

Constraints and Cautions in Formulating Recommendations to a Service, Especially in the Context of an External PASS or PASSING Evaluation

Wolf Wolfensberger and Susan Thomas

Abstract

PASSING assessments of human services by outside teams have as their primary purpose evaluation, not charting a course for the future of the evaluated service. Evaluation teams should be cautious in making recommendations to a service for several reasons, and recommendations that are made should usually be framed within a broad context, rather than specific to isolated ratings in the evaluation instrument.

One of the fruits of service evaluation, such as with the PASS (Wolfensberger & Glenn, 1973, 1975) or PASSING tools (Wolfensberger & Thomas, 1983, 1988) is the recommendations that evaluators are able to formulate to the assessed service. Once one gets to know a service and its strengths and weaknesses in an objective fashion, one usually has some well-reasoned and fact-supported ideas of how the service could be improved, and one may then formulate these ideas into discrete recommendations. However, there are a number of constraints to, or cautions about, making recommendations of which one needs to be aware, and which may not be brought out clearly in the literature that may be made available to the typical service evaluator. Below, we sketch six such constraints or cautions.

The Overriding Purpose of an External Evaluation is Evaluation, Not Generation of Alternatives

Teams of outsiders who are brought in to conduct an objective evaluation of a service should keep in mind that their first and overriding function and purpose is to evaluate a service, not to improve it. That latter task is first, foremost, and ultimately the responsibility of those parties who are in charge of the service, and in a position to actually make changes in it. Also, all sorts of changes can and should be made by a service without necessitating a formal evaluation, or a further evaluation if such had been conducted earlier. In fact, there are all sorts of tools and procedures other than service quality assessment devices that can serve the change process, or were even designed to do so.

We are most certainly not saying that evaluation of the current quality of a service is not useful or desirable in charting a future course for it, but that the first and overriding function of an evaluation is feedback on how things are, now. We will have more to say about this further below.

Recommendations Are Normatively on Riskier Validity Grounds Than Feedback on "What Is"

An instrument for service evaluation is usually structured around the determination of the current realities in the service. This is certainly the case with the aforementioned PASS and PASSING evaluation

instruments. This means that an excursion by an assessment team into any other domain automatically entails a heightened risk to the validity of that excursion. For instance, in evaluating a service, a team often inquires into the history of a service. But because the evaluation tool is not a historiographic one, anything that the team concludes about the history of a service has a lower chance of being valid than does anything the team concludes about what the reality of the service is now, at the time of the evaluation. Similarly, a tool designed to measure or record primarily "what is" does not provide an equally solid basis for determining what is needed in the future. Because recommendations deal with what might or should occur in the future, the risk that they might be invalid is higher than the validity risk of describing what quality features exist now, especially since the future always brings with it many things that cannot be known in advance, and that may totally invalidate any predictions or recommendations made about it.

The Validity Risk of Recommendations Increases When Evaluations Are Conducted Mostly for the Benefit of Evaluation Team Members

Considering that validity risk for recommendations is higher than validity risk for giving feedback on "what is," the validity risk for recommendations increases yet more--indeed, vastly more-- in a training evaluation, such as those that take place as part of training workshops in the use of the PASS or PASSING tools. In such evaluations, the team members are generally not selected to have relevant competencies needed to assess a particular type of service, nor are the team members (other than the team leader) very experienced with the assessment tool, nor do teams usually have as much time to get to know the service as they would in a real assessment by an experienced team. Rather, the training evaluation is conducted primarily so that the trainee team members can learn the evaluation tool and its application, as well as what constitutes good and poor quality of service in general. Thus, training teams are not in as good a position to make recommendations to a service as are "real" evaluation teams that are conducting an evaluation primarily for the benefit of the service being assessed, and that are constituted and structured accordingly.

When PASSING is the Only Evaluation Tool Used, The Evaluation Team May Not be in a Good Position to Go Beyond Programmatic Issues

Since unlike PASS and many other evaluation tools, PASSING looks only at programmatic issues (i.e., things that are done to address the needs of the people being served, and that affect how they are valued by others), teams using only PASSING in order to assess a service are hardly in a good position to make any recommendations other than programmatic ones. However, much service change may be required on the non-programmatic level, e.g., in terms of service administration or finance. For instance, PASSING evaluation teams often discover that a client grouping is maladaptive: it may be too big, too heterogeneous in terms of abilities and needs, there may be maladaptive intra-group modeling and imitation, etc. These are all programmatic shortcomings that interfere with the clients getting what they need. However, any such programmatic weaknesses may be created by other non-programmatic problems, such as regulations, funder requirements, lack of comprehensiveness of an entire service system, etc. For example, there may be so few other kinds of services locally available that a single service ends up taking in everyone in need. There may not be other local services that have greater or lesser structure that allow them to take in people who have greater or lesser degrees of need than the service being assessed can adequately address. And so on. The lack of needed service options, and the pressure to accept clients who do not need what the service is structured to offer, are non-programmatic issues. These may be very real, and clearly do impact on programmatic issues, but they are all issues which members of a PASSING evaluation team are not

necessarily trained to address, and are certainly not trained to assess as part of their training in the use of the PASSING instrument. And not only may such issues be beyond the scope of competency of a PASSING team, but they certainly go beyond the scope of the PASSING instrument.

Evaluation Teams Rarely Bear the Consequences When Faulty Recommendations are Implemented and Then Prove Faulty

Evaluation teams come and go while services tend to endure. Members of external evaluation teams may even come from far away, and none may be anywhere near by the time the recommendations they have made to a service get implemented. Thus, if it turns out that the recommendations were unsound or even harmful, it is rarely the team members who have to bear the burden of the consequences of their faulty judgments.

Additionally, evaluation team members--at least for PASS and PASSING assessments--work as a group, a team. And it is well-known that groups are willing to take greater risks than any one of their members might do as individuals. This is called the "risky shift" phenomenon, i.e., people in groups are generally prepared to do riskier things than when they function alone. This phenomenon can be good or bad, depending on the circumstances. As a result of this group dynamic, or because of limited sophistication by team members, a team may be willing to make recommendations that are on shaky ground, that may not be well thought-out, and that may prove very risky indeed to the assessed service and the people it serves. Again, neither the team as a whole, nor its individual members, have to bear the consequences--which may be outright disastrous--when the risky recommendations they have made backfire for the service and its recipients.

Recommendations Need to be Made in a Way That is Contextual

Evaluation teams often make lists of recommendations as if each one were freestanding, or as important as any other recommendation. This is particularly apt to happen if teams make recommendations for each, or almost each, separate rating issue in an evaluation instrument. However, several context issues need to be considered in making recommendations.

1. It should be made utterly clear what the team considers to be overarching, minor, and in-between recommendations.
2. Often, a recommendation would only be helpful if certain other recommendations were also implemented. For instance, changes in program content may only be truly helpful if the client grouping is also changed to be more coherent, so as to benefit more precisely from the new content. Thus, the relationship of recommendations to each other needs to be made clear.
3. Relatedly, it is not uncommon for recommendations to be mutually contradictory. For instance, in order to address one deficiency, a team may recommend something that is either opposite to, or incompatible with, what it recommends in response to another deficiency. Potential contributors to mutually incompatible recommendations are the following. (a) They are formulated piecemeal, and/or in response to single rating issues. (b) Many recommendations are made, and/or are scattered throughout the report, so that team members lose oversight of how the various recommendations would affect each other. (c) Lack of sophistication and mastery by evaluators can contribute to incompatible recommendations, and such shortcomings are not overcome by more training if the talent is lacking.

At any rate, it is important that teams review all specific potential recommendations, put them into larger context (e.g., how some depend on others), and either make them compatible with each other, or point out how one improvement would incur a cost somewhere else, i.e., what the trade-offs would be. Teams often do not do this, at least not in written evaluation reports to the service, particularly in instances where the recommendations are scattered throughout the report narrative.

If teams do not thoroughly review their recommendations, and either eliminate or reconcile incompatible ones, or explain to the service the costs of trying to implement one recommendation that is mutually incompatible with another, then the evaluation report, the entire evaluation team, the evaluation itself, and even the evaluation instrument may lose credibility in the readers' eyes. The same may happen if readers see recommendations that do not take into account the cost (in the broadest sense of that term) of implementation to the service, or recommendations that are outright absurd. Because readers will dismiss such recommendations, they may also dismiss the validity of the entire evaluation. However, if the report does acknowledge that some recommendations are apt to have a high cost to the assessed service, and perhaps even discusses this issue sympathetically, then a loss of credibility is less likely to occur.

Conclusion

Recommendations must be thought of as a way of converting the evaluation findings into potentially useful action strategies. But they are not on the same level of imperative for an evaluation team as the actual conduct of a good evaluation, nor can one feel as confident about the recommendations a team makes as about its findings. This is not to say that teams should not make recommendations; indeed, a service assessment typically is expected to generate recommendations, almost always does, and many such recommendations may be extremely useful. But teams should learn to be more conscious about the risks involved in making recommendations, and therefore also be more cautious, especially since some agencies may implement some recommendations without due consideration of what might be lost as well as gained thereby.

References

- Wolfensberger, W., & Glenn, L. (1973). *PASS (Program Analysis of Service Systems): A method for the quantitative evaluation of human services. Handbook & Field Manual* (2nd ed.). Toronto: National Institute on Mental Retardation.
- Wolfensberger, W., & Glenn, L. (1975). *PASS (Program Analysis of Service Systems): A method for the quantitative evaluation of human services. Handbook & Field Manual* (3rd ed.). Toronto: National Institute on Mental Retardation.
- Wolfensberger, W., & Thomas, S. (1983). *PASSING (Program Analysis of Service Systems' Implementation of Normalization Goals): Normalization criteria and ratings manual* (2nd ed.). Toronto: National Institute on Mental Retardation.
- Wolfensberger, W., & Thomas, S. (1989). *PASSING (Programme d'analyse des systèmes de services application des buts de la valorisation des rôles sociaux: Manuel des critères et des mesures de la valorisation des rôles sociaux* (2ième éd.). (M. Roberge, trans.; J. Pelletier, adap.). Toronto: l'Institut G. Allan Roeher & Les Communications Opell.

The proper citation for this article according to the American Psychological Association Publication Manual (3rd edition) is:

Wolfensberger, W. & Thomas, S. (1994). Constraints and Cautions in Formulating Recommendations to a Service, Especially in the Context of an External PASS or PASSING Evaluation. SRV-VRS: *The International Social Role Valorization Journal*, 1(2), 3-6.