

Protocol for assessing Investment Risk – with regard to the likelihood that an organization is producing Social Value

**Version SocServ 1.0 (for
social/human service agencies)**

For use by External Evaluators

**8th Revision
Revised July 20, 2009**

**David E. K. Hunter, Ph.D.
HUNTER CONSULTING, LLC
Hamden, CT USA**

**Steve Butz
Social Solutions, Inc.
Baltimore, MD USA**

Introduction

This tool is designed to enable external evaluators *trained in its use* – with explicit attention paid to achieving and assuring inter-rater reliability – to assess the risk that a **nonprofit social or human services agency**¹ is not able to deliver measurable social value². This should provide meaningful information to nonprofit funders who care about whether the money that they bring to a given nonprofit is likely to be of benefit to society.

The information developed through this assessment is informed by the approach to understanding social value described in our 2008 white paper *Guide to Effective Social Investing*.³

This assessment looks at qualitative and quantitative data to assign a level of risk to the funding of a given nonprofit organization. By **high risk investment** we mean that funding a nonprofit, as it exists at the moment, is not likely to produce social value. We recognize fully that such high risk, developmental investing is essential to the nonprofit sector and to our society, inasmuch as it stimulates and supports innovative efforts to grapple with intractable societal challenges and provides the resources necessary to launch start-up projects or build up young or relatively small or weak organizations that seek to address them. **We believe that high risk investing brings with it high levels of investor responsibility** – that it is incumbent on the investor to become involved with the organization and bring nonfinancial resources to it (technical assistance, access to other funders, etc.) in order to help it develop organizational depth, sustainability, effectiveness, and reliability. **Failure to do so will more often than not mean that the organization will remain at its current level of development and be unlikely to produce social value** – which at the very least means the investment is largely a waste of money and, for intended beneficiaries, the services they get will have high opportunity costs.

In contrast, a **low risk investment** generally will involve funding more mature organizations with highly developed performance management capacities. (Some people call this mezzanine investing.) Doing so is also essential to society and the nonprofit sector, for it strengthens the sector's ability to do effective work right now, addressing pressing problems where delay means continued human suffering or perpetuating structural social inequalities and the associated problems this engenders for marginalized or disadvantaged groups, families, and individuals. **Such investing amounts to purchasing social value, and requires little of the investor beyond making good selections.**

In Guide we identified three performance domains – tactical data use, strategic data use, and program value – that together define the social value of an agency's work. We further identified two indicators for each domain. These domains and indicators are:

¹ Organizations working in different domains would need to have a tool with significantly different questions.

² An agency's social value consists of the measurable improvement(s) in some aspect of the world that an organization produces. It is not the work that an organization does – it is the *results* of that work.

³ A copy in pdf format can be downloaded from the website of the *Working Group for Effective Social Investment* at: <http://www.alleffective.org/>.

1. **Tactical Data Use Domain**
 - a. Data Integrity Indicator
 - b. Outcomes Focus Indicator

2. **Strategic Data Use Domain**
 - a. Making Essential Adjustments Indicator
 - b. Relating Staff Efforts to Outcomes Indicator

3. **Program Value Domain**
 - a. Capacity to Deliver Program/Services with Fidelity Indicator
 - b. Program Impact Data Indicator

We believe that the six indicators that define the three performance domains constitute those aspects of an organization that should be looked at closely when a social investor is attempting to determine the likelihood that an organization is able to generate social value reliably and sustainably.

To assess these indicators we have developed a short but comprehensive set of simple and straightforward questions that, when answered and scored, produce ratings of organizations in each of the three performance domains. ***Organizations that rate high (in the red) in one or more domains are the riskiest investments.*** Of course, as we discussed above, that does not mean they should be avoided. But it does mean that a social investor should proceed with open eyes and the understanding that simply giving such organizations money most likely will not generate much in the way of social value. Such organizations will need considerable time and significant nonfinancial support – and a high degree of involvement from investors – to help them improve to the point where they are reliable sources of public good. These kinds of high risk social investments (with great opportunity costs) can be considered the equivalent to venture capital investments in start-ups. Examples include many beloved grass roots, community-based, neighborhood, and Faith-based organizations. And ***we propose that the intended results of such a high risk investment be the development of the organization itself***, not (in the foreseeable future) any outcomes it seeks for the people it serves. We call these ***social venture investments***.

Investments in organizations with domain ratings mostly in the medium range (yellow) are somewhat less risky, but most will require no less by way of vigilance, nonfinancial support and investor involvement if they are to mature into high social value institutions. Again, the opportunity costs are significant and should not be underestimated. But, then too, they offer investors the chance to roll up their sleeves and help those nonprofits that they find attractive do what is necessary to become high performing organizations – a real contribution to society when done successfully.

By now it will come as no surprise that ***we see only those organizations in the low risk range (green) in all three domains as “blue chip” social investment opportunities*** – where investors’ dollars are essentially purchasing proven social value and there is little need for further investor efforts. However, even with such organizations there is a need for periodic reassessment. We believe that by applying this method initially, and at least annually thereafter, investors can

usefully track the investment types (“blue chip,” to “social venture”) and risk levels across their social investment portfolios, plan the kind and degree of their involvement with the nonprofits in which they invest, and calculate how best to make future investments align with and enhance the social value that they care about.

But to be clear: ***An organization’s rating should be seen as something that captures only a snapshot of the organization at a single point in time.*** Organizations that are defined as “risky” investments using the rating tool on a given date may, in fact, be on an upward trajectory over a two or three year period that a social venture investor might reasonably expect to see realized. These organizations are perhaps the best long run “bets” of the social venture investor, provided they continue to improve and when appropriate undertake rigorous program evaluations. Of course, in the same line of thinking, a “risky” organization trending the wrong way through periodic use of the assessment tool should be avoided altogether – unless the investment is made very intentionally to arrest or reverse the organization’s decline (in which case there should be a time limit after which no further investments are made if no improvement is evident).

While we have designed this tool for use by external evaluators, in our view it also can and indeed should be used by the leaders of nonprofit social service agencies to assess their own organizations. Thoughtful, honest, judicious answers to the 26 question will result in a clear understanding of an agency’s strengths and weaknesses with regard to its capacity to deliver social value – and also provide its leaders with clearly identified targets for organizational development.

It is in that spirit that we offer this tool for public use...while hoping that at the same time social investors will look for external evaluators to use it to map the field of social service agencies.

Finally, we do not regard this tool as a finished product. While we have tested many of its questions with several hundred nonprofit agencies, it is very much a work in progress. And we hope that others will use it as a springboard for refining their thoughts about how best to assess the ability of social service agencies to deliver solid social value reliably, measurably, and sustainably.

*David Hunter
Steve Butz*

August 2009

Protocol for assessing Investment Risk with regard to the likelihood that an organization is producing Social Value

Nonprofit Investment Risk Assessment

For use by an external evaluator

Used with the permission of the authors

Date: _____

Nonprofit organization being assessed: _____

Location: _____

Evaluator / Analyst: _____

Instructions: *The external evaluator will base his or her ratings on data collected during a site visit to the organization being assessed. At a minimum, the ratings should be based on:*

- a review of the organization’s internal performance management (including internal measurement and evaluation) data
- a review of all external evaluation reports
- discussions with the executive director, chief operating officer or director of programs
- discussions with all program directors
- discussions with the development director (and/or anybody who is responsible for reporting data to funders and other external parties)
- discussions with front line staff – at least one person from each program

Summary of Assessment Findings⁴

Summative Domain Ratings	Investment Risk Level Low, Medium, High
Domain I: Tactical Data Use	
Domain II: Strategic Data use	
Domain III: Program Value	
Overall Organizational Rating	

⁴ See pages 22 – 28.

Rationale for Overall Investment Risk Rating⁵

⁵ Beyond that, we would suggest the following standard: A “low investment risk” rating can only be achieved if an organization has received a “low” rating in all 3 domains; a “high investment risk” rating is assigned if any of these domains has received a “high” rating. All other combinations of domain ratings should result in a “medium investment risk” rating.

Section I
Score all Items⁶

1. The organization measures and analyzes those outcomes for clients that it believes are produced by staff activities/services (including negative outcomes when these become apparent).

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following data were reviewed to make this assessment:

List the specific outcomes and indicators the organization measures and tracks.

⁶ Please Note: Scales vary in the direction of scores based on Items' contents.

2. The organization tracks service utilization and is able to identify utilization trends.

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following data were reviewed to make this assessment:

--

3. In the past year, the organization has made a significant change (restructuring, adding/eliminating program[s], program redesign, change[s] to service frequency/intensity) due to knowledge it gained from positive and negative performance data it collected and analyzed.⁷

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

--

⁷ Changes made for other reasons are not pertinent to this assessment.

4. Some of the data the organization collects are never or rarely reported on, and may be unnecessary, because they are not related to its mission.

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
4	3	2	1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

List the specific data sets you reviewed in relation to this question.

5. There are things that the organization believes it does very well, which it never collects data about or reports on (and therefore can't justify its beliefs regarding them).

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
4	3	2	1

The following interviews and associated findings were used to make this assessment:

Be specific in listing the things the organization believes it does very well, but about which it collects no data.

6. When it is time to generate annual reports or requests for information from funders or other key constituencies, the organization spends _____ amount of time⁸ just getting the data into a usable format.

In respondents' opinions, <i>in each case</i> data for such reports take at least a month or more to pull together	In respondents' opinions, <i>in each case</i> data for such reports can be pulled together in 2 weeks to a month	In respondents' opinions, <i>in each case</i> data for such reports can be pulled together in 1 - 2 weeks	In respondents' opinions, <i>in each case</i> data for such reports can be pulled together in less than 1 week
4	3	2	1

The following interviews and associated findings were used to make this assessment:

Indicate specific estimates by each person interviewed regarding how much time the organization requires to produce and respond to such requests.

⁸ This question is deliberately phrased in a subjective way – it rests on the assumption that organizations which feel burdened by having to spend what they consider to be a great amount of time to producing data outputs are not likely to have a system for generating data that is useful for day to day performance management.

7. The organization is opportunistic in pursuing revenues

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
4	3	2	1

The following interviews and associated findings were used to make this assessment:

What this means is that the organization has a history of, and an intention to continue, pursuing and accepting grants, contracts and other funds largely regardless of whether the activities they support are well aligned with its mission and competencies. Indeed, it will, if necessary, hire new staff to provide the services called for in new contracts/grants if it doesn't already have the capacity to do the work. Please discuss how your observations and discussions informed your rating.

8. The organization is strategic in pursuing revenues.

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following interviews and associated findings were used to make this assessment:

What this means is that the organization generally will not apply for, and will decline grants, contracts and other funds that in any way involve activities not aligned with its mission and competencies. Please discuss how your observations and discussions informed your rating.

9. Growth priority: the organization has dedicated _____% of its discretionary resources to grow in scale through replication, expansion of current program size, etc. over the coming 1 year.

Over 50%	35 – 50%	26 - 335%	No more than 25%
4	3	2	1

The following interviews and associated findings were used to make this assessment:

10. The organization has dedicated _____% of its discretionary resources to maintain the quality and fidelity of our program(s)

Over 50%	35 – 50%	26 - 335%	No more than 25%
1	2	3	4

The following interviews and associated findings were used to make this assessment:

11. The organization's performance management database records and tracks (check all that apply):

- a. client demographics and baseline data ___
- b. service utilization ___
- c. outcome data ___
- d. staff activities ___

4 items checked	3 items checked	2 items checked somewhat	0 or 1 item checked
1	3	3	4

The following data were reviewed to make this assessment:

12. The organization's front line staff members are held accountable in documented performance reviews for clients' progress toward achieving targeted outcomes.

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following interviews and associated findings were used to make this assessment:

This will require a review of personnel policies and procedures. If these are vague or not available in regard to this matter, respondents' assertions that staff are held accountable for client outcomes should be tempered with judicious skepticism. Please discuss what led to your final rating of this item.

13. The organization’s front-line staff and their supervisors use aggregated client data in scheduled and documented meetings to manage performance

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

This is essential to ensure fidelity to their service models or programs.

14. The organization systematically tracks client outcomes (positive and negative) until it knows whether participants have achieved benefits that are sustained no less than a full year after service (not just short term and intermediate outcomes).

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following data were reviewed to make this assessment:

Indicate the long term outcomes that are tracked in the organization’s data system, and/or are documented in external evaluation(s); list the indicators that are used, and the time period for which the organization tracks such data after participants have completed its program(s).⁹

⁹ This also is a way to look at how deeply the organization understands basic performance management concepts – such as, in this case, long term out comes.

15. Staff at all levels of the organization are held accountable for timely and accurate data entry.

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

This will require a review of personnel policies and procedures. If these are vague or not available in regard to this matter, respondents' assertions that staff are held accountable for integrity and timeliness of data entry should be tempered with judicious skepticism. Please discuss what led to your final rating of this item.

16. The organization routinely conducts data integrity audits and mandates corrective actions as indicated.

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

17. Supervision of front-line staff is grounded in reviews of the outcomes achieved by the clients they are serving.

Agree strongly	Agree somewhat	Disagree somewhat	Disagree strongly
1	2	3	4

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

This will require a review of personnel policies and procedures. If these are vague or not available in regard to this matter, respondents' assertions that staff supervision is grounded in client outcome data should be tempered with judicious skepticism. Please discuss what led to your final rating of this item.

Section II

Give the rating that most closely describes the organization

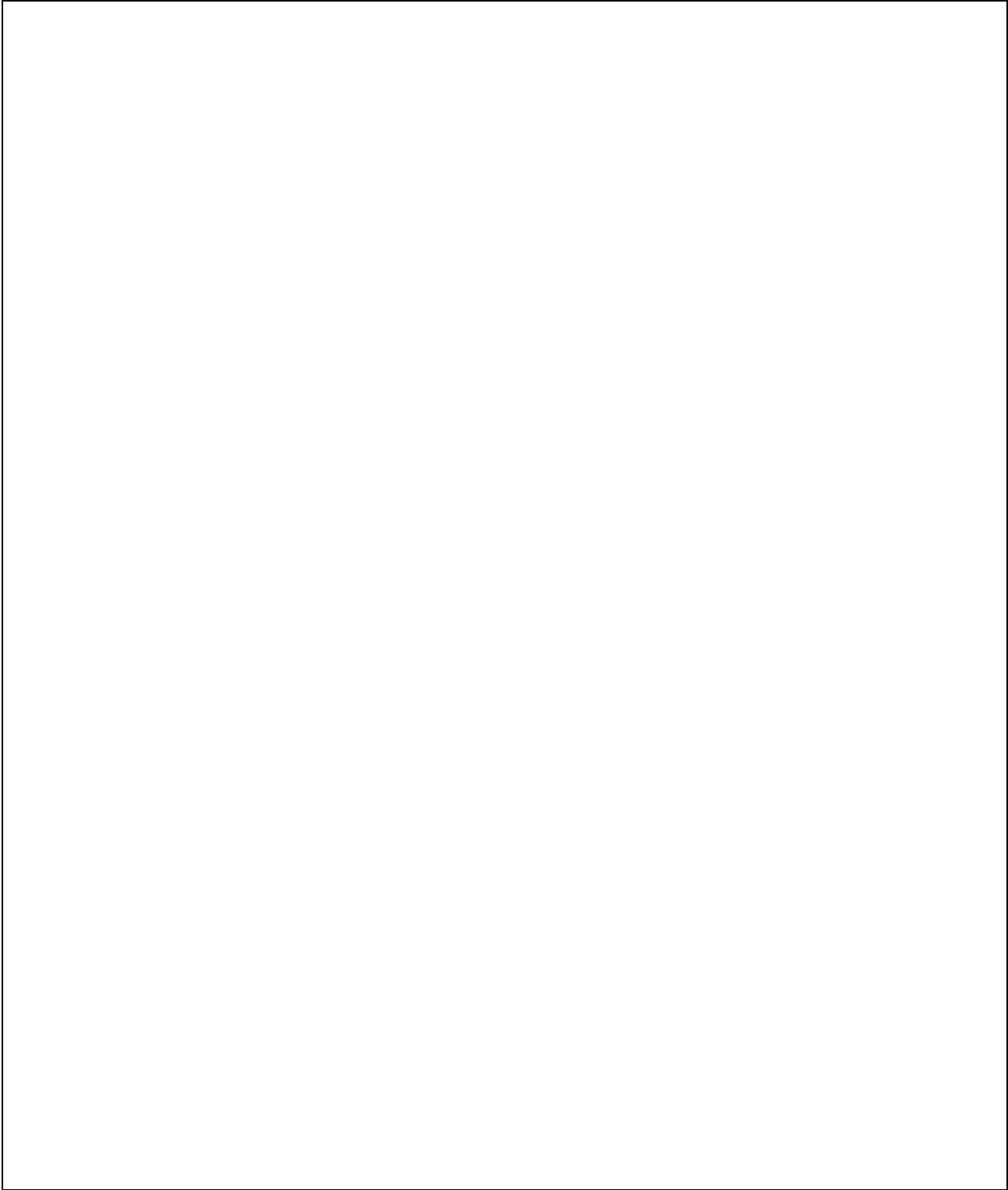
18. The likely positive¹⁰ impact of the organization’s program(s) has been:

Proven in at least one randomized control group study by an external evaluator at a 90-95% degree of confidence.	Demonstrated through internal performance data (outcomes) that have been analyzed by external evaluator(s) and compared to outcomes obtained for carefully matched groups of non-participants with at least a 90-95% degree of confidence.	Demonstrated through internal program data that have been benchmarked against similar outcomes for the general population (or key groups) through the use of public data sets.	Suggested by comparing program participant outcomes to duration and dosage of their program participation.	Suggested by tracking program participant outcomes.	Suggested by collecting anecdotes about program participants – or just assumed without collecting any pertinent information at all.
1 or 2 <i>use the higher number if findings were less statistically valid</i>	3 or 4 <i>use the higher number if findings were less statistically valid</i>	5 or 6 <i>use the higher number if findings were less statistically valid</i>	7	8	10

Please be sure to fill out the requests for supportive information on the following three pages.

¹⁰ **Negative impact data should, if they point to serious threats to program participants’ lives and prospects, immediately result in a rating of “high risk” investment...trumping all other considerations.** Where some negligible negative impacts are noted, these should be weighed against whatever positive impacts are established...and might result in a decision to rate the organization as a “medium level risk” investment, with the idea that with some hard work the negative impacts can be eliminated and the positive ones strengthened or amplified.

The following interviews and associated findings were used, and data were reviewed, to make this assessment:



Please list the outcomes or impacts were documented by the evaluation(s) you reviewed, and state your confidence in the data and findings, as well as your opinion of their significance (statistical significance, effect size) and social importance:

Please develop and discuss the rational for your finale rating of this question:

Section III

For each item, choose the answer which best reflects the organization

19. The organization's program(s) are evidence-based

- a. None are
- b. One out of several is
- c. Several but not all are
- d. All are¹¹

Score as follows: a = 4, b = 3, c = 2, d = 1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

¹¹ This may be one program for a single-service agency.

20. The organization has a theory of change featuring:

- a. A well delimited domain for action (e.g., a target population)
- b. One or more codified program(s) – with clear elements/activities, personnel competencies, implementation standards, performance standards
- c. Measurable outcomes for programs intended to produce outcomes

Score as follows: no elements are present = 4, one element present = 3,
two elements present = 2, three elements present = 1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

21. If the organization were asked to identify the percentage of clients with certain key demographics (e.g., gender, race, income level) who were enrolled in its program(s) last year, how quickly could it produce accurate data?

- a. There is no way to know (it has never reported on that or the information is missing/incorrect)
- b. In a week or more
- c. In a few days
- d. In one day or less

Score as follows: a = 4, b = 3, c = 2, d = 1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

22. If the organization were asked to assemble all of the previously recorded performance information (e.g., case notes, enrollment data, attendance/service utilization information, outcomes) on any given individual it has served in its program(s), it would be:
- a. Impossible – The organization doesn't know where all the information is, and some of it may have been lost or misplaced
 - b. Difficult – The information is in multiple locations and formats
 - c. Not Very Difficult – The information is all in the same place, but is in hard copy form only
 - d. Easy – the information is all in the same place and is in digital form

Score as follows: a = 4, b = 3, c = 2, d = 1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

23. The performance data collected by the organization can best be characterized as:

- a. There is no reason to believe their data are accurate
- b. Their data are possibly accurate (there seem to be efforts to promote data accuracy, but these are inconsistent or not very effective when compared to good practice benchmarks)
- c. Their data are likely to be accurate (there are clear policies and efforts to promote data accuracy, but these are still inadequate when compared to good practice benchmarks)
- d. Their data are certainly accurate

Score as follows: a = 4, b = 3, c = 2, d = 1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

24. How detailed are the participation/service utilization and progress reports for each program participant?

- a. The organization does not record this type of data on program participants
- b. The organization reports on enrollment
- c. The organization reports on enrollment and attendance at each event
- d. The organization reports on enrollment/attendance and time spent with each program participant
- e. The organization reports on enrollment, attendance, and personal changes (improved attitude, reduced risk of violence, etc.) for each program participant.

Score as follows: a = 5, b = 4, c = 3, d = 2, e = 1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

25. The organization measures client outcomes...

- a. Never or inconsistently
- b. At the end of a program (or when clients' exit from a program)
- c. At the beginning of a program (or when clients enroll in a program – i.e., at baseline) and at the end of the program (or when clients' exit from program)
- d. At specified intervals in addition to the beginning and the end
- e. At each client contact

Score as follows: a = 5, b = 4, c = 3, d = 2, e = 1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

26. To the extent that the organization does measure client outcomes, how easily can it aggregate the data?

- a. The organization doesn't know (it does not measure outcomes, report on outcomes, or the information is missing/incorrect)
- b. In a week or more
- c. In a few days
- d. In one day or less

Score as follows: a = 4, b = 3, c = 2, d = 1

The following interviews and associated findings were used, and data were reviewed, to make this assessment:

Indicator Scoring Tool
 Scoring System Developed by
 David E. K. Hunter, Ph.D.
 Hunter Consulting LLC

Domain I: Tactical Data Use

Indicator 1: Data Integrity

Pertinent Assessment Items	Evaluator's Assessment Score
#5	
#6	
#15	
#16	
#19	
#20	
#21	
#22	
#23	
Total Score	

Indicator 1: Data Integrity – Total Score Range	Indicator 1: Numeric risk rating
9 - 10	1
11 - 12	2
13 - 15	3
16 - 19	4
20 - 23	5
24 - 27	6
28 - 30	7
31 - 33	8
34 - 36	9
37 - 38	10

Indicator 2: Outcomes Focus

Pertinent Assessment Items	Evaluator's Assessment Score
#1	
#7	
#8	
#9	
#10	
#11	
#12	
#14	
#17	
#20	
#23	
#24	
Total Score	

Indicator 2: Data Integrity – Total Score Range	Indicator 2: Numeric risk rating
12 - 15	1
16 - 19	2
20 - 23	3
24 - 27	4
28 - 31	5
32 - 35	6
36 - 39	7
40 - 43	8
44 - 46	9
47 - 49	10

Domain II: Strategic Data Use

Indicator 3: Making Adjustments

Pertinent Assessment Items	Evaluator's Assessment Score
#3	
#7	
#8	
#9	
#10	
#13	
Total Score	

Indicator 3: Making Adjustments – Total Score Range	Indicator 3: Numeric risk rating
6 - 7	1
8 – 9	2
10 - 11	3
12 - 13	4
14 - 15	5
16 - 17	6
18 - 19	7
20 - 21	8
22 - 23	9
24	10

Indicator 4: Relating Staff Efforts to Outcomes

Pertinent Assessment Items	Evaluator's Assessment Score
#1	
#12	
#17	
Total Score	

Indicator 4: Relating Staff Efforts to Outcomes – Total Score Range	Indicator 4: Numeric risk rating
3	1
4	2
5	3
6	4
7	5
8	6
9	7
10	8
11	9
12	10

Domain III: Program Value

Indicator 5: Capacity to Deliver Program/Services with Fidelity

Pertinent Assessment Items	Evaluator's Assessment Score
#1	
#3	
#5	
#8	
#9	
#10	
#11	
#12	
#13	
#17	
#20	
#21	
#22	
Total Score	

Indicator 5: Capacity to Deliver Program/Services with Fidelity – Total Score Range	Indicator 5: Numeric risk rating
19 - 21	1
22 - 24	2
25 - 28	3
29 - 33	4
34 - 37	5
38 - 42	6
43 - 46	7
47 - 49	8
50 - 51	9
53 - 52	10

Indicator 6: Program Impact Data

Pertinent Assessment Items	Evaluator's Assessment Score
#18	
#19	
#20	
Total Score	

Indicator 6: Program impact data – Total Score Range	Indicator 6: Numeric risk rating
3	1
4	2
5	3
6	4
7	5
8 – 10	6
11 - 13	7
14 - 15	8
16 - 17	9
18	10

ASSESSMENT RESULTS

Indicator	Level of Investment Risk for Indicator Ratings: LOW = rating of 1 - 3 MEDIUM = rating of 4 – 7 HIGH = rating 8 - 10
Domain I: Tactical Data Use	Domain Summative Risk Rating (low, medium, high): ¹²
Indicator 1: Data Integrity	Indicator numeric risk rating:
Indicator 2: Outcomes Focus	Indicator numeric risk rating:
Domain II: Strategic Data Use	Domain Summative Risk Rating (low, medium, high)
Indicator 3: Making Essential Adjustments	Indicator numeric risk rating:
Indicator 4: Relating Staff Efforts to Outcomes	Indicator numeric risk rating::
Domain III: Program Value	Domain Summative risk rating (low, medium, high):
Indicator 5: Capacity to Deliver Program/ Services with Fidelity	Indicator numeric risk rating:
Indicator 6: Program Impact Data	Indicator numeric risk rating:

¹² Use the two Indicator Ratings of each domain to calculate a Domain Summative Risk Rating as follows: the combination of 2 Indicator scores in the low range = a Domain Summative Risk Rating of LOW; and the combination of 2 Indicator scores in the high range = a Domain Summative Risk Rating of HIGH; any other combination of 2 Indicator scores = a Domain Summative Risk Rating of MEDIUM.