

Issues in Risk Assessment in Child Protective Services

Judith S. Rycus, PhD, MSW

Ronald C. Hughes, PhD, MScSA

Center for Child Welfare Policy, North American Resource Center for Child Welfare

Introduction

In recent years, the child welfare field has faced mounting moral and political pressures to improve its effectiveness and accountability and to demonstrate its public value. In response, child welfare organizations throughout North America have expended millions of dollars to develop, implement, and institutionalize formal risk assessment systems. However, the literature continues to raise provocative and disturbing questions about all aspects of risk assessment technology and implementation. The question remains whether the results have been worth the investment.

In many jurisdictions, estimates of future risk are still being made largely on the basis of personal opinion and judgment. We continue to rely on tools that lack reliability and validity while believing that these tools standardize and greatly improve decision making. We create idiosyncratic adaptations of existing models for our own use, and we support large-scale and costly implementation initiatives without sufficiently managing the overall impact on the service system. We conduct studies to validate models that are fundamentally flawed at the outset. We mandate the use of protocols that make little sense to the work force and are often abandoned in frustration by the staff who must use them. Striving for improved accountability, we “hard code” entire risk assessment models and instruments into child welfare information systems, further cementing our reliance on this technology and creating potentially insurmountable challenges when changes are needed. And, because large-scale change has historically been so difficult for many organizations, it may ultimately be easier to support ineffective, even potentially harmful, technologies rather than change them, both because of the financial investment already made and because an overburdened work force cannot sustain another large-scale change.

Unfortunately, perhaps due to the many other seemingly intractable problems facing the child welfare field, we appear to have a collective vulnerability to the promises of untested and unproven risk assessment models and technology. Further, although individual researchers and practice jurisdictions have worked to refine and improve risk assessment technologies, the child welfare profession has yet to reach consensus on a plan of action to incorporate the strongest and most promising of these into practice, or to confront the many remaining issues and challenges.

In 2002, the Center for Child Welfare Policy of the North American Resource Center for Child Welfare (NARCCW) undertook an extensive risk assessment initiative, which consisted of the following: 1) a review and analysis of the literature on risk assessment in child welfare; 2) a two-day colloquy attended by researchers, academicians, risk assessment system designers, and child welfare practitioners to identify and explore key issues in risk assessment theory, technology, and implementation; and 3) a review of state, provincial, and agency risk assessment models. This article presents an abridged version of the NARCCW policy white paper that resulted from this initiative.*

Part I: Issues in Risk Assessment in Child Protective Services

Many issues and concerns have been raised and discussed in the child welfare research and practice literature, underscoring the conceptual and operational complexity of risk assessment as a practice technology. Many of these issues can be subsumed under the following six major themes.

A. There is lack of agreement regarding the proper scope and purpose of risk assessment technology in child welfare assessment and case planning.

All risk assessment models encompass four common components: 1) the broad categories (criteria) to be assessed; 2) behavioral descriptors that define and operationalize these criteria (also known as measures); 3) procedures and calculations for determining various levels of risk; and 4) standardized forms to uniformly capture and record this information.

However, existing risk assessment models differ greatly in their scope, their stated purposes, the relative importance or weight assigned to various factors, and the mechanics of gathering, organizing, and interpreting information (Cash, 2001; Pecora, Whittaker, Maluccio, & Barth, 2000; Cicchinelli, 1995; Wells, 1995; English & Pecora, 1994; Doueck, English, DePanfilis, & Mootte, 1993; Wald & Woolverton, 1994). Risk assessment models range on a continuum from a discrete, “point-in-time” assessment of the likelihood of future harm, to case management tools that promote an overarching attention to risk, and its reciprocal, safety, in a variety of contexts and at different decision-making points in the case planning and service delivery process.

Formal risk assessment technology was originally intended to help workers estimate the likelihood of future recurrences of serious child maltreatment in families (Baird, Ereth, & Wagner, 1999; Schene, 1996; Curran, 1995; English & Pecora, 1994). Some risk assessment systems, particularly actuarial models, still adhere to this discrete objective (Baird & Wagner, 2000; Johnson, 1996). In this context, risk assessment’s unique purpose is to evaluate families during the intake assessment and to classify them into groups on the basis of the assessed likelihood of future maltreatment. This information helps workers determine which family cases should be opened and transferred within the agency for more in-depth assessment and subsequent protective services. As only one component of a broader continuum of case management and safety assurance strategies, “point-in-time” risk assessments help assure that agencies focus attention on families in which a future recurrence of maltreatment is most likely. Lower-risk families who have service needs can then be referred to other community providers with reasonable confidence that future child maltreatment is not likely to occur.

At the opposite end of the continuum are risk assessment models intended to serve as overarching systems of data collection, analysis, and decision making throughout the life of a case (Pecora et al., 2000; Doueck et al., 1993; Cicchinelli & Keller, 1990). In these

models, the stated purposes of risk assessment include prioritizing cases for services, identifying a family's individual service needs, informing case plan development, allocating services and resources, reassessing progress, documenting risk reduction, informing reunification decisions, guiding case closure, and establishing work load standards. Because of this disparity in fundamental concepts, premises, and scope, it is often questionable whether professionals discussing risk assessment are even talking about the same thing (Cicchinelli, 1995; Wells, 1995; Cicchinelli & Keller, 1990).

Confusion Between Risk Assessment and Family Assessment

There is also considerable confusion among child welfare practitioners about the difference between risk assessment and family assessment (Pecora et al., 2000; Schene, 1996; Wells, 1995; Cicchinelli, 1995; English & Pecora, 1994; Doueck et al., 1993; Wald & Woolverton, 1994). In contrast to risk assessment, the purpose of family assessment is to identify and explore, in considerable depth, the unique complex of developmental and ecological factors in each family and its environment that may contribute to or mitigate maltreatment. Family assessment data should be used primarily for case planning purposes, to enable the identification and delivery of the most effective interventions to address maltreatment and to prevent its reoccurrence. However, despite significant differences in purpose, scope, and depth between risk assessment and family assessment, many agencies attempt to use a single "hybrid" instrument to do both, resulting in a variety of problems that include the following: truncating the assessment to fit within limited time frames; superficial assessments and sparse, boilerplate case plans; subjecting all families at intake to a level of scrutiny that may exceed the level necessary to simply determine the likelihood of future harm; wasting caseworkers' time; and increasing the likelihood of bias and error. In such situations, neither standardized risk assessments nor in-depth family assessments are effectively completed, and a preponderance of casework decisions may continue to be made largely on the basis of individual clinical judgment (Gambrill & Shlonsky, 2000; English & Pecora, 1994; Rossi, Schuerman, & Budde, 1966, as cited in Baird et al., 1999).

B. Fundamental concepts, premises, terminology, and measures have not always been well defined or articulated, are often applied in an idiosyncratic manner, are highly inconsistent among risk models, and in some cases, are simply inaccurate. This creates ambiguity, confusion, and contradiction, and it greatly increases the likelihood of error and bias in risk ratings and subsequent practice decisions.

Confusing Language

As professionals have implemented formal risk assessment models into practice, and as organizations have modified risk models to meet their perceived unique circumstances, a confusing array of new language has been developed (Wells, 1995; Pecora et al., 2000). Idiosyncratic terminology has been coined by child welfare agencies, national child welfare organizations, national resource centers, researchers, academicians, and marketing strategists. The wide discrepancies in language increase the difficulty in understanding what is already an inherently complicated technology.

Examples of some of the terms used to represent risk are "risk elements," "risk factors," "risk influences," "risk contributors," "safety threats," "present danger," "threats of serious harm," "imminent

danger," "emerging dangers," "future danger," "immediate need for a safety intervention," "family concerns," "risk correlates," and "cluster elements." Developers have also coined language to represent the intervening factors that mitigate risk, including "family strengths," "safety factors," "protective capacities," "buffering factors," "positive factors," "compensating factors," "protective influences," and "factors offsetting risk" (English & Pecora, 1994; Schene, 1996; Holder & Morton, 1999; Wagner, Johnson, & Caskey, 1999; Gambrill & Shlonsky, 2000; Holder & Lund, 1995; Pecora, English, & Hodges, 1995). Diverse terms are often used in an interchangeable or idiosyncratic manner, sometimes within a single model or document (Holder & Morton, 1999). Further increasing the confusion, the term "safety factors" is frequently used to represent factors that *compromise* safety rather than factors that *promote* it.

In addition, the language used to describe risk assessment concepts and models is often unclear and confusing. Some models attempt to differentiate, for example, between risk "factors" or "influences" and risk "elements," suggesting that one is a subset or more discrete delineation of the other (Ohio Department of Job and Family Services, 1995; New York State Department of Social Services, 1994). One source contends that the "risk field" is made up of "forces," each force being a "complex assemblage of characteristics, factors, qualities, and aspects known as elements..." (The Child At Risk Field (CARF), cited by Cicchinelli & Keller, 1990, p. 15). Another contends that "correlates for family concerns receive added weight because they reinforce cluster elements... because correlates interact with causal factors" (Ohio Department of Job and Family Services, 1995, p. 43). One final example suggests that "Danger is present when there is a threat or likelihood of serious harm. What constitutes a threat? A threat may be a condition, behavior, thought, feeling, or perception" (National Resource Center for Child Maltreatment, 2002, p. 2).

Possible reasons for this proliferation of idiosyncratic language include lack of understanding of the importance of standardizing both concepts and language in risk assessment models as well as attempts by change agents to adapt a model for local use or by developers to establish a market niche. Unfortunately, lack of clarity in language creates unnecessary confusion, interferes with our ability to communicate fundamental concepts and principles, and compromises our ability to do comparative research between risk models.

Criteria and Measures

There are equally challenging problems related to the criteria or measures used to assess risk and to quantify it at various levels. For example, little standardization of assessment criteria can be found among currently used risk assessment models (Lyons et al., 1996; Cicchinelli & Keller, 1990). One comparative study found that no factors were common to all the risk models examined, and about 40% of the criteria were unique to a single model (Lyons et al., 1996). Risk assessment models also have wide variations in their numbers of criteria, ranging from a low of about six to a high of about fifty (Lyons, Doueck, & Wodarski, 1996; Cicchinelli & Keller, 1990). In some models, the primary criteria are further divided into more discrete subcategories, thereby creating dozens of individual measures. Many models fail to differentiate among risk factors for physical abuse, neglect, and sexual abuse, even though contributors and dynamics are often different for these types of maltreatment (Gambrill & Shlonsky, 2000; Schene, 1996; English & Pecora, 1994). Most measures have not been empirically tested or their reliability and validity are not supported by research (Pecora et al.,

cont'd on page 4

2000; Johnson, 1996; Doueck et al., 1993; Cicchinelli, 1995; McDonald & Marks, 1991). One study of eight risk assessment models determined that fewer than half of the 88 measures in these models had been empirically tested, much less validated, before being implemented into practice (McDonald & Marks, 1991).

The measures in many risk assessment models are also constructed in a manner that creates confusion, thereby undermining the instrument's reliability (Pecora et al., 2000; Wells, 1995; Cicchinelli, 1995; Wald & Woolverton, 1994; Selltiz, Wrightsman, & Cook, 1976). Measures are often poorly defined, nebulous and ambiguous, overly global, illogical, and very subjective. Some are quite simply inaccurate. These measures often leave considerable room for interpretation by different raters, and at times, the descriptors that delineate the various degrees of risk are virtually indistinguishable, making it possible to score the very same behaviors at more than one risk level.

One common practice is to develop measures that distinguish among levels of risk by creating a continuum, in which some variation of "none" or "a little" anchors one end and "a lot" anchors the other. Examples of such quantitative rating continuums include the following:

- minor, moderate, serious, severe, extreme
- safe, fairly safe, unsafe, very unsafe, extremely unsafe
- marginally, moderately, very, extremely
- isolated, sporadic, repeated
- has a history of, occasionally, sometimes, often

In an example of this practice, a set of measures to rate physical hazards in the home described a "minor gas leak" as a moderate risk and a "severe gas leak" as a high risk (Ohio Department of Job and Family Services, 1997). How would a rater determine how much gas escaping into a particular room, over what period of time, would warrant recategorizing a leak from *minor* to *severe*? These measures also ignore the fact that in typical circumstances, gas leaks have the potential to kill and should be considered, *de facto*, high risk.

These measures presume that changes in the amount, extent, or frequency of a behavior or condition represent gradations of harm that are meaningful in determining the existence or potential of maltreatment. In other words, a little exposure is less risky than more exposure, and both are less risky than a lot of exposure. Although this may be true in some instances, very often it is not. These measures also fail to designate exactly how much is *a lot* or *a little* and, instead, leave this differentiation to the discretion of the caseworker. Further, the words used to describe the behaviors and conditions at the various risk levels are often not easily differentiated; in fact, some descriptors, such as *severe*, *serious*, and *extreme*, are essentially synonymous. It would be similarly difficult to determine exactly how many incidents would have to occur before *isolated* behavior became *sporadic*.

At times, measures are not supported by empirical data. In one set of measures, a child who is propositioned or pressured to have sex, but in which no sex occurs, is rated at moderate risk. Only if the perpetrator has physically involved the child in a sexual act or exploitation does the risk become high (Ohio Department of Job and Family Services, 2000; Washington State Department of Social and Health Services, 2001). According to empirical data, both conditions may represent a high risk of recurrence of sexual abuse. Grooming behaviors by perpetrators typically include a sequence of esca-

lating and more intrusive sexual involvement over time, any of which represent a significant threat of continuing and future harm as long as the perpetrator has unhindered access to the child victim (Salter, 1988, 1995).

The previous examples may appear extreme, but, in fact, reflect alarmingly common problems in risk assessment measures. They illustrate a critical point. Reliable and valid measures are the cornerstone of any effective risk assessment instrument. Well-constructed measures promote consistency and accuracy in ratings, whereas ambiguous and poorly defined measures promote individualistic, inconsistent, and potentially biased interpretations.

Confusion Among Risk, Safety, and Substantiation

The recent national emphasis on child safety has spawned the development of a variety of new safety assessment instruments. However, child safety is not a new concept in child welfare, nor are safety assessments a recent invention (DePanfilis & Scannapieco, 1994). Child safety has always been, and remains, the mission and defining principle of the child welfare profession, and child welfare professionals have been assessing children's safety as long as there has been a child welfare profession.

The stated goal of recently developed safety assessments is the accurate and timely identification of children who are "unsafe" (i.e., currently being maltreated, have very recently been maltreated, or are in circumstances where they are likely to be maltreated in the immediate future) (Wagner et al., 1999). Attention to safety issues allows agencies to develop very short-term plans, referred to as "safety plans," to stabilize family situations or to make alternative placement arrangements so children can be protected until a more in-depth family assessment and service plan can be completed. Toward this end, the data collected in safety assessments tend to cluster around three fundamental questions.

- 1) *Has the child been recently maltreated, is the child currently being maltreated, or is the child at risk of imminent harm?*
Safety assessments are intended to accurately identify children who have recently been or are currently being maltreated, or are at risk of imminent harm, and to determine the nature and type of harm, its severity, and its potential consequences for the child.
- 2) *What additional family and environmental factors may increase the likelihood of harm in the near term?*
Safety assessments attempt to identify family and environmental factors that could potentially escalate, resulting in imminent, continuing, or increasing harm to children.
- 3) *Are there strengths and protective factors in the family that can mitigate maltreatment and assure the child's safety?*
Safety assessments were originally developed to prevent unnecessary out-of-home placements by identifying family and community resources that could stabilize volatile situations and protect children in their own families (DePanfilis & Scannapieco, 1994). The objective was to prevent emergency removal and foster care placement, which, themselves, can subject children and other family members to serious emotional trauma.

Though the objectives of safety assessment are fairly clear, considerable confusion remains about the relationship among safety assessment, risk assessment, investigation, and the substantiation of maltreatment.

Abusis Inibi

Safety assessments are, in fact, a form of risk assessment. However, they are concerned only with risk of severe harm in the near term or, as the Latin appellation indicates, “abuse near at hand,” rather than the likelihood of harm at some time in a more protracted future. Special emphasis on this subclass of risk assessment is not only justifiable, but a necessary correlate of risk assessment, because the two most important variables in defining risk—the likelihood of harm and the potential severity of such harm—are both very high when children are *unsafe*.

Safety assessments reflect the *a priori* assumption that we are most concerned with severe maltreatment that is *inibi*—that has just happened, is happening, or is imminent. However, although it is justifiable, even necessary, that we carve off this class of potentially severe and imminent risk for special and urgent consideration, to suggest that safety assessment is qualitatively different from risk assessment will only cause additional confusion and discontinuity.

Confusion Between Safety Assessment and Investigation/Substantiation

Upon close scrutiny, the objectives and activities of safety assessment appear to be equivalent to those of child protective services investigations, the substantiation of maltreatment, and the assessment of imminent risk—albeit repackaged and renamed. In this context, substantiation refers to the formal process of determining whether an alleged incident of child maltreatment occurred, as well as the nature, severity, and circumstances of such maltreatment. Safety assessment, stripped of ideology and rhetoric, essentially combines substantiation of maltreatment and emergency case planning. Items on safety assessments routinely probe for information about existing unsafe environmental conditions, negligent or abusive parenting practices, and conditions that currently compromise a child’s health or well-being. The specific measures in safety assessments typically include the physical, emotional, and behavioral indicators of various types of maltreatment; descriptions of potentially harmful familial, environmental, and social conditions; and the extent and type of harm a child has already experienced (Wagner et al., 1999; Salovitz, 1993; Ontario Association of Children’s Aid Societies, 2000; Ohio Department of Job and Family Services, 2003; New Brunswick Department of Health and Community Services, 1999).

Rhetoric contends that safety assessment is categorically different from the investigation and substantiation of maltreatment, but this sends a contradictory message. Workers are admonished not to *investigate* or to *substantiate*, as these are viewed as unfriendly and disempowering to families. Yet, the preponderance of items on safety assessments were designed to identify and document prior, current, and continuing abuse or neglect.

The current focus on safety assessment is a legitimate reemphasis of the importance of child welfare’s fundamental responsibility—child safety—and, it deserves the emphasis it has received from recent federal policy and action. However, it is problematic to suggest that the substantiation of abuse and neglect is unrelated to risk and safety assessment, safety planning, case planning, and documentation of outcomes. Moreover, in no way does the substantiation of maltreatment preclude a developmental, empowering, and family-centered approach to practice (Rycus & Hughes, 1998). Child welfare workers can strengthen and preserve families, and they can help them prevent future maltreatment by assuring that information about

precursor conditions to prior maltreatment and associated risk factors drive case plans and service interventions that enable families to grow and change.

C. There are serious methodological problems in the design and development of many risk assessment technologies and models and, also, in much of the research designed to evaluate and validate them. This not only affects the reliability and validity of the models, but also results in the communication of inaccurate information about their methodological soundness to the practice field.

Reliability and Validity

Effective formal risk assessment is based on sound scientific principles and statistical methods (Ruscio, 1998; Johnson, 1996; Blenkner, 1954). Two fundamental research principles, reliability and validity, underlie any assessment of the relative effectiveness of different risk assessment models (Ruscio, 1998; Johnson, 1996; Cicchinelli, 1995).

Reliability can be broadly defined as the degree to which a particular measure yields *consistent* results. One type, known as inter-rater reliability, refers to whether different people using the same criteria will reach the same conclusions from the same information. This is most relevant in formal risk assessment, in which the goal is to standardize the collection and interpretation of case-related information by different workers in different places and at different times. High inter-rater reliability reduces error and bias.

Validity in risk assessment generally refers to the degree to which an instrument can *accurately* categorize or classify families into different levels of risk. Thus, for a child welfare risk assessment instrument to be valid, the families it has identified as *high risk* should, as a group, maltreat their children significantly more often than the group of families identified as *low risk*, and the group of families identified as *moderate risk* should fall clearly in-between.

A formal risk assessment model’s reliability and validity provide the “litmus test” of its effectiveness. The higher a model’s reliability and validity, the more likely it is to promote the *consistent* collection of *accurate* information about the condition being examined, ultimately promoting *consistent* and *accurate* conclusions regarding potential risk (Macdonald, 2001; Johnson, 1996). Conversely, risk models that lack reliability or validity formalize and sustain the collection of inconsistent and inaccurate data, which results in faulty decision making using this data (Macdonald, 2001; Gambrill & Shlonsky, 2000; Ruscio, 1998; Wald & Woolverton, 1994).

Actuarial and Consensus Models

There are two primary types of formal child welfare risk assessment models: actuarial models and consensus models (sometimes referred to as matrix models). Actuarial models are common in many professional disciplines to formally estimate outcomes, such as who is most likely to have a heart attack or to survive one. Tables used by insurance companies to establish insurance premiums are also examples of actuarial instruments. Actuarial instruments are typically used because research has repeatedly demonstrated their superiority over clinical judgment in accurately estimating the likelihood of future outcomes (Macdonald, 2001; Gambrill & Shlonsky, 2000; Baird & Wagner, 2000; Baird et al., 1999; Ruscio, 1998; Grove & Meehl, 1996; Dawes, Faust, & Meehl, 1993; Dawes, 1993).

cont'd on page 6

Actuarial models use standardized statistical procedures to identify the specific criteria, and their combined effects, that have the greatest power to discriminate among groups of people in the future occurrence of a particular outcome. Criteria are formalized into standardized assessment protocols only *after* the relationships among the variables have been quantified and thoroughly tested. Further, the ratings of individual criteria and the scoring of an overall risk level are dictated by the previously determined statistical weighting of these previously identified associations (Macdonald, 2001; Gambrill & Shlonsky, 2000; Ruscio, 1998; Johnson, 1996). The presence of these variables in families in certain specific combinations can be said to increase the likelihood (but not to guarantee) that maltreatment will reoccur (Baird & Wagner, 2000). The greater the statistical association between the combined variables in the instrument and the occurrence of subsequent maltreatment, the greater the capacity of the instrument to consistently and accurately classify families into various levels of risk.

Consensus models, by contrast, rely on a preponderance of professional agreement about which variables or conditions are most highly associated with recurrences of child maltreatment (Pecora et al., 2000; Wald & Woolverton, 1994). Although a large body of professional literature describes and documents the individual, family, and environmental conditions found to be *associated* with child maltreatment, these factors are often not tested to determine their capacity to *estimate the likelihood* of future maltreatment (Lyons et al., 1996; Cicchinelli & Keller, 1990). Further, there are usually no empirical data regarding how the various factors interact or how they should be weighted and scored (Lyons et al., 1996; English & Pecora, 1994). Thus, consensus risk assessment models do not lend themselves to the use of numerical scoring systems.

The Fallacy of Consensus

Considerable confusion exists in the child welfare field about what constitutes consensus. Many people incorrectly interpret consensus to mean the negotiated opinions of whatever group of professionals is convened to develop or revise a risk assessment model. Ad hoc committees of practitioners are asked to present and discuss their judgments and opinions and to try to reach agreement on which criteria, definitions, and rating methods *work best for them*. Referring to this process as “generating consensus,” “further refining the model,” or “addressing our unique circumstances” gives apparent validity to a process that is notoriously subject to error and bias (Macdonald, 2001; Gambrill & Shlonsky, 2000; Ruscio, 1998; Dawes, Faust, & Meehl, 1989; Meehl, 1992, as cited in Ruscio, 1998; Cicchinelli & Keller, 1990). Even as the child welfare profession espouses the necessity of evidence-based practice, critical case decisions continue to be made using risk assessment instruments based not on evidence, but on the personal opinions of a variety of informants with differing degrees of expertise.

It is troublesome that many jurisdictions adapt risk assessment models, in whole or in part, without assuring that their changes are empirically based and without testing for reliability and validity. The literature delineates a variety of potential reasons for this practice (DePanfilis, 1996; Johnson, 1996; Cicchinelli, 1995; Cicchinelli & Keller, 1990). Many users revise models to make them shorter, simpler, or more easily understandable. Some believe that staff members’ “buy-in” to a model depends on their participation in the model’s development and their agreement with the final product (Cicchinelli & Keller, 1990). Some believe their agency’s circumstances to be so unique as to warrant an individualized model. And

some equate any form of standardization as a rigid mandate that undermines individuality and creativity. Even though the majority of idiosyncratic revisions in risk models are presumably well intentioned, these changes often further undermine a model’s reliability and validity.

Comparative Research of Risk Assessment Models

Extensive research has been conducted in a wide variety of practice fields, including child welfare, on both actuarial and consensus-based decision-making models. These studies have repeatedly demonstrated the superior reliability, validity, and performance of actuarial models over consensus-based models in estimating the likelihood of future outcomes (Macdonald, 2001; Gambrill & Shlonsky, 2000; Baird & Wagner, 2000; Baird et al., 1999; Ruscio, 1998; Falco & Salovitz, 1997, as cited in Gambrill & Shlonsky, 2000; Grove & Meehl, 1996; Dawes et al., 1993; Dawes, 1993).

The preponderance of research literature continues to raise serious questions about the reliability and validity of most of the risk assessment models and instruments currently used by child welfare agencies (Macdonald, 2001; Pecora et al., 2000; Gambrill & Shlonsky, 2000; Baird et al., 1999; Lyons et al., 1996; Schene, 1996; Camasso & Jagannathan, 1995; English & Pecora, 1994; McDonald & Marks, 1991; Wald & Woolverton, 1994; Cicchinelli & Keller, 1990). In practice, many child welfare professionals are making decisions about children and families with little more accuracy than flipping a coin, yet believing they are using technologies that reduce subjectivity and bias and that increase the quality of their decisions.

There is also skepticism among researchers about the soundness of much of the research conducted to test the reliability and validity of risk assessment models (Pecora et al., 2000; Camasso & Jagannathan, 2000; Gambrill & Shlonsky, 2000; Baird et al., 1999; Schene, 1996; Lyons et al., 1996; Curran, 1995). Evidence-based practice stresses that research should be competently constructed and executed, reported findings should be supported by the data, and research methods should be accurately described. It also calls for full disclosure of methodological problems or other constraints that potentially skew the results or limit the generalizability of findings and conclusions (Gambrill, 2000; Gambrill & Shlonsky, 2000; Lyons et al., 1996). Risk assessment research often does not adhere to these guidelines (Gambrill & Shlonsky, 2000). Unfortunately, child welfare practitioners may believe a study’s claims of reliability or validity whether or not this conclusion is warranted.

D. A variety of systemic, bureaucratic, and individual barriers impede the large-scale implementation of formal risk assessment technologies by child welfare agencies.

As is true with any large-scale change initiative, systemwide implementation of formal risk assessment requires a significant allocation of time, work, and resources; and its success depends upon strong and continuing organizational commitment and support in the face of many deterrents and barriers.

The literature has identified multiple problems related to implementation of formal risk assessment at the local agency level (DePanfilis, 1996; Curran, 1995; Cicchinelli, 1995; Cicchinelli & Keller, 1990). In some organizations, workers vary greatly in their use and interpretation of risk assessment models, even though the models ostensibly standardize decision making (Gambrill &

Shlonsky, 2000; Cicchinelli, 1995). Although many workers do use standardized risk protocols to help guide their decisions, many others use risk rating instruments simply to record conclusions and decisions they have already made by other means, including personal clinical judgment (Gambrill & Shlonsky, 2000; Schene, 1996; English & Pecora, 1994; Fluke, 1993, as cited in Lyons, Doueck, & Wodarski, 1996, and in Johnson, 1996; Cicchinelli & Keller, 1990). Many caseworkers consider formal risk assessment a burdensome, bureaucratic, and unwarranted increase in an already heavy work load. Many staff perceive it to be an administrative mandate, rather than a necessity to promote unbiased, accurate, and relevant decision making. Some workers view formal risk assessment as an unwarranted intrusion into families, and they may abandon or short-cut the assessment when they encounter resistance from family members.

Another prevalent barrier to implementation is lack of training in the prerequisite clinical competencies for effective assessment (Pecora et al., 2000; Schene, 1996, 6; Curran, 1995; Cicchinelli, 1995; Doueck et al., 1993, p. 442; Cicchinelli & Keller, 1990; Wald & Woolverton, 1990). Assessment in human services is a very complicated activity. It requires high levels of skill in critical thinking, observation and listening, interviewing, information gathering, and data analysis and synthesis. Caseworkers must also master the specialized knowledge needed to recognize and assess certain conditions. For example, caseworkers who are unable to recognize indicators of substance abuse, or who don't understand its behavioral dynamics, cannot accurately assess its presence or extent in families. Assessments can also be rendered inaccurate by lack of cultural knowledge, or by workers' inability to recognize how their personal culture, values, and beliefs can obscure their interpretations and conclusions about families. Caseworkers without thorough training can produce assessments with frighteningly inaccurate conclusions, even when they appear to be asking the right questions and properly recording the information. Finally, many workers are better at collecting information than at synthesizing it, using it effectively to inform casework decisions, or documenting it accurately to enhance both planning and accountability (Schene, 1996; Fluke, 1993; Cicchinelli & Keller, 1990).

Yet, in spite of the inherent complexity of assessment in human services, risk assessment training often consists primarily of policy briefings, a description of the risk model, an explanation of its criteria and measures, and instruction in how to complete the protocol and record the data. Many staff do not receive sufficient training in fundamental and, significantly more important, core assessment skills. Much risk assessment training has been likened to teaching airline pilots how to complete a preflight checklist before taking off, without ever having taught them navigation, meteorology, or even the essentials of flying a plane. Yet, many jurisdictions continue to expect two or three days of training on a risk assessment model to fully prepare staff to implement it consistently and accurately.

Moreover, risk assessment models are often superimposed on preexisting case management systems without thoughtful consideration of their "fit." This contributes to repetition, duplication, and even contradiction in procedures and forms for the collection, recording, and management of case-related information (Cicchinelli, 1995; Doueck et al., 1993; Cicchinelli & Keller, 1990). Finally, the common organizational dynamics and barriers that often undermine other change initiatives also impede the implementation of risk as-

essment. As a result, complete and successful implementation of formal risk assessment at the local level continues to be elusive (English & Pecora, 1994).

E. It is often expected that formal risk assessment activities should serve a variety of administrative, political, and systemic functions in child welfare organizations that have little to do with making accurate protective decisions for children.

The child welfare literature describes a variety of ways formal risk assessment is expected to improve child welfare practice. Among these are improving workers' decision making at all stages of casework; improving the quality and consistency of services to families; improving the case referral and case management process; providing a forum for case discussion and supervision; delineating child welfare practice standards; increasing agency accountability; demonstrating agency accountability to the public; reducing agency liability; improving court presentations; compensating for inexperienced staff and the effects of turnover; helping manage workloads; and providing a framework for case documentation (Schene, 1996; DePanfilis, 1996; Wells, 1995; Cicchinelli, 1995; Doueck et al., 1993; Cicchinelli & Keller, 1990).

Formal risk assessment is a single technology with the limited purpose of estimating, with acceptable accuracy, which children in our communities are most likely to be maltreated. Maintaining unrealistic expectations for formal risk assessment can actually deter policy makers, administrators, legislators, and potential funders from seeking and developing more appropriate strategies to address the many organizational, community, and direct practice problems that plague contemporary child welfare, thus ultimately increasing, rather than decreasing, the potential of future harm for high-risk children (Cicchinelli, 1995; Wald & Woolverton, 1994).

F. A number of ethical and legal issues related to risk assessment have not been fully addressed.

There are currently no risk assessment technologies that can predict with certainty that child maltreatment will reoccur, even in families identified to be at high risk (Gambrill & Shlonsky, 2000; Lyons et al., 1996; Dawes et al., 1989). Some formal risk assessments can accurately categorize families into high-risk, moderate-risk, and low-risk groups, on the basis of the statistical likelihood of a reoccurrence of maltreatment at some time in the future. This is the best that current research and technology have to offer. Given these realities, it is difficult to see how one could justify opening a nonvoluntary protective services case based entirely upon risk assessment findings in the absence of substantiated abuse or neglect.

Even so, child welfare professionals in some states and agencies have considered shifting the focus of child protection from the investigation and substantiation of a past incident of child maltreatment to risk assessment, which is "future-oriented" and not "aimed at proof or disproof of specific allegations of past maltreatment" (Schene, 1996; Doueck et al., 1993). Common arguments to support this action include: substantiation isn't family friendly; it focuses attention on pathology rather than strengths; it dwells on a family's past behaviors rather than growth and change; in many families, maltreatment never reoccurs; substantiation is too subjective a concept to be meaningful; and substantiation sets up a confrontational, rather than collaborative, relationship between families and the agency.

cont'd on page 8

However, a clear and well-documented indication of child maltreatment may be the only legal and ethical justification for intrusion by child protective services into the private dynamics of family life. Without such documentation, intrusion into families may violate parental rights legislation and federal civil rights law.

It must also be understood that even reliable and valid risk assessment technologies serve a limited purpose in the broader context of child protective services, and ethical and potential legal liabilities may result if these limitations are not acknowledged. Claiming, either by design or ignorance, that formal risk assessment will achieve what it cannot, creates potential liabilities that the child protection system can ill afford. The stakes increase greatly when risk assessment protocols used by agencies are neither reliable nor valid. Agencies place themselves in a precarious legal position by claiming that their decision making is based on standardized, validated risk assessment protocols when it is not (Curran, 1995). A state or provincial child protective service system that endorses or mandates a formal risk assessment model that it knows, or should know, is potentially harmful to children and families is at risk of significant legal liability. If children are harmed as a result of faulty decisions based on these models, agencies may be subject to legal remedies.

Summary

Despite all good intentions and hard work, formal risk assessment may not have significantly improved services to children and families and, in some cases, may actually have had a harmful impact. We must collectively reevaluate our options, identify and capitalize on our strengths, and implement strategic measures that will promote the most ethical and effective use of risk assessment technologies to assure equitable and legitimate protective decisions for abused and neglected children and their families.

Part II: Recommendations

A. There is lack of agreement regarding the proper scope and purpose of risk assessment technology in child welfare assessment and case planning.

- A1) Formal risk assessment should be considered one tool in a broader, structured process of safety assessment and safety planning, family assessment, case planning, decision making, and ongoing risk analysis throughout the life of a case.
- A2) Formal risk assessment should be used by intake assessment caseworkers to guide decisions about whether children and their families should receive ongoing protective services from the agency; whether they should be diverted to other community service providers; or whether they should be closed at the intake level.
- A3) Agencies should not attempt to use "hybridized" instruments as both a formal risk assessment and a family assessment. Formal risk assessment requires measures that can accurately estimate the likelihood of future occurrences of child maltreatment. Family assessment requires measures that guide the collection of data to identify family needs, strengths, and dynamics. These goals, criteria, methodologies, and uses of data are sufficiently different to warrant two different instruments and processes.
- A4) Because of frequent differences in many of the family dynamics associated with physical abuse, sexual abuse, and neglect, both risk assessment and family assessment protocols should incorporate and assess those criteria that are most relevant for each type of child maltreatment.

B. Fundamental concepts, premises, terminology, and measures have not always been well defined or articulated, are often applied in an idiosyncratic manner, are highly inconsistent among risk models, and in some cases, are simply inaccurate. This creates ambiguity, confusion, and contradiction, which greatly increases the likelihood of error and bias in risk ratings and subsequent practice decisions.

- B1) The child welfare field should establish standardized and consistent terminology to represent all components and facets of the formal risk assessment process. All models should utilize the same terms for the same concepts and elements, including risk factors, protective factors, criteria, and measures.
- B2) The identification and substantiation of recent or current maltreatment, and the assessment of risk of imminent maltreatment, should be clearly stated objectives for all safety assessments.
- B3) Safety assessment should not replace formal risk assessment. Both are essential components of a structured continuum of decision making, but their purposes are different and the data are used toward different ends. Safety assessment evaluates both abuse in inibi (i.e., the presence of recent or current maltreatment and the potential for imminent maltreatment) and factors in the family and community that can help mitigate maltreatment. With these data, children at risk of imminent harm can often be protected within their own families and communities, thereby minimizing family disruption and placement trauma. Formal risk assessment should follow safety assessment to discern the likelihood of a recurrence of maltreatment. These data help agencies determine which families should receive ongoing protective services from the agency and at what level of intensity this should occur.
- B4) Safety plans should be developed for all children found to be recently or currently maltreated, or those in volatile and unstable situations in which they are at imminent risk of severe harm. Safety plans should focus only on assuring children's protection in the immediate term. Safety plans should not substitute for formal case plans. Case plans should be developed only after completion of a comprehensive, individualized family assessment that provides relevant information to guide the selection and provision of ongoing services.

C. There are serious methodological problems in the design and development of many risk assessment technologies and models, and also in much of the research designed to evaluate and validate them. This not only affects the reliability and validity of the models, but also results in the communication of inaccurate information about their methodological soundness to the practice field.

- C1) All formal risk assessment protocols should be empirically derived—developed on the basis of findings and conclusions of well-designed and -implemented research. All criteria and measures should be pre-tested and determined to have the requisite levels of reliability and validity prior to being used in any risk assessment protocol. The structure for data analysis, scoring, and ranking should also be based on scientific and statistical procedures that promote the highest possible levels of reliability and validity. Criteria and measures in risk assessment instruments must be clearly defined and measurable and must leave as little room as possible for bias and misinterpretation.

C2) Considering the current state of formal risk assessment technology, child welfare agencies should use reliable and valid actuarial risk assessment models for formal risk assessment in all child protective service cases.

C3) Consensus decision-making models based on credible empirical data, and that include relevant and clearly-articulated measures, may be appropriate tools to guide the ongoing clinical assessment of safety and risk, family assessment, and service planning. However, consensus-based models should not be used to estimate the likelihood of future occurrences of maltreatment in place of actuarial decision-making technologies, which have higher reliability and validity.

C4) Formal risk assessment models and instruments should be developed or modified only in collaboration with professionals who have specialized expertise in the construction, evaluation, and validation of such instruments. This responsibility should not be delegated to ad hoc committees of practitioners and administrators without such support.

D. A variety of systemic, bureaucratic, and individual barriers impede the large-scale implementation of formal risk assessment technologies by child welfare agencies.

D1) Systemwide implementation of formal risk assessment should be viewed as large-scale system change and should be guided by fundamental principles of change management. Agencies must make the commitment to support and sustain the use of risk assessment technologies over time.

D2) Because of the inherent complexity of assessment in human services, and the high level of skill needed to gather and interpret assessment information, safety and risk assessments are best performed by highly skilled caseworkers with specialized training and prior child welfare experience. Although these functions are typically, and appropriately, performed by intake assessment caseworkers, many agencies assign newly hired caseworkers who have little training or practice experience to work in intake units. Lack of worker skill in interviewing and assessment will undermine even the most reliable and valid of protocols. It would be helpful if job classifications and salary levels for assessment caseworkers were upgraded to reflect these higher prerequisite qualifications.

D3) Comprehensive training in prerequisite core-level assessment and interviewing competencies should always precede training in the use of specific risk assessment models or protocols. Training should also be provided for supervisors who are assigned responsibility to monitor their staff's assessment activities. Coaching and educational supervision need to be supported by all local agencies to promote the transfer of learning and skill mastery.

D4) Risk assessment models and forms should not be "hard coded" into computerized child welfare information systems. Information systems must be sufficiently flexible to accommodate rapid changes in risk assessment criteria, measures, and scoring.

E. It is often expected that formal risk assessment activities should serve a variety of administrative, political, and systemic functions in child welfare organizations that have little to do with making accurate protective decisions for children.

E1) Agencies should not use formal risk assessment instruments for purposes other than that for which they were developed—to estimate the likelihood of a future recurrence of child maltreatment in families. More

appropriate technologies must be developed to address other organizational and systemic needs, including case planning, public relations, quality assurance, communication, supervision, workload management, and monitoring and recording. Risk assessment cannot substitute for formal systems of data collection and recording to assure accountability.

F. A number of ethical and legal issues related to risk assessment have not been fully addressed.

F1) The child welfare profession must acknowledge and address the potential legal and ethical liabilities of continuing to use untested or unproven formal risk assessment models.

F2) Considering the limitations of even the most well-developed, reliable, and valid risk assessment technologies, agencies should not rely on risk assessment as the sole, or even the primary, resource to justify their casework and child placement decisions. Investigation with confirmed findings of abuse and neglect must remain the primary justification for opening nonvoluntary cases for child protective services.

* The unabridged version of the NARCCW policy white paper, "Issues in Risk Assessment in Child Protective Services," may be downloaded in PDF format from the NARCCW web site (www.narccw.com). Printed, bound copies may be requested via e-mail to JRycus@ihs-trainet.com, or by calling (614) 251-6000.

About the Authors

Judith S. Rycus has a doctorate in psychology from The Ohio State University and a master's degree in social work from Wayne State University. Dr. Rycus is Program Director of the Institute for Human Services, Columbus, Ohio, and is Senior Policy Analyst for the Center for Child Welfare Policy of the North American Resource Center for Child Welfare.

Ronald C. Hughes has a doctorate in psychology from The Ohio State University and a Master of Science in Social Administration from Case Western Reserve University. He is the Director of the North American Resource Center for Child Welfare's Center for Child Welfare Policy, and the Institute for Human Services.

Drs. Rycus and Hughes have coauthored numerous publications, including the four-volume *Field Guide to Child Welfare* (Child Welfare League of America, 1998). They have provided extensive consultation and technical assistance to child welfare systems throughout North America in implementing large-scale, competency-based inservice training programs for child welfare staff.

cont'd on page 10

BIBLIOGRAPHY

Bibliography**

- Baird, C., Ereth, J., & Wagner, D. (1999). *Research-based risk assessment: Adding equity to CPS decision making*. Madison, WI: Children's Research Center.
- Baird, C., & Wagner, D. (2000). The relative validity of actuarial and consensus-based risk assessment systems. *Children and Youth Services Review*, 22(11/12), 839-871.
- Baird, C., Wagner, D., Healy, T., & Johnson, K. (1999, November-December). Risk assessment in child protective services: Consensus and actuarial model reliability. *Child Welfare*, LXXVIII(6), 723-749. Blenkner, M. (1954). Predictive factors in the initial interview in family casework. *Social Science Review*, 28, 65-73.
- Camasso, M., & Jagannathan, R. (1995, September). Prediction accuracy of the Washington and Illinois risk assessment instruments: An application of receiver operating characteristic curve analysis. *Social Work Research*, 19(3), 174-183.
- Camasso, M., & Jagannathan, R. (2000). Modeling the reliability and predictive validity of risk assessment in child protective services. *Children and Youth Services Review*, 22(11/12), 873-896.
- Cash, S. J. (2001). Risk assessment in child welfare: The art and science. *Children and Youth Services Review*, 23(11), 811-830.
- Cicchinelli, L. F., & Keller, R. A. (1990, June). *A comparative analysis of risk assessment models and systems: Final report*. Grant No. 90-CA-1302. Lakewood, CO: Applied Research Associates.
- Cicchinelli, L. F. (1995, Winter). Risk assessment: Expectations and realities. *The APSAC Advisor, Special Issue on Risk Assessment*, 8(4), 3-8.
- Curran, T. F. (1995, Winter). Legal issues in the use of CPS risk assessment instruments. *The APSAC Advisor, Special Issue on Risk Assessment*, 8(4), 15-20.
- Dawes, R. (1993, June 9). Finding guidelines for tough decisions. *Chronicle of Higher Education*, p. A40.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1993). Statistical prediction versus clinical prediction: Improving what works. In G. Keren & C. Lewis (Eds.), *Handbook for data analysis in the behavioral sciences: Methodological issues*. Hillsdale, NJ: Erlbaum.
- DePanfilis, D. (1996). Implementing child maltreatment risk assessment systems: Lessons from theory. *Administration in Social Work*, 20(2), 41-59.
- DePanfilis, D., & Scannapieco, M. (1994, May-June). Assessing the safety of children at risk of maltreatment: Decision-making models. *Child Welfare*, LXXIII(3), 229-245.
- Doueck, H. J., English, D. J., DePanfilis, D., & Moote, G. T. (1993, September-October). Decision-making in child protective services: A comparison of selected risk-assessment systems. *Child Welfare*, 72(5), 441-453.
- English, D., & Pecora, P. (1994, September-October). Risk assessment as a practice method in child protective services. *Child Welfare*, 73(5), 451-473.
- Falco, G., & Salovitz, B. (1997). *Clinical versus actuarial risk assessment in child protective services: Results from recent research in New York*. Paper presented at the 11th Annual CPS Roundtable on Risk Assessment, San Francisco, CA.
- Fluke, John. (1993). Evaluation of the Pennsylvania approach to risk assessment. In T. Tatar (Ed.), *Seventh national roundtable on CPS risk assessment: Summary of highlights*. Washington, DC: American Public Welfare Association.
- Gambrill, E. (2000, Fall). Honest brokering of knowledge and ignorance. *Journal of Social Work Education*, 36(3), 387-397.
- Gambrill, E., & Shlonsky, A. (2000). Risk assessment in context. *Children and Youth Services Review*, 22(11/12), 813-837.
- Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law*, 2(2), 293-323.
- Holder, W., & Lund, T. (1995, Winter). Translating risks to positive outcomes. *APSAC Advisor, Special Issue on Risk Assessment*, 8(4), 20-24.
- Holder, W., & Morton, T. D. (1999). *Designing a comprehensive approach to child safety*. Duluth, GA: Child Welfare Institute, and Aurora, CO: ACTION for Child Protection.
- Johnson, W. (1996). Risk assessment research: Progress and future directions. *Protecting Children*, 12(2), 14-19.
- Lyons, P., Doueck, H. J., & Wodarski, J. S. (1996). Risk assessment for child protective services: A review of the empirical literature on instrument performance. *Social Work Research*, 20(3), 143-155.
- Macdonald, G. (2001). *Risk assessment and decision making. In Effective interventions for child abuse and neglect: An evidence-based approach to planning and evaluating interventions* (pp. 269-284). Chichester, UK: Wiley.
- McDonald, T. P., & Marks, J. (1991). A review of risk factors assessed in child protective services. *Social Services Review*, 65, 112-132.
- National Resource Center on Child Maltreatment. (2002). *Briefing paper for the Ohio safety and risk assessment workgroups*. Duluth, GA: Author.
- New Brunswick Department of Health and Community Services. (1999). *Risk management system*. Fredericton, NB: Author.
- New York State Department of Social Services. (1992, 1994). *New York State risk assessment and services planning model*. [Publication #1160]. New York: Author.
- Ohio Department of Job and Family Services, Office for Children and Families. (1995, 1997, 2001). *Family risk assessment model (FRAM)*. Columbus, Ohio: Author.
- Ontario Association of Children's Aid Societies. (2000). *Risk assessment model for child protection in Ontario*. Toronto, Ontario: Author.
- Pecora, P. J., English, D. J., & Hodges, V. G. (1995, Winter). Selected multicultural guidelines for child maltreatment risk assessment. *APSAC Advisor, Special Issue on Risk Assessment*, 8(4), 24-27.
- Pecora, P. J., Whittaker, J. K., Maluccio, A. N., & Barth, R. P. (2000). *The child welfare challenge: Policy, practice, and research (2nd ed.)*. New York: Aldine De Gruyter.
- Rossi, P., Schuerman, J., & Budde, S. (1996). *Understanding child maltreatment decisions and those who make them*. Chicago: University of Chicago, Chapin Hall Center for Children.
- Ruscio, J. (1998, May). Information integration in child welfare cases: An introduction to statistical decision making. *Child Maltreatment*, 3(2), 143-156.
- Rycus, J. S., & Hughes, R. C. (1998). *Field guide to child welfare, I-IV*. Washington, DC: Child Welfare League of America.
- Salovitz, B. (1993). *New York State risk assessment and service planning model: A review of the development process*. Sixth National Roundtable on CPS Risk Assessment. APWA, June 1993. Cited in Wagner, Johnson, & Caskey (1999).
- Salter, Anna C. (1988). *Treating child sex offenders and victims: A practical guide*. Newbury Park, CA: Sage.
- Salter, Anna C. (1995). *Transforming trauma: A guide to understanding and treating adult survivors of child sexual abuse*. Thousand Oaks, CA: Sage.
- Schene, P. (1996). The risk assessment roundtables: A ten-year perspective. *Protecting Children*, 12(2), 4-8.
- Selltiz, C., Wrightsman, L. S., & Cook, S. W. (1976). *Research methods in social relations (3rd ed.)*. New York: Holt, Rinehart & Winston.
- Wagner, D., Johnson, K., & Caskey, R. (1999, June). *Family Independence Agency of Michigan; Safety assessment validation report*. Madison, WI: Children's Research Center.
- Wald, M. S., & Woolverton, M. (1990, November-December). Risk assessment: The emperor's new clothes? *Child Welfare*, 69(6), 483-511.
- Washington State Department of Social and Health Services. (2001). *The practice guide to risk assessment: Kids come first, and safety assessment*. Seattle: Author.
- Wells, S. (1995, Winter). Introduction. *APSAC Advisor, Special Issue on Risk Assessment*, 8(4), 1.

** This list includes only sources cited in the above document. A complete listing of sources used in the NARCCW risk assessment initiative is included in the unabridged version of the policy white paper.

