Challenges in Evaluation of Intervention Programmes with Children: Reflections from a Research(er) Perspective

Gabriela Mişca*

Abstract. *This paper will review some of the challenges that researchers, evaluators and designers of intervention programme for vulnerable children may encounter. The challenges considered in the present paper are around methodological issues, specifically the issues of validity of outcome measures ; and service users' involvement, particularly children's perspectives, in research evaluation. The author draws on extensive experience of conducting research on evaluating the outcomes of interventions for looked after children in the UK and the article invites the reader to reflect upon some of these challenges.*

Keywords : evaluation research, methodology challenges, stakeholders

Setting the scene : Why research evaluation of intervention programmes for looked after children ?

Evidence based practice occupies a central place in current social work practice worldwide (Adams, Dominelli and Payne, 2009). In respect to children and young people's services, rigorous research evaluation(s) of programmes for children and young people is high Government's agenda in the UK (DCSF, 2010). Different policy initiatives for looked after children in the UK (for various examples from England see http : //www.dcsf.gov.uk/ everychildmatters/resources-and-practice) are evaluated at the stage of piloting and full implementation is conditional on positive results of such evaluations. Usually, such evaluations are commissioned by Government and undertaken by research teams/organisations, both private or University based.

Recently, a „Centre for Excellence and Outcomes in Children and Young People's Services" has been created (www.c4eo.org.uk) with the aim of providing Local Authorities and their Children's Trust partners in the UK with evidence that they can rely on. It is pledged that this will include : the best research available ; the most robust data ; examples of effective local practice, drawn from across the country (C4EO, 2010). The stress is on

* Keele University, School of Public Policy and Professional Practice & Research Institute for Life Course Studies, Staffordshire ST5 5BG, UK, Tel : +44 (0)1782 734290, Fax : +44 (0)1782 734069, E-mail : g.m.misca@keele.ac.uk

research rigour in demonstrating what it will be „deemed" as „best practice" and claims of excellence in the field will be required to be backed up by robust data.

This trend has put considerable pressure on independent providers of intervention programmes for looked after children to move towards considering the evaluation of their work and engaging with research on various levels. The Local Authorities are „buying" the services of such independent providers of programmes for looked after children and, as such, are highly interested in the evidence of success of the programmes they offer, as well as in their cost-effectiveness.

For example, in recent years there has been a growth in the number of independent agencies providing specialist/therapeutic treatment in residential and/or foster setting for such children in the UK. Children and young people entering public care today in the UK are likely to have suffered abuse and neglect and they present elevated levels of emotional and behavioural difficulties. While the need for efficient therapeutic service provision for such children and young people has been acknowledged (Ward *et al.*, 2004) there are a small number of „looked after children" with a high level of support needs that local authorities have difficulty providing best care using „mainstream" placements. This includes children who are considered to be „hard to place" because of the severity of their emotional and behavioural difficulties, assumed to stem from their early experiences of abuse and neglect and/or the high member of placements experienced since entering care. Seizing the gap in the provision of specialist services for these children by local authorities, several independent agencies across the country are now offering various residential or foster care programmes for such children. However, notwithstanding the proliferation of such agencies, there is little evidence of whether and to what extent these placement achieve better outcomes for children and if so which kind of supports are crucial to children's recovery. Examination of such evidence is of importance to local authorities who pay enhanced fees for these services and to agencies for monitoring their services.

The author draws on wide-ranging experience in undertaking evaluation research of both foster and residential care based intervention programmes for looked after children in the UK and highlights some of the challenges encountered, inviting the reader to reflect on such challenges and ways to overcome them.

Methodological challenges : Validity of methods used for evaluating outcomes

This section will provide an overview of measurement validity relevant to designing and constructing instruments used in social research with special reference to evaluating outcomes of intervention programmes.

What is the validity of an assessment instrument ?

Validity is vitally important when designing instruments for evaluation as well as interventions such as treatments and programmes. Validity is a term used to describe the extent to which an instrument measures what it claims to measure. In basic terms, an assessment method/instrument's validity concerns what the instrument measures and how well it does so (Anastasi and Urbina, 1997) and as such an evaluation of its merits should be guarded in accepting the instrument's name as sound proof of what it claims to measure.

To use a relevant analogy, when translating the concept of „recovery from child abuse and neglect" into carrying out a real intervention – that is „a recovery programme for

children who suffered abuse and neglect" – any independent evaluators need to be concerned about how well the translation was made by the programme designer/developer. Moreover, when assessing the merits of a method of evaluating the outcomes of such intervention, attention needs to be paid to the alignment between the concept(s) supporting the intervention, the programme/intervention itself and the measurement of the outcomes (Macdonald, 2001).

Validity is necessary but not sufficient on its own as a condition for a measure to perform satisfactorily. A second consideration is reliability, which refers to the ability of the measure to produce the same results under the same conditions (Field, 2009). For example, if an instrument is reliable a respondent should interpret the meaning of the questions the same way each time is administered. Ways to demonstrate reliability involve empirical data (such as test-retest reliability ; Field, 2009).

Demonstrating validity : What kinds of validity ?

Validity of an instrument or method of evaluation is a matter of degree, an attribute that exists along a continuum, not a simple question of „valid or invalid" (Colton and Covert, 2007). Thus, there are a variety of categories that are used to describe the validity of measures. However, it is important to elucidate that although there are distinctive types of validity (as shown below) they are all means of answering the same question : „is the instrument measuring what is supposed to measure ?".

For example, *face validity* reveals the degree to which an instrument appears to be an appropriate measure for its intended purpose (Colton and Covert, 2007). In face validity, an expert looks at an instrument to see whether „on its face" seems a good translation of the construct it aims to measure. It is essentially a judgement call and as such it may be seen as a less rigorous approach than other types of validity, such as criterion validity which is based on empirical data collected (see below). Nevertheless, any research process relies on expert and peer judgments. Face validity, particularly when provided by an independent evaluator, can provide useful information about the instrument as a whole and the extent the instrument meets its intended purpose (Gilbert, 2008).

In *construct validity* the instrument is checked against the relevant theoretical concept that is meant to operationalize and used to determine to what extent the measure covers the full range of the concept's meaning and the accuracy with which the theoretical construct is reflected in the instrument (Colton and Covert, 2007). Constructs or concepts cannot be observed directly, therefore tangible and observable ways to measure these have to be found (Macdonald, 2001) and this process is refer to as operationalization.

Operationalizing constructs in the social sciences is not always a straightforward process. Therefore a major threat to the instrument's validity (and hence the information collected using that instrument) is when respondents interpret an item or question in their own way and respond to that interpretation which may be very different from the intended meaning. The theoretical underpinning of an intervention programme is of vital importance and as such the reliability of the theoretical foundation of an intervention programme should be integral part of the overall evaluation process.

Qualitative approaches to assessing instrument validity, such as these above, are evaluative. They are based on literature reviews of the relevant topic and provide evidence to the extent of which the instrument is measuring the intended constructs. Experts review the instrument and based on their professional judgment indicate whether each item appears to measure what is intended to measure.

However, there are ways of establishing instrument validity that rely on empirical data and as such offering a „hard" test for their validity. For example, in establishing *criterion validity* a comparison is made between a measure and an external standard and it is particular important if an instrument will predict future behaviours, such as how individuals will perform as a result of receiving an intervention (Colton and Covert, 2007). Usually this is accomplished by comparing the scores of the measure to be validated against the score of a measure that is already validated.

In *predictive validity*, the measure's ability to predict something that it should be theoretically able to predict is evaluated (Gilbert, 2008). This involves statistical analyses on data collected using both the instrument to be validated and a well established instrument. It is recognised as strong evidence for/against a proposed instrument's validity as it is based on actual data analyses.

In *concurrent validity*, the measure's ability to distinguish between groups that it should theoretically be able to distinguish between is assessed. This type of validity involves comparing data collected using the instrument to be validated with different groups that are theoretically considered to be different, to ascertain if the instrument is able to distinguish between the two groups/populations (Field, 2009).

In *convergent validity*, the degree to which the operationalization is similar to (converges on) other operationalizations that it theoretically should be similar to is assessed (Colton and Covert, 2007). It usually involves comparison between two instruments based on data collected with both. Correspondingly, in *discriminant validity*, the degree to which the operationalization is not similar to (diverges from) other operationalizations that it theoretically should be not be similar to is examined (Field, 2009). This also involves comparisons of data collected with two measures that theoretically are considered to be different.

This section highlighted some of the methodological challenges in research evaluation of intervention programmes, specifically outlining the ways for establishing an instrument's validity in evaluating outcomes and their relative strength dependant on whether the validity is reviewed qualitatively or quantitatively (using empirical data collected). The challenges underline the necessity that research evaluations are undertaken by multi-disciplinary teams so that various expertises can be drawn into the evaluation process – for example, if an evaluation aims to assess the cost-effectiveness of an intervention.

„Service user" and „stakeholder" challenges

„Independent programme evaluation" : For whom and who is paying for it ?

It is often disputed, albeit not always in an open manner, that „evaluation research" is somewhat inferior to the „blue sky", academic research. There are several assumptions that fuel such connotations associated with evaluation research. For example, sceptics may ask how independent is an evaluation research that is commissioned and paid for by the Government or an independent agency upon own practice ?

The author argues that the answer to such question(s) is positioned in the understanding of the research process itself and the ethical code of research practice to which researchers adhere to. Nevertheless, there are often challenges that researchers need to overcome when a conflict of interest may appear, such in the case of an independent evaluation paid for by the agency to be evaluated itself. Depending on the agency's familiarity with the research

process, ethics and overall research purpose, it may not be uncommon that the understanding of what evaluation research does is that it will show that „what we do works" (after all, they are paying for it?).

In the current climate of diminishing and increasingly competitive funding for academic research, particularly in the UK, it may not be wise to dismiss evaluation research – particularly in the social work area where research is acutely needed for both practice and academic benefits (Corby, 2006; Misca, 2009). Instead, researchers may find that they need to invest time and resources in „coaching" the agency in understanding of what research does (and does not) and how it operates. It may be an important priority that at the stage of research negotiations, open discussions take place about each party's expectations *vis-à-vis* the research process and outcomes, and protocols are put in place for the eventuality that the outcome of research evaluation is not the one expected by the commissioning agency (i.e.: in the case of Government funded evaluation – will the initiative be cancelled?)

Involving service users and stakeholders in the research evaluation

It is often overlooked in evaluation research who the „service users" are from the research point of view. Is the „service user" the agency and/or Government that commissions/pays the research and thus intend to use its outputs? Are the service users the people working in the researched agency or the recipients of the services (i.e.: the children who are receiving an intervention)? Are the ultimate „service users" the academic/research community, as a piece of research is expected to contribute to the overall knowledge base?

There are no definitive answers to these dilemmas. However, considering carefully the „stake holding" of each category of participants, may influence the design of the evaluation research. For example, a „participatory research design" and „action research" designs (Flick, 2009) may be more successful when an agency invite a research evaluation into own practice. In such instances, involving the agency and its service users in the entire research process will not only „smooth the way" but will provide an important learning experience for the participants as well as researchers. In the case of Government commissioned evaluation, the research may be perceived as an audit process by the participants and thus in turn be received with great resistance hindering the progress of research. However, recent Government initiatives favour the participatory style research and service user involvement. In addition, it is often considered good practice to involve all the „stakeholders" in the research process, for example by establishing a research advisory group that will have representative from the categories mentioned above, so all parties can have an input in the research process.

Involving children in evaluation research

Linking back to the first section of this paper and to the overarching theme of evaluation of interventions with vulnerable children, this last section will highlight some of the challenges in involving children in evaluation research (Fraser *et al.*, 2004; Greene and Hogan, 2005; Greig *et al.*, 2007).

There have been considerable advances in research involving children as active participants over the past decade. However, balancing children's rights *versus* adult's rights is not an easy act. The dominant social construction of childhood is still that of dependency and „not knowing"; a dependency that lessens with age but nevertheless, assuming that adults will make decisions and act in the best interest of the child. The children's rights approach

challenges the assumption that adults are always more competent in arriving at best decisions for children. However, the adults making decisions can disregard a child's view by making the assumption that a child lacks competence (Aldgate *et al.*, 2006).

The social construction of childhood as dependency has been challenged by sociological studies of childhood since 1990s (James and Prout, 1997) and alternative ways of constructing childhood have been put forward. The „Every Child Matters" agenda in the UK (www.everychildmatters.gov.uk) urges the promotion of children's participation in planning for their own lives and collectively. The concept of active citizenship gives young people responsibility for their own behaviour which in practice means giving young people proper information for decision-making and allowing them to participate in decisions that affect them (McLeod, 2008).

Conclusion

The present paper has been intended as an invitation to researchers and practitioners to reflect upon some of the challenges of the evaluation research, both generally and in relation to evaluation research involving vulnerable children. The methodological challenges reflected upon here concern the validity of the outcomes measures and the alignment between theoretical underpinning – intervention programme – evaluation. Furthermore, the author considered the implications of various categories of stakeholders and service users in the evaluation research with a particular view towards involving children in evaluation research.

Evaluation research it is not an uncomplicated process, however. The sources of knowledge are contrasting : while a traditional practice, relies on accumulated subjective experience with individual cases over many years (Sudbery *et al.*, 2009) ; evidence-based practice and evaluation research requires well-designed and standardized measures and the emphasis is on „the data show that...". Assumptions about outcomes in traditional practice are based on beliefs reinforced by practice, whereas evidence-based practice starts from a position of scepticism that beneficial outcomes must be empirically demonstrated, not assumed (Chaffin and Friedrich, 2004). Ultimately, evidence-based practice aims to bring services in line with the best available science and promote practices that have been demonstrated to be safe and effective.

As a researcher, the author believes that both research and practice should be challenged and critically evaluated. There are tremendous opportunities from mutual learning from the process of research evaluating practice, as research is challenged too, to develop new designs and methods of evaluation and integrate the practice knowledge in essentially shaping the knowledge base.

Note

1. Since the writing of this article, the new UK government department responsible is the Department for Education, http : //www.education.gov.uk/

References

Adams, R., Dominelli, L. and Payne, M. (eds.) (2009) *Social Work : Themes, Issues and Critical Debates*, Basingstoke : Palgrave Macmillan.

Aldgate, J., Jones, D., Rose, W. and Jeffery, C. (eds.) (2006) *The Developing World of the Child*. London : Jessica Kingsley.

Anastasi, A. and Urbina, S. (1997), *Psychological testing*, Prentice-Hall International.

Ardila Centre for Excellence and Outcomes in Children and Young People's Services (2010). Available : www.c4eo.org.uk (last visited : 31/08/2010).

Chaffin, M. and Friedrich, B. (2004) Evidence-based treatments in child abuse and neglect, *Children and Youth Services Review*, 26 (2004), 1097-1113.

Colton, D. and Covert, R. (2007) *Designing and constructing instruments for social research and evaluation*, San Francisco : Wiley & Sons.

Corby, B. (2006) *Applying Research in Social Work Practice*, Berkshire : Open University Press.

Department for Children Schools and Families (2010), *Every Child Matters*, Available : http : // www.dcsf.gov.uk/everychildmatters/ (last visited 31/08/2010).

Field, A. (2009) *Discovering Statistics Using SPSS*, 3rd Edition, London : Sage.

Flick, U. (2009) *An Introduction to Qualitative Research*, 4th Edition, London : Sage.

Fraser, S., Lewis, V., Ding, S., Kellett, M. and Robinson, C. (2004) *Doing Research with Children and Young People*. London : Sage.

Gilbert, N. (2008) *Researching Social Life*, 3rd Edition, London : Sage.

Greene, S. and Hogan, D. (2005) *Researching Children's Experience : Approaches and Methods*, London : Sage.

Greig, A.D., Taylor, J. and MacKay, T. (2007) *Doing Research with Children*, 2nd Edition, London : Sage.

James, A. and Prout, A. (1997) *Constructing and Reconstructing Childhood : Contemporary Issues in the Sociological Study of Childhood*, 2nd Edition, London : Routledge.

Lewis, V., Kellett, M., Robinson, C., Fraser, S. and Ding, S. (2004) *The Reality of Research with Children and Young People*, London : Sage.

Macdonald, G. (2001) *Effective Interventions for Child Abuse and Neglect : An Evidence-Based Approach to Planning and Evaluating Interventions*, Chicester : Wiley.

McLeod, A. (2008) *Listening to Children : A Practitioner's Guide*, London : Jessica Kingsley.

Misca, G. (2009) Perspectives on the Life Course : Childhood and Adolescence. In Adams, R., Dominelli, L. and Payne, M. (eds.), *Social Work : Themes, Issues and Critical Debates*, Basingstoke : Palgrave Macmillan, 116-128.

Sudbery, J. Shardlow, S. and Huntington, A. (2009) To Have and to Hold : Questions about a Therapeutic Service for Children, *British Journal of Social Work*, 1-19.

Ward, H., Holmes, L., Soper, J. and Olsen, R. (2004), *Costs and Consequences of Different Types of Child Care Provision*, Loughborough : Centre for Child and Family Research.