the scale that constitutes a reliable change (See page 10). As a further convenience, the numbers reported are rounded to make comparisons of results obtained with the reliable change amount easier to use. This follows the conventions used by Ogles, Melendez, Davis & Lunnen (1999).

² Although Jacobson and Truax (1991) employed the test-retest reliability for calculation of RCI, some researchers argue that Cronbach's Alpha or another parameter of internal consistency is probably the most theoretically consistent approach since the theory behind this is classical reliability theory (Evans, Margison, & Barkham, 1998). A test-retest reliability measure always includes not only simple unreliability of the measure but also any real changes in whatever is being measured. This means that internal reliability is almost always higher than test-retest and will generally result in more people being seen to have changed reliably. In this report we employed the Cronbach's Alpha in the calculation

How to calculate Clinical Significance

Adopting the third method recommended by Jacobson and Truax (1991), we calculate the clinical significance with the means and standard deviations from both the clinical group and the non-clinical population to establish a cutting score for this purpose.

$$\text{Cutting score} = \frac{(M_{\text{clinical}} \times SD_{\text{norm}}) + (M_{\text{norm}} \times SD_{\text{clinical}})}{(SD_{\text{norm}} + SD_{\text{clinical}})}$$

Where M_{clinical} and M_{norm} are the mean scores of the clinical group and the non-clinical population respectively. SD_{clinical} and SD_{norm} are the standard deviations of the clinical group and the non-clinical population respectively.

A Worked Example

In order to illustrate the calculation of the Reliable Change Index, we use the Symptom Distress Scale from Adult Consumer Form A in the following calculation.

Data use for calculating Reliable Change for Symptom Distress Scale from Adult Consumer Form A

Standard deviation of Symptom Distress Scale at initial assessment	14.93
Cronbach's Alpha for the Symptom Distress Scale (15 items)	.93
Standard error of measurement for Symptom Distress Scale [$SD_1 \times \sqrt{(1-r)}$]	3.95
Standard error of the difference between the two test scores [$SE_{\text{diff}} = \sqrt{2(S_E)^2}$]	5.59

Setting $RCI > 1.96$ to achieve a reliable change:

$$\frac{x_2 - x_1}{SE_{\text{diff}}} = RCI > 1.96$$

$$x_2 - x_1 > 1.96 \times SE_{\text{diff}}$$

$$x_2 - x_1 > 10.95$$

Results showed that a difference greater than 10.95 between the two assessments would show a reliable change for the Symptom Distress Scale for Adult Consumer Form A.

For the illustration of the calculation the Clinical Significance, we use information of the Problem Severity scale of the Ohio Scales published by Ogles, Melendez, Davis, & Lunnen (2000).

Data use for calculating Clinical Significance Change for Problem Severity scale from the Ohio Scales (parent version)^a.

Mean score of Problem Severity scale from community sample (parents)	13.28
Standard deviation of Problem Severity scale from community sample (parents)	10.01
Mean score of Problem Severity scale from clinical sample #2 (parents)	35.43
Standard deviation of Problem Severity scale from clinical sample #2 (parents)	19.72

^a Data adopted from Ogles, Melendez, Davis, & Lunnen (2000), p42

$$\begin{aligned} \text{Cutting score for the Problem Severity scale (parent version)} &= \frac{(13.28 \times 19.72) + (35.43 \times 10.01)}{(10.01 + 19.72)} \\ &= 20.74 \end{aligned}$$

An Illustration with Outcomes Instruments

A further explanation below (see Figure 1) involves data for a random selected sample of 30 individuals from the production data of Adult Consumer Form A with valid Symptom Distress scores at initial and 6-month assessments. The scatter plot shows results of these 30 cases with the initial scores on the x-axis and 6-month results on the y-axis.

The area between line A and C indicates score differences between 11 and -11. Results falling within this boundary indicate that the changes between the two time points may be due to random measurement error. Results on the right side of line C are those scores with lower 6-month scores and at the same time the magnitude of the difference is larger than 11 (lower Symptom Distress indicates improvement). We refer to this as *reliable improvement*. Individuals with these scores can be classified as having positive reliable changes in their Symptom Distress Scale. Results on the left side of line A were those cases having 6-month scores higher than their initial scores. This indicates deterioration in their Symptom Distress Scale and this deterioration is also unlikely due to random measurement error. We call this *reliable deterioration*. Observation of reliable deterioration should serve as a red flag to clinicians that things are not going well in treatment.

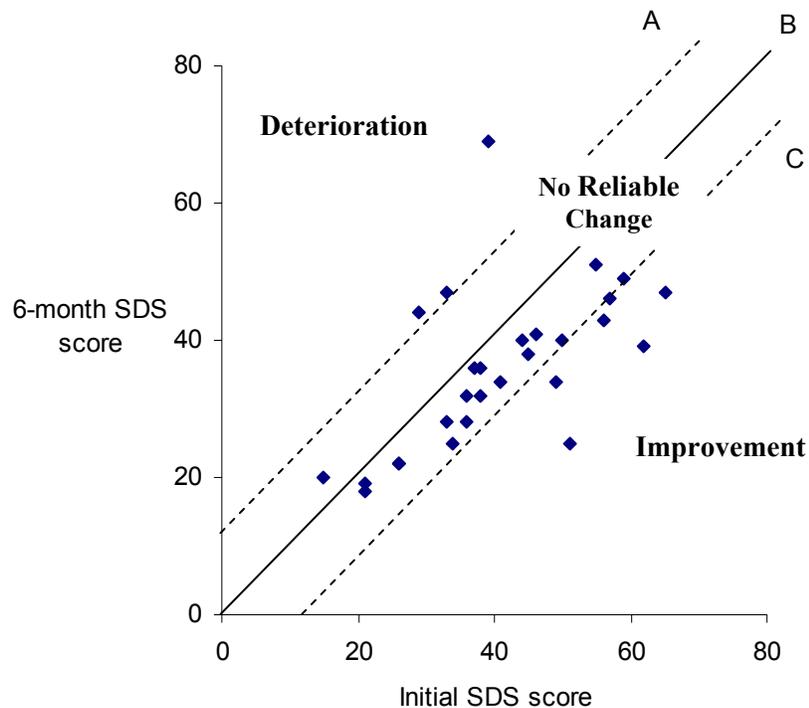


Figure 1. Scatter plot of Symptom Distress Scale (SDS) scores at initial and 6-month assessments.

A simple aggregate reporting technique based on the Reliable Change is to count the number and calculate the percent of cases that make reliable improvement, no reliable change, or reliable deterioration. In the above illustration, 20% of the cases showed reliable improvement (a decrease of 11 or more from the initial assessment to 6-month assessment), 10% of the cases showed reliable deterioration (an increase of 11 or more from the initial assessment to 6-month assessment), and the remaining 70% of the cases showed some positive or negative changes but their changes were not considered statistically reliable.

<i>Instrument</i>	<i>Scale</i>	<i>N</i>	<i>Reliable Improvement (%)</i>	<i>No Reliable Change (%)</i>	<i>Reliable Deterioration (%)</i>
Adult Consumer Form A	Symptom Distress	30	20.0	70.0	10.0

Figure 2 below shows another randomly selected sample of 32 individuals with valid Problem Severity Scale scores at initial and 6-month assessments from the Ohio Scales (Parent version). This illustration presents the use of clinical significance in addition to the reliable change we have discussed in the above example. The scatter plot shows results of these 32 cases with the initial scores on the x-axis and 6-month results on the y-axis³.

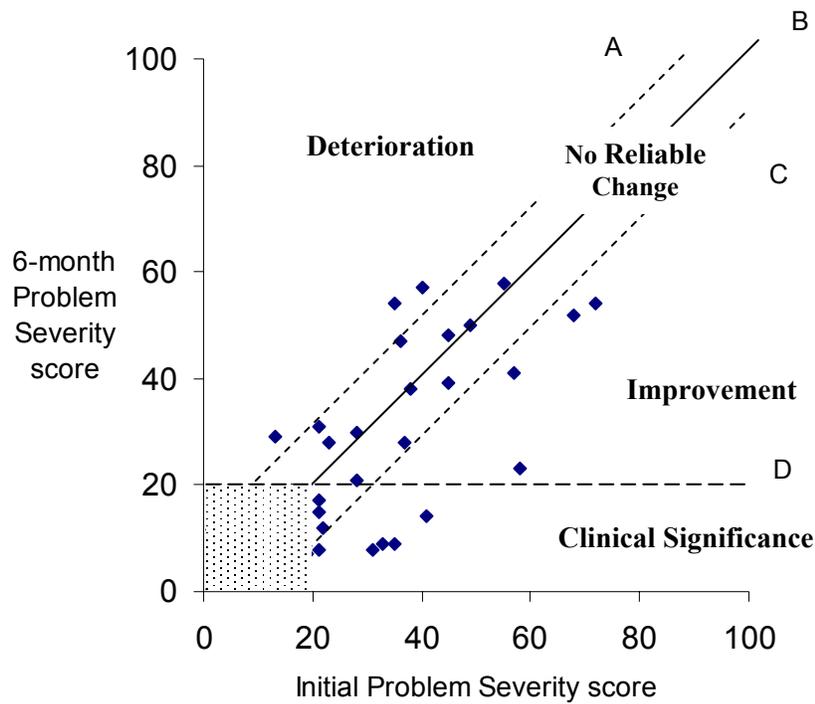


Figure 2. Scatter plot of Problem Severity Scale (Parent version) scores at initial and 6-month assessment.

The area between line A and C indicates score differences between 10 and -10. Results falling within this boundary indicate that the changes between the two time points may be due to random measurement error. Results on the right side of line C showed *reliable improvement* in that the scores' difference is greater than 10 (lower Problem Severity indicates improvement). Results on the left side of line A were those cases having 6-month scores higher than their initial scores and their differences are greater than 10. This indicates a *reliable deterioration* in Problem Severity and that this deterioration is unlikely due to random measurement error.

In addition, the scores that fall below the dashed line D are cases with their scores less than 20 (clinical significance cutting score) at their 6-month assessment. The cases in the gray area are records with initial Problem Severity score greater than 20 but have a 6-month Problem Severity score less than 20, which indicates a positive clinical significance (crossing from dysfunctional group to normative group).

³ Ogles notes that some cases may look worse at 3 months than at the initial time period for cases open longer than one year. It may be appropriate to use the lower of the initial or 3-month score as the baseline measure.

The area under the dotted area to the left of the gray area shows cases with both their initial and 6-month Problem Severity scores below 20. These cases are within the normative range at both the initial and the 6-month assessments. Since their scores have not migrated from dysfunctional group to the normative group, their changes will not be considered as clinically significant.

Records under the grayed area AND at the same time below line C are cases that show both clinical significance and reliable change.

For the Problem Severity scales from the Ohio Scales (parent version), 34.4% of this sample indicated reliable improvements (a decrease of 10 or more from the initial assessment to 6-month assessment), and 50% did not show a reliable change. The remaining 15.6% showed reliable deterioration (an increase of 10 or more between the two assessment time points). Twenty five percent of the cases show clinical significance in their Problem Severity score between the 6-month and initial assessment. Taking both criteria of reliable change and clinical significance together, 15.6% of the cases in this illustration are considered as showing a reliable and clinically significant change.

<i>Instrument</i>	<i>Scales</i>	<i>N</i>	<i>Reliable change (%)</i>	<i>Clinical Significance (%)</i>	<i>Reliable and Clinical Significant Change(%)</i>	
Ohio Scales (Parent version)	Problem Severity Scale	32	<i>Reliable Improvement^a</i>	34.4		
			<i>No Reliable Change</i>	50.0		
			<i>Reliable Deterioration^b</i>	15.6		
			<i>Clinical Significance^c</i>		25.0	
			<i>Not Clinical Significance</i>		75.0	
					<i>Reliable and Clinical Significant Change</i>	15.6

^a A change of 10 or more between 6-month assessment from initial assessment.

^b A change of -10 or more between 6-month assessment from initial assessment

^c An initial score above 20 and a 6-month score at 20 or below.

Discussion

With the Reliable Change Index available, researchers and clinicians can check the level of change in individual clients in order to assess whether a change identified is large enough to be reliable or may only be due to random measurement error. A summary of the Reliable Change and some Clinical Significance scores for Outcomes instruments were presented in page 10. It is important to keep in mind that change less than the reliable amount

may still be important and clinically meaningful -- standard statistics of significance can be used to detect such effects in the aggregate.

A simple aggregate reporting technique based on the Reliable Change is to count the number and calculate the percent of cases that make reliable improvement, no reliable change, or deterioration.

For the Clinical Significance, an aggregate report can present the number and calculate percent of cases that are in functional and dysfunctional ranges, and report how many and what percentage of cases move from the dysfunctional range to functional AND make a Reliable Improvement.

In reviewing results from various assessments over time, cases that show both reliable and clinically significant improvement would certainly fall into the successful category for positive clinical case reviews. At the same time, those cases with reliable deterioration certainly need the attention of the service provider and may signal the need for an in-depth service review. In handling of lack of positive response to treatment, Harmon, et al. (2005) suggested service providers may review various aspects of the treatment such as therapeutic alliance, the readiness to change of clients, social support resources, diagnostic formulation, medication, and providing feedback to clients.

There may also be individuals with Reliable Improvement who do not achieve the Clinical Significance criteria. These individuals may already be in the normative range before treatment or may have had scores that were so far in the dysfunctional range that their Outcomes scores do not move below the normative cutting score. Service providers may examine these cases with reference to their migration from dysfunctional to normative range and set up realistic goals for these cases. Similarly, cases may achieve clinical significance but their change may fall below the Reliable Change range. These cases may fall around the boundary of the dysfunctional range and normative range and they can cross the boundary with a small improvement. However, since their improvement is less than the reliable change criterion, their improvement may be unreliable and there may be a higher chance to relapse (Evans, 1998).

The use of a Reliable Change Index in treatment monitoring is not the only way to test for progress, and a change score less than the Reliable Change amount may still have clinical meaning. Some treatment modalities set expectations based upon clinical research, and clinicians are encouraged to consider the Outcomes scores obtained from the frame of reference of their treatment modality.

The clinical cutting score between the functional and dysfunctional range is but one example of cutting scores. Other cutting scores may be established that identify levels of severity. Barker, Barron, McFarland, & Bigelow (1994) devised three levels of functioning for the Multnomah County Community Functioning Assessment Scale (CFAS): High (63 and up), Medium (48 to 62) and Low (47 and down). A calibration between the CFAS and the Adult Provider A Community Functioning Scale allowed the identification of the equivalent three levels of severity on the Adult Provider A Community Functioning Scale (Healy, 2005). A percentage equivalency method was used to devise equivalent scores. The equivalent scores on the Community Functioning Scale were: High ≥ 43 ; Medium ≥ 30 and < 43 and Low < 30 . Using similar methodology in a calibration between the Child and Adolescent Functioning Assessment Scale (CAFAS) and the Ohio Scales Worker Problem Severity and Functioning scales, Ogles and Healy (2003) identified 4 clinical cutting scores. For the Problem Severity scale, the ranges are none: 0-9, mild: 10-19, moderate: 20-36, severe: 37-52, and extreme: 53 and up. For the Functioning scale, the ranges are: none: 66 and up, mild: 51-65, moderate: 35-50, severe: 23-34, and extreme: less than 23. The caveats about movement across clinical cutting score between functional and dysfunctional also apply to these cutting scores.

In this report, the data needed for determining Clinical Significance is available only for the Ohio Scales Problem Severity and Functioning scales. Future research is needed to document results against the functioning normative population and a clinically-identified group for the adult assessment instruments in the Ohio Mental Health Consumer Outcomes System.

Reliability Change Index for Outcomes Instruments

Instrument	Respondents	Collection Times	Scales	Range	Positive Direction	Reliable Change	Clinical Cutting Score
Adult Consumer A (ACA)	Adult Consumers who are thought to have severe mental illnesses	Initial, Six months, Annually, At Termination	Symptom Distress	15-75	Lower	11.0	
			Empowerment	1-4	Higher	0.4	
			Overall Quality of Life	1-5	Higher	0.8	
			Financial QOL	1-5	Higher	1.0	
Adult Provider A (APA)	Providers of those consumers who complete the Adult Consumer A Form	Initial, Six months, Annually, At Termination	Community Functioning	11-55	Higher	4.0 ^a	
Adult Consumer B (ACB)	Adult Consumers who are thought to have less severe mental illnesses	Initial, At Termination	Symptom Distress	15-75	Lower	11.0	
			Overall Quality of Life	1-5	Higher	0.8	
			Financial QOL	1-5	Higher	1.0	
Ohio Scales - Agency Worker	Providers of Youth 5 to 17 years	Initial, Six months, Annually, At Termination	Functioning	0 -80	Higher	8.0 ^b	50.0 ^b
			Problem Severity	0-100	Lower	10.0 ^b	20.0 ^b
Ohio Scales - Parent	Parents of Youth 5 to 17 years	Initial, Six months, Annually, At Termination	Functioning	0 -80	Higher	8.0 ^b	50.0 ^b
			Problem Severity	0-100	Lower	10.0 ^b	20.0 ^b
			Satisfaction	4-24	Lower	6.0	
			Hopefulness	4-24	Lower	6.0	
Ohio Scales - Youth	Youth 12 to 17 years	Initial, Six months, Annually, At Termination	Functioning	0 -80	Higher	8.0 ^b	60.0 ^b
			Problem Severity	0-100	Lower	10.0 ^b	20.0 ^b
			Satisfaction	4-24	Lower	6.0	
			Hopefulness	4-24	Lower	5.0	

^a Findings adopted from Healy (2005).

^b Findings adopted from Ogles, Melendez, Davis, & Lunnen (1999).

References

- Barker, S., Barron, N., McFarland, B., & Bigelow, D. (1994). *Multnomah Community Ability Scale: User's Manual*. Portland, OR: Western Mental Health Research Center.
- Christensen, L. & Mendoza, J. L. (1986). A method of assessing change in a single subject: an alteration of the RC index. *Behavior Therapy* 17: 305-308.
- Evans, C. (1998). *Reliable and clinically significant change*. <http://www.psych.org/stats/rcsc.htm>.
- Evans, C., Margison, F. & Barkham, M. (1998) The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence Based Mental Health* 1:70-72.
- Harmon, C., Hawkins, E.J., Lambert, M., Slade, K., Whipple, J.L. (2006). Improving outcomes for poorly responding clients: The use of clinical support tools and feedback to clients. *Journal of Clinical Psychology*, 61(2): 175-185.
- Healy, J. (2005). Establishing reliability and validity for the Adult Provider A Functioning Assessment Scale. Stark County Community Mental Health Board, Unpublished Manuscripts.
- Jacobson, N. S., Follette, W. C. & Revenstorf, D. (1984). Psychotherapy outcome research: methods for reporting variability and evaluating clinical significance. *Behavior Therapy* 15: 336-352.
- 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(1): 12-19
- Lunnen, K. M. & Ogles, B. M (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychotherapy*, 66 (2): 400-410.
- McGlinchey, J. B., Atkins, D. C., & Jacobson N. S. (2002). Clinical significance methods: Which one to use and how useful are they? *Behavior Therapy*, 33: 529-550.
- Ogles, B. M., & Healy, J. (2003). Calibrating the agency worker rated Ohio Scales with the Child and Adolescent Functional Assessment Scale. In Roth, D. and Lutz, W. (Eds.) *New Research in Mental Health*, (16). 249-255. Columbus, OH: Ohio Department of Mental Health.
- Ogles, B. M., Melendez, G., Davis, D.C. & Lunnen K.M. (2000). The Ohio youth problem, functioning and satisfaction scales technical manual. Columbus, OH: Ohio Department of Mental Health, Office of Program Evaluation and Research. (<http://www.mh.state.oh.us/oper/outcomes/instruments/os.technical.manual.pdf>)
- Ogles, B.M., Melendez, G., Davis, D.C. & Lunnen K.M. (1999). The Ohio Youth Problems, Functioning and Satisfaction Scales (Short Form) User's Manual. Columbus, OH: Ohio Department of Mental Health, Office of Program Evaluation and Research. (<http://www.mh.state.oh.us/oper/outcomes/instruments/os.user.manual.pdf>)
- Wise, E. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendation for future directions. *Journal of Personality Assessment*, 82 (1) 50-59.

Appendix: Standard Deviation and Reliability Coefficient of Outcomes Instruments

Instrument	Scales	Range	N	Mean	Standard Deviation	Cronbach Alpha	Reliable Change
Adult Consumer A (ACA)	Symptom Distress	15-75	104,335	38.50	14.93	.93	11.0
	Empowerment	1-4	96,508	2.66	.34	.84	0.4
	Overall Quality of Life	1-5	102,023	2.98	.79	.87	0.8
	Financial QOL	1-5	100,479	2.23	1.10	.89	1.0
Adult Provider A (APA)	Community Functioning Ability	11-55					4.0 ^a
Adult Consumer B (ACB)	Symptom Distress	15-75	97,986	39.39	15.23	.94	11.0
	Overall Quality of Life	1-5	95,236	2.94	.77	.87	0.8
	Financial QOL	1-5	93,358	2.09	1.03	.88	1.0
Ohio Scales - Agency Worker	Functioning	0 -80					8.0 ^b
	Problem Severity	0-100					10.0 ^b
Ohio Scales - Parent	Functioning	0 -80					8.0 ^b
	Problem Severity	0-100					10.0 ^b
	Satisfaction	4-24	67,644	9.99	4.97	.81	6.0
	Hopefulness	4-24	110,229	10.33	4.18	.76	6.0
Ohio Scales - Youth	Functioning	0 -80					8.0 ^b
	Problem Severity	0-100					10.0 ^b
	Satisfaction	4-24	46,111	8.16	4.64	.88	6.0
	Hopefulness	4-24	64,882	11.93	4.41	.77	5.0

^a Findings adopted from Healy (2005).

^b Findings adopted from Ogles, Melendez, Davis, & Lunnen (1999).