



# Routine outcome measurement in mental health service consumers: Who should provide support for the self-assessments?



Marc Gelkopf<sup>a</sup>, Shira Pagorek-Eshel<sup>a,b,c,\*</sup>, Tom Trauer<sup>d,e,1</sup>, David Roe<sup>a</sup>

<sup>a</sup> Department of Community Mental Health, Faculty of Social Welfare and Health Sciences, University of Haifa, Mount Carmel, Haifa, Israel

<sup>b</sup> Department of Social Work, Zefat Academic College, Zefat, Israel

<sup>c</sup> Bob Shapell School of Social Work, Tel Aviv University, Tel Aviv, Israel

<sup>d</sup> St Vincent's Hospital, Melbourne, VIC, Australia

<sup>e</sup> The University of Melbourne, Melbourne, VIC, Australia

## ARTICLE INFO

### Article history:

Received 9 December 2013

Received in revised form 3 February 2015

Accepted 3 February 2015

Available online 20 February 2015

### Keywords:

Mental health assessments

Outcome measurement

Mental health services

Assessment biases

Internal vs. external assessment

## ABSTRACT

This study examined whether mental health community service users completed outcome self-reports differently when assessments were supervised by internal vs. external staff. The examination of potential differences between the two has useful implications for mental health systems that take upon themselves the challenge of Routine Outcome Measurement (ROM), as it might impact allocation of public resources and managed care program planning. 73 consumers completed the *Manchester Short Assessment of Quality of Life* (MANSA), a shortened version of the *Recovery Assessment Scale* (RAS), and a functioning questionnaire. Questionnaires were administered, once using support provided by internal staff and once using support provided by external professional staff, with a one-month time interval and in random order. A MANOVA Repeated Measures showed no differences in outcomes of quality of life and recovery between internal and external support. Functioning scores were higher for the internal support when the internal assessments were performed first. Overall, except for the differences in functioning assessment, outcome scores were not determined by the supporting agency. This might indicate that when measuring quality of life and recovery, different supporting methods can be used to gather outcome measures and internal staff might be a good default agency to do this. Differences found in functioning assessment are discussed.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

It has been well-established that mental health consumers should and can take part in Routine Outcome Measurement (ROM; Linhorst & Eckert, 2002; Oakley, Malins, Riste, & Allan, 2011; Trauer, 2010). Consumers' involvement in the evaluation process is consistent with several models of program evaluation and can enrich evaluations. Furthermore, the actual process of participation can empower consumers (Linhorst & Eckert, 2002).

Since taking part in ROM requires cognitive and attentional abilities, consumers might benefit from support, namely, assistance

tailored to their specific needs and requests (i.e., explaining the meaning of a certain word, reading the questions aloud, using the computer keyboard and mouse). However, it remains unclear whether the way in which the support is provided might affect the reporting of outcomes.

The use of internal service providers to aid data collection has several advantages. First, they usually know the clients and the context within which the evaluation is conducted (Fitzpatrick, Sanders, & Worthen, 2003; Love, 1991; Spaulding, 2008). Second, they might be sensitive to cognitive and attentional abilities and subtle cues regarding whether questions were understood. Third, internal supporters are likely to be cheaper and available, leading to better long-term implementation of ROM, thereby improving service provision (Australian Health Ministers, 1992).

Nonetheless, legitimate concerns can be raised as to whether providers, who might have a stake in the outcome of the evaluation, should be involved. Indeed, they might pressure consumers to answer questionnaires in a certain fashion, for

\* Corresponding author at: Department of Social Work, Zefat Academic College, Jerusalem 11 st., Zefat, Israel. Tel.: +972 77 4410468/+972 507250621.

E-mail addresses: [shirap@zefat.ac.il](mailto:shirap@zefat.ac.il), [shira.pagorekeshel@gmail.com](mailto:shira.pagorekeshel@gmail.com) (S. Pagorek-Eshel).

<sup>1</sup> Deceased 10 July 2013.

economic, administrative, or personal reasons (Mathison, 1999). Previous research has also reported that clients tended to overstate psychopathology and rate themselves worse in aspects such as well-being when they wished to gain eligibility for services. Conversely, they might rate their mental health as better when seeking discharge or to please their service providers (Bilsker & Goldner, 2002; Choi & Pak, 2005).

Although the literature has focused mainly on potential biases of internal service providers, bias might also occur with external support. For example, a client might use the opportunity to speak to an external administrator, sent by the “authorities,” to overstate grievances about service provision. Furthermore, the encounter between the client and an external supporter with whom the client is not familiar might cause suspicion and anxiety and therefore undermine the evaluation process (Linhorst & Eckert, 2002).

While ROM is increasingly becoming a requirement by planners of mental health services (Trauer, 2010), the examination of potential differences between internal and external support might have useful implications for service providers taking up the challenge of ROM.

Therefore, the purpose of this study was to test whether discrepancies existed in outcome data when support was procured by internal service providers vs. external professional assistance. As the literature suggests, bias might occur in cases of internal as well as external support. We hypothesized that no differences would be found between outcomes following internal and external support.

## 2. Methods

The present study was part of a pilot study of the Israeli National Outcome Measurement Project of Community Rehabilitation Services (CRS; Roe, Lachman, & Mueser et al., 2010). These services are provided by law to all people with a psychiatric disability and include housing, vocational, social, and educational services (Roe et al., 2010). Inclusion criteria were being adult and eligible for CRS following a diagnosis of severe mental illness, and at least a 40% psychiatric disability determined by a medical committee. Exclusion criteria, as determined by the client's case manager within the rehabilitation service, were being in an acute psychotic state, hospitalization (and thus absent from the premises), current violent behavior, lack of basic cognitive ability required to answer a questionnaire or to give informed consent.

Ethical approval was obtained from the Ministry of Health. Written informed consent was obtained from the study participants.

The self-report questionnaires consisted of three outcome measures, chosen, developed in cooperation with local practitioners, and repeatedly pilot tested over the course of 6 months with numerous groups of consumers ( $n = 220$  in 20 different services).

The chosen measures represent the key variables in which outcomes of the rehabilitation process are to be expected and are considered important by both providers and consumers of rehabilitation services (Fossey & Harvey, 2001). The measures have been widely used in previous outcome studies of adult mental health services (Royal College of Psychiatrists, 2010).

The Manchester Short Assessment of Quality of Life (MANSA; Priebe, Huxley, Knight, & Evans, 1999) is a 16-item instrument assessing quality of life focusing on satisfaction with life as a whole and within life domains (employment, vocational, social, familial, accommodation, leisure, financial situation, physical and mental health). For this study, we included the 12 items assessing subjective satisfaction ( $\alpha = .76$ ). We used a shortened 5-item version of the Recovery Assessment Scale (RAS; Corrigan, Salzer, & Ralph, 2004) that included items related to one's ability to cope

with life, hope, asking for help when needed, coping with mental health problems, feeling good about oneself ( $\alpha = .80$ ). These items were rated on a 5-point Likert scale. Finally, we included a 10-item functioning questionnaire, rated on a 4-point Likert scale developed specifically for this project (Roe & Gelkopf, 2012). Participants were asked to rate their functioning level in several domains such as personal hygiene, house chores, finances (e.g., paying bills on time), mobility (e.g., going from place to place independently), coping with daily problems (e.g., solving problems), medication intake, medical supervision, social life (e.g., participating in social activities), community participation, and work (or studies) ( $\alpha = .83$ ). All instruments were previously tested in an additional pilot project and showed satisfactory reliability and validity (Roe & Gelkopf, 2012). Software was used to fill out the questionnaire on computers.

Two methods of assistance for completing the questionnaire were evaluated: support by an internal service staff member and support by an external, trained individual, who was not personally acquainted with the consumer.

Each service participating in the study had up to three mental health case managers available to provide the internal support, depending on which of them knew the clients best. A pool of up to three external professionals for each service was available to provide the external support. All questionnaire administrators had a degree in social sciences or health sciences, or worked in the rehabilitation field. None of them had consumed rehabilitation mental health services. Both the internal and external teams had received basic training in questionnaire administration, although the external administrators were more experienced. Some administrators were men and some were women.

In both internal and external support provisions, the type of support provided was according to the client's request. For example, sometimes a client asked the meaning of a certain word, or asked for someone else to click the answers on the computer if the software was intimidating. Sometimes, a client wanted to discuss the interpretation of a question with a supporter (for further elaboration on types of support, see Section 2).

Both types of supporters underwent similar training, which included explaining the study to the client, learning the software used for completing the questionnaires, and being taught strategies of providing support while maintaining maximum consumer independence (Morrell-Bellai & Boydell, 1994). The supporters received a written manual with specific guidelines on how to deal with different situations that might occur during the evaluation process. For example, if the consumer wished to stop filling out the questionnaire, the supporter was guided in to encourage the client to continue (e.g., “you've done a great job so far and have only a few more questions to answer”). If the consumer was reluctant to proceed, the supporter was to thank him/her for cooperating, with an invitation to continue at another time; if the consumer did not understand a sentence, the supporter would repeat the sentence, rather than provide an answer, etc. The manual was built on basic principles of questionnaire administration, on our own experience of training during the pilot studies, and on our own and other researchers' studies on the rehabilitation of people with mental illness. The manual also included a unified “dictionary”, which included a bank of words and clear instructions on how to explain their meaning in case a consumer had difficulty understanding certain words (for example: “frequency” should be explained as “how many times”). The administrators were asked to fill in a written report, which included the type of support the consumers received and whether they encountered additional problems not addressed in the manual. The written reports further verified the manual's consistency with the consumers' and supporters' needs.

The study was conducted between May and August 2012 in 16 different services of various types in the Haifa region. To ensure randomization as well as participation of all modes of services, we listed all the services in the region and sorted them by type (housing, vocational, social, and educational) (Roe et al., 2010). From this list, 16 agencies for the housing, vocational, and social services were drawn out of a box, as well as the only local educational service.

Participants within agencies were sampled randomly (by coin tossing) from the list of consumers who had already met eligibility criteria until the quota of five individuals for each agency was reached. In seven services only, four individuals were interviewed due to their absence during one of the two interviewing slots. Consumers were told that the questionnaire included questions about different areas of their lives. They were asked to read the explanations before each question and to choose the answer that best described their experience. They were also informed about the number of questions in the questionnaire and the amount of time it should take to fill it out. Consumers were guided to try to answer all questions fully and honestly and were told that if any of the questions were unclear they were to feel free to ask for clarifications. These instructions were given both orally and in writing.

Questionnaires were completed twice by each consumer ( $n = 73$ ), once using internal support and once using external support with an interval of one month. Based on a coin-tossing procedure, we randomly divided the services so that for half the services, which made up half the sample, the internal support was provided first, and for the other half, external support was provided first. The administration of questionnaires was conducted at the consumer's service site. All services were applied 'as usual' during this one month period, and no additional services were provided.

### 3. Results

Of the 73 participants who filled out the questionnaire at the two time-points, 19 (26.0%) were women, 45 (63.40%) were single, 12 (16.90%) were married or in a long-term relationship, 14 (19.70%) were divorced or widowed, and all were Jewish. Fifty-one (71.80%) had up to 12 years of schooling, five (7.0%) had a full matriculation certificate, and 15 (21.10%) had more than 12 years of academic or professional education. Mean age was 37.81 (SD = 12.3).

Support included word clarification, technical aid, or verbal encouragement to complete the questionnaire (38.81%,  $n = 28$ ); filling out the answers on behalf of the consumer after the consumer read the questions (27.86%,  $n = 20$ ), or both reading the questions aloud as well as filling out the answers (23.88%,  $n = 18$ ). In 9.59% of cases ( $n = 7$ ), supporters remained available in the room, while consumers completed the questionnaires without requesting aid. The pattern of support that was provided to consumers at both time-points was identical.

A Repeated Measure MANOVA was conducted. The dependent variables were the quality of life, functioning, and recovery measures. The independent variables were support type – internal or external – and sequence of type of support during the procedure.

As can be seen in Table 1, the interaction group (internal/external)  $\times$  support sequence revealed no significant difference between internal and external support. However, the measures  $\times$  support type  $\times$  sequence (of the type of support) interaction was significant. To examine which measure was accountable for this result, a Repeated Measure ANOVA was conducted for each measure separately.

Analyses showed a difference only on the functioning measure ( $F = 14.36$ ,  $df = 1$  and  $70$ ,  $p < .001$ ; Wilks'  $\lambda = .83$ , partial  $\eta^2 = .17$ ). A second Repeated Measure ANOVA was conducted to examine

**Table 1**

A repeated measure MANOVA model of measures, type of support and sequence.

Effect	$\Lambda$	$F$	$df1$	$df2$	Partial $\eta^2$
Measures	.71	13.45***	2	66	.29
Quality of life; Recovery; Functioning.					
Measures <sup>a</sup> sequence of type of support during the procedure	.96	1.24	2	66	.04
Support type internal or external	.96	2.44	1	67	.04
Support type <sup>b</sup> sequence	.95	3.38	1	67	.05
Measures support type	.94	1.94	2	66	.06
Measures <sup>c</sup> support type <sup>c</sup> sequence	.83	6.78**	2	66	.17

<sup>a</sup>  $p < 0.05$ .

<sup>b</sup>  $p < 0.01$ .

<sup>c</sup>  $p < 0.001$ .

whether this difference existed in both sequences of support. The results showed a significant difference only when external support was given prior to internal support ( $F = 14.93$ ,  $df = 1$  and  $29$ ,  $p < .001$ ; Wilks'  $\lambda = .66$ , partial  $\eta^2 = .34$ ), but not vice versa ( $F = 1.19$ ,  $df = 1$  and  $41$ ,  $p = .281$ ; Wilks'  $\lambda = .97$ , partial  $\eta^2 = .03$ ). In this case, the functioning score was 15% higher for the internal support than for the external support.

### 4. Discussion

This study was the first to compare ROM data collected with the support of external and internal administrators. The question regarding whether to use internal or external administrators arose from the growing demand to implement ROM in mental health services while having a limited understanding of the costs and benefits of depending on each of those administration modes. The use of internal support could have several advantages such as low cost and high availability which may lead to a better long-term implementation of ROM and thereby improving service provision (Australian Health Ministers, 1992). Yet external support could be considered as more objective and less prone to bias compared to internal support (Spaulding, 2008).

In the current study, most results revealed no discrepancies in outcome scores according to the type of support. The findings may indicate that consumers, although presumed to be potentially influenced by the type of support they received, were consistent in their self-reports regardless of who helped them. This consistency raises the question of whether the presumption of consumers' susceptibility is indeed evidence-based or a product of an underlying preconception of mental health consumers (Linhorst & Eckert, 2002). Further studies are needed to examine this issue.

Nevertheless, one significant difference was found in functioning reports. When external assessments came first, the internal functioning self-assessment was comparatively better. It can be hypothesized that towards external administrators, consumers might wish, at first, to make their plight worse, possibly to express some grievance (Bilsker & Goldner, 2002; Choi & Pak, 2005). Interestingly, this would not happen if consumers had already been 'cued' by an internal assessment. As this is a preliminary study, this suggestion should be confirmed in future studies. In short, the current research suggests that no differences exist between the modes of support, apart from functioning being worse when external support is provided first.

### 5. Limitations

Although this study has several strengths, such as a within-control design and random group assignment, it did not provide for the assessment of the influence of specific report biases such as

social desirability, institutional 'pressure,' or a tendency to express grievances to external observers (or alternatively, the fear of expressing grievances to internal staff). Furthermore, we did not compare the perception of support helpfulness. We recommend that future studies examine whether the type of support might interact differentially with different measures such as potentially more bias-sensitive measures, e.g., service use satisfaction and therapeutic alliance. Indeed, it is possible that quality of life would not be affected by support type, but that satisfaction with service provision would. Additional factors such as consumers' educational level, religiosity and ethnicity might also interact differentially with type of support and should also be examined in future studies. Finally, similar studies should be performed in other settings where organizational culture might differentially affect the way in which consumers respond to internal and external administration of questionnaires.

## 6. Conclusions

Outcome scores were generally not affected by the type of support received. This might have useful implications for mental health systems, as it allows flexibility in selecting the type of support used in the evaluation process. Nevertheless, in the measurement of functioning, internal support might be preferable although further studies are needed to (a) replicate the finding, and (b) elucidate why this might be the case.

If different forms of assessment yield equivalent responses, it may be concluded that fears of bias in internally supported ratings are exaggerated, and that the results are likely to be trustworthy. This might have important implications because internal assessments, conducted by staff who are directly involved in the consumer's care might be more likely to be used in the actual delivery of care in light of considerations of availability and cost (Graham et al., 2001). Furthermore, providers and policy makers have also made the point that outcome measures should be an integral part of the delivery of care (Jacobs & McDaid, 2010; Kelley & Bickman, 2009). Therefore, the next step in the process of using internal support would be to make consumers an integral part of the feedback process to their personal case managers, the services and policy makers. This would help provide quality control as well as rational and evidence-based decision-making tools for recovery and service planning.

## Financial support

A joint Grant by the Ministry of Health 45980 and the Tauber Foundation was obtained for this project

## Disclosures

None for any author.

## Acknowledgements

This study as part of the large National Outcome Measurement Project of Community Rehabilitation Services that was funded by the Laszlo N. Tauber Family Foundation and the Israeli Ministry of Health. We would like to thank the organizational support provided by the, Mrs. Noa Joseph for her statistical support, Mrs. Adi Telem and Mrs. Ronit Uliel Mirkin for their work in the field in organizing this study, as well as Mr. Aid Rohanna for supervising the data collection process.

## References

Australian Health Ministers (1992). *National Mental Health Policy*. Canberra: Australian Government Publishing Service.

- Bilsker, D., & Goldner, E. M. (2002). Routine outcome measurement by mental health care providers. Is it worth doing? *Lancet*, 23, 1689–1690.
- Choi, B. C. K., & Pak, A. W. P. (2005). *A catalog of biases in questionnaires preventing chronic disease*, 2, 1–13.
- Corrigan, P. W., Salzer, M., & Ralph, R. (2004). Examining the factor structure of the Recovery Assessment Scale. *Schizophrenia Bulletin*, 30, 1035–1041.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2003). *Program evaluation: Alternative approaches and practical guidelines* (3rd ed.). Boston, MA: Pearson.
- Fossey, E. M., & Harvey, C. A. (2001). A conceptual review of functioning: Implications for the development of consumer outcome measures. *Australian & New Zealand Journal of Psychiatry*, 35, 91–98.
- Graham, C., Coombs, T., Buckingham, W., Eagar, K., Trauer, T., & Callaly, T. (2001). Consumer perspectives of future directions for outcome self-assessment. In *Report of the consumer consultation project*. Wollongong, NSW: Centre for Health Service Development, University of Wollongong.
- Jacobs, R., & McDaid, D. (2010). Performance measurement in mental health services. In P. C. Smith, E. Mossialos, & I. Papanicolas (Eds.), *Performance measurement for health system improvement: Experiences, challenges and prospects*. Cambridge: Cambridge University Press.
- Kelley, S. D., & Bickman, L. (2009). Beyond outcomes monitoring: Measurement feedback systems in child and adolescent clinical practice. *Current Opinion in Psychiatry*, 22, 363–368.
- Linhorst, D. M., & Eckert, A. (2002). Involving people with severe mental illness in evaluation and performance improvement. *Evaluation & the Health Professions*, 25, 284–301.
- Love, A. J. (1991). *Internal evaluation: Building organizations from within*. Newbury Park, CA: Sage.
- Mathison, S. (1999). Rights, responsibilities, and duties: A comparison of ethics for internal and external evaluators. *New Directions for Evaluation*, 82, 25–34.
- Morrell-Bellai, T. L., & Boydell, K. M. (1994). The experience of mental health consumers as researchers. *Canadian Journal of Community Mental Health*, 13, 97–110.
- Oakley, K., Malins, G., Riste, L., & Allan, J. (2011). Consumer participation in service evaluation and quality improvement: Key ingredients for a system to deliver national indicators. *Australasian Psychiatry*, 19, 493–497.
- Priebe, S., Huxley, P., Knight, S., & Evans, S. (1999). Application and results of the Manchester short assessment of quality of life (MANSA). *International Journal of Social Psychiatry*, 45, 7–12.
- Roe, D., & Gelkopf, M. (2012). *Report of the Israeli national outcome measurement pilot project of community rehabilitation services*. Haifa: The Center for Community Mental Health Research, Practice, and Policy, University of Haifa.
- Roe, D., Lachman, M., & Mueser, K. T. (2010). A decade of the Israeli psychiatric rehabilitation law. *Israeli Journal of Psychiatry*, 47, 164–165.
- Royal College of Psychiatrists (RCP) (2010). *Recommended outcome measures for use in adult psychiatry*. London: RCP.
- Spaulding, D. T. (2008). *Program evaluation in practice: Core concepts and examples for discussion and analysis*. San Francisco, CA: Jossey-Bass.
- Trauer, T. (2010). *Outcome measurement in mental health: Theory and practice*. Cambridge: Cambridge University Press.

**Marc Gelkopf** is an associate Professor at the Department of Community Mental Health at the University of Haifa, Israel. He has written more than 80 peer-reviewed articles on subjects related to mental health, drug use, trauma and psychiatric rehabilitation. Together with Prof David Roe and in collaboration with the Ministry of Health and the Tauber foundation they have set up the National Rehabilitation Outcome Project which aims to provide for ongoing assessment of all rehabilitation service users in Israel. Marc Gelkopf also leads international delegations in disaster stricken areas to provide for psychological stress relief.

**Shira Pagorek-Eshel** received a Postdoctoral scholarship at the Center for Mental Health Research, Practice and Policy at the Department of Community Mental Health, the University of Haifa, Israel. She is a lecturer at the Department of Social Work at Zefat Academic College and the Bob Shapell School of Social Work at Tel Aviv University. Her research and teaching focus on outcome evaluation, traumatic stress and social work methods.

**Tom Trauer** who has passed away on the 10th of July 2013, was a leading researcher in psychiatric outcome measurement and advocate for mental health service development. His 2010 book, *Outcome Measurement in Mental Health: Theory and Practice*, and his more than 200 peer-reviewed articles on this subject advanced the field significantly. His work with colleagues in Australia and across the world led to numerous collaborations to improve mental health service delivery across the Asia Pacific region. Tom was an estimated colleague and friend. He has significantly contributed to the conception and writing of the current manuscript.

**David Roe** is a licensed clinical psychologist and Professor at Department of Community Mental Health at the University of Haifa. His research has been funded by local and international sources, including NIMH, The Israeli Ministry of Health, The Israel National Institute for Health Services Research and Health Policy, the Israeli National Insurance Institution and the Tauber and Rich foundations. He has published over 120 peer reviewed articles and co-authored two books. He serves as deputy editor of the *Israeli Journal of Psychiatry*, associate editor of *BMC Psychiatry* and *Journal of Mental Health* and is on the editorial board of several other journals.