Children and Youth Services Review 32 (2010) 1483–1486

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth

Measurement issues in home visitation: A research note

Craig W. LeCroy [a,*], Judy Krysik [b,1]

[a] Arizona State University, School of Social Work, 340 N. Commerce Park Loop, Suite 250, Tucson, AZ 85743, United States
[b] Arizona State University, School of Social Work, Mail Code 3920, 411 N. Central Ave., Suite 800, Phoenix, Arizona 85004-0689, United States

ARTICLE INFO

Article history:
Received 30 January 2010
Received in revised form 24 May 2010
Accepted 25 May 2010
Available online 31 May 2010

Keywords:
Measurement
Home visitation
Child abuse and neglect prevention
Outcome
Evaluation
Outcome assessment
Program evaluation

ABSTRACT

The field of home visitation to prevent child abuse and neglect provides a good example of issues in outcome measurement that have not undergone sufficient critique. Some frequently used measures being used as outcome assessments have not been specifically designed for outcome measurement and therefore have limitations. In particular, some of these measures are not well equipped to document changes that are intended to result from intervention programs. Outcome measurement in home visitation can be improved with more attention to measurement issues.

The measurement of concepts that cannot be measured directly is replete with methodological challenges. Although outcome measurement in research and program evaluation has improved in recent years, problems remain that have yet to be addressed. The field of home visitation to prevent child abuse and neglect provides a good example of issues in outcome measurement that have not undergone sufficient critique. Research has focused on substantiated CPS cases as an outcome indicator, however, increasingly this is being recognized as a seriously flawed measure (Harding, Galano, Martin, Huntington, & Schellenbach, 2007; Olds, Eckenrode, & Kitzman, 2005). Researchers confronting the challenges of using CPS reports have sought alternative measures for outcome assessment that represent constructs associated with child abuse. For instance, measures of maternal depression and quality of the home environment are used as alternatives to CPS reports. In a well-known comprehensive review of home visitation programs, Gomby (1999) noted that over 100 different measures were used to evaluate outcomes in various studies of home visitation. Some of the more common standardized outcome measures used in the evaluation of home visitation include: the Home Observation for Measurement of the Environment (HOME) instrument (e.g., used in Duggan et al., 2004), the Bayley Test of Maternal Responsiveness (used in Heinicke et al., 2000), the Conflict Tactics Scale (used in DuMont et al., 2008), the Parenting Stress Index (e.g., used in Kemp et al., 2008), the Beck Depression Inventory (Ammerman et al., 2007), and the North Carolina Family Assessment Scale (used in de la Rosa, Perry, Dalton, & Johnson, 2005). Although all of these researchers are interested in home visitation outcomes in programs with similar goals, they have selected very different outcome measures to examine program impact.

In spite of the plethora of constructs considered and measures being used, a central thesis of this paper is that some of the frequently used measures selected for outcome evaluation are not relevant for use in an applied setting (Ogles, Lambert, & Fields, 2002). At the core of the issue is that some of these measures were not specifically designed as *outcome* measures and, therefore, have limitations when used to document client change from intervention programs. For example, some measures were created within cross sectional studies and, therefore, the reliability and validity estimates are limited to one point in time. Some measures were developed to measure one concept such as parenting stress, or to predict the risk of child abuse and neglect, and were not designed to measure program outcomes *per se*. Recent meta analyses in the field (Geeraert, Noortgate, Grietens, & Onghena, 2004; Sweet & Appelbaum, 2004) have also raised similar measurement issues. The following article will discuss some of these limitations and suggest why measures designed as outcome measures are preferable to some of the frequently used measures used in home visitation evaluations.

1. Home visitation and the measurement of outcomes

Since the 1980s there have been a number of outcome studies of home visitation programs aimed at preventing child abuse and

* Corresponding author.
E-mail addresses: craig.lecroy@asu.edu (C.W. LeCroy), judy.krysik@asu.edu (J. Krysik).
[1] Tel.: +1 602 496 0086; fax: +1 602 496 0960.

0190-7409/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.childyouth.2010.05.005

neglect (e.g., Barth, 1991; Fraser, Armstrong, Morris, & Dadds, 2000; Olds, Henderson, Tatelbaum, & Chamberlain, 1986). Although a plethora of measures have been used in these studies, some measures have been used more than others. Two instruments in particular, the Parenting Stress Index or PSI designed for the early identification of parent/child problems that impede normal development (Abdin, 1995), and the Child Abuse Potential (CAP) Inventory designed as a screening tool to identify risk of child physical abuse (Milner, 1994) have been widely used. Some studies have shown positive effects with the PSI (e.g., Fraser et al., 2000), whereas other studies have shown no effects (e.g., Duggan et al., 1999). Similarly, the CAP Inventory has shown mixed effects with some outcome studies reporting no differences between treatment and control groups (e.g., Barth, Hacking, & Ash, 1988; Black, Dubowitz, Hutcheson, Berenson-Howard, & Starr, 1995; Gondolf, 1997; MacMillan et al., 2005), and other studies showing positive effects (e.g., Ethier, Couture, Lacharité, & Gagnier, 2000).

Unfortunately, when changes are not detected on well-established outcome measures, it is often concluded that the program did not work. For example, Gondolf (1997) who used the CAP Inventory to evaluate the effects of child abuse prevention classes for court-ordered women concluded, "the findings fail to endorse the addition of parenting classes for woman batterers" (p.15). To demonstrate that an intervention is effective, however, one must measure change. A plausible alternative hypothesis to program failure then, is that the measure was inadequate in detecting the change that occurred, or it was measuring the wrong constructs. Hill and Lambert (2004) describe sensitivity to change as the degree that an instrument accurately reflects client changes that occur following participation in the intervention. If a measure is not good at detecting program-related changes, this might explain why the findings are mixed.

Of course, it is certainly plausible that the intervention did not work and reliably shows no change—and this is the conclusion most researchers arrive at when examining outcome studies that show no change. Further, variation in program implementation and training is likely to influence outcomes (LeCroy & Whitaker, 2005). The point of this article is to encourage critical examination of all findings and suggest that measurement issues may also be relevant and need to be carefully examined when interpreting findings. If the measure shows no change, it is unclear if the lack of change reflected by the measure is due to a property of the measure (method variance) or to the intervention. In measurement theory, a measure is assumed to be composed of the true score and random error. Method variance refers to the most concerning type of error, and that is systematic bias or systematic error that also may be captured in a measure (Woszczynski & Whitman, 2004). As noted by Eddy, Dishion, and Stoolmiller (1998), "the topic of method variance has been neglected in intervention studies" (p. 61). These researchers demonstrate that for almost half of the measures used in the Adolescent Transitions Program, method variance accounted for approximately 50% of the variance. They conclude: "because method variance is likely a very large part of the variance in many psychological measures caution is warranted when drawing conclusions from data" (Eddy et al., 1998, p. 61). An examination of measures used in home visitation suggests method variance may be a methodological issue of some significance.

A lack of findings may be, in part, due to the sensitivity of the outcome measure to detect change. In one of the few studies to compare multiple methods of measuring change, Andrews and Dishion (1994) collected observational data of parent's ratings of negative engagement with their children and found significant reductions following the intervention. However, when they asked parents to respond to a 5-point interval scale about how well they monitored the behaviors of their child, they did not obtain statistically reliable effects. As this study demonstrates different measures of the same or related constructs may be more or less sensitive to the measurement of change.

2. Common measures used in home visitation research

2.1. The Parenting Stress Index

The PSI is a good example of how evaluations of home visitation programs may be attempting to measure outcomes that are less amenable to change. The PSI consists of 120 items with six child subscales, seven parent subscales, and a total scale score. An examination of the PSI reveals some items that relate to underlying personality and temperament characteristics. For example, the five items listed in Table 1 form the PSI child-related subscale for mood. These items may tap more directly into heritable traits that are believed to be stable and less amenable to change over time (Seligman, 2007). In recent years, evolutionary psychologists have been convincing in their studies of how extensive heritable traits are, and how resistant they are to change (Pinker, 2002). Furthermore, personality and temperament, even if not heritable traits, are often consider difficult to change. Seligman (2007) addresses this topic directly by discussing the research literature on what problems are amenable to change and what problems are not likely to change with psychological intervention. Intervention research has not examined this issue with much depth; yet this analysis can help us better understand intervention outcomes.

2.2. The CAP Inventory

The CAP Inventory is a widely used measure in home visitation studies. This instrument provides an example of two aspects that are problematic when attempting to measure change. For example, the 11 items from the CAP Inventory presented in Table 2 illustrates items that reflect static rather than dynamic characteristics. Having been abused by one's parents, for example, or having a physical handicap are static characteristics. That is, no amount or type of intervention is going to change one's status on these items over time. Although these items may be predictive of an outcome in the general population such as potential for child physical abuse, using them to evaluate a secondary prevention program in which the majority of participants are likely to have experienced them is implicitly contributing to the use of "outcomes" that cannot show change. Indeed, the presence of parental risk factors such as a childhood history of abuse is commonly used to determine eligibility for home visitation programs.

The measurement of change is best when based on dynamic rather than static constructs. Many of the items in the CAP Inventory represent more static constructs and do not contain any reference to time, whereas a dynamic construct includes items that show patterns of change over time. For example, the question, 'Have you ever been arrested?' is a static indicator. Rewritten with a reference to time, e.g., 'Have you been arrested in the past six months?' the question becomes a dynamic indicator. If our interest is in reflecting change in a construct over time then the manner in which the construct is measured should be able to capture such change (see Stoolmiller, 1994).

Measures used for outcome assessment are often missing an assessment of how the items contribute directly to a reflection of change. Bereiter (1963) issued sage advice long ago—create change items to create reliable measures of change. He suggested that when a measure is used at two points in time, the raw difference score for

Table 1
Example items from the Parenting Stress Index mood subscale.

My child cries and fuses.
My child seems to cry or fuss more often than most children.
When playing, my child doesn't often giggle or laugh.
My child generally wakes up in a bad mood.
I feel that my child is very moody and easily upset.

Table 2
Example items from the Child Abuse Potential Inventory.

I enjoy having pets.
I have a child who is clumsy.
My telephone number is unlisted.
I sometimes worry that I will not have enough to eat.
My parents did not really care about me.
I have a child who is sick a lot.
I have a physical handicap.
As a child I was abused.
I have a child who is slow.
Right now I am deeply in love.
As a child I was knocked around by my parents.

each item in the measure should be computed for each subject and then the set of item difference scores could be used to examine the value of the measure as an indicator of change. When the majority of items change in the same way across time, then the measure is a good measure of change. Alternatively, one could compute the weighted average or slope of the repeated measure for use as a variable in outcome analysis.

In fairness to the developers of the PSI and CAP Inventory, they developed measures that were more concerned with screening and assessment rather than documenting change. It is the application of these measures that has been problematic, rather than the measures.

Cultural relevance is a second issue in outcome evaluation that is illustrated by the CAP Inventory. Items such as "I enjoy having pets," may reflect an upper-income and ethnic bias. Many of the participants in home visitation programs have extremely low incomes, are transient, and have difficulty providing for their family's basic needs. For many, the reality of poverty is not conducive to pet ownership. Some participants of home visitation, for instance, those with a Mexican heritage, typically do not favor house pets. A second item from the CAP Inventory, "my telephone number is unlisted," also reflects a middle-class bias and does not take into consideration that a number of families rely on cellular telephones which have unlisted numbers. Fit with the program context is also an important consideration. The item shown in Table 2 that states "right now I am deeply in love," although undoubtedly intended to reflect stability and decreased risk for child physical abuse in the general population, tends to represent misplaced trust and elevated risk in a population of parents who fall in and out of love frequently with partners whom they barely know.

3. Conclusions and recommendations

In summary, the more items an outcome measure includes that relate to heritable traits, static conditions, and that are culturally biased or a poor fit with the program context, the less likely it will demonstrate program impact. This suggests that for some programs, such as home visitation, to most effectively measure program outcomes there is a need to use or develop outcome measures that are tailored to the context and are explicitly designed for the purpose of outcome evaluation. The measure should pertain to what the program purports to change. Additionally, instruments used for program evaluation, beyond increased relevance and sensitivity to change, should yield information that is immediately useful in practice. That is, the best outcome instruments would assist practitioners to identify areas of need and concern, as well as areas of strength that could be used in treatment planning.

One of behavioral psychology's significant contributions has been the development of objective and specific measures of change (Eddy et al., 1998). It is important to not lose this emphasis as outcome evaluation and evidence based practice becomes main stream. A vital need in intervention research is critical examination of measures. The effectiveness of interventions cannot be discovered without well designed measures that are sensitive to change. Lambert's work (Lambert, 2007; Vermeersch et al., 2004) provides a model as the OQ®-45.2 outcome instrument was designed for repeated measurement and the ability to show change. The authors instruct those completing the measure by stating: "Looking back over the last week, including today, help us understand how you have been feeling." This measure has been shown to produce total scores that are sensitive to change in clients over short time periods and can predict outcomes documenting improvers, no changers, and deteriorators—providing practitioners with feedback directly relevant to their work. Although there are books with measures and related psychometric information (see for example Fischer & Corcoran, 2007) and review articles that list commonly used measures and related psychometric information on specific issues (see for example DeVoe & Kantor, 2002), the consideration of measures for use in outcome evaluation should go beyond a standard review of validity and reliability. Evaluators would do well to consider criteria such as sensitivity to change, heritable traits, static versus dynamic indicators, method variance, and program context. With increased attention to measurement issues we can improve the quality and value of home visitation research.

References

Abdin, R. L. (1995). *The Parenting Stress Index.* Odessa, FL: Psychological Assessment Resources.

Ammerman, R. T., Bodley, A. L., Putnam, F. W., Lopez, W. L., Holleb, L. J., Stevens, J., et al. (2007). In-home cognitive behavior therapy for a depressed mother in a home visitation program. *Clinical Case Studies*, *6*, 161—180.

Andrews, D. W., & Dishion, T. J. (1994). The microsocial structure underpinnings of adolescent problem behavior. In M. Lamb, & R. Ketterlinus (Eds.), *Adolescent problem behavior* (pp. 187—207). Hillsdale, NJ: Erlbaum.

Barth, R. P. (1991). An experimental evaluation of in-home child abuse prevention services. *Child Abuse and Neglect*, *15*, 363—375.

Barth, R. P., Hacking, S., & Ash, J. R. (1988). Preventing child abuse: An experimental evaluation of the Child Parent Enrichment Project. *Journal of Primary Prevention*, *8*, 201—217.

Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. Harris (Ed.), *Problems in measuring change* (pp. 24—36). Madison: University of Wisconsin Press.

Black, M. B., Dubowitz, H., Hutcheson, J., Berenson-Howard, J., & Starr, R. H., Jr. (1995). A randomized clinical trial of home intervention for children with failure to thrive. *Pediatrics*, *95*, 807—814.

de la Rosa, I. A., Perry, J., Dalton, L. E., & Johnson, V. (2005). Strengthening families with first-born children: Exploratory story of the outcomes of a home visitation intervention. *Research on Social Work Practice*, *15*, 323—338.

DeVoe, E. R., & Kantor, G. K. (2002). Measurement issues in child maltreatment and family violence prevention programs. *Trauma, Violence, and Abuse*, *3*, 15—39.

Duggan, A., McFarlane, E., Fuddy, L., Burrell, L., Higman, S. M., Windham, A., et al. (2004). Randomized trial of a statewide home visiting program: Impact in preventing child abuse and neglect. *Child Abuse & Neglect*, *28*, 597—622.

Duggan, A. K., McFarlane, E. C., Windham, A. M., Rhode, C. A., Salkever, D. S., Fuddy, L., et al. (1999). Evaluation of Hawaii's Healthy Start Program. *The Future of Children*, *9*, 66—90.

DuMont, K., Mitchell-Herzfeld, S., Greene, R., Lee, E., Lowenfels, A., Rodriguez, M., et al. (2008). Healthy Families New York (HFNY) randomized trial: Effects on early child abuse and neglect. *Child Abuse & Neglect*, *32*, 295—315.

Eddy, M. J., Dishion, T. J., & Stoolmiller, M. (1998). The analysis of intervention change in children and families: Methodological and conceptual issues embedded in intervention studies. *Journal of Abnormal Child Psychology*, *26*, 53—69.

Ethier, L. S., Couture, G., Lacharité, C., & Gagnier, J. P. (2000). Impact of a multidimensional intervention programme applied to families at risk for child neglect. *Child Abuse Review*, *9*, 19—36.

Fischer, J., & Corcoran, K. (2007). *Measures for clinical practice and research: A sourcebook*, 4th ed Oxford: Oxford University Press.

Fraser, J. A., Armstrong, K. L., Morris, J. P., & Dadds, M. R. (2000). Home visiting intervention for vulnerable families with newborns: Follow-up results of a randomized controlled trial. *Child Abuse and Neglect*, *24*, 1399—1429.

Geeraert, L., Noortgate, W. V., Grietens, H., & Onghena, P. (2004). The effects of early prevention programs for families with young children at risk for physical child abuse and neglect: A meta-analysis. *Child Maltreatment*, *9*, 277—291.

Gomby, D. S. (1999). Understanding evaluations of home visitation programs. *The Future of Children*, *9*, 27—43.

Gondolf, E. W. (1997). *An experimental evaluation of child abuse prevention courses for court-ordered women batterers.* A report for the Child Trust Fund of Pennsylvania. Harrisburg, PA: Pennsylvania Department of Welfare.

Harding, K., Galano, J., Martin, J., Huntington, L., & Schellenbach, C. J. (2007). Healthy Families America effectiveness: A comprehensive review of outcomes. *Journal of Prevention and Intervention in the Community*, *34*, 149—179.

Heinicke, C. M., Goorsky, M., Moscov, S., Dudley, K., Gordon, J., Schneider, C., et al. (2000). Relationship-based intervention with at-risk mothers: Factors affecting variations in outcome. *Infant Mental Health Journal*, *21*, 133—155.

Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (pp. 84—137)., 5th ed. NY: Wiley.

Kemp, L., Harris, E., McMahon, C., Matthey, S., Vimpani, G., Anderson, T., et al. (2008). Miller early childhood sustained home-visiting (MECSH) Trial: Design, method and sample description. *BMC Public Health*, *8*, 1—12.

Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, *17*, 1—14.

LeCroy, C. W., & Whitaker, K. (2005). Improving the quality of home visitations: An exploratory study of difficult situations. *Child Abuse and Neglect: The International Journal*, *29*, 1003—1013.

MacMillan, H. L., Thomas, B. H., Jamieson, E., Walsh, C. A., Boyle, M. H., Shannon, H. S., et al. (2005). Effectiveness of home visitation by public-health nurses in prevention of the recurrence of child physical abuse and neglect: a randomized controlled trial. *Lancet*, *365*, 1786—1793.

Milner, J. S. (1994). Assessing physical child abuse risk: The child abuse potential inventory. *Clinical Psychology Review*, *14*, 547—583.

Ogles, B. M., Lambert, M. J., & Fields, S. A. (2002). *Essentials of outcome assessment.* NY: Wiley.

Olds, D., Eckenrode, J., & Kitzman, H. (2005). Clarifying the impact of the Nurse–Family Partnership on child maltreatment: Response to Chaffin (2004).*Child Abuse & Neglect*, *29*, 229—234 letter to the editor.

Olds, D. L., Henderson, C. R., Tatelbaum, R., & Chamberlain, R. (1986). Improving the delivery of prenatal care and a randomized trial of nurse home visitation. *Pediatrics*, *77*, 16—28.

Pinker, S. (2002). *The blank slate: The modern denial of human nature.* NY: Penguin.

Seligman, M. E. (2007). *What you can change and can't.* NY: Vintage.

Stoolmiller, M. (1994). Antisocial behavior, delinquent peer association, and unsupervised wandering for boys: Growth and change from childhood to early adolescence. *Multivariate Behavioral Research*, *29*, 263—288.

Sweet, M., & Appelbaum, M. I. (2004). Is home visiting an effective strategy: A meta-analytic review of home visiting programs for families with young children. *Child Development*, *75*, 1435—1456.

Vermeersch, D. A., Whipple, J. L., Lambert, M. J., Hawkins, E. J., Burchfield, C. M., & Okiishi, J. C. (2004). Outcome questionnaire: Is it sensitive to changes in counseling center clients? *Journal of Counseling Psychology*, *51*, 38—49.

Woszczynski, A. B., & Whitman, M. E. (2004). The problem of common method variance in IS research. In A. B. Woszczynski, & M. E. Whitman (Eds.), *The Handbook of Information Systems Research* (pp. 66—77). Hershey, PA: Idea Group Inc.