Program evaluation for child and family services

What can be done?

Frank Ainsworth

This article is about program evaluation for child and family services. It sets out to offer some basic frameworks for thinking about program evaluation and about the issue of program effectiveness. A rationale for the emphasis on effectiveness is identified and then linked to three areas of possible measurement. These areas, changes in user/client condition, quality of services provided and user/client satisfaction, are then considered in more detail. Finally, it is argued that service users/clients will gain from program evaluation exercises. The evaluation of services contributes potentially to an improvement in the effectiveness of child and family services so service users/clients obtain benefit from such evaluations.

Frank Ainsworth, Ph.D.
Lecturer, School of Community Studies
Edith Cowan University
Joondalup Campus, Perth, WA 6027

Increasingly both government and non government not-for-profit child and family service agencies are faced with demands that they evaluate their programs. The demand is for evidence that these programs are effective. In this context effectiveness generally means demonstrating that the aims of a program have been achieved. These demands are part of a demand for an increased accountability which has been accentuated by the implementation of a tendering and contract awarding approach to the provision of human services (Ainsworth, 1994; 1996). This approach has been adopted by all state and territory authorities to some degree. It is the mechanism they use to purchase services for children and families from the not-for-profit sector. Demands for increased accountability and for service evaluations also come from agency boards of management of not-for-profit organisations. This stems from an acknowledgment that the culture within which not-for-profit organisations now operate is dramatically different from that of a few years ago. Issues of accountability, program evaluation, and evidence of service effectiveness now dominate the agenda. Some commentators think that this agenda leads to a neglect of the needs of service users. This article argues a different position. It is proposed that the emphasis on accountability, program evaluation and service effectiveness actually benefits service users.

WHAT IS A PROGRAM?

As an organised response to an identified need, a service program should be carefully designed. A program has to have a clear focus and defined boundaries. It is not a set of loosely related or jumbled pieces of practice. At the design stage, knowledge from theory, research and practice experience needs to be carefully considered. Based on a mature and consistent integration of these sources of knowledge, a clear program aim – sometimes called program goals or objectives – then has to be defined (Ainsworth, 1991). A program may of course have more than one aim but each aim must be clear, compatible with other aims and there should be no overlap. If goals are not compatible then they will not be achievable. If aims overlap it may be impossible to tell which aim is being achieved. An example of incompatible and over-lapping aims might be as follows. An emergency accommodation program might have two aims, to provide short term, emergency accommodation, and to offer permanent tenancies to needy families. In this situation the first aim may be in conflict with the second aim. To offer permanent tenancies to needy families would almost certainly diminish the capacity of the program to provide emergency accommodation. Alternatively, the program may aim to assist families to enter the private rental market while at the same time offering subsidised accommodation that makes it disadvantageous for families to rent unsubsidised private accommodation.

After defining the aims of a program, the next step is to develop a clear structure and a set of activities that are to be followed in order for the program aims to be achieved. The activities can be regarded as the program curriculum. For example, the above emergency accommodation program will need a suitable building for the program and policies and practice guidelines that state how the program is to be accessed and by whom. The activities to be pursued to achieve the program aims also have to be specified.

Increasingly both government and non government not-for-profit child and family service agencies are faced with demands that they evaluate their programs. The demand is for evidence that these programs are effective.

A program also needs staff who will work toward the program aims using the program activities and policy and practice guidelines to achieve these aims. The evaluation of the program effectiveness is then based on the extent to which the program aims are achieved. An evaluation is not only targeted at the outcome for the individual service user, although this is important. Another target for an evaluation is the outcome for the group, that is, all users, in a given time period. What is sought is evidence that the program consistently delivers a positive outcome for most, if not all, service users. In this sense a program is a bit like a theatre program. It has three components: it represents an outline of what the performance is about; a series of scenes or sets of activities that is the mechanism for the delivery of the performance; and a promotional device by which another audience or group of service users may be attracted (Fulcher & Ainsworth, 1981). An internal consistency between all of these aspects

– program aim, structure, activities, and policies and practice guidelines – must exist. Any inconsistency will make for a poor performance and lead to an ineffective or less effective program.

WHAT IS EFFECTIVENESS?

Service effectiveness is for human service agencies what the profit motive is for business organisations (Patti, 1987). It is the *raison d'être* for their existence. This justification stems from the fact that human service agencies are in the business of 'changing people and/or their social conditions' (Patti, 1987, p.9). To exist without a prime commitment to providing effective services essentially means that a program has ceased to respond to the needs of service users. Fortunately, all personnel in child and family service organisations are likely to agree that the programs they provide should be effective. In simple terms a program that makes a positive difference is effective while a program that makes a negative difference or no difference is ineffective. The question that must be faced by service delivery organisations is, if a program makes a negative or no difference, why does it exist?

Patti (1987) suggested that human service programs can be assessed or evaluated on three dimensions: changes in user/client status or condition; the quality of the service provided; and the level of satisfaction of the service users. To use the same example, an emergency accommodation program should make a positive difference and change a client's condition. This type of program has to be responsive to immediate need and quickly provide accommodation in a manner that respects the vulnerability of those applying for the services. When this is the case, it can be judged as a high quality and effective service. If an emergency accommodation program relieves a service user's distress, and if this is done in a timely and sensitive manner, the client's satisfaction with the service is also likely to be high. Again the service can be judged to be effective.

However, program evaluation requires that we go beyond simple judgments made by program personnel. Indeed, all three of these dimensions, changes in

user/client status or condition, the quality of the service provided and the level of satisfaction of the service users, represent some types of service outcome that are capable of being measured. Service outcomes should reflect the stated aim of a program and in that respect can be viewed as the product of the services provided. Measuring such outcomes for evaluation purposes is one step toward demonstrating the effectiveness of a human service program.

TYPES OF EVALUATION

At this point it is worth noting five possible types of program evaluation and the different purposes for which each may be used. These are:

- process evaluation;
- evaluation for program management purposes;
- evaluation of the design of a program;
- evaluation for development purposes;
- impact or outcome evaluation.

(Owen, 1993)

Each of these types of evaluation are used differently although they are not necessarily always used completely separately.

Process evaluation

Process evaluation is commonly used early in the life of a program to establish the extent to which the program has been implemented according to the original design. Qualitative data is collected to check how a program is going, to identify unanticipated difficulties and to see if any minor adjustments to the program are necessary. Process evaluation does not seek a major redesign of the program. However, it may lead to the formulation of a series of questions that may result in a more rigorous examination of the program. An example of this type of evaluation is a review of the operation of a new substance abuse program that aims to teach heroin users how to avoid overdosage, conducted at the end of the first six months.

Evaluation for program management purposes

Program management evaluation occurs when agency managers want additional information about the functioning of a

particular program. For example, this may be about a component of a program that is new or experimental, or it may be about specific cost issues. Essentially, this type of evaluation involves taking a closer look at an aspect of a program so that informed decisions that enhance the development of the program can be made. A domestic violence prevention program that is significantly over budget might be subject to this type of inquiry or evaluation. The additional information from this type of evaluation does have the potential to raise issues for further investigation by other forms of evaluation.

Evaluation of the design of a program

Evaluation of program design is a more fundamental questioning of the way a program is operating. It involves taking a serious look at the relevance of a program. An example might be an evaluation of an existing intensive foster care program because of the high number of placement breakdowns that appear to stem from the increasingly difficult population the program must serve. This type of evaluation may lead to confirmation of the appropriateness of the program, the disestablishment of a program, or a radical shift in design.

Evaluation for development purposes

Evaluation for program development may be necessary if a long established program is seen as of diminished relevance. For example, a drift of service users to the suburbs and a decline in the enrolment at a central city parent education program might lead to the need for this type of evaluation. The results of such an evaluation may confirm the need for this particular type of program or it may recommend its disestablishment or its redevelopment in a more contemporary format and/or at a new location.

Outcome or impact evaluation

Outcome evaluation is used to assess the impact of a well established and stable program. That is to say, what difference has this program made to the lives of those it has served. Outcome evaluation is closely associated with accountability issues. This type of evaluation is inappropriate when a program is still being implemented or

when a program is unstable (Pecora, Seelig, Zirps & Davies, 1996). The likely audiences for outcome or impact evaluations are government funders, charitable trusts and other philanthropic organisations, community leaders and the media (Jacobs, 1988). Outcome evaluations focus on the linkage between the stated aims of a program and what was delivered at the end of the service period. A relevant example is those family preservation services that claim that they can prevent children from being placed in out-of-home care (Kinney, Haapala, Booth & Leavitt, 1990). An outcome evaluation would take this aim and examine the extent to which it has been achieved. Outcome evaluations usually rely heavily on quantitative data and make programs accountable in performance terms. It is this form of evaluation that finds most favour in a contracting or purchase of service environment. It asks for evidence that 'the program did what it said it would do'. This issue is at the heart of purchase of service agreements or contract arrangements. It responds to the value for money question by seeking information about what was provided and how it changed the lives of the service users. Without data about the outcome of services it is hard to persuade funders to continue to purchase services or award a contract because there is no clarity about what they are purchasing.

WHAT CAN BE MEASURED?

Particularly with outcome-based evaluation (Schalock, 1995), the issue is what outcomes can we measure and how? Contrary to popular belief, it is possible to measure the outcome of many child and family service activities. For programs to do so there is a need to change attitudes and develop new skills.

The necessary skills are in relation to psychometric measurement and the use of advanced statistical techniques. Used carefully and applied sensitively these techniques can allow a program to demonstrate that the outcome of program participation for service users is significant and was not achieved by chance. Alternatively, agencies will need to hire individuals with these

skills or a consultant to help them with these issues.

In a contracting or purchase of service environment, outcome evaluation and the evidence of program effectiveness it provides is not an option to be occasionally indulged. It is a necessary part of operating in such an environment and staying in the child and family services business. Earlier it was noted that Patti (1987) suggested that human service programs could be assessed or evaluated on three dimensions: changes in user/client status or condition; the quality of the service provided; and the level of satisfaction of the service users. Each of these dimensions provides great opportunities for outcome measurement and is worthy of detailed consideration.

To exist without a prime commitment to providing effective services essentially means that a program has ceased to respond to the needs of service users.

Changes in user/client status or condition

Changes in relation to user/client status or condition are probably the easiest items to be measured. We already have numerous reliable and validated instruments, that is, questionnaires, indexes, inventories, rating scales, that have been developed by human service personnel for this purpose (Cross & McDonald, 1995; Faul, Hanekom & van Niekerk, 1997; Fisher & Corcoran, 1994a; 1994b; Hudson, 1996; McCroskey & Meezan, 1997; McCubbin, Thompson & McCubbin, 1996). For example, the Walmyr instruments (Hudson, 1996) number more than 20 and include a multi-problem screening inventory, an index of marital satisfaction, an index of family relations, a child's behaviour rating scale, an index of drug involvement, an index of peer relations, a partner abuse scale: physical abuse. A scoring manual and computer software

are also available for use with these particular instruments. Other authors provide a sourcebook of over 320 measures for practice with reports on the reliability, validity and availability of each instrument (Fisher & Corcoran, 1994a, 1994b).

Program evaluation means that practitioners and managers alike will have to learn new skills. If this is achieved and the appropriateness of the program is demonstrated, then arguments that services are intrusive, make matters worse, or are not needed will be difficult to sustain. Equally difficult to sustain will be the reticence of governments to fund these services.

Instruments that are relevant for a particular child and family services program can be selected from these extensive collections, and service users can be asked to complete an instrument prior to the commencement of service. For example, a child and family services program that provides family counselling services might ask families to complete the Walmyr index of family relations (Hudson, 1996) prior to the first counselling appointment. This is generally referred to as the pre-test. This would provide baseline data against which to measure any changes that occur while the family is participating in the program. An index of this type produces a score for each service user. At various points across the service period the test can be readministered to map a service user's progress. At the end of the service user's participation in the program another test score can be obtained. This is generally referred to as the post-test. It is then possible to compare the pre-test and post-test

scores and produce a hopefully positive change score. Once a program has collected a number of scores of this type, the aggregated data can be used to establish if the change is statistically significant. If the results show statistical significance, this evidence can be used as the basis of claims for program effectiveness. Quantification of service outcomes in this way is now essential for child and family service programs. The pre- and post-test method is just one of a number of ways to obtain measurement of change.

Quality of service

The quality of service provided to program users can also be quantified although in a somewhat different manner. One way to achieve this is for practitioners and managers to identify best practice standards (Pietrzak, Ramler, Renner, Ford & Gilbert, 1990; Pecora, Seelig, Zirps & Davies, 1996). Once agreed these standards can then be used as the basis of an audit of a program. These practice standards and the linked quantification effectively become benchmarks when checked against comparable service situations in other agencies (Pecora & Seelig, 1996). They can then be used to judge the quality of the services that are provided. A typical standard for a child and family service program might be the requirement that all potential participants referred to a program for neglectful parents are seen within 24 hours of referral. The benchmark might be that this should occur for 90% of referrals. By diligent record keeping it is possible to quantify the number of times this practice standard was breached or observed. The percentage of referrals where this was achieved indicates the extent to which the program reached the required standard. If the percentage is below the set benchmark, the indication is that the required service quality is not being achieved. When a breach of standards has occurred a program is then required to deal with this issue before the next program audit. If the breach is serious, then a special audit may be called or the regular monthly or yearly pattern may be maintained. Quantification of service quality in this way is also essential for child and family service programs. Indeed, in some places child and family service agencies who wish to be accredited and

eligible to tender for service contracts must have in place quality improvement mechanisms of this type.

Level of satisfaction

Measuring the level of satisfaction of the service users is also something that can be done. It is possible to ask service users a series of questions about how they experienced the program and the benefits they gained. The more specific these questions are the better. It is also best to ask questions about each component of a program separately. Just like other instruments, satisfaction questionnaires have to be carefully designed and validated. Currently, many of those handed out to users at the end of a period of service have not been validated and therefore are of little use for evaluation purposes. Essentially, what they provide is feedback to program staff which cannot be relied upon to give a clear picture about the effectiveness of the service that was provided. Accordingly, child and family service programs should not rely on information about user satisfaction when it has been obtained using inadequately developed instruments.

At this point some cautionary notes about the use of measurement instruments are needed. When searching for an instrument to use for evaluation purposes it is unusual to find one that is a perfect fit for a particular evaluation. When faced with this situation many practitioners seem to think that they will achieve their best results by designing a new instrument. Extreme caution needs to be exercised as the process of designing and validating an instrument is technically difficult and time consuming. There are many poorly designed instruments which have not been validated and they are of limited value. The message is – if at all possible use an existing, tried and tested instrument. This will allow you to produce results which will have credibility and value. If you are forced to consider designing a new instrument before you start then, throughout the exercise, make sure that you have access to high level technical assistance.

A further note of caution relates to the use of part of an instrument or sub-scale from a validated instrument. When searching for an instrument it is not

uncommon to find an instrument that contains a number of sub-scales. Some of the sub-scales may be attractive and others less so and it is tempting just to use the attractive parts. The issue is that the sub-scales may not have been validated separately from the total instrument. To use a sub-scale that has only been validated as part of a total instrument means that the items you are choosing to use are of limited value. By using only one part of the total scale you can no longer be sure of the reliability and validity of the results. If sub-scales have been validated separately then there is freedom to use them independently.

Finally, reliance on data from only one of the three areas of effectiveness is unwise. Ideally, an evaluation of a child and family services program will include measures from each of these areas: changes in user/client status or condition, the quality of the services and the satisfaction of the service users. This is the standard to aim for and reach.

CONCLUSION

Program evaluation deserves to be embraced enthusiastically by child and family service programs. While program evaluation may demonstrate that participation in a program was correlated with changes in a user/client status or condition, in research terms this does not demonstrate a causal linkage (Pedazur & Pedazur Schmelkin, 1991). Nevertheless, establishing this correlation is important as it provides support for the continuation of the service program that has been evaluated.

Program evaluation means that practitioners and managers alike will have to learn new skills. If this is achieved and the appropriateness of the program is demonstrated, then arguments that services are intrusive, make matters worse, or are not needed will be difficult to sustain. Equally difficult to

sustain will be the reticence of governments to fund these services. Evidence of appropriateness and some measure of effectiveness is a powerful ally when child and family service providers are bidding for service contracts. Let it also be clear that the users of child and family service programs benefit from the demand that programs be evaluated. In the long term the evaluation of programs will lead to users receiving services that are known to work. Children and families that enter the human service system deserve services that are known to work effectively. They should not have to put up with services that have not demonstrated their effectiveness (Ainsworth, 1997). You and I would not want to accept ineffective services. Why should they! □

REFERENCES

Ainsworth, F. 1991, 'The development of a residential program for adolescents', The Child and Youth Care Administrator, 3, 2, 21-24.

Ainsworth, F. 1994, 'Contracting arrangements in children's services', Children Australia, 19, 2, 11-14.

Ainsworth, F. 1996, 'Surviving and thriving in a contracting environment', Children Australia, 21, 1, 23-26.

Ainsworth, F. 1997, Family centred group care: Model building, Aldershot: Ashgate Publishing.

Cross, T. & McDonald, E. 1995, Evaluating the outcome of children's mental health services: A guide for the use of available child and family outcome measures, Boston, MA: Judge Baker Children's Center.

Faul, A C., Hanekom, A. J. & van Niekerk, C. 1997, Workshop manual, Perspective Training College, Silverton.

Fisher, J. & Corcoran, K. (eds), 1994a, Measures for clinical practice. Vol. 1. Couples, families and children, (2nd edn.), New York: Free Press.

Fisher, J. & Corcoran, K. (eds), 1994b, Measures for clinical practice. Vol. 2. Adults, (2nd edn.), New York: Free Press.

Fulcher, L. C. & Ainsworth, F. 1981, 'Planned care and treatment: The notion of programme', in F. Ainsworth & L. C. Fulcher, Group care for children: Concept and issues, London: Tavistock Publications.

Hudson, W. 1996, Walmyr assessment scales scoring manual, Tampe. AZ: Walmyr Publishing Company. (Available from Measures, PO Box 58, Osborne Park, W. Australia 6917),

McCroskey, J. & Meezan, W. 1997, Family preservation and family functioning, Washington, DC: Child Welfare League of America.

McCubbin, H., Thompson, A. I. & McCubbin, M. A. 1996, Family assessment: Resiliency, coping and adaptation, Madison. WI: University of Wisconsin Publishers.

Owen, J. M. 1993, Program evaluation. Forms and approaches, Sydney: Allen and Unwin.

Jacobs, F. H. 1988, 'The five tiered approach to evaluation: Context and implementation', in H. B. Weiss & F. H. Jacobs (eds), Evaluating family programs, New York, Aldine de Gruyter.

Kinney J., Haapala D., Booth C. & Leavitt S. 1990, 'The Homebuilders model', in J. K. Whittaker, J. Kinney, E, M. Tracy, & C. Booth (eds), Reaching high risk families: Intensive family preservation services, New York: Aldine de Gruyter.

Patti, R. J. 1987, 'Managing for service effectiveness in social welfare: Toward a performance model', in R. J. Patti, J. Poertner & C. A. Rapp (eds), Managing for service effectiveness in social welfare organisations, New York: Haworth Press.

Pecora, P. J., Seelig, W. R., Zirps, F. A. & Davis, S. M. (eds) 1996, Quality improvements and evaluation in child and family services, Washington, DC: Child Welfare League of America.

Pecora, P. J. & Seelig, W. R. 1996, 'Selected quality improvement implementation issues', in P. J. Pecora, W. R. Seelig, F. A. Zirps & S. M. Davis (eds), Quality improvements and evaluation in child and family services, Washington, DC: Child Welfare League of America.

Pedazur, E. J. & Pedazur Schmelkin, L. 1991, Measurement, design and analysis, Hillsdale, NJ: Lawrence Erlbaum Associates.

Pietrzak, J., Ramler, M., Renner, T., Ford, L. & Gilbert, N. 1990, Practical program evaluation: Examples from child abuse prevention programs, Newbury Park, CA: Sage Publications.

Schalock, R. L. 1995, Outcome based evaluation, New York: Plenum Press.