Double Jeopardy: Risk Assessment in the Context of Child Maltreatment and Domestic Violence

Aron Shlonsky, MSW, MPH, PhD
Colleen Friend, PhD, LCSW

Investigations of child maltreatment often involve domestic violence, but there is little guidance about how to properly assess risk in such cases. Empirically validated risk assessment tools have been used successfully in child welfare and, to a lesser extent, in cases involving domestic violence, but these have generally not been utilized in tandem. Using the allegation of child maltreatment as the entry point for services, this paper proposes a nested risk assessment framework whereby risk of both child maltreatment and domestic violence are considered simultaneously using two different standardized instruments.


KEY WORDS: child abuse, domestic violence, risk assessment.

Responding to child maltreatment (CM) is far more complicated than keeping children “safe” or “protected” from their own parents. The twin goals of safety and permanence imply that caseworkers must consider both the safety and ultimate well-being of the child. That is, at each decision point, caseworkers must weigh the potential for harm if nothing is done (i.e., leaving the child in a potentially abusive home) with the risk that intrusive actions aimed at child protection will, ultimately, prove to be harmful (i.e., unnecessarily separating a child from their parent). This is no simple equation and the stakes are high. Yet the combination of severe consequences, the inherent difficulty of making accurate assessments, and differences in skill levels among Children’s Protective Services (CPS) workers is a set-up for unreliable case decision-making (Shlonsky & Wagner, 2005).

CPS has begun to deal with this complicated decision-making context by using various assessment tools (Rycus & Hughes, 2004; Shlonsky & Wagner, 2005). These include one or more of the following:

1. Safety assessments: consensus-based lists of factors thought to be related to the likelihood of immediate harm.
2. Actuarial risk assessments: empirically derived estimations of the likelihood of maltreatment recurrence over time.
3. Structured contextual assessments: detailed appraisals of individual and family functioning.

© The Author 2007. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org.
The combination of such approaches allows caseworkers to more simply and reliably assess whether children might be safe if left in the home (safety assessment), generates a more reliable and valid prediction of the likelihood of future harm (actuarial risk assessment), and compiles detailed information that can be used to develop an individualized case plan (contextual assessment). Nonetheless, children and families who are reported for maltreatment often present with multiple problems spanning several service systems, each carrying its own risk of harm. Among the most serious of these is domestic violence (DV). Although DV is often included as an item in safety and risk assessments, the intersection of these two threats to children may necessitate an expanded course of action.

This paper is conceptualized in the context of responding to CM allegations. That is, it assumes that the entry point for co-occurring CM and DV cases is a CM allegation. From this perspective, the literature is reviewed with respect to the prevalence of DV and how it is linked to CM. Next, we examine the challenges in making predictive assessments in both DV and child protection, positing that a nested or layered risk classification system offers the greatest potential to assist caseworkers in making service decisions. Key to this nested approach is the integration of safety and risk assessment information with a detailed assessment of child and family functioning. This should include consideration of the survivor’s perception of risk and the potential for long-term harm that could accompany a range of responses from either a child’s placement or removal from the home, as well as the child’s remaining in the home. Finally, we suggest that engaging in the process of evidence-based practice (EBP) encompasses the use of these two elements (risk and contextual assessment) and extends to the identification and continued evaluation of services for both CM and DV. At the outset, we concede that the science of predicting human behavior, especially when it comes to violence, is complex, risky, and not likely to be mastered. Nevertheless, there is a public expectation that vulnerable children and parents will be protected from repeated assault and that state intervention is both necessary and acceptable to prevent such injury (Finkelhor, 1990). Though imperfect, this integrated approach appears to hold promise for minimizing harm and providing effective services.

**DV and CPS: Scope and Consequences**

The terms DV and intimate partner violence can be used interchangeably to represent a pattern of battering or abusive acts in the context of an intimate relationship. DV spans a continuum of severity and includes physical, sexual, and emotional abuse (Roberts, 2001). In the second family violence survey, conducted in 1985, 16% of American couples (married and cohabitating) reported experiencing at least one episode of physical violence over the course of the relationship (Straus & Gelles, 1986).

Each year in the United States, at least 3–8 million women of all races and classes are battered by an intimate partner (Roberts, 1998; Straus, Gelles, & Steinmetz, 1980). As disconcerting as these figures are, they likely underestimate the true prevalence of single and multiple acts of DV. Such approximations are based upon self-report, and survey respondents may be reluctant to disclose events that harbor feelings of shame or embarrassment. Indeed, Straus and Gelles (1986) estimated that only 14.8% of victims officially report DV incidents. Experts generally agree that women are more likely than men to be seriously physically injured by DV because of men’s greater use of force and severity of tactic (Barnett, Miller-Perrin, & Perrin, 1997). Further compounding these gender differences, women are far more...
likely to experience an injury as a result of assault than men (Straus, 1993).

Although there is still controversy as to whether DV is “bi-directional” involving aggression by both parties, the issue is relevant to a discussion of DV and CM for two important reasons: most reports of child abuse and neglect are made against women (American Association for Protecting Children, 1988; Gelles & Cornell, 1990); and battered women sometimes mistreat their own children (Casanueva, 2005; Ross, 1996; Walker, 1984). Women’s involvement in violence (not merely as responders, but as initiators) has been documented in over 100 studies (Straus, 1999), yet this seemingly intractable finding is at odds with the dominant DV advocacy paradigm that sees women only as victims (Dutton & Nicholls, 2005). This lack of clarity has caused tension between DV advocates and CPS, the latter operating from the standpoint that the child is always the victim.

Estimates of the number of American children exposed to DV vary greatly and are also calculated from data in national surveys. Based on earlier calculations that 3 million American households experienced at least one incident of interpersonal violence in the past year (Straus et al., 1980), Carlson (1984) estimated that 3.3 million children per year are at risk of exposure to parental violence. In their latest (1985) national survey, Straus and Gelles (1990) found that 30% of parents self-reported that their children witnessed at least one incident of physical violence over the course of that marriage. Although this estimate includes incidents that have a wide range of severity, some of which would not be considered by CPS to qualify as CM, the magnitude of the problem in the general population is of grave concern.

Children exposed to parental violence are frequently the victims of co-occurring maltreatment. This co-occurrence has been investigated in single site, clinical samples, and shelter samples of abused women and their physically abused children, with rates of co-occurrence ranging from 30 to 60% (Appel & Holden, 1998; Edleson, 1999). In Canada, where exposure to DV is treated as a maltreatment category, approximately 34% of substantiated cases in 2003 involved exposure to DV and 28% of indicated maltreatment reports were substantiated primarily for DV exposure, making DV the second most common form of substantiated maltreatment (Trocmé et al., 2005).

Clearly, children are at risk of abuse from both adults in the household. As already discussed, the risk of CM from battered mothers is an important consideration when discussing risk assessment. It is equally important to note that in households where the male batterer abuses his partner, batterers may also physically abuse the child. Estimates of such co-occurrence range from 47 to 80% (Hart, 1992; O’Keefe, 1995), making it imperative that DV advocates and child protection workers understand and come to terms with both forms of abuse (Mills et al., 2000). Several studies and scholars have identified child protection workers’ tendency to hold the mother to a higher standard of responsibility than her partner in protecting her children (Davidson, 1995; Davis, 1995; Magen, 1999; Mills, 2000). DV advocates propose that this is a gender bias, and the differential perception of the role of the battered woman and the batterer has led to friction between the two service systems (Beeman, Hagemeisten, & Edleson, 1999; Hartley, 2004; Saunders & Anderson, 2000). Edleson (1999) aptly notes that the CPS system may lack the authority to hold a male batterer accountable if he is not the father of the children. As Hartley (2004) correctly points out, not all reported CM cases in families with DV are inaccurate in their assessment of failure to protect. In some cases, 1

---

1A child has directly or indirectly (e.g., observed physical injuries or heard the violence) witnessed violence occurring between a caregiver and his/her partner (Trocmé et al., 2005).
children are also being physically abused or neglected by both parents. Thus, DV and CM (including neglect) can be two simultaneously occurring events (Hartley, 2004). Having found a surprisingly high level of neglect by mothers in families with severe DV, she argues for a continuing shift from a view of mother’s failure to protect to a view that recognizes the need for interventions focusing on the circumstances that endanger both mother and child (Hartley, 2004).

New information about DV in the context of CPS is also emerging from The National Survey of Child and Adolescent Well Being. This survey begins with a U.S. national probability sample of children investigated for abuse and neglect between October 1999 and December 2000 and follows them for the next 3 years. Casanueva, Foshee, and Barth (2004) used these data to investigate emergency room visits by children. Although the survey is limited to the primary caregiver’s self-report of DV, and only a few caregivers were willing to acknowledge that their child’s injury was due to DV, mothers’ reports of current, severe DV were positively associated with children’s use of the hospital emergency room. The authors went on to find that maternal depression (a key factor associated with child neglect) and lack of supervision (an element of child neglect) were also associated with children’s injuries. They concluded that the identification of current, severe DV in the home and depression among mothers would help prevent future injuries to children. Taken as a whole, Casanueva (2005) and Hartley’s (2004) work supports an earlier finding made by a Los Angeles Juvenile Court in In re Heather A. (1997). Here the Los Angeles Court of Appeals supported the lower court’s finding that children’s exposure to DV, if only secondary, constituted neglect on the part of the battering father.

Best practice for families affected by both DV and CM calls upon advocates and child protection workers to “see double”; meaning they need to draw from knowledge and understanding of both perspectives (Fleck-Henderson, 2000). But “seeing double” comes with its own set of impediments, relative to the way these families enter and behave in the two systems. In the DV track, women may self-report and voluntarily remain for “services” after the violence becomes intolerable. On the other hand, entrance into the CPS system is typically not self-initiated. Services are generally involuntary and the child’s removal is feared by most families.

Children who witness DV can experience a broad range of harmful responses including behavioral, emotional, or cognitive problems that may follow them into adulthood (Edleson, 1999; Felitti, 1998; Grove, 1999; Nicholson v. Williams, 2002). When children both witness and experience abuse, they are more likely to exhibit severe behavior problems than children who only witness DV or children who are not exposed at all (Hughes, 1988), making effective intervention all the more important. Despite the increased risk of poor outcomes, some children display remarkable resiliency in the face of exposure to violence. Such resilience may be moderated by the level of violence, degree of exposure, child’s exposure to other stresses, and his/her innate coping skills (Edleson, 1999). On the other hand, Grove (1999) attributes this resiliency to children being able to talk about the problem and the presence of another adult who can both mediate the experience and promote coping, which would coincide with the findings of resiliency studies (Werner, 1995; Werner & Smith, 1992). Canadian researchers found that 26% of the children in their school sample could be classified as resilient, despite their exposure to DV (Wolfe, Gafee, Wilson, & Zak, 1985). Although not immediately obvious, such findings have serious implications for responding to DV in the context of CM. A U.S. district court judge found these arguments of resiliency to be persuasive when he ordered New York City’s Administration for Children’s Services to stop removing children solely
because they saw their mother being beaten (Nicholson v. Williams, 2002). This challenge to a common practice in one of the largest Public Child Welfare agencies in the country put the entire CPS system on notice that their decisions about removal had to adequately protect the rights of the nonabusing parent and consider the overall well-being of the child.

The Challenge of Prediction

The challenges posed in making protective services risk determinations have been detailed elsewhere (Gambrill & Shlonsky, 2001; Wald & Woolverton, 1990), as have risk decisions in DV response (Cattaneo & Goodman, 2005; Dutton & Kropp, 2000). However, few studies have integrated the two areas. DV and CM assessment share many of the same methodological issues in terms of predicting risk and making subsequent service decisions. Specifically, the discovery of CM and DV begs the questions: will it happen again if nothing is done? What are the consequences if it does recur? How might my actions, as a worker, forestall this eventuality? Who is my client—the child, the battered parent, the abusing parent, or all three? At the agency and policy level, what can we do to make sure that we are expending scarce resources only on cases where CM and/or DV are most likely to recur? How can we tell whether services are effective?

Cognitive Biases and Thinking Errors

Clearly in both fields, clinical prediction of risk is marked by cognitive biases and thinking errors, resulting in decisions that tend to have limited predictive validity (Dawes, 1994; Grove & Meehl, 1996). The sheer volume of observed information, the speed in which decisions must be made, and the pressure to “get it right” can influence a worker’s assessment of risk (Shlonsky & Wagner, 2005). Yet there is little evidence that, in the face of such demands, workers can make reliable and valid predictions of future events. In fact, the opposite is likely true, even for those armed with good information and experience (Dawes, 1994; Dawes, Faust, & Meehl, 1989). One of the major reasons for this shortcoming involves the inability of most people to accurately weigh and combine large amounts of disparate and often conflicting information, prompting the worker to select factors for the decision that have no relationship to the behavioral outcome being forecast (Faust, 1984; Shlonsky & Wagner, 2005). For example, in a CM case, an investigative worker might understandably focus on a parent’s combativeness with them rather than on their overall parenting skills. There are situations in which experts can quickly and accurately make judgments (Klein, 1998), but these rarely involve long-term predictions of human behavior.

Fortunately, formal risk assessment measures have been developed in both child protection (Rycus & Hughes, 2003) and DV services (Cattaneo & Goodman, 2005; Dutton & Kropp, 2000) in order to combat the shortfalls of unassisted clinical judgment. These tools are designed to guide decision makers to those characteristics and observed behaviors that best predict the event of interest. Although there is still some debate about whether tools based on a consensus of experts (consensus based) or that employ statistics to generate an optimal combination of factors that predict the event (actuarial) are more predictive, actuarial instruments tend to perform at least as well as consensus based tools and almost always outperform unassisted clinical judgment (Dawes, 1994; Grove & Meehl, 1996). Certainly, this is the case in child protection, where the most rigorous of studies testing actuarial and

---

2Validity (internal): The degree to which a tool measures its stated purpose. Validity (predictive): The degree to which a tool accurately predicts future events.
consensus-based tools favor the actuarial approach (Baird, 1999, 2000).³

Laying this argument aside, then, what other issues should be considered? Why not merely find an actuarial tool that works for both CM and DV, implement it, and be done? If only the world were that simple. Although decisions informed by evidence (in this case, validated tools) promise to be better than decisions based on other sources, their predictive capacity is quite limited due to the near impossible task of predicting human behavior, as well as the difficulty of accurately predicting events with a low base rate of occurrence (e.g., femicide, child death by maltreatment)⁴. In other words, tools can only go so far. In addition, there are several methodological and contextual factors that must be addressed when considering both CM and DV. Finally, actuarial tools are designed for a very specific purpose: making an optimal classification of risk (e.g., low, medium, high). They are not inclusive of all risk factors and there is no guarantee that risk factors are causal for recurrence rather than links in a chain originating elsewhere. That is, the factors contained in a risk assessment instrument cannot be used to develop a comprehensive service plan.

The Tools and Their Capacities

Risk assessment tools for DV have been under development and in use for at least the last decade (Fein, Vossekuil, & Holden, 1995), but there has been somewhat limited success in predicting recidivism (Hilton & Harris, 2005). Two commonly used and validated instruments are the Danger Assessment (DA) and its revision (DA2) (Campbell, 1995; 2004), and the Spousal Assault Risk Assessment or SARA (Kropp, Hart, Webster, & Eaves, 1995). The DA and DA2 are measures designed to predict the risk that a woman will be killed (femicide) by her partner. The DA was validated retrospectively on a small sample, calling its properties into question and presenting some interpretive problems (Dutton & Kropp, 2000). Acknowledging these limitations, Campbell (2004) recruited a larger and more diverse, multisite sample and revised the instrument based on her findings (2004). All but one of the 15 yes/no items were significant predictors of intimate partner femicide, and the nonpredictive item (perpetrator’s suicidality) was retained due to its theoretical relationship with femicide. Five items were added and a few combined and otherwise modified. The DA2 (available at http://www.dangerassessment.com/WebApplication1/pages/product.aspx) contains 20 items and is reported to have acceptable reliability ranging from 0.74 to 0.80.⁵ Given that the DA2 is predicting lethality, a fairly rare event, there are concerns about its ability to simultaneously identify women at risk of femicide and women who are not at risk. That is, as sensitivity (ability to detect women who will be killed) is increased, the specificity (ability to predict women who will not be killed) decreases.⁶ For example, in Campbell’s (2004) study, a cut-off score of four produced a sensitivity of 83.4%, meaning that 83% of the women who were killed were correctly identified.
identified retrospectively.\textsuperscript{7} The trade-off for such a sensitive instrument is a specificity of 39.2%. As a result of this statistical dilemma, the number of false positives (number of women incorrectly predicted to be killed) is very high. This does not mean that the instrument is not valuable or well constructed but, as discussed below, it does raise philosophical and political questions about where the “bar” is set.

Although the DA and DA2 are important factors for intimate partner femicide, this represents a small (albeit important) part of all DV assaults. The most common forms of family violence are “minor” violent acts, and those acts are performed by both genders (Straus & Gelles, 1990). The SARA, on the other hand, is a consensus-based clinical checklist of 20 factors clustered into five areas. The SARA’s original purpose was to structure and enhance professional judgments about risk (Dutton & Kropp, 2000). Similar to actuarial tools in use in child protection (Wagner & Johnson, 1999), the SARA allows for clinical overrides in order to incorporate some level of clinical judgment into risk decisions (Dutton & Kropp, 2000). Although the SARA’s interrater reliability is reported to be high and its internal consistency moderate, evidence of predictive validity (the ability of the tool to predict DV) is modest (Heckert & Gondolf, 2004). In addition, it is unclear whether the SARA’s psychometric properties have been tested on a CPS sample.\textsuperscript{8} The SARA’s 20 factors each have a range of response categories consisting of three items: 0 (absent); 1 (subthreshold), and 2 (present). Each of these items is totaled and risk of DV is said to increase as the score increases but, unlike an actuarial approach, there appear to be no pre-established cutpoints to establish low, moderate, or high degree of risk. Unlike the DA, the SARA is completed by a caseworker and, ultimately, yields an estimation of harm rather than lethality. Although they differ in the severity of what they seek to measure, the good news is that both the SARA and the DA2 share certain factors, indicating that there may be reasonable convergence between the two. The measures also appear to have fairly good reliability and are easily completed. Nonetheless, overall predictive validity of both tools remains modest.

The Ontario Domestic Assault Risk Assessment is an actuarial tool developed for use by police officers conducting DV investigations and, in this context, appears to predict recidivism better than the SARA (Hilton et al., 2004). Hilton and Harris (2005) describe the tool as a 13-item scale consisting of DV history, general criminal history, threats and confinement during the most recent assault, children in the relationship, substance abuse, and victim barriers to support. Similar to other actuarial tools, each item is binary (0, 1), and the total score is used to generate a probability of recidivism. This tool holds promise for a number of reasons. First, its psychometric properties (Hilton et al., 2004) appear to be similar to other actuarial tools used in different fields. Somewhat related, its simple, easy to use structure will likely increase the reliability of DV risk ratings and, by extension, the validity of such predictions. Moreover, the tool was designed for police investigations, and such inquiries have at least some similarity to CM investigations in terms of their immediacy and inherently coercive nature. Nevertheless, like the DA and SARA, the Ontario Domestic Assault Risk Assessment has not been normed on a CPS sample.

Child protection safety and risk assessment tools have also been in use for some time (Fluke, Edwards, Bussey, Wells, & Johnson, 2001; Johnson, 1984; McDonald & Marks, 1991; Wald &

\textsuperscript{7}Because Campbell was investigating lethality, her informants were often mothers, sisters, and friends of the decedent.

\textsuperscript{8}Psychometric properties: The set of constructs, including various forms of reliability and validity, that reflect upon a tool’s viability.
Woolverton, 1990), though the quality of the measures and the integrity of their application vary. In general, these tools are designed to predict the risk of immediate harm (safety assessment) or risk of maltreatment recurrence over time (risk assessment). The safety assessment is usually completed shortly after the initial contact with the family, and the risk assessment is usually completed toward the end of the investigation period. Unfortunately, most of the early tools lacked sufficient predictive validity to be of much use in the field (Lyons, Doueck, & Woodarski, 1996). More recently, however, safety and risk assessment tools have been successfully used in the field to more accurately contend with unsafe situations and high-risk families (Fluke et al., 2001; Johnson, 2004; Wagner & Johnson, 2003). However, these instruments also suffer from an inability to predict at high levels of accuracy for the same problems detailed above (i.e., high sensitivity and low specificity). There is some evidence, though, that a well constructed, easily scored actuarial instrument can be effectively used in the field. Following up on the retrospective validation of a similar tool (Baird & Wagner, 2000), Wagner and Johnson (2003) and Johnson (2004) conducted a prospective validation of the California Family Risk Assessment using a sample using over 7,000 CPS cases from a variety of California counties. Each tool was completed by trained workers in the field during the course of their investigation. They found that the instrument maintained its psychometric properties indicating that, with proper training, the instrument transfers well to the field.

The Challenge of Measuring and Defining Outcomes

The prediction of CM is made difficult in the face of vague definitions and outcome measures (Gambrill & Shlonsky, 2000; Wald & Woolverton, 1990), and this likely translates into the DV sphere as well. Arguably, physical and sexual abuse can be more readily defined and classified in terms of severity than other forms of maltreatment. However, child neglect, the most common form of maltreatment in the United States (U.S. Department of Health and Human Services, 2004) is subject to widely ranging definitions and cutpoints (measurable point beyond which one can say neglect has occurred) across studies (Zuravin, 1999). Defining DV itself might be an easier task, but defining when DV becomes child abuse is another matter. Although there are those who would argue that witnessing DV is a form of CM, and to some extent they may be right, this is not always a viable reason for mandated services and, ultimately, removing a child from their family or home (Nicholson v. Williams, 2002). At what point does CPS become involved in the response to DV? If we base this on emotional harm to the child, how is this measured? The subtleties involved may make the creation of valid cutpoints untenable. The presence of children who appear to be resilient to some of the measurable effects of DV (Edleson, 1999) indicates that children may react differently to similar types of exposure to violence. What is not clear is whether these same children would remain resilient if they were removed from the care of their parents. That is, if resilience is a confluence of personal and situational factors, a change in situation might result in a change in resilience. If resilience involves personal coping strategies, insight capacity, and parental relationship, then removal might compromise or overwhelm the individual’s capacity to maintain these so-called traits.

Many risk assessment tools use substantiation or indication (social work finding that maltreatment has occurred) as the sole measure of maltreatment recurrence with the acknowledgment that it is limited to known recurrence. For instance, there are an unknown number of
children who are maltreated but are not reported to CPS (English, Marshall, & Orme, 1999). Similarly, there may be a surveillance effect (families receiving CPS services are under increased scrutiny), and such children may be reported more often than would otherwise be expected (Fluke et al., 2001; Lindsey, 1994). Practically speaking, however, substantiation remains the best measure available for reabuse. In addition, valid instruments that measure risk rereport, child injury, and foster care placement have been developed and can be used to inform the decision-making process (Wagner & Johnson, 2003; Johnson, 2004). For example, a high-risk rating for a child on the injury scale may inform a service decision differently than a high-risk rating for rereport. DV studies have a similar problem in that they largely rely on subsequent police reports to measure recurrence, though there have been studies that use victim self-report as well (Dutton & Kropp, 2000).

Reliability and validity of the tools is also a challenge. DV and CM risk assessment tools range in quality, and it is exceedingly important to ascertain a tool’s psychometric properties. However, even the best tools have limitations. Risk of DV and CM is not static. That is, risk likely changes over time in CM cases (DePanfilis & Zuravin, 1998) and about half of DV incidents are single occurrences (Dutton & Kropp, 2000). Thus, we may be observing an escalation or de-escalation at any given moment in time. If escalation is always assumed at the point of risk assessment, the false positive rate might be very high, whereas if escalation is not assumed the number of false negatives might be high (Gambrill & Shlonsky, 2000).

In addition, attempts to make simple (yes/no) predictions of whether a child will be reabused are problematic. For instance, the California risk assessment instrument, while meeting key standards for reliability, is unable to predict maltreatment recurrence at acceptable levels if it is constrained to simply predict whether maltreatment will recur (more detail presented below and in Shlonsky & Wagner, 2005). Again, this is due to the near impossibility of trying to predict complex human behavior. Thus, even the best risk assessment tools should not be used as the sole decision-making device, but a good actuarial classification system can be used to reasonably inform service decisions.

**Goal and Role Confusion**

The co-occurrence of DV and CM raises some serious questions about the very nature of services involving children and families. Much of the emphasis in child protection is focused on keeping children safe and facilitating a permanent home. Yet, a child is less likely to be safe if the parent is not safe. Clearly, the welfare of children depends upon the welfare of parents. Likewise, a response to DV that does not consider issues of CM that go beyond DV (i.e., that the assaulted parent may also be abusive or neglectful) errs in the other direction. For the purposes of this paper, we are only focusing on children who are reported for maltreatment. Even with this smaller population, a number of different types of risk are present when factoring in the occurrence of DV. These risks generally fall into two categories, risk of harm to the child and risk of harm to the parent, and include:

1. CM that is not directly DV involved.
2. CM as a direct result of DV.
3. Child emotional harm as a result of observing DV.
4. Parent physical harm as a result of DV, potentially limiting the parent’s ability to meet the child’s needs.
5. Parent emotional harm as a result of DV, potentially limiting the parent’s ability to meet the child’s needs.
These overlapping risks pose considerable challenges to both measurement and service response. Actuarial models of risk assessment are statistically derived sets of factors that estimate the likelihood of an event. The items themselves are not necessarily causal. That is, their presence may predict an event without actually causing it. Although it seems logical that DV is both a risk factor and causal for maltreatment recurrence, most tools use overall maltreatment recurrence as a benchmark, not whether it recurred in the context of a DV incident. That is, the presence of the risk factor of DV indicates that some children are probably reabused as a direct result of DV between partners, but this is a subset of the larger group of children who are reabused for other reasons. Thus, predicting maltreatment is not predicting DV, and vice-versa.

Additionally, items on child protection risk assessment instruments often ask questions about whether there is currently DV in the home or whether the primary caregiver has a history of DV. What is generally not asked is whether the child was physically or emotionally injured during a DV episode. This is a critical point of inquiry, otherwise there may not be a child protection issue. The relationship between prior violent acts (presumably including physical abuse of children) as well as battery while pregnant have been established as markers for femicide (Campbell, 1995). Child injury during a DV incident likely indicates a level of severity that should not be ignored. Along these same lines, consideration should be given as to whether a parent was injured as part of the DV issues that brought the family to the attention of CPS.

Integrated Assessment Strategies: A Proposed Solution

Despite the fact that actuarial prediction is likely to produce results that are better than clinical decisions alone, the reality is that we are currently unable to predict either CM recurrence or DV lethality or injury at sufficient levels to make outright statements about whether either will occur in the future. There are just too many unexplained factors and the phenomena being predicted occur too infrequently to attain great accuracy. To illustrate this point, Shlonsky and Wagner (2005) combined the four classifications (low, moderate, high, very high) of the California risk assessment instrument into two risk classifications forming a simple (yes/no) prediction of whether maltreatment would recur. Although this configuration predicted at levels slightly greater than chance, the rate of false positives (predicting that individuals would reabuse when they, in fact, did not) was exceedingly high. For such a low base rate of recurrence, the best prediction would be that it would not happen. Similarly, the DA2, although clearly reliable in the sense that it predicts lethal DV quite a bit better than chance alone, it suffers from the same inability to make an outright (yes/no) prediction (Campbell, 2004 website referencing psychometric qualities). The limited predictive capacity of high-quality tools means that the best we can do is to develop classification systems that categorize people into varying degrees of risk and tailor the intensity of the response according to these groupings. In other words, we make a statistically informed guess about what will happen in the future and respond accordingly. Given the level of accuracy of risk assessment tools in these fields, a forensic conclusion would never say more than, “this family is at higher risk than most other families” for one or both of these outcomes.

---

9The California risk assessment tool and other Children’s Research Center measures do distinguish between physical abuse and neglect as outcomes.
With this limitation in mind, actuarial approaches categorize individuals and/or families into graded levels of risk. Examples of this approach in child protection are the Michigan Actuarial Model which was validated retrospectively (Baird & Wagner, 2000; Baird, Wagner, Healy, & Johnson, 1999) and the California Actuarial Risk Assessment which has now been validated prospectively (Johnson, 2004; Wagner & Johnson, 2003). These models consist of a short set of questions, mostly binary, that have been found to separately predict abuse and neglect. Again, despite its limitations, this actuarial model clearly differentiates level of risk for resubstantiation, subsequent child placement, and child injury (Figure 1). As level of risk increases, the percentage of children experiencing these outcomes increases. Children classified in the highest risk categories have a higher likelihood of experiencing these events, whereas children classified in the lower risk levels have a lower likelihood. The model does not claim to be right every time, nor is it intended to be the sole source for decision making. The risk assessment tool simply assigns a level of risk relative to other cases (Shlonsky & Wagner, 2005). If an instrument cannot adequately distinguish between risk categories, then it cannot serve as a decision aid. That is, if high-risk cases end up recurring as often as moderate risk cases, the decision maker would not gain any information from the tool. A comparison of this approach (Baird & Wagner, 2000) to two commonly used consensus-based tools found that the actuarial tool differentiated between risk levels whereas the two expert-driven models struggled to distinguish between risk levels.

**Nested Risk Assessment**

The presence of related yet separate risk constructs (in this case, maltreatment recurrence and DV recurrence) requires careful consideration with respect to risk instrumentation and application. One of the problems with assessment instruments is their implementation in the field. Instruments that are too long or too difficult to complete are unlikely to be used by practitioners. Clients, too, especially involuntary clients, may not engage with a social worker who asks them countless questions contained on an endless instrument. Thus, a comprehensive risk assessment instrument that covers all areas of risk would be ill-advised. There are statistical as well as practical concerns. How do two instruments interact to alter risk? That is, are all children who are at high risk for DV recurrence also at high risk for CM? Perhaps so, depending on the definition of maltreatment. But is the converse true? Are all cases at high risk for CM recurrence also at high risk of DV recurrence? Clearly not. DV may not have occurred the first time making an assessment of recurrence somewhat nonsensical. If we are functioning within the CPS realm, it would seem that the primary assessment of risk should be CM in all its forms.

A nested approach to risk assessment, with risk of CM recurrence as the first-order assessment, has the potential to optimally employ more than one type of risk assessment instrument. That is, a hierarchy of instruments, beginning with a
maltreatment recurrence measure and moving to other assessment instruments as needed, would provide valuable information for making key service decisions. In child protection, one common approach is to screen cases in for investigation, assign a service priority (i.e., immediate or more delayed response), conduct a safety assessment, determine whether the maltreatment occurred (substantiation decision), complete a risk assessment, and decide whether to open a case for services. This is followed by a contextual assessment and the development of a service plan (see, e.g., Wagner & Johnson, 2003). This approach can be enhanced by conducting a DV risk assessment at various points along this continuum if there is an indication that DV is a current and ongoing issue for this family (Figure 2): If the original allegation includes issues of DV or DV is discovered during the safety assessment, a joint assessment for risk of DV might also be conducted focusing specifically on the immediate risk of harm or danger from DV (e.g., DA2). At the end of the investigation period, the original allegations are found to be substantiated or indicated (the maltreatment occurred), unsubstantiated (insufficient evidence), or unfounded (the maltreatment did not occur). At this point, a child protection risk assessment is completed prior to the decision about whether or not to open a case for services. The decision is informed by the level of risk as well as caseworker input and agency guidelines. If opened for services (ranging from referral to child placement) and DV has been identified in the child protection risk or contextual assessments as a current family issue, a DV screener for general risk of DV recurrence could be administered and the information used for case planning purposes.

Table 1 presents an example of a framework for service decisions based upon risk level of both CM recurrence and DV. These responses are suggestions only. Risk assessment should not dictate service response due to the issues.
touched upon in this chapter and in greater detail elsewhere (Gambrill & Shlonsky, 2000; Shlonsky & Wagner, 2005; Wald & Woolverton, 1990). Especially with mandated services, decisions should be made by carefully weighing risk assessment information and clinical judgment. Due to political considerations and population dynamics, individual agencies may decide on a different set of responses. At the outset, it is acknowledged that most DV instruments have not been extensively tested and, to our knowledge, have not been normed on a CPS sample. This framework is merely a suggestion and any instruments used in this context should meet basic psychometric standards as well as being rigorously evaluated once implemented.

Beginning with the primary assessment for maltreatment recurrence, low-risk cases would result in referrals to services only. The main function of the CPS system is to keep children safe from maltreatment. Low-risk families, despite the likelihood of having fairly serious problems, should generally not be forced to receive such services. High-risk cases, on the other hand, call for joint evaluations and greater intensity of services. High-intensity services might range from voluntary family preservation services to child placement. If a case is rated as having a high likelihood of maltreatment recurrence but is classified as low risk for DV, then the mix of services would not include DV prevention support. Thus, scarce DV resources would be conserved for families with the highest likelihood of having a subsequent DV incident. As risk for both CM and DV increases, so too does the intensity of the service mix.

A classification of risk, whether obtained from a consensus or actuarial assessment, estimates the probability that an event will occur among families with similar characteristics. It is not a perfect predictor, nor is it a cookbook for service decisions. Certainly, it is not a substitute for sound professional judgment, and the finding should not be the sole basis for a case decision. Appropriate use in the field requires that workers understand how actuarial risk assessments work, know the limitations of the estimates they make, and receive the training and policy guidance necessary to employ them effectively in the field (Shlonsky & Wagner, 2005). An important component of both the California Actuarial Tool and the SARA is the presence of an agency and clinical override feature. This option allows case-workers to upgrade the risk level (generally in consultation with their supervisor) in order to respond to information that may not be accounted for in the risk assessment instrument.

### TABLE 1. Service Decisions Based on Risk of Recurrence for CM and DV

<table>
<thead>
<tr>
<th>Recurrence risk</th>
<th>Child protection risk</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV risk</td>
<td>Low</td>
<td>Community referral only, no DV referral</td>
<td>Community referral only, no DV referral</td>
</tr>
<tr>
<td>Low</td>
<td>Community referral only</td>
<td>Community referral only, DV prevention referral</td>
<td>Community CW referral, high intensity DV prevention services</td>
</tr>
<tr>
<td>Medium</td>
<td>Community referral only</td>
<td>Community referral only, DV prevention referral</td>
<td>Community CW referral, high intensity DV prevention services</td>
</tr>
<tr>
<td>High</td>
<td>Community referral only, DV prevention referral</td>
<td>Community referral only, DV prevention referral</td>
<td>Community CW referral, high intensity DV prevention services</td>
</tr>
</tbody>
</table>
However, this feature should be used sparingly. The very structure of a good actuarial instrument would suggest that, on average, clinical overrides will result in less accuracy. This is not to say that a clinical override used on an individual family will always be the wrong decision. It simply means that, over time, the instrument will be correct more often than the clinical decision maker.

The Integration of Actuarial and Clinical Approaches

Despite the advantages of using actuarial tools (e.g., more reliable and accurate assessment of risk), there are clear limitations, some of which have been detailed here. Perhaps the greatest limitation of actuarial approach is that its intended use, assessment of risk, tells us nothing about people except how likely they are to act in a certain way. They are not designed to obtain a detailed understanding of family dynamics and functioning, and they are certainly not designed to be the sole basis of a treatment plan (Shlonsky & Wagner, 2005). Actuarial and clinical judgment must be integrated with the client’s perception of the situation to make prudent decisions about the type and scope of services offered to children and families. This combination offers the greatest opportunity for improving casework decisions.

A comprehensive, contextualized family assessment identifies and clarifies relevant strengths and needs at the individual, family, community, and societal level (Gambrill, 1997), explicates the reasons the family came into contact with the CPS system, and provides insight into the type and scope of services that might be necessary to prevent maltreatment and DV recurrence. An example of such an integration is the Children’s Research Center’s Structured Decision-Making approach. As detailed in Shlonsky and Wagner (2005) and in various state reports (see http://www.nccd-crc.org), the actuarial risk assessment tool is used to help agencies establish the intensity of services. However, case planning relies on a structured assessment of ‘Family Strengths and Needs’ that is completed after the risk assessment and is used to organize clinical assessment findings. This consensus-based assessment is sometimes completed as part of a case or family group decision-making conference, allowing families the opportunity to more fully participate in the assessment and case planning process, and includes such elements as substance abuse, mental health, DV, physical health, family relationships, housing, and social support. Standardization makes worker assessments more reliable, furnishes a brief format for documenting case notes, supplies additional criteria for classifying cases based on prioritized service or treatment needs, and provides useful information for constructing fundamental progress indicators.

The Link to EBP as a Process

This integrated approach to risk assessment can be seen as the beginning of the full-scale implementation of the process of EBP (Shlonsky & Gibbs, 2004) in CPS (Wagner & Shlonsky, 2005). As outlined for evidence-based medicine by Sackett, Richardson, Rosenberg, and Haynes (1997) and adapted for the helping professions by Gibbs (2003), EBP is the integration of current best evidence, clinical expertise, and client state/preferences. This integration is achieved through the process of posing an answerable question, querying a database in order to find current best evidence, evaluating evidence found, and applying it to client and clinical context (Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). Thus, EBP is more than simply the application of an intervention that has some evidence of effectiveness. Rather, it is a process that allows agencies and practitioners to truly
take account of what is known about both the client and the challenges they face.

The nested risk assessment approach described in this paper fits within the EBP conceptual model when visualized as a recursive cycle rather than a single event. Using a more recent conception of the evidence-based medicine model by Haynes, Devereaux, and Guyatt (2002), risk assessment can be seen as an entry point, targeting scarce resources to clients at highest risk (Figure 3). Moving counterclockwise around the circle, a search is conducted for current best safety and risk assessment instruments for use in child protection. Relevant data sources on current best evidence include the Cochrane and Campbell Collaborations (we are referring to both), Medline, Psycinfo, CINAHL, Social Services Abstracts, Social Work Abstracts, and others. Next, the contextual assessment uses clinical expertise to elicit key strengths and needs as well as client preferences as movement is made toward service provision. If, during the investigation process, current DV or a history of DV is discovered, current best evidence is again sought with respect to DV assessment tools (this process would work equally well if other problems such as depression or child behavior problems were discovered). At this stage, service decisions are made with consideration of risk level on both tools (perhaps using a predefined matrix similar to Table 1), family circumstances and preferences, and agency mandates. This stage should include a search of the literature for the current best evidence given the family’s specific problems. Again, rather than simply throwing services at unwilling clients, consideration of the family’s individual and group functioning, their preferences for providers or service type, and any barriers to service that might exist should be carefully weighed and, to the extent possible, used to modify services provided.
There is some debate as to whether conducting a detailed search with every client is realistic given the time constraints faced by caseworkers in the field (Mullen, Shlonsky, Bledsoe, & Bellamy, 2005). Initial searches by caseworkers will be more time-consuming but will amount to updates as problems faced by families are encountered for a second time. There may also be ways for the agency to anticipate the challenges faced by their clients, conduct specific searches for current best evidence with respect to risk assessment tools and effective services, and begin obtaining or developing such resources for use by caseworkers. Some large CPS agencies have already formed in-house special research units to assist line workers and policy makers (e.g., Administration for Children’s Services in New York City and Department of Children and Family Services in Los Angeles). These could provide the necessary infrastructure for an EBP approach at the site or broader agency level. This would not preclude the need for continued searches and revisions of the assessment and service constellation due to the quickly changing state of evidence. Nonetheless, the anticipation of assessment and service provides a solid evidence base upon which to guide service decisions.

Conclusion

Our examination of risk assessment in the context of CM and DV has led us to see a number of pressing needs. First, there should be more cross-disciplinary work. To their credit, states such as Massachusetts have pioneered joint CPS and DV case assessments (Aron & Olson, 1997) but far more needs to be done. DV risk assessment instruments must be normed on CPS populations in order to use these instruments with greater confidence or modify them. For instance, although the SARA is not necessarily a predictive instrument (Dutton & Kropp, 2000), it does contain sets of risk factors that can be developed into an actuarial instrument built upon CPS cases.

Good assessment tools and the skills to use them are meaningless if services are not effective at ameliorating the problems that bring families to the attention of CPS. After multiple batterers’ treatment program evaluations, the most recent national analysis called for improved program evaluations and concluded that it was too early to abandon the concept, and too early to believe we have all the answers (Jackson, Feder, Davis, Maxwell, & Taylor, 2003). Similarly, the literature is clear in recommending group treatment and various components for battered women’s counseling, but program evaluations have been scarce to nonexistent (Lipchik, Sirles, & Kubicki, 1997). In other words, we are not sure what works for whom and at what point. Critics of the current service approaches for DV argue that Western feminist ideology has been the driving force behind the menu of services offered battered women, but that this has been done without adequate evaluation that these approaches lead to enhanced safety (Mills, 2003). Regarding children, Cunningham and Baker (2004) have identified only 11 evaluations of children’s treatment programs for DV exposure in the published literature, none addressing treatment effectiveness. Thus, we have a small set of DV-specific services that consist largely of shelter care, none have been adequately evaluated, and where viable programs for families who decide to remain intact exist, they need to be better publicized. These shortcomings must be addressed by moving beyond standard DV service provision, perhaps toward a harm reduction approach.

Given that Risk assessments tend to be abuser focused and a review of the literature of Cavenaugh and Gelles (2005) found that most male offenders in the low to moderate category do not escalate over time. They make a case for matching these typologies with treatment
interventions, much like the stages of change approach (Prochaska, DiClemente, & Norcross, 1992). Batterers appear to be a heterogeneous population as opposed to the homogeneous, ever-escalating group typified by current approaches to treatment and intervention. Thus, we need to find specific strategies that are effective with particular risk groups. The danger of mismatching a batterer to treatment services, according to Cavenaugh and Gelles (2005) is that it is possible, and perhaps likely, that a batterer may complete a program without having his needs addressed. At its worst, a homogenous approach could undermine the victim’s future safety.

Understanding that battered women are (for the most part) keenly and uniquely aware of their own danger, we need to study how their knowledge can enhance the performance of risk assessment instruments. Perhaps alternative treatment approaches such as the work of Penell and Burford (2002) in Family Group Decision Making and the experiment in restorative justice Approaches for Batterers Treatment proposed and underway by Mills (2005) hold promise for improving prediction and reducing risk of recurrence by engaging the extended family and community members to monitor and provide acceptable resources for at-risk families.

Having explored the connection between CM and DV, as well as the challenges in making predictive assessments, we are advocating a nested risk assessment that considers CM recurrence first, and then proceeds with a DV risk assessment. Both of these then lead to a comprehensive and contextual family assessment which is the basis of connecting the family with appropriate services. Further, anchoring this within an EBP framework will help workers understand the limits as well as the strengths of risk assessment instruments, the proper use of contextual assessment measures, and the range of effective treatment options available to children and families. Integral to this approach, we recommend the following:

- Child protection workers need more specific, focused training in understanding risk assessment. They need to understand the terms discussed in this article (i.e., reliability, validity, sensitivity, specificity), as well as the current state of what Cash (2001) calls the art and science of risk assessment. Similarly, managers and policy makers must understand that there is no way to eliminate risk; there is only the minimization of harm through risk management (Gambrill & Shlonsky, 2001).

- In addition, child protection workers need more training in determining where and when to intervene and how to conduct interviews that are sensitive to the issues surrounding DV. Beyond prediction, the worker’s goal is to prevent recurrence of harm. On the whole, good risk assessment instruments outperform “clinical judgment” with respect to prediction, but there is a role for the worker in assessing the dynamic context of CM and DV. In particular, this will aid in the selection of appropriate treatment. Because instruments such as the DA rely so heavily on victim self-report, workers also need training in engendering a battered woman’s trust, as she may accurately perceive that honesty may put her at risk of losing her children.

- Reliably placing families into graded levels of risk can be readily accomplished with instruments such as the DA and the California family risk assessment tool. As identified earlier, these gradations may be useful in matching typologies to treatment.

- More research is needed to discover how DV and CM might interact to alter risk levels. For instance, they may have shared pathways that converge in child neglect.
We are just beginning to understand how such markers as children’s use of the emergency room, maternal depression, and severe DV are linked. Because we know that both CM (Lindsey, 1994) and DV correlate with poverty, the role of unemployment needs to be fully explored with respect to both risk assessment and prevention of recurrence.

- Effective services must be identified and made available for locally prevalent problems (Shlonsky & Wagner, 2005). Each agency should identify a core set of commonly needed services for the treatment and prevention of DV, CM, and their related problems. Where such services do not exist or cannot be found, old services should be evaluated and innovations sought using the EBP methods discussed here. In any case, the current state of knowledge (or lack thereof) should be acknowledged rather than ignored.

CPS workers face the monumental and often impossible task of trying to prevent maltreatment while keeping families together. The presence of another unpredictable and harmful family problem, DV, raises the stakes even higher. Risk assessment tools, despite their ability to predict future harm, are only the beginning of what is needed to prevent harm. Such tools must be integrated with a structured assessment of family functioning and a set of effective, individualized services geared toward addressing both concerns.

Acknowledgments

The authors gratefully acknowledge the substantial contributions made to this paper by Raelene Freitag (Children’s Research Center), Linda Mills (NYU), and Dennis Wagner (Children’s Research Center). An earlier version of this paper was published as Shlonsky, A., & Friend, C. (2007). Double Jeopardy: Risk Assessment in the Context of Child Maltreatment and DV. In D. W. Springer and A. R. Roberts (Eds.), Handbook of forensic mental health with victims and offenders: Assessment, treatment and research (pp.25–51). New York: Springer.

References


