Does Maltreatment Beget Maltreatment? A Systematic Review of the Intergenerational Literature

Terence P. Thornberry, Kelly E. Knight, and Peter J. Lovegrove

Abstract

In this paper, we critically review the literature testing the cycle of maltreatment hypothesis which posits continuity in maltreatment across adjacent generations. That is, we examine whether a history of maltreatment victimization is a significant risk factor for the later perpetration of maltreatment. We begin by establishing 11 methodological criteria that studies testing this hypothesis should meet. They include such basic standards as using representative samples, valid and reliable measures, prospective designs, and different reporters for each generation. We identify 47 studies that investigated this issue and then evaluate them with regard to the 11 methodological criteria. Overall, most of these studies report findings consistent with the cycle of maltreatment hypothesis. Unfortunately, at the same time, few of them satisfy the basic methodological criteria that we established; indeed, even the stronger studies in this area only meet about half of them. Moreover, the methodologically stronger studies present mixed support for the hypothesis. As a result, the positive association often reported in the literature appears to be based largely on the methodologically weaker designs. Based on our systematic methodological review, we conclude that this small and methodologically weak body of literature does not provide a definitive test of the cycle of maltreatment hypothesis. We conclude that it is imperative to develop more robust and methodologically adequate assessments of this hypothesis to more accurately inform the development of prevention and treatment programs.

Keywords

child maltreatment; intergenerational continuity; cycle of violence; research methodology

There is a common assumption, in both the popular and scientific literatures, that child maltreatment begets child maltreatment. That is to say, it is often assumed that children who are maltreated by their parents are at elevated risk for maltreating their own children when
they become parents. Whether this assumption is correct or not constitutes an important scientific and policy question for several reasons.

First, child maltreatment—behavior that can include physical abuse, neglect, sexual abuse, or psychological abuse of children, as well as their co-occurrence (Child Abuse Prevention and Treatment Act, 1974)—is, unfortunately, all too common in American society (see Smith & Ireland, 2009, for a review). The National Child Abuse and Neglect Data System, which collects and analyzes annual Child Protective Services (CPS) data from states, indicates that about 12.1 children per 1,000 (about 900,000 children) have a substantiated report of maltreatment (U. S. Department of Health and Human Services, 2008). Girls and boys are approximately equally likely to be maltreated (except for sexual abuse where the rates are higher for girls); neglect is by far the predominant form of substantiated maltreatment (64%), followed by physical abuse (16%), sexual abuse (8.8%), and emotional maltreatment (6.6%). The National Incidence Studies, which estimate maltreatment prevalence based on reports from a national sample of community professionals, find a higher rate of maltreatment of 23 per 1,000 children, or 2.3% of all children (Sedlak & Broadhurst, 1996). Maltreatment prevalence in community surveys is even higher, at 15% or more (e.g., Straus & Gelles, 1986; Thornberry, Ireland, & Smith, 2001). In these data, neglect is also the most typical form of child maltreatment, followed by physical abuse and then sexual abuse, but there is much overlap between types of maltreatment (Crittenden, Claussen, & Sugarman, 1994; McGee, Wolfe, & Wilson, 1997; Smith, Thornberry, & Ireland, 2004).

Second, there is an abundant literature demonstrating the negative consequences of maltreatment for the victim’s later development (Gilbert et al., 2009). Maltreated children suffer from childhood developmental deficits including externalizing behaviors, disruptive behavior, academic problems, and depressive symptoms (see Cicchetti & Lynch, 1993, 1995, and Trickett & McBride-Chang, 1995, for reviews). Maltreatment increases the likelihood of antisocial behavior in adolescence (Ireland & Widom, 1994; Smith & Thornberry, 1995), internalizing problems (e.g., Kaplan, Pelcovitz, & Labruna, 1999), off-time transitions such as teen pregnancy (Elliott, Avery, Fishman, & Hoshiko, 2002; Smith, 1996; Widom & Kuhns, 1996), and poorer cognitive and school functioning (Kendall-Tackett & Eckenrode, 1996; Perez & Widom, 1994). In the longer term, longitudinal studies have shown that experiencing maltreatment is a significant risk factor for crime and violence (Fagan, 2001; Ireland & Widom, 1994; Rebollon & Van Gundy, 2005; Smith, Ireland, & Thornberry, 2005; Widom, 1989b), later alcohol and drug use and arrest (Ireland & Widom, 1994; Widom, Ireland, & Glynn, 1995), and a range of adult mental health problems (Widom, DuMont, & Czaja, 2007).

Finally, all of these consequences exact a substantial toll on the broader community, ultimately compromising the public health of American society (Daro, Edleson, & Pinderhughes, 2004; Leeb, Paulozzi, Melanson, Simon, & Arias, 2008). Indeed, it has been estimated that the cost of maltreatment exceeds 100 billion dollars per year (Wang & Holton, 2007). As a result, as the Centers for Disease Control and Prevention have noted, there is a pressing need for “a better understanding of the developmental pathways and social circumstances that contribute to perpetration [of maltreatment to] enhance the
development of effective primary prevention programs and guide refinement of existing prevention programs” (National Center for Injury Prevention and Control, 2002, p. 7).

One commonly-suggested developmental pathway to maltreatment perpetration starts with maltreatment victimization. Earlier victimization is often viewed as a risk factor for, and potential cause of, subsequent perpetration of maltreatment. If this is true, then unless we understand how to break the “cycle of violence” (Widom, 1989b) subsequent generations are not only at increased risk of being maltreated but also of suffering the same negative sequelae of maltreatment, such as delinquency, drug use, and mental health problems, as were their parents. Given the possibility of these cascading consequences from generation to generation, it is imperative to have solid empirical information on which to base effective intervention programs. In particular, we need to know the strength of the association between maltreatment victimization and later perpetration, as well as moderating influences and mediating pathways that help explain the intergenerational relationships.

Despite the importance of this issue and the plausibility of the cycle of maltreatment hypothesis, there are, in fact, very few methodologically rigorous prospective studies that assess its validity. The purpose of this article is to present a systematic review of the literature examining intergenerational continuity in maltreatment, assess its methodological rigor, and identify the design elements that are needed for accurately testing this hypothesis. In particular, we first identify the basic methodological criteria that a study should meet when attempting to test this hypothesis. We then describe the methods we used to identify empirical studies to include in the review and, following that, we evaluate the literature by applying the criteria we established to each study and summarize the findings with respect to the level of intergenerational continuity in maltreatment. Finally, we discuss the implications of our findings for future research in this area.

**Methodological Criteria for Testing the Hypothesis**

Studies investigating intergenerational continuity of maltreatment are necessarily nonexperimental. As is the case with all observational research, studies in this area struggle to address a variety of methodological obstacles as they attempt to draw valid conclusions. As we review this literature, it is helpful to start with a template of what a strong methodological study on this topic would look like. In creating this template, we draw upon earlier work by Ertem, Leventhal, and Dobbs (2000), who reviewed the literature on the intergenerational continuity of physical abuse, establishing eight methodological criteria that studies should meet. Their approach was to describe the core strengths of an experimental design and then use those strengths to identify the methodological features that a good observational study of this hypothesis should have.

We extend the work by Ertem et al. (2000) in two ways. First, instead of limiting our review to physical abuse, we consider the broader category of child maltreatment, including investigations of physical abuse, sexual abuse, emotional abuse and neglect, as well as their co-occurrence. Second, we add new methodological criteria to be used in the evaluation of the studies. In particular, in order for a study to provide valid evidence for testing the cycle of maltreatment hypothesis it should meet the following 11 standards:
1. **A sample that is representative of a general population**

The cycle of maltreatment hypothesis, as typically posed, applies to the general population. Ideally, therefore, the sample used to test the cycle of maltreatment hypothesis should be drawn from a clearly defined, general population and should be selected using probability sampling techniques. Studies of subpopulations, for example those conducted to identify important moderators of this relationship, should use probability sampling techniques and should take care to carefully define the subpopulations so that proper generalizations can be made. Evidence taken from highly selected, nonrandom samples—such as clinical samples—provides biased estimates that are likely to overestimate the rate at which intergenerational continuity of maltreatment occurs. While oversampling along the lines of occurrence of maltreatment may enhance the prevalence of maltreatment in the data and ensure adequate statistical power, when this happens sample weights should be used to ensure that the findings can be referred back to a representative population in order to properly evaluate the intergenerational hypothesis.

2. **A satisfactory participation rate and low levels of attrition**

Refusal and attrition rates rarely occur in a random fashion and often the most at-risk individuals and families are the ones who are also most likely to refuse participation or drop out of a study (Thornberry, Bjerregaard, & Miles, 1993). Studies that do not actively attempt to recruit and retain participants from at-risk populations run the danger of biasing the results by relying on a less disadvantaged and more prosocial group of individuals (Thornberry et al., 1993). Furthermore, if high attrition and low participation occur in studies using a matching design where comparison and “treatment” groups are used, the non-random nature of these processes may render the two groups incomparable, thereby biasing observed differences between the groups.

3. **Maltreated and non-maltreated sample members in the focal generation**

Some studies identify a sample of maltreated individuals and then calculate the degree to which maltreatment is observed in either the subsequent or the prior generation. Doing so is tantamount to sampling on the dependent variable, a strategy that can severely bias proper inferences (Geddes, 1990). Including only maltreated individuals is likely to provide overestimates of continuity across generations, and findings will not be generalizable to a larger population. It is essential to include both maltreated and not maltreated individuals in the primary sample. This allows us to also observe whether or not individuals in one generation who were not victims of maltreatment eventually become perpetrators, thereby providing the necessary comparison statistic.

4. **Assessment of the maltreatment status for the comparison group**

It is possible for studies evaluating the intergenerational hypothesis to compare maltreated and not maltreated groups recruited and followed in different ways. It is essential to verify, to the extent possible, that the comparison group participants were not maltreated. Failure to do so will cloud comparisons between treatment and comparison groups, since the application of the “treatment” would not be limited to the experimental group. We recognize that the assessment of maltreatment status for the comparison group will be done with some
degree of error, as is true of all assessments of maltreatment, but it is essential that the study attempt to do so with as much validity as possible.

5. Controls for antecedent factors that may cause spurious relationships

In any study, it is important to reduce the potential for an observed relationship to be caused by endogeneity. Using simple correlations to evaluate the intergenerational hypothesis does not provide an adequate test for the hypothesis because it cannot control for a variety of potential confounding factors that should be held constant for valid inferences to be made. Not controlling for such factors is likely to lead to an overestimation of the direct effect of maltreatment victimization on maltreatment perpetration. In studies that compare a maltreated group to a non-maltreated group, it is important to ensure that the two groups are matched on as many pertinent factors as possible. If this is not done, the potential for the observed difference between the two groups due to endogeneity is high. Among the studies we review in this article, there was variability in the number and variety of control variables used in analyses, and in the number of factors across which maltreatment and control groups were matched. In our review this standard was satisfied if the study attempted to account for selection effects to any degree by controlling for, or matching on, more than simply race and gender since maltreatment is known to be directly affected by a variety of other variables (Belsky, 1980).

6. Prospective data

Many studies of intergenerational continuity in maltreatment rely on retrospective reports to measure maltreatment victimization. Retrospective measures of maltreatment have some distinct advantages—for example, they are more cost-effective and they do not entail legal mandatory reporting requirements (Tajima, Herenkohl, Huang, & Whitney, 2004). Moreover, retrospective reporting of maltreatment has been found to have moderate correspondence with prospective reports (Smith, Ireland, Thornberry, & Elwyn, 2008; Widom, Weiler, & Cottler, 1999). Retrospective reporting of maltreatment is, relatively speaking, more concordant with prospective measures (Widom & Shepard, 1996) than is true of other phenomena such as mental health problems or family conflict (Henry, Moffitt, Caspi, Langley, & Silva, 1994) and Hardt and Rutter (2004) found that retrospective and prospective measures differed less when the maltreatment being reported was especially serious. Nevertheless, there are several sources of measurement error for retrospective reports, which are likely to lead to a high rate of false negatives, that is, the underreporting of maltreatment. Hardt and Rutter (2004) find that retrospective reporting can be affected by current mood of the respondent, false recall, respondent forgetting (possibly due to infantile amnesia), and after-the-fact assignment of meaning to memories. Another reason might be that respondents whose trauma has not been fully resolved might not report instances of earlier maltreatment due to dissociation (Ouimette, Read, & Brown, 2005). In general, those who have empirically compared prospective and retrospective measures tend to conclude that findings relying on retrospective measures of maltreatment should be viewed with caution (Henry et al., 1994; Smith & Ireland, 2009) and there is a growing consensus that “reliance on retrospective report in particular can often lead to overestimation” of intergenerational continuity (Cicchetti & Valentino, 2006, p. 139; see also Hunter & Kilstrom, 1979, and Pears & Capaldi, 2001). Furthermore, retrospective measures are
particularly suspect when asked across long recall periods (Widom, Raphael, & DuMont, 2004), which is often the case in this literature. As a result, we assert in our criteria that studies using prospective measures more likely provide a more accurate estimation of the cycle of violence hypothesis than those which rely on retrospective measures.

7. Different reporters of maltreatment for each generation

A basic premise of all science is that measures of the independent and dependent variables are independent of one another. In an experiment this is ensured by the manipulation of the independent variable and the observation of the dependent variable. In testing the cycle of maltreatment hypothesis it can be ensured, or at least enhanced, if the measures are based on different sources, for example, the members of each generation report on their own behavior or survey data are used for one and official data for the other. Reliance on a single reporter who reports on his/her own perpetration and on the perpetration of maltreatment by his/her parents obviously violates this assumption and the “bias that can result... is recognized increasingly as a substantial problem” (Pears & Capaldi, 2001, p. 1443; see also Knutson & Schartz, 1997). Although it is plausible that perpetrating parents might overreport victimization as a way of explaining their own maltreatment to interviewers (Tajima et al., 2004), it is also plausible that they may underreport their childhood experiences due to respondent embarrassment, desire to protect those who perpetrated maltreatment against them, and a desire to forget the victimization (Della Femina, Yeager, & Lewis, 1990). Of these, the latter has been found to occur more frequently (Della Femina et al., 1990) and, although underreporting may lead to more conservative estimates of the rate of transmission across generations, the results are nonetheless biased. If a study uses different reporters or methods to measure maltreatment victimization and perpetration in subsequent generations, this would constitute a primary effort to reduce some of the bias associated with using one method of measurement.

8. The same exposure period for treatment and comparison groups

When comparing outcomes between two groups, it is important that the outcome in question be measured similarly for all respondents. For example, a study might observe the children of maltreated parents until late adolescence, and compare them to the children of non-maltreated parents who were observed until only age 12. This methodology would likely result in an overestimation of differences in the rates of maltreatment between the groups. It is therefore vital that maltreatment in the treatment and control groups is measured over the same age ranges.

9. Follow-up over an extended portion of the life course

The exposure period for maltreatment is from birth to age 18, although maltreatment is quite rare after age 15 (Thornberry et al., 2001). Studies with short exposure time, for either or both generations, are likely to substantially undercount the prevalence of maltreatment and therefore to misrepresent the level of intergenerational continuity in maltreatment. In our review, we considered follow-ups of less than five years to be of questionable validity.
10. Measures of maltreatment with proven validity and reliability

Survey studies should use a measure of maltreatment in both generations that has been demonstrated to properly measure the occurrence of maltreatment. Examples of this include the Conflict Tactics Scale (Straus, 1979) and the Childhood Trauma Questionnaire (Bernstein, Fink, Handelsman, & Foote, 1994). Measures with poor psychometric properties are likely to produce misleading estimates of the prevalence of maltreatment in each generation and therefore of the level of intergenerational continuity. Using established scales with strong psychometric properties also improves the comparability of a study’s findings to those of other researchers. Studies based on official records of maltreatment should assess, to the extent possible, the consistency and completeness of the reporting system.

11. A clear definition of maltreatment

In many cases, the exact definition of maltreatment used, and so the meaning of what has been measured, is obscure. In order for others to better interpret a study’s findings, it is important to stipulate exactly what constitutes abuse or maltreatment—whether or not it is measured as a degree of seriousness, number of occurrences, or prevalence; whether the maltreatment is physical abuse, sexual abuse, or neglect, and exactly what constitutes each of these; for what ages the abuse is measured; and who can potentially be perpetrators.

In applying these criteria to the studies we used two general conventions. First, the criterion had to be met for both generations. For example, if retrospective data were used to measure maltreatment in either generation or if there was questionable validity and reliability of the measure used for either generation, the study was considered not to have met that particular criterion. Second, if there was no discussion or treatment of a particular methodological issue, for example participation and retention rates, the study was considered not to have that that criterion. In other words, methodological adequacy was not assumed.

No doubt, there are other methodological criteria that one could identify for assessing these studies. Certainly, if a study had other severe limitations beyond these central methodological requirements, we might call the evidence of that study into question even further. But if a study met all 11 criteria, it would be able to provide robust evidence about the cycle of maltreatment hypothesis. The evaluation of the cycle of violence hypothesis, it should be noted, is just the first step of the study of intergenerational transmission of maltreatment. A study’s ability to examine potential mediators and moderators of this observed association is also important, but somewhat secondary to the prior question of whether there is actually a relationship to mediate. The primary concern of this review is the believability of estimates of the direct effect of parents’ victimization on subsequent perpetration.

Method for Review

We conducted a literature review of English language studies assessing intergenerational maltreatment in childhood and adolescence using a technique akin to snowball sampling. We did not restrict the publication date range. The titles and abstracts of approximately 1,000 articles and books were reviewed from computerized databases (e.g., psycINFO, Sociological Abstracts), the internet, and bibliographies using search terms such as
intergenerational maltreatment or cycle of abuse. Of these, 194 were deemed potentially relevant and obtained for further review. However, 147 articles were eliminated because (a) child maltreatment as defined here was not measured (for example, the study focused on harsh parenting or physical punishment that did not rise to the level of abuse), (b) maltreatment was measured in only one generation, (c) participants were not yet parents (e.g., in several studies such as Litty, Kowalski, & Minor, 1996, and Rodriguez & Price, 2004, college students were sampled and completed the Child Abuse Potential Inventory [CAPI], Milner & Wimberley, 1979, 1980), or (d) the sample size was extremely small (n < 5). In addition, (e) qualitative articles, (f) conference papers, (g) review articles, (h) duplicate studies using the same data already reviewed, and (i) dissertations written before 2004 were not included. Somewhat related to (c), studies using Milner and Wimberley’s (1979, 1980) CAPI were included only if the inventory was used as a proxy measure for a parent’s perpetration of maltreatment in the second generation. In the end, 47 articles were retained for in-depth review. These included the 10 studies of physical abuse reviewed by Ertem et al. (2000).

Given the relatively small number of available studies we decided to be as inclusive as possible in our review, allowing for variability across studies on three important dimensions. First, we allow for variability in how maltreatment is conceptualized in the adjacent generations. Some studies measure a focal participant’s maltreatment victimization as a child and then that person’s maltreatment perpetration as an adult. Other studies measure a focal participant’s maltreatment as a child and then measure their children’s victimization during childhood. We recognize that these are slightly different versions of intergenerational continuity but, since they both assess the continuity of maltreatment across adjacent generations, we include both types of studies to be comprehensive. When appropriate we make these distinctions explicit but, in general, maltreatment victimization refers to maltreatment victimization of a parent (in the first generation) and maltreatment perpetration refers to later maltreatment perpetration by that person or to maltreatment that occurs to his/her child(ren) (in the second generation).

Second, we include studies of physical, emotional, and sexual maltreatment, as well as neglect. This approach is consistent with the observation that most instances of maltreatment involve multiple types (see, for example, Crittenden et al., 1994; McGee et al., 1997; Smith et al., 2005). Indeed, some have argued “that it may not be appropriate for research to attempt to focus on a single form of maltreatment, since other types of abuse are often present as well” (Grayson, 2010, p. 2; Higgins &McCabe, 2000; Richmond, Elliott, Pierce, Aspelmeier, & Alexander, 2009). While there may be somewhat distinctive risk and protective factors for the different types of maltreatment as well as some distinct mediators that would help account for intergenerational continuity, the investigation of antecedents and of mediators and moderators is not the focus of this investigation.

Relatedly, we did not include studies that focused on harsh parenting. While harsh parenting is associated with a variety of negative outcomes for children (e.g., Gershoff, 2002), it is distinct from maltreatment, primarily in its degree of seriousness. For instance, Barker et al. (2008) define harsh or reactive parenting as when a mother or father feels angry, raises his or her voice or shouts, or spanks a child who is being fussy. On the other hand, definitions of
child maltreatment stress the abject seriousness of behavior. Federal legislation, for example, defines maltreatment as any recent act or failure to act on the part of a parent or caretaker that results in death, serious physical or emotional harm, sexual abuse or exploitation, or an act or failure to act which presents an imminent risk of serious harm (U.S. Department of Health and Human Services, 2008). Thus, we include all studies measuring behaviors that could be properly considered maltreatment, regardless of type of maltreatment.

Third, there are several sources of data to measure maltreatment that are used in this literature. Some studies rely on official indicators, such as Child Protective Services records, others rely on survey measures, and yet others rely on a mix of indicators across the generations. We include studies using all these measurement types to be as comprehensive as possible. None of these measurement strategies is perfect and all have limitations. Including studies with different approaches provides some degree of content validity to our conclusions.

Overall, therefore, we present a general assessment of the cycle of maltreatment hypothesis rather than focus on one aspect of the literature, for example, one type of maltreatment like physical abuse or one measurement source, like substantiated incidents. We recognize that this introduces some heterogeneity into our study but at the same time it provides the broadest assessment possible of the level of intergenerational continuity.

**Review of the Evidence**

To help organize our review, we summarize articles in four major categories based on their research design. The first category uses official data to measure maltreatment in both generations. The second group, by far the largest, uses survey data to measure the parent’s history of maltreatment and official or archival data to measure the child’s maltreatment. The third group consists of studies that use survey data to measure maltreatment in both generations—some with different reporters and some with a single reporter. Finally, we review studies that use the CAPI (Milner & Wimberley, 1979, 1980) to measure maltreatment in one or both generations. Even though the CAPI does not measure actual maltreatment, it does focus on the potential to engage in maltreating behaviors and not simply poor parenting practices. This variable is different from those that focus on actual behavior but, given the generally small literature, we include these studies and do so in a separate section so that their results are not conflated with the others studies that assess actual child maltreatment.

The table in the Appendix lists the 47 studies that are included in our review along with each of the 11 methodological criteria that we described above. For each criterion we indicate whether, in our judgment, the study met the criterion, scored as a 1, or failed to meet the criterion, scored as a 0. In arriving at these judgments we proceeded in the following way. First, after developing an initial rating form all three authors read and rated a selection of the articles. We then met to discuss them and, based on those discussions, we refined the rating form and the standards for using it. Each article was then read and rated by at least two of the authors. Any differences in the ratings were identified, discussed, and, based on that
discussion, we arrived at a consensus rating. If there were any disagreements at that point, which was quite rare, the judgment of the senior author was used. The final ratings also appear in the Appendix.

Official Data Studies

We identified only one study that used official data to measure maltreatment in both the parent and child generations. Widom (1989a) employed a prospective, cohort design to investigate the consequences of child abuse and neglect on a variety of outcomes, including the perpetration of maltreatment during the adult years. Her study started with a sample of 908 individuals with a confirmed case of maltreatment victimization when the person was 11 years of age or younger based on court records and a matched sample of 667 individuals with no record of maltreatment. The indicator of the perpetration of maltreatment is based on being arrested for child abuse anytime between the ages of 18 and 32. Based on these data, Widom did not find a significant relationship between a history of being maltreated and the likelihood of perpetrating maltreatment during the adult years.

The study is based on a large sample of maltreated youth with a large, matched comparison group. It ensured that the members of the comparison group did not have a record of being maltreated, both groups were followed over a long portion of the life course, the parent’s maltreatment victimization clearly preceded their own perpetration of maltreatment, both the maltreatment and comparison groups were followed for equal periods of time and in an equal manner for assessing the perpetration of maltreatment. Based on these and other design features, this is the strongest study in the area and, in fact, meets 10 of our 11 criteria.

Unfortunately for the specific assessment of the cycle of maltreatment hypothesis, the study suffers from a fundamental flaw. The indicator of perpetration was being arrested for child abuse, an indicator with low validity. Maltreatment is typically handled by referral to Child Protective Services, or a similar social service agency, not arrest. The overall prevalence (1%) is so low there is virtually no variability in the measure and, as a result, the rates of perpetration are 1.1% for the abuse or neglect group and 1.0% for the comparison group. Based upon the lack of variation in the measure this study cannot provide evidence, one way or the other, for the cycle of maltreatment hypothesis. That is unfortunate, as the overall study by Widom has informed our understanding of the consequences of maltreatment for a variety of other outcomes (e.g., Widom, 1989a; Widom & Maxfield, 2001).

Parent Self-Report and Child Official Data Studies

In our review of the literature assessing the intergenerational hypothesis, we found 31 studies in which maltreatment in the parent or first generation was assessed using self-report of the parent’s history of abuse (invariably based on retrospective accounts) and the measurement of maltreatment in the child or second generation is based on CPS records or other official data sources. We begin by describing in some detail the seven studies that, in our judgment, are the methodologically strongest based on our criteria. We note, however, that even these studies meet only about half of the basic design requirements.

Dixon, Browne, and Hamilton-Giachritsis (2005a) examined a large cohort of 4,351 families with newborn children in Essex, England. Information on the parents’ history of
maltreatment was collected by nurses who provided visiting health services. They asked both parents to jointly complete one self-report questionnaire about their childhood histories of abuse and the analysis is based on this joint parental measure. Thirteen months after the birth of the second generation child, information about the perpetration of maltreatment was collected. They found that 6.7% \((n = 9)\) of parents with a childhood history of abuse were referred to child protection for perpetrating maltreatment against their own children, compared to \(0.4\% \(n = 18\) of parents who were not abused in childhood, revealing a significant difference between abused and non-abused parents (Fisher’s Exact, \(p < .001\)).

Three variables partially mediated intergenerational continuity—having been a parent under 21 years of age, a history of mental illness, and residing with a violent adult. These factors explained 53% of the intergenerational effect. In follow-up studies, these authors found that poor parenting styles (Dixon, Hamilton-Giachritsis, & Browne, 2005b) and financial solvency and social support (Dixon, Browne, & Hamilton-Giachritsis, 2009) also mediated the intergenerational cycle of maltreatment.

Important strengths of this study include a large, generalizable sample, prospective data for the second generation, and because the sample is a birth cohort, the assessed exposure period for maltreatment is the same for all children. Last, perpetration was assessed for families with abused and non-abused parents and meditational effects were examined. On the other hand, substantial limitations apply that reduce the ability of this study to test the cycle of maltreatment hypothesis. These include the use of retrospective reports of the parents’ history of abuse based on measures with no proven validity or reliability, difficulty of assessing the fact that the parents discussed and completed the questionnaire jointly, and the relatively long and varying recall periods across respondents. Also, no clear definition of maltreatment was provided. Perhaps most importantly, the follow-up period is quite short for the child generation, only 13 months long. That is clearly a small portion of the overall exposure period and longer follow-up could alter the nature of the relationship.

Egeland, Jacobvitz, and Sroufe (1988) examined a high-risk subsample of low socio-economic women participating in a mid-western prenatal clinic \((n = 161)\). Using retrospective reports collected when their children were 4 years old, 29% of mothers \((n = 47)\) reported maltreatment experiences in childhood. Multiple in-home and laboratory observations over a 64-month follow-up period (which were later verified by CPS records) revealed that 38% \((n = 18)\) of the abused mothers (11% of total sample) perpetrated maltreatment against their own children. The rate of maltreatment for the non-abused mothers was reported in a subsequent study (Egeland, Jacobvitz, & Papatola, 1987) to be only 7% (as measured over a 3-year period). Interestingly, they also examined mothers who “mistreated” their children, defined as clear-cut abuse, suspected maltreatment, or abandonment. Under this definition, 70% \((n = 33/47)\) of maltreated mothers and 47% \((n = 54/114)\) of mothers not maltreated in childhood “mistreated” their own children.

This study find support for the intergenerational hypothesis and is the only study to use observational data to assess mothers’ possible perpetration. In addition, different reporters collected data on each generation (i.e., parents’ self-reported history and observations of parents assessed by research staff), from maltreated and non-maltreated participants. However, the small non-representative sample, the study’s reliance on mothers’
retrospective self-reports of childhood maltreatment, and the lack of analytic techniques to adequately control for potentially spurious factors limit the findings.

Sidebotham, Golding, and the ALSPAC Study Team (2001) assessed a cohort of children born over 17 months in Avon, UK, between 1991 and 1992 (n = 14,138) to test the effect of mothers’ and fathers’ childhood maltreatment separately on the likelihood of their children’s victimization. Both parents’ maltreatment was retrospectively measured using separate postal questionnaires, covering whether each had a history of ever being sexually abused or subjected to physical or emotional cruelty by their parents. Their children’s maltreatment was measured by searching through official child welfare records to determine whether they were either involved with, or confirmed as a victim in, an investigation of physical or sexual maltreatment or neglect to the age of 6. The authors found that 162 of 14,138 children in the birth cohort sample had been maltreated. Bivariate analyses identified several significant differences in rates of maltreatment: 24% of maltreated children were born to mothers with a history of sexual abuse, compared to 5% of non-investigated children; 10% of maltreated children were born to mothers whose parents were physically cruel, compared to 3% of non-investigated children; experiencing emotional cruelty was not significantly related to their children’s maltreatment. For fathers, only physical cruelty was significantly different for the maltreated and non-investigated children: 11% of maltreated children had fathers whose parents were physically cruel, compared to 5% of the non-investigated group. However, in multivariate regressions that controlled for relevant background variables (maternal age less than 20, maternal history of child guidance, mother’s father absent in childhood, mother’s highest educational level, and mother’s psychiatric illness excluding depression), considerably less support for the intergenerational hypothesis was found; only maternal sexual abuse remained significantly related to the children’s maltreatment to age 6. Net of background factors, none of the maltreatment indicators for the fathers were significantly related to their children’s maltreatment. Note, however, that this null finding may be of less importance because the majority of abuse of young children is not perpetrated by fathers.

Overall, this study has several benefits—for example, the sample came from a general population, each generation’s maltreatment history was collected using different reporters, and several important control variables were included. But, the response rates (14% to 86%) in the study were highly variable, depending on the respondent, assessment, and question, and a disproportionate number of maltreating families, especially fathers, did not return the postal questionnaires. The retrospective surveys assessed exposure to physical cruelty and emotional cruelty which are poorly defined and are perhaps a liberal proxy for parent’s history of abuse; even so, consistent evidence supporting the intergenerational hypothesis was not found.

Renner and Slack (2006) used a sample of 1,005 low-income women who were randomly selected from the 1998 Temporary Assistance to Needy Families caseload in Illinois to examine the relationship between mothers’ history of being maltreated and their involvement in maltreatment of their own children. The mothers ranged in age from 22 to 42. CPS reports were used to determine perpetration of maltreatment, regardless of whether or not the reports were substantiated. The mothers’ retrospective report of childhood maltreatment included instances of physical punishment or abuse, serious neglect, and
attempted or actual sexual abuse (by anyone). CPS records were checked for reports of perpetration from 1980 to 2002. Controlling for family of origin characteristics using logistic regression, a mother’s childhood physical maltreatment increased her likelihood of later maltreatment perpetration—but only for the measure *leaving a child at risk for harm* \((OR = 2.04)\). No other significant relationships were found for the 16 combinations of childhood maltreatment and forms of adult perpetration.

This study found minimal support for intergenerational continuity of maltreatment. Indeed, finding one significant effect is close to what one would expect by chance (1 in 20) at the 95% confidence level. Strengths include collecting information on perpetration over 22 years and guaranteeing that mothers were the actual perpetrators of maltreatment. However, limitations include a sample of low-income women not representative of the general population, the combination of substantiated and unsubstantiated reports of maltreatment, and different recall periods for the retrospective reports given the wide age distribution of the mothers.

Thompson (2006) examined an urban, high-risk sample (based on areas of the city with the highest rates of CPS cases; location not reported) of 220 families with infants in their first year of life. Using retrospective self-report questionnaires of mothers’ victimization, half reported experiencing physical abuse and a third sexual abuse. In 58% of all cases, an official CPS report documented a complaint of maltreatment perpetration against the family, although not in all cases was the perpetrator the mother. Bivariate analyses found that mothers’ physical and sexual victimization predicted their children’s risk of maltreatment. However, the significance of this effect disappeared when mothers’ marital status, depressive symptoms, and adult experiences of physical victimization were included in the analysis.

This study had several noticeable strengths, including respondents in both generations who were maltreated and non-maltreated, different reporters assessing maltreatment across generations, and controls for the more proximal factors in the mothers’ lives (e.g., adult victimization) which appear here to mediate the intergenerational effect. Important limitations, however, include a small, non-representative sample residing in areas with very high rates of maltreatment based on CPS reports. Also, the study relied entirely on retrospective self-reports of the mothers’ victimization. Perhaps most noteworthy is the relatively short exposure period for maltreatment in the second generation, a period lasting less than one year.

Altemeier, O’Connor, Sherrod, Tucker, and Vietze (1986) assessed intergenerational continuity in a subsample of white mothers \((n = 927)\) attending an urban prenatal clinic for low income families between 1975 and 1976. Using open-ended interview questions, mothers who indicated that they had been beaten up by a caregiver and hit on body parts other than hands, buttocks, or legs for punishment were categorized as abused. Substantiated child abuse and neglect in the families of these mothers was determined using official reports from statewide protective agencies four years after the initial interview. Altogether, 21 cases of abuse and 38 cases of neglect were documented. Chi-square tests revealed no significant differences in abuse or neglect (assessed separately) between families with

_Trauma Violence Abuse._ Author manuscript; available in PMC 2014 May 27.
mothers who were and were not battered as children, and so support for the intergenerational hypothesis was not found.

This study had satisfactory participation and minimal attrition, included maltreated and non-maltreated participants in both generations, used different reporters to collect data for each generation, and had equal exposure periods for all participants. However, a non-generalizable low-income sample was selected, analytic techniques did not control for antecedent factors, and children’s maltreatment was followed for only the first four years of life.

Berlin, Appleyard, and Dodge (2011) recruited 351 pregnant women receiving prenatal care from public and private practices, as well as an additional 148 women recruited from flyers posted in the community—a small Southeastern city. The sample ranged in age from 12 to 41. Using the Parent-Child Conflict Tactics Scale, retrospective histories of mothers’ maltreatment during childhood revealed that 9.6% (n = 48) of the sample experienced physical abuse and 10.6% (n = 53) neglect. Allegations and substantiations of maltreatment collected from county records at 26 months showed that 8% (n = 40) of the sample’s children had experienced abuse or neglect. The level of intergenerational continuity in maltreatment was rather modest. The mother’s history of neglect was not significantly related to offspring victimization and the correlation of mother’s physical abuse and offspring victimization was significant but the coefficient is only .10; in other words only 1% of the variance is explained. Importantly, however, this significant association was maintained in multivariate models controlling for several background factors. In addition, mothers’ social isolation and aggressive response biases (but not mothers’ mental health problems or hostile attributions) mediated continuity for physical abuse.

This study has several important strengths. Prospective data from county social service records were collected on the child generation, important controls were included in the analyses, and valid and reliable measures and clear definitions of maltreatment were used. However, the study was based on a convenience sample and maltreatment in the child generation was only observed for the first 26 months of life. A longer exposure period for the child’s maltreatment would further strengthen the validity of these findings. The study also relied entirely on retrospective accounts of the mother’s childhood victimization with relatively long and varying recall periods that range from a few years to several decades.

In addition to these seven studies there are 24 other investigations that rely on survey data to measure the parent generation’s history of maltreatment and official data to measure the child generation’s history of maltreatment. Using the methodological criteria outlined above they are, in our judgment, weaker than the seven studies just reviewed in detail. Of these studies, 20 found at least some support for the intergenerational hypothesis (Avery, Hutchinson, & Whitaker, 2002; Baldwin & Oliver, 1975; Caplan, Waters, White, Parry & Bates, 1984; Coohey & Braun, 1997; Faller, 1989; Folsom, Christensen, Avery, & Moore, 2003; Goodwin, McCarty, & DiVasto, 1981; Haapasalo & Aaltonen, 1999; Healy, Kennedy, & Sinclair, 1991; Hunter & Kilstrom, 1979; Kim, Trickett, & Putnam, 2010; Korbin, Anetzberger, & Austin, 1995; Kotelchuck, 1982; Leifer, Kilbane, Jacobsen, & Grossman, 2004; McCloskey & Bailey, 2000; Oates, Tebbutt, Swanston, Lynch, & O’Toole, 1999;
Smith & Adler, 1991; Smith & Hanson, 1975; Wolock & Horowitz, 1979; Zaidi, Knutson, & Mehnn, 1989). In contrast, 4 did not find any support (Banyard, Williams, & Siegel, 2003; Deblinger, Hathaway, Lippmann, & Steer, 1993; Estes & Tidwell, 2002; Zuravin, McMillen, Depanfilis, & Risley-Curtiss, 1996). Unfortunately, most of these studies met few of the design criteria making it difficult to ascertain evidence for or against the intergenerational hypothesis. For example, several studies did not select samples representative of the general population (Folsom et al., 2003; Haapasalo & Aaltonen, 1999; McCloskey & Bailey, 2000), had no comparison group (Caplan et al., 1984; Faller, 1989; Zuravin et al., 1996), assessed maltreatment using different exposure periods (Baldwin & Oliver, 1975; Coohey & Braun, 1997; Korbin et al., 1995), did not include a clear definition of maltreatment (Avery et al., 2002; Goodwin et al., 1981; Wolock & Horowitz, 1979), and all of them relied on retrospective data, often with long recall periods to measure maltreatment in the parent generation. In several studies, the measure of maltreatment came from hospital records and these samples were then matched with samples from the general population (e.g., Banyard et al., 2003; Deblinger et al., 1993; Estes & Tidwell, 2002; Hunter & Kilstrom, 1979; Kotelchuck, 1982; Oates et al., 1999; Smith & Adler, 1991; Smith & Hanson, 1975). In addition, we found studies that assessed maltreatment using public and private agency referrals (Korbin et al., 1995), case records from a domestic violence shelter (Avery et al., 2002), and a combination of official and unofficial sources, such as referrals from doctors, social welfare organizations, families, and the public (Baldwin & Oliver, 1975).

Overall, 31 studies assessed maltreatment in the first generation using retrospective self-reports of parents’ history of abuse and documented maltreatment in the second generation using official data sources. None of these studies met all eleven of our criteria. The seven strongest studies were discussed in detail, and support for the intergenerational hypothesis was quite mixed. Two studies found general support for the hypothesis (Dixon et al., 2005a; Egeland et al., 1988), one study found support for a mediated model (Thompson, 2006), three additional studies found support for only one specific type of maltreatment—sexual abuse (Sidebotham et al., 2001), risk of harm (Renner & Slack, 2006), and physical abuse (Berlin et al., 2011)—but not for other forms of maltreatment, and one did not find any support for the intergenerational hypothesis (Altemeier et al., 1986). Also, in all cases the effect size was quite modest. Note, however, that even though these 7 studies were the strongest in the category, they only met on average 6 of the 11 methodological criteria and none met more than 7. Noticeably, all of them rely on retrospective reports, sometimes with rather long recall periods, to assess the parents’ history of maltreatment and, with the exception of Renner and Slack (2006), had relatively short follow-up periods in which to assess the perpetration of maltreatment. Both of these are serious methodological limitations. In addition, 24 methodologically weaker studies were reviewed and even though 20 found support for the intergenerational hypothesis, each of these studies failed to meet a majority of the basic methodological criteria needed to adequately assess intergenerational continuity in maltreatment.
Parent Self-Report and Child Self-Report Studies

Pears and Capaldi (2001) is the only study we reviewed that collected survey data independently from respondents from both the parent and child generations. The data for this study were taken from 109 boys from Eugene, Oregon, who were enrolled in the Oregon Youth Study for whom one or both biological parents were present through their eighteenth year. The sample of parents consisted of 106 mothers and 73 fathers. These parents were asked to report retrospectively on their own victimization of physical maltreatment during childhood using the Assessing Environments-III questionnaire (AE-III). When the second generation participants were approximately 21 years old, they were asked to report on their physical maltreatment victimization during childhood using a modified version of the AE-III. Using hierarchical regression analysis controlling for pertinent child and parent factors, this study found that parents’ history of physical abuse significantly predicted a higher likelihood of the child also being physically maltreated. As such, the study estimated a rate of intergenerational continuity of 23%.

Relatively speaking, this study was methodologically strong, and it fulfilled 7 of the 11 methodological criteria. Different respondents from successive generations responded to surveys that included a measure of maltreatment with proven validity and reliability. Also, the study measured maltreatment across a substantial period of the developmental life course for both parents and children. However, this study also had some limitations, including the fact that it relied upon a small and non-representative sample, and that the survey data on maltreatment were collected retrospectively in both generations.

Single Respondent Self-Report Studies

We reviewed eight articles whose data on intergenerational continuity of maltreatment was obtained from a single reporter (Ball, 2009; Finkelhor, Moore, Hamby, & Straus, 1997; Frias-Armenta, 2002; Herrenkohl, Herrenkohl, & Toedter, 1983; Heyman & Smith Slep, 2002; Ney, 1988; Straus, 1979; Tutty, 1999). In these studies, an individual respondent answered questions concerning his or her own maltreatment victimization, as well as questions about his or her perpetration of maltreatment. Significant relationships between the maltreatment measures are found in seven of these eight articles.

Every article in this category had problematic methodological elements. In addition to having a single reporter there was typically a long time lag between maltreatment perpetration and survey administration. For the reasons outlined earlier, we should be concerned that a study that uses data taken from a single reporter may provide biased estimates of rates of intergenerational continuity of maltreatment. Authors of the articles in this category more often than not acknowledge that having a single reporter is a limitation of their studies, but this drawback is frequently not mentioned by later review articles and those relying on these studies to make assertions of a strong intergenerational link in maltreatment (e.g., Browne & Herbert, 1997; Egeland, 1993). Some of the studies in this category used a sample that is not generalizable to a well-defined population (e.g., Ball, 2009; Frias-Armenta, 2002; Herrenkohl et al., 1983; Tutty, 1999) and some did not use measures of maltreatment with proven validity and reliability (e.g., Herrenkohl et al., 1983; Heyman & Smith Slep, 2002; Ney, 1988; Tutty, 1999). Although these studies find significant...
associations in maltreatment in subsequent generations, their findings should not be cited as strong evidence for intergenerational continuity.

**Child Abuse Potential Inventory Studies**

We also reviewed six studies that did not directly measure maltreatment perpetration by parents, but instead administered a proxy measure, typically the Child Abuse Potential Inventory (CAPI; Milner and Wimberley, 1979, 1980) for use as an outcome (Caliso & Milner, 1992; DiLillo, Tremblay, & Peterson, 2000; Doumas, Margolin, & John, 1994; Lounds, Borkowski, & Whitman, 2006; Narang & Contreras, 2005; Ornduff, Kesley, Bursi, Alpert, & Bada, 2002). The CAPI is designed to estimate the likelihood of a parent physically abusing a child, and can be used as a screening tool to identify families at risk for maltreatment.\(^1\) All of these articles except one found a significant relationship between prior maltreatment victimization by parents—physical, sexual, or emotional abuse, or a combination of these types—and higher adult CAPI scores. Lounds et al. (2006) found that being a victim of neglect did not significantly increase CAPI scores for mothers. They did find, however, that, controlling for respondents’ propensity to give more socially desirable responses (Paulhus, 1991), maternal history of victimization was significantly associated with higher scores on the Mother-Child Neglect Scale (Lounds, Borkowski, Whitman, and the Centers for Prevention of Child Neglect, 2004), which measures mothers’ potential to neglect children.

These studies, however, have a variety of methodological problems. The most basic, of course, is that the studies do not measure actual maltreatment, only the potential for engaging in maltreatment. Also, all the studies examined in this category used retrospective items asking parents about maltreatment they experienced during childhood; the recall lag of victimization reporting was typically long (according to our estimations, some studies entailed an average lag of 20 years or more). All of these studies relied on a single reporter, using measures of victimization and potential perpetration taken from the same respondent. There were also several sampling limitations associated with these studies—many had small sample sizes (e.g., Caliso & Milner, 1992; Lounds et al., 2006; Narang & Contreras, 2005), others were taken from non-representative samples (e.g., DiLillo et al., 2000; Doumas et al., 1994, Ornduff et al., 2002), and some utilized data collection techniques typically associated with low participation rates, such as mailed questionnaires and public advertisements (e.g., Doumas et al., 1994).

**Discussion**

There is a very common assumption that maltreatment begets maltreatment. People who were maltreated during their childhood are expected to be at greater risk of maltreating their own children than are people who were not maltreated while growing up. This article reviewed both the substantive findings and the methodological quality of the research studies that have tested this hypothesis. What have we learned?

---

\(^1\) We limited the review of CAPI studies to samples of parents (who therefore had an opportunity to maltreat a child) and eliminated several studies based on college students since they merely assessed the potential to maltreat a child should they have a child in the future.
Substantively, most studies report support for the cycle of maltreatment hypothesis. That is, most report a significant relationship, albeit a rather modest one, between a history of maltreatment and the likelihood of perpetrating maltreatment or the likelihood of having a maltreated child. While certainly not universal, that is the typical conclusion found in the literature. When the methodological rigor of this body of research is taken into account, however, the level of support for this hypothesis becomes, in our view, much less certain.

To assess the quality of the research on this topic, we established 11 methodological criteria that a strong study of the cycle of maltreatment should meet. In our view, our criteria are not overly rigorous. Indeed, they do little more than systematically enumerate the basic elements of research methods. They call, for example, for the use of representative samples, clear conceptual definitions, measures with validity and reliability, the use of prospective data, the use of different respondents or data sources for each generation, adequate observation periods, and control for obvious confounding variables either by matching or by statistical control. Despite the basic nature of these criteria, the vast majority of the 47 studies in our review meet fewer than half of these methodological expectations.

Perhaps more important than the number of criteria that a study meets is the nature of the methodological standards that are violated. Several criteria are particularly salient. This body of research overwhelmingly relies on retrospective assessments of maltreatment, typically with long recall periods, despite the well-known psychometric weaknesses of retrospective measures (Henry et al., 1994; Pears & Capaldi, 2001; Widom et al., 2004). Few studies are based on representative samples and many rely on highly selected clinical samples. Many studies rely on a single reporter to assess both their own maltreatment and the maltreatment in either the prior or subsequent generation. And most studies have relatively short follow-up periods, assessing the child generation’s maltreatment only during a small portion of the overall possible exposure period of 18 years; obviously, whatever statistical relationship was observed during a relatively small window can change with continued follow-up. Studies of the cycle of maltreatment hypothesis typically fail to meet several of these core criteria. Taken together, this failure leads to the primary conclusion of our systematic review: the assessment of the cycle of maltreatment hypothesis rests on a very shallow body of scientific evidence that suffers from fundamental methodological weaknesses.

We recognize that not everyone would agree with all of our judgments about these criteria for each of the studies. For example, other scientists might be more satisfied with the clarity of the conceptual definitions or the psychometric adequacy of the measures. We are hard-pressed, however, to imagine that a reasonable reviewer would come to the conclusion that this is a strong, or even adequate, body of scientific evidence. Again, few studies meet even half of these criteria and the most damaging limitations—retrospective data, single reporters, highly selected samples, and short follow-up periods—are, unfortunately, frequently occurring among these studies.

Overall, when methodological adequacy is taken into account, it is very difficult to reach any firm conclusion about the cycle of maltreatment hypothesis. This review identified nine studies that, in our view, appear to be stronger than the others. Widom’s research is by far...
the best, meeting 10 of the 11 criteria, but unfortunately suffers from a fundamental flaw for testing this particular hypothesis—the lack of variability in the perpetration of maltreatment measure. The other eight studies are those that were reviewed in some detail in the previous sections. If we focus just on these nine studies, four find general support for the hypothesis (Dixon et al., 2005a; Egeland et al., 1988; Pears & Capaldi, 2001; Thompson, 2006), three find very limited support for only one type of maltreatment (Berlin et al., 2011; Renner & Slack, 2006; Sidebotham et al., 2001), and two find no support for the hypothesis (Altemeier et al., 1986; Widom 1989a). In other words, the weight of the evidence is quite mixed. Thus, the general notion that maltreatment begets maltreatment that is often drawn from this literature is primarily based on the other, methodologically weaker, studies. We concur with Ertem et al.’s (2000) conclusion that there is likely to be a significant association between maltreatment in the first generation and maltreatment in the second generation but, at the present time, there is insufficient scientific evidence to draw a definitive conclusion about the cycle of maltreatment hypothesis. In commenting on the general acceptance of the notion that maltreatment begets maltreatment, Belsky stated that “there are few in the scientific community who would embrace such remarks… most scholars are all too aware of the inherent limitations of the available database” (1993, p. 415). Unfortunately, the database does not appear to have improved markedly in the ensuing time period. It is important to emphasize that when more well-designed studies are conducted is quite possible that there may be even stronger evidence in favor of the cycle of maltreatment hypothesis than is currently assumed. In other words, the methodological limitations of the existing literature preclude firm conclusions in either direction. In many ways, the issue of intergenerational continuity and maltreatment is an open question at the present time.

Why then does there seem to be such a general acceptance of the cycle of maltreatment hypothesis in the scientific literature? For example, Thompson claims that, “there is much evidence to support the intergenerational transmission hypothesis” (2006, p. 58). Our view is obviously quite different. One possible explanation for this inconsistency is found in the number of review articles. In conducting our literature search for this review, which resulted in 47 original studies of the cycle of maltreatment hypothesis, we also uncovered 31 publications that reviewed the literature in this area. In other words, there are almost as many review pieces as there are original studies. Twelve articles or chapters specifically reviewed the evidence about intergenerational continuity in maltreatment and the rest presented more general reviews about the topic of child maltreatment, including a discussion of intergenerational continuity. Thus the findings from a relatively small and, in our view, methodologically weak number of studies appear to reverberate throughout the literature via a disproportionately large number of review articles, many of which do not evaluate the methodological rigor of the studies.

As discussed in the introduction, child maltreatment is a serious individual and public health concern in American society. Victims of maltreatment have been shown to suffer from a number of externalizing and internalizing problems throughout adolescence and adulthood. Recent evidence also suggests that this relationship is likely to be causal (Thornberry, Henry, Ireland, & Smith, 2010). Given these consequences, it is imperative to halt the repetition of maltreatment across successive generations. That, of course, requires a solid
scientific understanding of the phenomenon in order to develop effective prevention programs (National Center for Injury Prevention and Control, 2002). Based on this review, however, the existing literature does not seem to provide the solid scientific basis needed for this task, even with respect to the most basic issue—the level of intergenerational continuity.

Future research efforts should be based on stronger research designs, that is, ones that meet the basic methodological criteria laid out here and by Ertem et al. (2000). Several specific suggestions come to mind. First, future studies should rely on multiple reporters or data sources to measure maltreatment in the adjacent generations. For example, the participants in each generation could respond to survey questions about their own maltreatment experiences. Alternatively, studies could rely on archival records, for example from Child Protective Services, to measure maltreatment victimization or perpetration. And, of course, different sources could be used across the generations. Relatedly, all of these measures of maltreatment are problematic to some extent (Smith et al., 2008) and all have relatively serious limitations. Multi-method/multi-agent approaches to measurement (Patterson et al., 1992) should be developed for this area to provide more robust assessments of maltreatment.

Second, there should be less reliance on retrospective designs especially ones with long recall periods. Studies using Child Protective Service records do not have this limitation as the dates of the maltreatment incidents are typically available. There are also several validated interview measures, for example, the Conflict Tactics Scale and the Childhood Trauma Questionnaire to capture self-reports of maltreatment with relatively short recall periods in prospective study designs. There are ethical issues associated with collecting these data but since all human research participant issues are a matter of balancing benefits and costs, this issue warrants re-examination in an effort to provide more rigorous scientific evidence for an admittedly serious public health concern. If retrospective data are used it would be better to have shorter and uniform recall periods, for example, by using them in prospective studies with cohort designs.

Third, there are unfortunately few studies in this area that use probability samples drawn from a clearly defined population. Many studies of the cycle of maltreatment rely on highly selected groups not selected using a probability sample so that it is impossible to weight the sample appropriately. Community samples should become the norm and not the exception. Moreover, if there is a need to include more “at risk families” in the study to measure this relatively rare event, that can be accomplished in the framework of probability samples by using appropriate stratifying techniques. In addition, it will be helpful to investigate this relationship in various subpopulations to help identify important moderators of intergenerational continuity in child maltreatment.

Fourth, it is important to control more adequately for potential confounding variables in assessing the cycle of maltreatment hypothesis. In survey designs this means identifying and measuring more than two or three antecedent variables and then holding them constant in multivariate models to more precisely assess the impact of maltreatment victimization on the subsequent perpetration of maltreatment. In matching designs it means to have a fuller set of matching variables to make the maltreated and the comparison group as similar as possible and to use appropriate multivariate modeling to account for residual differences. There is
also a relatively new set of techniques, such as propensity score models (Rosenbaum & Rubin, 1983), to control for selection effects but to our knowledge they have not been used in this research arena.

In general, the methodological criteria proposed here and by Ertem et al. (2000) are not in any way exotic or esoteric. As noted above, they do little more than recommend the use of basic research methods—representative samples, independent measures, prospective designs, etc.—that are appropriate for the investigation of this particular topic. Movement towards their use is essential if we are to have robust scientific evidence about the cycle of maltreatment hypothesis.

Given the importance of this complex research topic both for understanding human development and for establishing effective prevention we first need to understand the nature and strength of the relationship between a history of maltreatment victimization and the likelihood of maltreatment perpetration. If indeed there is clear and compelling evidence of intergenerational continuity, we then need to understand the mediating processes that link the generations in this regard. We also need to understand the moderating influences that yield discontinuity or intergenerational resilience to better inform the design of prevention programs. Last, we need to push forward the evaluative research on the intergenerational hypothesis by not only assessing the basic criteria discussed above but also by including a more thorough analysis of the biases and implicit assumptions that are beyond the purview of this methodological review (e.g., a consideration of the best factors on which to match, the use of multiple sources of data for each respondent, the differences between harsh parenting and abuse, the perpetrator’s relationship to the child, and the various forms of maltreatment and their distinct intergenerational pathways). We clearly have the methodological knowledge to move this area of inquiry forward and it is imperative that both funding agencies and the research community begin to do so.

Acknowledgments

Support for this project has been provided by the Office of Juvenile Justice and Delinquency Prevention (2006-JW-BX-0074) and the Centers for Disease Control and Prevention (R01CE001572). Points of view or opinions in this document are those of the authors and do not necessarily represent the official position or policies of the funding agencies.

References


Browne, KD.; Herbert, M. Preventing family violence. Chichester, UK: Wiley; 1997.


*Trauma Violence Abuse*. Author manuscript; available in PMC 2014 May 27.


Geddes B. How the cases you choose affect the answers you get: Selection bias in comparative politics. Political Analysis. 1990; 2:131–150.


Grayson J. Maltreatment: Long-Term Effects. Virginia Child Protection Newsletter. 2010 Winter;87


Patterson, GR.; Reid, JB.; Dishion, TJ. Anti social Boys. Eugene, OR: Castalia Publishing Company; 1992.


Widom CS, DuMont KA, Czaja SJ. A prospective investigation of major depression disorder and comorbidity in abused and neglected children grown up. Archives of General Psychiatry. 2007; 64:49–56. [PubMed: 17199054]


Biography

Terence P. Thornberry’s primary interest is in developmental, life-course criminology.

Kelly E. Knight’s primary interest is in life-course criminology and the family.

Peter J. Lovegrove’s primary interest is in intergenerational patterns of crime and delinquency.
## Appendix: Evaluation Criteria Met by Each Reviewed Article

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Altemeier et al. (1986)</td>
<td>SR, CPS</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>2. Avery et al. (2010)</td>
<td>SR, Official</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>3. Baldwin &amp; Ollier (1975)</td>
<td>SR, Official</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>4. Black (1980)</td>
<td>Same Respondent</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>5. Bree et al. (2001)</td>
<td>SR, CPS</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>6. Cave &amp; Milner (1962)</td>
<td>CAPS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>7. Caplice et al. (1984)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>8. Coley &amp; Brot (1995)</td>
<td>SR, CPS</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>9. Dillinger et al. (1993)</td>
<td>SR, Official</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>10. Diller et al. (2000)</td>
<td>CAPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>11. Dunn et al. (2002a)</td>
<td>SR, CPS</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>12. Dunn &amp; Miller (1985)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>13. Elgin et al. (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>14. Faller (1990)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>15. Faller (1998a)</td>
<td>SR, CAPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>16. Faller &amp; Kilstrom (1979)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>17. Faller &amp; Kilstrom (1979)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>18. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>19. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>20. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>21. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>22. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>23. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>24. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>25. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>26. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>27. Faller &amp; Kilstrom (1980)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ney (1988)</td>
<td>Same Respondent</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Oates et al. (1999)</td>
<td>SR, Official</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Ornduff et al. (2002)</td>
<td>CAPI</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Pears &amp; Capaldi (2001)</td>
<td>SR, 38</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Smith &amp; Adler (1991)</td>
<td>SR, Official</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Smith &amp; Hanson (1975)</td>
<td>Official</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Straus (1979)</td>
<td>Same Respondent</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Tutty (1999)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Tutty (2006)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Tutty (1995)</td>
<td>Same Respondent</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Wallace (1980)</td>
<td>Official, Official</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Widom &amp; Berenson (1979)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Widom &amp; Berenson (1986)</td>
<td>SR, CPS</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Zaidi et al. (1989)</td>
<td>SR, CPS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

Note: SR = Self-Report; CPS = Child Protective Services; CAPI = Child Abuse Potential Inventory; Official = Official Record (Arrest, Hospital, etc.) G1 = Generation 1 (Parent); G2 = Generation 2 (Child)