Give us the tools and we will finish the job:
evidence-based practice and user evaluation in mental health services

Dr David M Rea

A paper for presentation at the 5th biennial conference of the European Evaluation Society, Seville, Spain, October 10-12th 2002

Address for correspondence: Dr David M Rea, Centre for Health Economics and Policy Studies, School of Health Science, University of Swansea, Singleton Park, Swansea, SA2 8PP, UK
Tel: +44 1792 295314
Fax +44 1792 295643
http://www.healthscience.swan.ac.uk/
d.m.rea@swansea.ac.uk

1 The author gratefully acknowledges support from service users in Swansea, Swansea User Network, Swansea Social Services and Swansea NHS Trust.
Abstract

Attempts to evaluate medical and health care services have generated a new technology of assessment. This technology requires the generation of data, collection and ordering through information systems, and a variety of tools with which evaluation can be carried out. In this context, governments and third-party payers are increasingly demanding evidence-based practice. But, they are also demanding user evaluation of services.

However, evidence-based practice and user evaluation imply differing viewpoints and differing criteria of accepted evidence. As a consequence, no gold standard has yet been developed. As may be anticipated, evaluation methods that emphasise the validity of the research instrument fail to involve service users because validity is corrupted by such involvement. Users are the intended object of the evaluation, not the agent. This is carried to extremes where they are users of mental health services. On the other hand, evaluation methods that involve service users – self-evaluation or user-focused methods – frequently seek to avoid the statistical and technical complexity required to establish validity. The paper is illustrated by comparing (two) examples of service evaluation of mental health services. It suggests that the tools are of minor significance compared with the benefits gained by involving service users in the evaluation.
Give us the tools and we will finish the job: evidence-based practice and user evaluation in mental health services

‘Give us the tools and we will finish the job’
Winston Churchill’s radio address to President Roosevelt 9th Feb 1941

Introduction

Churchill’s appeal to the US president in 1941 implied that Britain had already demonstrated the will and the courage to conduct a war against fascism, and all that what was lacking was sufficient weaponry. The tools Churchill asked for are something apart from us as human beings. Tools are something we use to carry out what we will. They are commonly regarded as external to ourselves, just as technology is perceived as a threat. An alternative position would be that it is our use of technology that makes us human, tools enhance our abilities and alter the nature of our being positively. But, frequently, we regard tools and technology as alien to humanity. Technologies remain externalities, things we use to achieve our will.

British public policy today is engaged in a new search for tools – tools that will enable government to ensure public services are effective and good value for money. These tools need to offer comparability across services, they need to accurately inform public debate and policy making at national and local levels. Old tools are being dusted down and re-examined. New tools are being created. The enemy is clear – it is inefficiency and its ally ineffectiveness.

In health and social care, the old measures of effectiveness, indicators of longevity, morbidity, and social deprivation (Jarman, 1983; Meltzer et al., 1995; ONS, 2001; Townsend et al., 1988), are being put to new uses - to examine whether improvements have occurred as a result of public policy. A sufficient ‘evidence base’ is made a requirement of national standards for treatment and service delivery (DoH, 1999; DoH, 2001; Welsh Assembly Government, 2002).
Changes in employment rates among people with mental health problems are to be used as an indicator or the success of anti-stigma strategies. It is in Key Action 3 of the National Service Framework for mental health services in Wales (Welsh Assembly Government, 2002).

As the discussion below indicates, the status of evidence and causality is frequently contentious.

However, new tools have been required because there are questions about effectiveness. What is the desired effect of a service? Whose judgement should be used to judge effectiveness? Carpet-bombing is no longer regarded as an effective means of waging war, and public services cannot simply be provided and funded on any assumptions of effectiveness. Just as modern warfare demands the public broadcast of laser-guided missiles hitting their targets with no collateral damage, new measures are required to ensure services are effectively targeted – at those in greatest need, or where intervention can be of greatest benefit.

The usefulness of these tools is contentious because allocation decisions are inextricably tied in with them. If resource allocations are not directly linked, then government has introduced new powers to take over the management of poorly performing public services. This is an obvious threat to the officers and politicians involved. For many, this means the technologies of assessment are required be as objective as possible. This is frequently interpreted as meaning assessment should be based on clinical or psychiatric indicators based on symptoms and functionality. Others accept that need and outcome are best understood from a service user perspective, such as whether a person feels ‘respected’ when dealing with health and social care professionals.

Measurement of effectiveness from the perspective of service users and caregivers is increasingly required by public policy. However, mental health
service users views have tended to be viewed as ill-informed, subjective, or corrupted (by the symptoms or by the treatment) (Barnes & Wistow, 1994). So, measures of effectiveness that incorporate service user views have often been regarded as invalid. Any tool that has proven validity will be attractive to decision makers.

Validity simply refers to whether the results ‘accurately reflect what is claimed they reflect’ (Gomm et al., 2000, p27) or whether there is a ‘correspondence between the measure… and the concept in question’ (Bryman, 1989, p58). But validity is not straightforward. ‘The fruits of test/retest reliability, construct validity and criterion validity… are often not definitive because they are capable of more than one interpretation’ (Bryman, 1989, p62). Moreover, the concepts being measured may necessitate precision that may be threatened once ‘non-experts’ take part in their definition. Face validity, for example, is often achieved by assembling a panel of experts to make a judgement. It is a ‘highly judgemental process and easily prone to error ’ (Bryman, 1989, p58).

Nevertheless, the attraction of ready-made, off-the-shelf, tools whose authors claim have already been validated is obvious. It corresponds with perceptions that tools are external to us. Like Britain in 1941, policy makers and decision makers are under considerable pressure. In 2002, the pressure is to justify existing practice or to actively develop different practices in the light of firm evidence. Government has strengthened its powers of review, creating and funding arms-length bodies such as the Audit Commission, Social Services Inspectorate, Joint Reviews, and the Centre for Health Improvement.

The alternative is to develop, in-house, the capacity to evaluate needs, service delivery and outcome in a way that involves service user and caregiver perceptions. There are also considerable pressures on local policy makers to involve service users and caregivers in assessment. For examples the National service framework for mental health services in Wales requires, under Key Action
2, evidence of periodic surveys of users’ experience in different areas of Wales (Welsh Assembly Government, 2002).

However, services rarely have the time, skills and expertise available to develop their own measures of effectiveness. And, even when they do, there are doubts about its comparability, reliability and validity. Comparability of results will be difficult because different things will have been measured using different methods. Reliability and validity will be difficult because it is unlikely that its use will be repeated exactly as before. Moreover, this alternative does not correspond to the view that tools are external. Instead, measurement and the development of measurement methods is particular to those concerned and affects them at a profound level. Indeed, recognition of this would be one of its main advantages.

This paper also argues that evaluation should empower people (Fetterman et al., 1996; Smith, 1998) and that, especially when dealing with vulnerable people, we cannot regard any method as valid unless it empowers them. That is, it should add to their capacity as people. Smith advocates empowerment evaluation ‘despite the limitations of reliability and validity’ (p149). This paper argues, more strongly, no method can be regarded as having validity unless it succeeds in empowering the people whose services are being evaluated. For want of a better term, we can call this ‘empowerment validity’.

Validity is normally concerned with preparation and implementation of an evaluation, with feedback and learning as less important parts of the process. Without due emphasis on preparation and implementation, an evaluation will be jeopardised (Forss et al., 2002). This paper argues that evaluation will also be jeopardised unless the service users have helped define the concepts being measured and taken part in defining the methods, it cannot be valid for them. No matter what the evaluators and the service providers think they have learned from an evaluation, service users and carers will not see the process as valid and they will not learn from it. Involving service users and caregivers has benefits for
the evaluation in terms of what providers, users and caregivers can all learn from each other. This is knowledge that is worth acquiring and can have considerable effects on planning and delivery of future services.

The first section briefly reviews the range of ready-made and validated measures of need and outcome. Then the paper examines an example from this range before outlining a model of evaluation that involved service users.

**Ready-made and validated measures of need and outcome**

**Research tools to measure needs assessment**

1. The MRC Needs for Care Assessment Schedule (NFCAS) (Reid, et al. 2001)
2. The Camberwell High Contact Survey (Reid, et al. 2001)
3. The Camberwell Assessment of Need (CAN) - (Leese, et al. 1998)
4. The Cardinal Needs Schedule (CNS) - (Lockwood & Marshall, 1999)

**Research tools to measure service user satisfaction**

5. The Verona Service Satisfaction Scale (VSSS) - (Leese, et al. 1998; Greenwood, et al. 1999)
6. Client Satisfaction Questionnaire (CSQ) - (Greenwood, et al. 1999)

**Research tools to measure other outcomes**

8. Health of the Nation Outcome Scales (HoNOS) - (Brooker, et al. 1997; Bruce, et al. 1999)
9. Knowledge about Schizophrenia Interview (KASI) - (Wray, 1994)
10. The PRiSM psychosis Study (PRiSM) - (Leese, et al. 1998).

In a sense, all these are measures of need. Although there are several kinds of outcome (McCallum, 1993), in general terms, a favourable outcome is a reduction in need. But, as many have pointed out, need is a contentious concept and not easily defined by clinical symptoms, social functioning, feelings of self-worth, or some combination. There are doubts that needs can be objectively measured (Percy-Smith, 1996).

Moreover, there is much less willingness now than there was previously for needs to be defined objectively. Some commentators have argued that the New Right is responsible for taking needs out of public debate because the concept requires resource allocation decisions to be taken by professionals and others who are not disinterested (Doyal & Gough, 1991). However, needs ought to be defined - at least in part - by service users and their caregivers.

What these measures listed above seek to achieve is to provide a measure of need(s). Sometime this requires an acceptance that needs judgements frequently rely on subjectivity. But that subjectively expressed needs can still be quantified and rendered calculable. That is, numerical expressions of subjectively expressed needs can be ranked, combined, added, etc. A need can be placed over or below another. One health state can be valued over or below another. So, a person’s anxiety state, for example, can be expressed using Likert scaling as very severe, less severe, and so on. Numbers can be assigned to these and the levels of need - or a reduction/increase over time calculated.

**CANSAS - a ready-made tool with proven validity**

The above point can be illustrated by the Camberwell Assessment of Needs Short Appraisal Schedule (CANSAS) which makes it possible to identify how many needs someone has from among 22 domains, and how many needs are unmet (Tunnicliffe, 2001). Then, once sufficient numbers of service users - and staff -
have been interviewed, it is possible to summarise how well needs are being met by one part or another of the service (such as a Community Mental Health Team). This information can then be rendered in tabular form to show, for example, that 8.5 (say) needs are being met in one area of service provision as opposed to 7.5 in another. It therefore serves as a performance indicator, a measure of variability, and as a means for identifying scope for improvement.

CANSAS is one of the more up-to-date tools, validated during the 1990s. It is widely used, probably because it is easy to use, requires no training, and can be completed within 30 minutes (Evans et al., 2000). This ease of use is undoubtedly achieved at the expense of user perceptions of need which are limited in their expression to a closed form of questioning (Evans et al., 2000, p384). Nevertheless, CANSAS is highly regarded as a tool for measuring need from a users’ perspective. It places importance on the subjective dimension of needs. Moreover, user conceptions of their needs are expressed more than in others (Wiersma et al., 1998).

**Empowerment evaluation**

We can contrast the approach above with another, the Avon Mental Health Measure (AMHM) which has emphasised user involvement from the outset. Starting from an assessment document used by Avon Social Services, it was adapted through a process of consultation with users of mental health services. As a result, the final layout was a jointly owned document and no single group has copyright. The AMHM was associated with a DoH funded project, Changing Minds, which encouraged the use of the measure as part of an ethos of social inclusion (Frost et al., 2002). Autonomous Local Action Groups were important to its development and use.

However, it remains focused on individual assessment, ensuring service users can take part meaningfully in their own care planning. The main advantage
is that it doesn’t simply look for deficits, but portrays the whole client. Services can ensure issues are focused and priorities addressed. Both statutory agencies and service users can see that the measure was developed in consultation with service users. Its major disadvantage is that when voluntary organisations take the lead, the statutory sector may be hard to engage. Service users and volunteers will feel less than empowered if the service providers are perceived as failing to respond adequately.

This is not a failure of the AMHM, as such, because the measure itself cannot transfer power and decision-making to the service user. The measure can still be regarded as a tool by the service providers, albeit well-intentioned. It is there as a tool to ensure better services and to assure external scrutinisers. To use Hickey & Kipping’s concept of a ‘participation continuum’ (1998), its use in this way does not ensure partnership or user control.

In Swansea, it is hoped to develop the AMHM as a measure of outcome. In order to ensure it is rooted in an empowerment model of evaluation, groups of mental health service users were involved from the beginning. As a result, the measure was altered considerably for local use before being introduced. We are now half-way through its one-year trial period and it is expected that it will be altered again afterwards.

It is hoped this will lead towards partnership working, where both service providers and service users can engage in service improvement resulting from a genuinely open dialogue. The adapted AMHM will not be the final result. The process by which it is adapted for use will need to encourage service users and providers to listen to and learn from each other. Crucially, it should enable service users to conceptualise their needs, rather than allow professional staff to define needs for them. It should also enable a common understanding of their needs to be shared.
These hopes will not be realised if there is insufficient drive to move the process forward. Nor will it occur if partners feel that other partners are not contributing sufficiently. Equally, it will not occur if any partner regards another as dominant. The difficulty with empowerment is that it involves a shift in power. To give up power to some degree is itself an act of power. Local providers may feel forced by the current policy context to rely on service user partnership. But the policy context is largely determined by central government. At present, the empowerment agenda looks as if it may be driven off the rails by concerns over dangerousness and risk. The Mental Health Bill, currently before Parliament, may undermine much of the strategy for developing and improving mental health services. It may prove difficult for service users to see themselves as partners in this context, and local providers may feel they cannot risk a partnership.

This list represents a formidable challenge to those involved in the evaluation process. Learning is not simply a matter of interpreting the data and adjusting service delivery accordingly. Instead, empowerment evaluation requires partnership and partners have to learn and develop their roles. They have to work on this explicitly. Trust has to be developed between partners. Understanding of each others limits, and their own limitations. Empowerment validity consists of the following:

1. Service users must have ample opportunity to express their expectations of the services they receive

2. Service users should have ample opportunity to participate in the design of any data gathering methods adopted, take part in data gathering, and in the interpretation of results. Their role should be acknowledged in any report or publication

3. Service users must have every assurance possible that service provision will be affected by the evaluation process
4. Service users are fully entitled to feedback from the evaluation that explains the findings and recommendations.

5. Service users are fully entitled to expect an explanation from providers of why changes to services have been introduced – whether they result from the evaluation process or from some other factors.

6. Service users are fully entitled to an explanation for any decision not to act on recommendations of an evaluation.

**Conclusion**

Empowerment evaluation has more to offer than evaluation methods that focus on the perfection of a measurement tool and their proper implementation. A validated tool will effectively assess needs and whether needs are being met adequately. If it is perceived as valid and reliable, then policy makers and decision makers may be able to argue the case for changes in service delivery with funding bodies, professionals, and service users. But evaluation processes that develop and educate both service users and providers will go further. Ultimately, it should develop an understanding among the policy-makers and fund holders that service providers are accountable to.

Empowerment evaluation is not simply about empowering the client, providing feedback to the providing organisation that has paid for the evaluation. That part of the process may itself be frequently overlooked. However, evaluation can and should go much further than that. We cannot regard an evaluation as valid unless the service users take some major part in defining the needs which the service is going to improve delivery on. Their perception of performance is required. This much is now recognised.

However, empowerment evaluation can fundamentally alter relationships between service provider, professionals, and service users. The model currently
dominating public policy is a model of accountability where service provision is accountable upwards and downwards (to government, electors, and service users). Involvement of service users and caregivers though an empowering evaluation process will enable service users and caregivers to own the services and be responsible to each other for the services provided. The work of an evaluator - and the service providers - is to enable sufficient numbers of service users to accept the challenges implied by this. This is an educational task - not just in the sense of transfer and acquisition of knowledge. It is educational in another sense - that such knowledge enhances capability and empowers confidence and action. Ultimately, developing empowerment is an educational task that implies a sharing of responsibility. Public services will need to alter considerably before this will occur.
References


Tunnicliffe, C. (2001), A profile of needs of clients receiving Community Mental Health Team Services: an empirical study using the Camberwell Assessment of Need, Carmarthen, Dyfed Powys Health Authority.

