



resource sheet

NATIONAL CHILD PROTECTION CLEARINGHOUSE

Evaluating child abuse and neglect intervention programs

Alister Lamont

Program evaluation is a key component to all child welfare and child protection intervention programs. A rigorous evaluation assists in improving service delivery and enables policy-makers and service providers to determine the effectiveness of an intervention program.

This Resource Sheet provides an overview of the key aspects of program evaluation. It outlines different evaluation types, identifies the key elements to developing a rigorous evaluation and highlights possible limitations in evaluating interventions in the child welfare sector.

Interventions in the child welfare and child protection system

Broadly, interventions in the child welfare and child protection system aim to prevent the occurrence or reoccurrence of child abuse and neglect and/or to mitigate the negative impacts of trauma, abuse and neglect on children. Such interventions are generally operated and funded by government and community welfare organisations. Intervention programs are commonly classified into three main levels: primary, secondary and tertiary. Primary interventions provide support and education before problems arise and are offered to everyone in the community (e.g., general media awareness campaigns and school-based education programs). Secondary interventions are targeted at families deemed to be at risk of child maltreatment, and provide additional support to help alleviate identified problems and prevent

further problems (e.g., home visiting programs and parent education programs). Tertiary interventions target families in which child abuse or neglect has already occurred (e.g., statutory child protection services and clinical services such as sexual abuse counselling or attachment therapy) (Holzer, 2007).

Why are evaluations important?

Program evaluations are an important mechanism for quality assurance. Without an evaluation, interventions are unable to measure their effectiveness. Essentially, evaluations are important for determining whether or not an intervention is of benefit to participants (Holzer, Higgins, Bromfield, Richardson, & Higgins, 2006). Evaluations may analyse how services are used, profile service users, assist in the ongoing improvement and refinement of program content and provide informed cost-benefit analyses (Tomison, 2000). Evaluation findings are of use for those planning a new service, practitioners providing services and policy-makers making decisions about public policy and program funding.

Types of evaluations

There are a number of evaluation types and methods. The three main types of evaluations are: process, impact and outcome evaluations. It should be noted that different types of evaluations are intrinsically linked and can be used either independently or together (Tomison, 2000).

Process evaluation

Process evaluation is concerned with how a service is delivered. Process evaluations may assess the timing of an intervention, where it is occurring, the costs involved, the services offered, who is participating and who is facilitating the intervention (Hall, 2009). The purpose of process evaluation is to identify areas that are working well and areas that may require or benefit from change in order to enhance service delivery.

Process evaluations may be used to answer key questions such as:

- What were the demographic and clinical characteristics of clients?
- Are all service providers administering the program in the same way?
- Has the program or training been implemented as planned (i.e., achieved model fidelity)?
- Has the intended population been targeted effectively?
- Have collaborative links with other programs or service providers been successfully established?

Process evaluations are unable to answer the question of whether or not an intervention has had an impact on participants.

Impact evaluation

Impact evaluations are the most common evaluation type (Administration for Children and Families, 2003). Impact evaluations measure whether an intervention has an effect on participants in accordance with the intervention's aims and objectives. For example, an impact evaluation on a parent education program would assess whether participating parents had improved their parenting skills, such as a decrease in the use of physical punishment in favour of less violent and consistent parenting practices (Holzer et al., 2006). Any change in participants is assumed to result from participating in the program.

Key message: Impact evaluations measure the effect a program has on participants.

Outcome evaluation

Outcome evaluation assesses the long-term effectiveness of an intervention program. The difference between an impact and an outcome

evaluation is that an impact evaluation looks at direct aims, such as improving parenting skills, whereas an outcome evaluation considers underlying goals, such as child abuse prevention (Holzer et al., 2006). Outcome evaluations may investigate whether the direct aims of a program are accurate. For example, outcome evaluations may assess whether the increases in parenting skills gained from a parent education program resulted in the reduction of child maltreatment incidences (the overarching goal or purpose of the program). Impact and outcome evaluations are often described interchangeably; however, only outcome evaluations consider the underlying and long-term goal of an intervention.

Key message: The difference between an impact and an outcome evaluation is that an impact evaluation looks at direct aims, such as improving parenting skills, whereas as an outcome evaluation considers underlying goals, such as child abuse prevention.

When is a program ready for evaluation?

Not all programs are ready to be evaluated; for example, if the program does not have a clear set of goals, adequate data are not available on program performance, there are no clear guidelines on how the program should be delivered, or there is disagreement among program stakeholders about the nature of the program (Hall, 2009).

In order to assess the readiness of a program for evaluation, an intensive analysis prior to evaluation should be undertaken. This may involve: reviewing program documents; conducting site visits to view the program in operation; interviewing program clients, facilitators and managers; and holding discussions with key stakeholders (Hall, 2009). A process evaluation can be excellent preparation for a subsequent impact or outcome evaluation. Once it is established that the program is ready for evaluation, researchers need to design a rigorous impact or outcome evaluation methodology.

Designing a rigorous impact or outcome evaluation

As the main goal of an evaluation is to indicate whether a program is effective or not, it is important that an evaluation is conducted appropriately. There

are three elements that represent the gold standard for a rigorous evaluation: pre- and post-test designs, a comparison group and follow-up testing.

Pre- and post-test designs

Pre- and post-test designs assess participants “before” and “after” a program in order to ascertain whether participants have changed according to program goals (Chalk & King, 1998). Testing participant skills before an intervention service (pre-test) allows for a useful comparison of tests made after service participation (post-test) and can tell us how much change has occurred during the time period (even though it cannot be determined whether the change was a result of the intervention or not). For example, at the end of a parenting education course, participants might rate highly on knowledge about the importance of reading to children to enhance child literacy but, without a pre-test, we don't know if the parenting education course increased their knowledge or they already knew about the importance of reading to children before they started the course.

Comparison group

A comparison group is a group of people who are not involved in the intervention program being evaluated. Comparison groups are used to compare the outcomes of service participants with non-participants. Using a pre- and post-test design on its own is insufficient for determining program impact, as other variables or factors that cannot be ruled out may have influenced the result. A program is considered effective if the outcomes for the program group were significantly better than those for the comparison group. For example, pre- and post-testing of participants in a grief and loss counselling program might show change over time, with a decrease in depression at the end of the intervention. Without a comparison group, however, it would not be known if their improvement was a result of the program or whether their depression would have gotten better over time without any intervention.

The most effective method for obtaining a comparison group is to randomly allocate people into the program being evaluated and the comparison group, which is commonly known as an experimental design or a randomised control trial.

Box 1: Quasi-experimental designs vs experimental designs

There are two main methods for obtaining a comparison group for evaluation: quasi-experimental designs and experimental designs. Quasi-experimental designs are the most common way of identifying a comparison group. The method may involve recruiting people from a service waiting list with similar characteristics and experiences to the group participating in the program or service. However, as individuals are not assigned randomly to the two groups, there can be differences between the participant and comparison groups before the start of the intervention service that were not apparent to researchers.

Often referred to as a randomised control trial, experimental designs randomly allocate people to the intervention program and the comparison group. Experimental designs are considered more rigorous than quasi-experimental designs as they have the strongest degree of experimental control over the program environment. The method therefore enables a better chance of determining the effect of participation in a program (Chalk & King, 1998; Nixon, 1997).

Follow-up testing

Follow-up testing is needed for assessing whether successful outcomes of an intervention extend beyond the short-term. For example, participants may have increased knowledge about parent strategies at the end of a two-day course. If, after one month, they have not retained the skills learnt at the course, the program is unlikely to be effective in enhancing parenting skills. To determine whether an intervention has a lasting effect, an evaluation will need to conduct follow-up assessments on the same outcome measures. This can be referred to as a longitudinal research design.

Avoiding design flaws

To avoid research design flaws, it is important that evaluations only measure the influence of the intervention program. Findings in an evaluation should not be swayed by other events or circumstances in participants' lives. Key questions to ask when evaluating outcome measures include:

- Are the participants also attending another program or service?

- Do participants differ in their social supports or family arrangements?
- Do participants have a greater motivation to change than non-program participants in the comparison group?
- Are there differences between program participants and non-program participants in terms of the severity of problems?
- Are people who are assessed as needing welfare services likely to be different from the general community in terms of personal characteristics such as income, education or health? (Holzer et al., 2006)

Action research

Action research is an alternative approach to program evaluation. It may be used as a form of process evaluation and consists of integrating basic research, program development and evaluation into a continuous, cyclical process (Tomison, 2000). Action research actively involves input from program participants to help improve practice. By using action research, service providers and evaluators can regularly update and change aspects of the intervention that are deemed ineffective. Action research cannot measure the outcomes or impact of program participation.

Key message: Action research integrates basic research, program development and evaluation into a continuous, cyclical process. Action research may be used during a process evaluation or to inform continuous improvement. Action research cannot measure the outcomes of program participation.

Preparing for an evaluation

Evaluations can be complex and difficult to implement. To optimise their effectiveness, service providers need to ensure that the intervention is adequately documented. In preparing for an evaluation, the following aspects of an intervention should be clearly identified:

- the problem being addressed;
- a clear description of the program and the program logic (including the assumptions used to identify the target population, program aims and assumptions about program outcomes);
- the stages of implementation;
- the client referral, screening and baseline assessment processes;

- the client population (characteristics of the participant population);
- the organisational structure of the program; and
- a statement of measurable objectives (Chalk & King, 1998; Hall, 2009; Tomison, 2000).

Program evaluation in the child welfare context

Developing an evaluation with a rigorous research design can be problematic when applied to an intervention program in the child welfare and child protection system. Ethical and practical measures must be considered. In relation to comparison groups, researchers need to consider whether it is ethical for a comparison or control group to be prevented from accessing services. For example, when targeting a specific group of disadvantaged families, it could be unethical to provide services to some families but not to others simply to gain a more rigorous comparison group. The counter-argument to this is that participation in a new program is also a risk, as it is not known whether the program will benefit or further harm participants.

An example of a program that had unforeseen negative consequences for participants was the Scared Straight program in the United States. The program involved young offenders visiting prisons in an attempt to deter future offending; however, a later review found that not only did it fail to deter future crime, it actually led to more offending behaviour for some young people (Petrosino, Turpin-Petrosino, & Buehler, 2002). To assist in minimising the risks to participants and non-participants of any new intervention program, it is important that evaluations have ongoing measures that monitor the potential positive and negative effects of the intervention.

Outcome evaluations with a randomised comparison group (experimental design) are the optimal form of program evaluation; however, for ethical reasons, randomised comparison groups are often unobtainable. The use of waiting list participants as a matched comparison group is generally seen as preferable by service providers (Farmer, Burns, Guiles, Behar, & Gerber, 1997).

Other ethical issues that evaluation researchers may need to consider include confidentiality issues, mandatory reporting and gaining informed consent (Holzer et al., 2006). There may also be numerous other factors affecting families deemed to be at

risk of child maltreatment that may not be readily controlled within an evaluation study (Tomison, 2000). Service providers attempting to evaluate a particular intervention service may not be able to “control” or “contain” the influence of participant circumstances beyond the program (such as changes in employment or home environment).

Conclusion

There is no “perfect” way to conduct a program evaluation in the real world. The most effective evaluations are those with a rigorous research design that assess the impact or outcomes of an intervention program. In developing an evaluation, researchers should be aware of the limitations within their research design in order to reduce the influences of such limitations and enhance the validity of the evaluation.

Additional reading and research available from the clearinghouse website

Holzer, P. J., Higgins, J., Bromfield, L. M., Richardson, N., & Higgins, D. J. (2006). *The effectiveness of parent education and home visiting child maltreatment prevention programs* (Child Abuse Prevention Issues No. 24). Melbourne: National Child Protection Clearinghouse. Retrieved 3 November 2009, from <www.aifs.gov.au/nch/pubs/issues/issues24/issues24.html>.

Tomison, A. M. (2000). *Evaluating child abuse prevention programs* (Child Abuse Prevention Issues No. 12). Melbourne: National Child Protection Clearinghouse. Retrieved 3 November 2009, from <www.aifs.gov.au/nch/pubs/issues/issues12/issues12.html>.

References

Administration for Children and Families. (2003). *The program manager's guide to evaluation*. Washington, DC: ACF, US Department of Health and Human Services. Retrieved 3 November 2009, from <www.acf.hhs.gov/programs/opre/other_research/pm_guide_eval/reports/pmguide/pmguide_toc.html>.

Chalk, R., & King, P. A. (1998). *Violence in families: Assessing prevention and treatment programs*. Washington, DC: National Academy Press.

Farmer, E. M. Z., Burns, B. J., Guiles, H. B., Behar, L., & Gerber, D. I. (1997). Conducting randomized clinical trials in children's mental health: Experiences and lessons from one venture. In C. T. Nixon & D. A. Northrup (Eds.), *Children's mental health: Vol. 3: Evaluating mental health services: How do programs for children "work" in the real world?* Thousand Oaks: Sage.

Hall, R. (2009). Evaluation principles and practice. In G. Argyrous (Ed.), *Evidence for policy and decision-making*. Sydney: UNSW Press.

Holzer, P. J. (2007). *Defining the public health model for the child welfare services context* (Resource Sheet No. 11). Melbourne: National Child Protection Clearinghouse. Retrieved 3 November 2009, from <www.aifs.gov.au/nch/pubs/sheets/rs11/rs11.html>.

Holzer, P. J., Higgins, J., Bromfield, L. M., Richardson, N., & Higgins, D. J. (2006). *The effectiveness of parent education and home visiting child maltreatment prevention programs* (Child Abuse Prevention Issues No. 24). Melbourne: National Child Protection Clearinghouse. Retrieved 3 November 2009, from <www.aifs.gov.au/nch/pubs/issues/issues24/issues24.html>.

Nixon, C. T. (1997). How does evaluation of mental health services for children work in the real world? In C. T. Nixon & D. A. Northrup (Eds.), *Children's mental health series: Vol. 3. Evaluating mental health services: How do programs for children "work" in the real world?* Thousand Oaks: Sage.

Petrosino, A., Turpin-Petrosino, C., & Buehler, J. (2002). “Scared Straight” and other juvenile awareness programs for preventing juvenile delinquency. *Cochrane Database of Systematic Reviews*, 2002(2).

Tomison, A. M. (2000). *Evaluating child abuse prevention programs* (Child Abuse Prevention Issues No. 12). Melbourne: National Child Protection Clearinghouse. Retrieved 3 November 2009, from <www.aifs.gov.au/nch/pubs/issues/issues12/issues12.html>.

Author

Alister Lamont is a Research Officer for the National Child Protection Clearinghouse at the Australian Institute of Family Studies.

Acknowledgement

The authors would like to thank Nick Richardson and Adam Tomison for their work on a previous edition of this Resource Sheet.



Australian Government
Australian Institute of
Family Studies

First edition, *Evaluating child abuse prevention programs* (Nick Richardson and Adam Tomison), published December 2004.

Second edition, *Evaluating child abuse and neglect intervention programs* (Alister Lamont), published November 2009.

ISSN 1448-9112 (Online)

© Commonwealth of Australia 2009

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968*, no part may be reproduced by any process without prior written permission from the Commonwealth. Requests and inquiries concerning reproduction and rights should be addressed to the Commonwealth Copyright Administration, Attorney-General's Department, Robert Garran Offices, National Circuit, Barton ACT 2600 or posted at <www.ag.gov.au/cca>.

The Australian Institute of Family Studies is committed to the creation and dissemination of research-based information on family functioning and wellbeing. Views expressed in its publications are those of individual authors and may not reflect Australian Government or Institute policy, or the opinions of the Editors or of the Director.

Australian Institute of Family Studies
Level 20, 485 La Trobe Street, Melbourne VIC 3000 Australia
Phone: (03) 9214 7888 Fax: (03) 9214 7839
Internet: <www.aifs.gov.au>