Midway through 2017, a well-known national publication announced a search for innovative solutions to three global problems (including global health problems). The application said “Tell us about the problem you are inspired to solve.” Here’s what I wrote:

Children with mental illness symptoms need practice using new, mature problem-solving strategies and tactics. Seeing a counselor once or twice a week in their office will never be enough practice for the vast majority of children! Parents and teachers don’t have the skill, patience, time or ability to deliver all of the practice opportunities that children with disabilities (Autism, ADHD) require to learn new skills successfully. Professionals who are trained and supervised can do this.

Then, they said “Provide a clear sense of how your idea will solve the problem identified.” Here’s how I responded:

A single licensed mental health professional can train and closely supervise 10 or more Masters-level mental health professionals who in turn each train and supervise 9 or more Bachelors-level professionals who deliver behavioral support and therapeutic guidance (opportunities to practice new skills) in the child’s own home, school and community. No trips to an office, ever. The services are entirely “evidence based” treatment procedures with decades of success behind them. And they’re free.

Finally, they said “Share what inspired you to create your idea.” Here’s what I said:

I’m a licensed psychologist and a certified school psychologist in Pennsylvania with 40 years of experience. My staff and I have been delivering treatment to children using the model I created and described above, with funding for 20, 30 or more hours of intensive, individualized treatment weekly per child. We’ve treated hundreds and hundreds of children successfully with this model.

They wanted a one minute video with the application. I made the mistake of including the logo of The Institute for Behavior Change in the video - and that was enough to disqualify my entry....

Over the past 35 years, the staff of the Institute for Behavior Change and I have been perfecting the treatment model I created when I was a newly licensed psychologist in 1981. That treatment model, which I call Effective Treatment in a Wraparound Cup™ has been examined by independent researchers from four different educational institutions since 2007. They all agree that the results are remarkable and worthy of publication. With well over 1,000 treatment plans studied, they reported that the probability that the positive changes they found in the lives of the children we treated had occurred due to chance was less than one in ten thousand. That’s called a powerful and “statistically significant” finding in the professional literature, but not everybody respects findings like that, so I continue to call attention to the good work we’re doing to the best of my ability.

The measurement of treatment outcome is crucial to the development and implementation of a treatment program in any discipline. Unless treatment delivery personnel know where they’re hoping to go, they can’t evaluate the course of treatment and can stray from the path that they intended to take. Some mental health treatment delivery personnel manage this challenge by...
not describing their intended destination explicitly. By providing only vague statements about “creating outcomes of social significance,” they give themselves as much freedom to claim success as possible, no matter what happens to the child being treated. Obviously, it is important for professional treatment providers to know their client’s pre-treatment experiential milieu before they start implementing any treatment program; this is called “baseline” recording. Once a baseline has been recorded, a professional mental health treatment provider can design a treatment plan that makes sense on paper and can be described to the people who will fund it, and can be implemented with high fidelity by the people who will deliver it. Without a written treatment plan, treatment delivery personnel are simply taking advantage of “regression toward the mean” (things will get better over time if they’re pretty awful now) or other functions that have nothing to do with the treatment plan or how well it was delivered.

The treatment plan has to be developmental. It’s not appropriate to start with “D” then go to “F” and then to “B” in an alphabetic continuum of options. The progression of treatment has to make sense given the developmental characteristics of the child who is receiving treatment. Children are different, but not so much so that just one (1) and only one (1) treatment plan alone will be helpful to them. As Harry Stack Sullivan said in 1947 “We are all simply more human than otherwise.” Therefore, a few basic components are essential, no matter who is being treated or whatever the condition that is being addressed.

1. The treatment plan has to be written down so that it can be implemented with fidelity.
2. The treatment providers have to be trained so they can implement the plan with fidelity.
3. The treatment outcome must be measured throughout the treatment delivery because
4. The client will respond to the treatment, positively or negatively, on an ongoing basis, so
5. The treatment plan has to be amended when it starts to derail (which is inevitable).
6. The recipient of treatment has to be involved in the outcome measurement, because only they can determine the quality of life that the treatment plan has created.
7. The providers of treatment should also be measuring treatment outcome on an ongoing basis so that both treatment provider and recipient agree on the outcome being created.

The Kossor Scale for Treatment Outcome Measurement was developed between 1996 and 2017 to achieve these seven objectives. By incorporating the input of both treatment providers and treatment recipients (or their parents if the recipients are young children), standardized measures of both the frequency and the severity of identified target behavior were created. Aberrant “target behavior” was operationally defined in binary terms (so that it could be rated as present or absent), which translated directly into a comprehensible treatment plan.

Properly trained professionals who were closely supervised by other adequately trained, resourceful professionals implemented the plan with fidelity. This allowed the treatment providers and the parents of the treatment recipients to record whether a “target behavior” occurred with nearly 100% accuracy.
Replacement (good) behavior emerged with increasing frequency and intensity as the aberrant behavior diminished, and was measured by other scales (including norm-referenced measures) in conjunction with the Kossor Scale, which is a criterion-referenced measure. Such measures are commonly used in school classrooms to measure student progress in learning a curriculum that is defined by an individual teacher for his/her current students.

The Kossor Scale has been described as “an annotated Likert (LIK-ert) scale,” but this is inaccurate because a Likert scale uses a continuum that is generically defined, usually having five points, and where the characteristics of one point are only generally related to those of adjacent points.

For example, an opinion poll is often designed as a Likert scale:

```
[ ]---------[ ]---------[ ]---------[ ]---------[ ]
Strongly Agree Neutral Strongly Disagree
```

In comparison, the Kossor Scale has the following characteristics:

- 10 points define the continua for each of two companion scales, scored simultaneously:
  - frequency
  - severity
- Each point is explicitly defined and is explicitly related to the adjacent points
- Each of the points describes a set of conditions unique to that point on the continua
- Aberrant “target behavior” is operationally defined in objective, binary terms, either:
  - The aberrant behavior did occur
  - The aberrant behavior did not occur

By following these principles, recording errors are minimized and the treatment plan itself can be written in language that directs treatment providers to respond to particular events in particular ways. As the results of their interventions are recorded, the trajectory of the treatment program is defined. Necessary adjustments can promptly be made to correct the trajectory so that aberrant behavior reduces while adaptive “replacement behavior” increases.

Like all Criterion Referenced measures, the Kossor Scale measures the life experience of a particular individual whose behavior is being monitored within the context of a particular experiential milieu. Norm Referenced measures compare the performance of the monitored person with the performance of other “similarly situated” persons from a “normative group” – none of whom is experiencing the exact same particular experiential milieu. Norm-referenced measures have enjoyed wide recognition within the behavioral literature, but criterion-referenced measures provide a perspective that norm-referenced measures simply cannot.

It is helpful to have both Norm Referenced measures (CARS-2, BASC-3, Vineland-3, CANS, ADOS, ATEC, etc.) in addition to at least one Criterion Referenced measure in the evaluation of a child’s treatment program if the aim is to evaluate the quality of life that the treatment produces. It is clear that norm-referenced measures are not sensitive to relatively small changes over relatively short periods of time.
However, criterion-referenced measures can be exquisitely sensitive to such changes if the target behavior is defined in sufficiently objective (i.e., binary) terms. The Kossor Scale for treatment outcome measurement can be applied with ease to treatment of any sort (podiatry, for example) where a measure of quality-of-life is sought. It excels as a parental assessment of progress for a child’s behavioral treatment program.

The treatment provider has to construct a treatment plan with statements operationally defining aberrant target behavior in binary terms (“kicking others” for example). The supervisor of the treatment program has to interview the treatment recipient (or the parent if the child is very young) once weekly as to their impression of the frequency and severity of the operationally defined target behavior. By recording the ratings of the treatment recipient (or his/her single designated representative) on a set schedule, a Single Subject with Repeated Measures experimental design is implemented.

When the supervisor of the treatment program interacts with the rating provider during the data collection process (comparing previous ratings with current ratings), both the reliability and the validity of the Kossor Scale can approach 100%. Moreover, in the course of implementing treatment outcome measurement with the Kossor Scale, the perspectives of the treatment recipient and the treatment provider necessarily move closer together and result in a more consistent and satisfactory treatment delivery process. Here is an example of the Kossor Scale used to evaluate the treatment outcome of a behavioral treatment plan for a child:

The Kossor Scale for treatment outcome measurement – child behavior Frequency rating

*Frequency Ratings: IN THE PAST WEEK* the target aberrant behavior occurred ...

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1</td>
<td>Nonexistent (0) or virtually nonexistent - less than once in a month and not worrisome (1)</td>
</tr>
<tr>
<td>2</td>
<td>Upper limit of normal (tolerable) for a child of the same approximate age</td>
</tr>
<tr>
<td>3-4</td>
<td>A few times in the past week and almost every week in the past month (3) to every single week (4)</td>
</tr>
<tr>
<td>5-6</td>
<td>Many times in the past week and almost every week in the past month (5) to every single week (6)</td>
</tr>
<tr>
<td>7-8</td>
<td>Several times per day (with breaks in-between incidents), almost every day (7) to every single day (8)</td>
</tr>
<tr>
<td>9-10</td>
<td>Constantly (never stops, or stops briefly before restarting), almost every day (9) to every single day (10)</td>
</tr>
</tbody>
</table>

The preceding scale can be adapted to address the behavior of an adult. It can also be adapted to measure the display of any aberrant behavior or symptom presentation whatsoever. The treatment provider simply needs to construct a continuum of frequencies between “nonexistent” and “constantly” and use clearly demarcated time frames for each of the points in-between.

A score of 0 or 1 defines the frequency of a functionally absent component. It is either not happening at all (0), or is so rare (1) that it is functionally absent from the life experience of the person whose treatment outcome is being measured. Likewise, in the Severity scale to be described shortly, a score of 0 or 1 defines a severity that has either no effect whatsoever (0), or has an effect so minimal that it is functionally ineffectual (1) in the life experience of the person whose treatment outcome is being measured. Each score on the 1-10 continuum of the Kossor Scale has a corresponding description that allows the rater to identify that score, to the exclusion of all others, as the single best score to match his or her perception of the frequency and severity of the behavior being assessed.
It is not possible for a score on the Kossor Scale to be assigned to more than one instance of a given behavior. This is true because the treatment outcome “target behavior” is defined in binary terms and neither they, nor the rating scales used to measure them overlap; they are all mutually exclusive. The rater also responds to questions about the severity of the behavior.

The Kossor Scale for treatment outcome measurement – child behavior Severity rating

Severity Ratings: IN THE PAST WEEK the target aberrant behavior was ......

0-1 The behavior is not worrisome. There are no (0) or almost no (1) negative consequences imaginable.

2 Upper limit of normal (tolerable) for a child of the same approximate age. Typically developing children of the same approximate age who you know are displaying behavior at this level and it doesn’t warrant professional attention. It is manageable, but may be very frustrating at times.

3-4 The child’s behavior is severe enough to worry you, almost every time (3) to every single time (4). Behavior isn’t responding to the “usual” interventions that work with typically developing children of the same approximate age who you know. Behavior requires outside help to manage or change. It is not obviously hurting the child or others, but if it persists, it will probably result in someone getting hurt, or the child is failing to be successful in meeting life’s challenges.

5-6 You characterize the child’s behavior as very serious, almost every time (5) to every single time (6). Outside help has been involved in the past and hasn’t helped, or isn’t consistently helping now, but there are more “good days” than “bad days” overall. Behavior is hurting the child or others when it happens, but it requires no more than “first aid” and no professional medical attention. If it persists or gets worse, it will certainly result in a need for professional medical attention.

7-8 You characterize the child’s behavior as alarmingly serious, almost every time (7) to every single time (8). There are relatively more “bad days” than “good days” overall. Behavior results in a need for professional medical attention above the level of “first aid” (including an Emergency Room visit to treat a relatively minor injury without hospital admission), no suicidal or homicidal intent is apparent and the child was never missing from an assigned area.

9-10 You characterize the child’s behavior as catastrophic because it was, or is potentially, life-threatening. A rating of 9 would mean that nobody died and the child’s behavior did not appear to be intended as a suicidal or homicidal act, or that an Emergency Room visit was necessary to treat a serious injury but did not result in a hospital admission (of the child or anyone else), or that the child was missing from an assigned area for up to 15 minutes. Special Incident Report required. Medicaid Managed Care Organization (MMCO) or other insurance company notification may be necessary.

A rating of 10 would mean that a successful or thwarted suicidal or homicidal act has occurred, or that an Emergency Room consultation resulted in a hospital admission (of the child or anyone else), or that the child was in some other life-threatening situation, that the child was missing from an assigned area for more than 15 minutes (over 4 hours requires MMCO notification). Special Incident Report required.

A score of 2 in this example is defined as “the upper limit of normal (tolerable) for a child of the same age.” Thus, biting objects might be rated “2” for a one-year old child who is teething, whereas a “3” or higher would be appropriate for a five year-old who is doing the exact, same thing.

The severity ratings are divided into 10 discrete intervals that do not overlap. The Kossor Scale uses a Behavior Record Form to collect these data so that they can be viewed easily and recorded efficiently, rather like a section of motion picture film recording a series of images.
The Kossor Scale for Treatment Outcome Measurement

Each week, the frequency and severity ratings are recorded on the Form and an average is computed for each target behavior domain. Target behavior is operationally defined at the top of the form so that the rater and the treatment provider are both exactly aware of the behavior that is being measured (not just the domain within which the target behavior exists).

The treatment provider describes the target behavior explicitly in a binary fashion in the Operational Definition. Physical aggression might be operationally defined as “hitting, kicking and spitting” and these would be understood to refer to acts directed against a person (not an object). Spitting on the floor would not constitute “physical aggression” but spitting on (or at) a person would. Defining target behavior in this explicit way makes it possible for treatment providers to focus conscientiously on improving treatment delivery and to apply treatment outcome measurements consistently. These are two essential components of Applied Behavior Analysis (ABA).

Applied Behavior Analysis in action – Effective Treatment in a Wraparound Cup®

The staff I supervise at The Institute for Behavior Change have been delivering Behavioral Health Rehabilitation Services (BHRS, still mistakenly called "wraparound services" in Pennsylvania) since 1997 to children under the age of 21 who are disabled and enrolled in Medicaid. The BHRS model I developed infuses “Full Fidelity Wraparound” methodology with Applied Behavior Analysis (ABA) procedures to deliver Effective Treatment in a Wraparound Cup® in homes, schools and other community settings under the scope of practice of licensed professional psychologists. At least one parent (or guardian) must be actively involved in the planning and delivery of our BHRS program.

The staff take outcome data every week from parents and use those data to improve the quality of the treatment process on an ongoing basis, in accordance with the evidence-based standards of ABA and Wraparound practice. A written Treatment Plan that describes and controls the child’s treatment program is always developed with input from the child, parent(s), teacher(s) and other adults who have roles in the child’s life. The child’s strengths, weaknesses, and treatment needs are reviewed on an ongoing basis by a Masters-level Behavior Specialist who consults with parents (and others, if necessary) at least once weekly to record data about the child’s progress.

A Bachelors level Therapeutic Staff Support (TSS) provider is usually assigned to work directly with the child to implement the child’s treatment plan on an intensive, one-to-one basis for several hours each week. A Mobile Therapist may meet with the child at home, in school, or elsewhere in the community to provide psychological counseling on one or more occasions each week. A licensed psychologist assumes full and complete professional responsibility for all services provided.

Because the Effective Treatment in a Wraparound Cup® model of BHRS I developed combines Full Fidelity Wraparound methodology with ABA principles and practices, it creates the most effective mental health treatment delivery modality possible for children in their homes, schools and communities. Independent researchers at UNC - Chapel Hill (2007) and Thomas Jefferson University (2010) found statistically significant associations between the delivery of BHRS by our staff and reductions in physical aggression, lack of environmental safety, noncompliance with
adult prompts, communication deficits and socialization deficits in more than 1,000 treatment plans for children between 2 and 19 years of age.

Other researchers from Villanova University (2012) and Immaculata University (2013) found comparable evidence, summarized by the truism “You can’t drive a tack with a toothpick.” The data from the Behavior Record Form is projected onto the Behavior Record Graph shown below.

When these data are plotted correctly, and annotated by the treatment provider, a clear understanding of the interaction between the child and the treatment program emerges.

**Titration Planning**

Funding agencies always seek to reduce or eliminate treatment services, as if that is the reason for their existence. One of their favorite tactics to reduce or eliminate treatment funding is the “titration plan.” This is usually marketed as a scheme where a soothsayer with a license to practice a profession channels Nostradamus and makes predictions (without any evidence whatsoever to support their prognostications), about the future needs of the patient. In the children’s mental health treatment business, the soothsayer often divines that the child will need 25% fewer hours of treatment “to reduce dependency on the staff” so that, by the end of the funding authorization period, the funding agency’s expenses are cut by 25% or more. The overarching goal is to spread treatment funding as far and wide as possible, albeit an inch deep. Certificates of Excellence and commemorative plaques are awarded to politicians for such service to the community.
The staff of the Institute for Behavior Change approach titration very differently. While they are reducing or eliminating unnecessary dependency on service providers, the method they use allows the child to receive precisely the amount of intensive, individualized treatment that he or she requires, as their troublesome behavior waxes and wanes, until the treatment plan is completed successfully. This method brings a healthy dose of reality into the treatment planning process.

When the child has the inevitable experience of frustration that exceeds his or her capacity for tolerance, and regression begins to occur, the treatment provider can judiciously come to the child’s aid because the provider is still part of the child’s life. An ethical treatment provider can prevent the catastrophe that occurs when the soothsayer’s “titration” plan is followed and the child’s treatment hours are cut by the magical ratio of 25%, then 50%, then 75% over the span of a few future months.

A TITRATION PLAN THAT MAKES SENSE

The Titation Plan is to "titrate" (reduce the intensity of) 1:1 intensive, individualized treatment (TSS support) provided to [child] as conscientiously as possible throughout the treatment period without depriving [child] of necessary support, especially in emergencies. The TSS provider will seek to decrease the level of prompting (move from direct and physical prompts to indirect and verbal or symbolic prompts) and will seek to increase the physical distance separating him/her from [child].

This will eliminate excessive dependency on the TSS provider and will assure that the TSS provider is available to [child] to deliver the least-restrictive and least-intrusive level of intervention necessary at any given time in the treatment period in accordance with the various “Best Practice” guidelines favored by auditors of behavioral treatment programs. The level of prompting typically used with the child, and the TSS provider’s typical distance from the child are recorded on the TSS Progress Note to facilitate evaluation of the progress and success of a Titration Plan that Makes Sense.

The treatment provider works 20 or more hours per week with the child while maintaining as much distance from the child as possible while delivering the treatment prompts, cues and other interventions necessary for learning and practice of new skills to take place. When the treatment provider has consistently been at maximum distance from the child for 2-4 weeks, it can safely be concluded that the termination of the treatment plan is justified. Then, the treatment provider begins to work with another child who needs 20 or more hours of treatment, and so on, and so on.

The Kossor Scale is measuring aberrant “target” behavior that is explicitly defined operationally, not "improvement” which is much more difficult to operationally define and is therefore more prone to error. It is much easier to define aberrant target behavior in binary terms – it is either present or absent. Accordingly, its measurement is much less prone to error.

By assessing treatment outcome in terms of the reduction of aberrant target behavior, it is much harder to declare a treatment program “successful” if it is definitely not successful by any reasonable standard of quality of life.
The Kossor Scale makes no pretention of objectivity; it is demonstrably objective. The items do not overlap. They are explicitly defined by the examiner in dialogue with the rater each time ratings are obtained. The possibility of interpretive error is eliminated in the Kossor Scale, as long as the rater and the examiner engage in a dialogue in the process of data collection.

Responses to the Kossor Scale rarely show a normal or quasi-normal distribution because the scale is measuring aberrant behavior. Aberrant behavior never occurs in a normal distribution; it peaks during certain conditions and ebbs during other conditions. Traditional behavioral theorists say that behavior is maintained or suppressed by the consequences that follow it.

By tracking the incidence and intensity (frequency and severity) of aberrant behavior that has been explicitly defined in binary terms, it is possible to track the occurrence of that behavior with a high degree of accuracy. When aberrant behavior ebbs, adaptive behavior can emerge (and it will do so, if it is rewarded – reinforced – in accordance with the basic principles of behavior analysis). Positive responses increase the probability that the behavior preceding them occurs.

Responses on the Kossor Scale may be summed because the scale is an interval scale, so the central limit theorem allows treatment of the data as interval data measuring a latent variable. Because the summed or average responses fulfill these assumptions, parametric statistical tests such as the analysis of variance can be applied. The literature commonly cites a minimum of 4 and preferably 8 items in the sum. Scores from at least 10 (and up to 13) items from the Kossor Scale for over 1,000 treatment programs were analyzed in studies of treatment outcomes conducted by independent researchers from four different educational institutions since 2007.

Non-parametric tests such as chi-squared test, Mann–Whitney test, Wilcoxon signed-rank test, or Kruskal–Wallis test are often used in the analysis of Likert scale data. These same tests can be applied to Kossor Scale data.

Frequency scores and Severity scores can be studied separately, or the average ratings in a baseline period can be compared with the average of ratings in the final four weeks of any treatment period. As with Likert scale data, Kossor Scale data can be analyzed via Consensus Based Assessment (CBA) to create an objective standard in domains where no generally accepted or objective standard exists. Consensus based assessment can be used to refine or validate generally accepted standards. In addition, Hierarchical Linear Modeling (HLM) has been used with a large database of Kossor Scale data and yielded highly significant (p<.001) findings of a powerful association between treatment exposure and aberrant behavior reduction. That’s an error estimate of less than one in ten thousand. The Kossor Scale might be considered an evolutionary enhancement of the Likert scale.

CONTACT

Steven Kossor
Licensed Psychologist
Certified School Psychologist
Founder & Executive Director of the Institute for Behavior Change
Founder & Director of the Network for Behavior Change
sakossor@ibc-pa.org
REFERENCES


The Kossor Scale for Treatment Outcome Measurement

140: 1–55.


