

**Risk and Safety Assessment in Child Welfare:
Instrument Comparisons**

Amy D'Andrade, Ph.D., BASSC Research Director

Amy Benton, M.S.W., BASSC Research Assistant

Michael J. Austin, Ph.D., BASSC Staff Director

July, 2005

Supported by the Bay Area Social Services Consortium and
the Zellerbach Family Foundation

The Center for Social Services Research (CSSR) in the School of Social Welfare at the University of California at Berkeley conducts research, policy analysis, program planning, and evaluation toward the improvement of the publicly supported social services. The focus of the Center is on populations who are considered needy or disadvantaged, including victims of child abuse and neglect, the chronically mentally ill, the aged, the medically indigent, and the poor.

Housed at CSSR, the Research Response Team of the Bay Area Social Services Consortium (BASSC) was organized in 1995 to respond rapidly to the emerging needs of county social service agencies for information for their changing environments. Structured reviews of the research literature are undertaken in close collaboration with agency administrators and program staff. BASSC was founded in 1987 and is composed of the Directors of Bay Area county social service and human service agencies, deans of the Bay Area graduate social work departments, and foundation representatives.

The BASSC Research Response Team would like to thank the members of the BASSC Risk and Safety Assessment Advisory Group: Will Johnson, Alameda County; Stacy Buchanan, Contra Costa County; Heather Ravani, Marin County; Earlene McClair, Monterey County; Liz Crudo, San Francisco County; Jerry Lindner, San Mateo County; Cindie Stoops, Santa Clara County; Ellen Timberlake, Santa Cruz County; Paul Bidinger, Solano County; and Margaret Ahern, Sonoma County. The involvement of these county staff members in shaping the study, defining search terms, and providing feedback on study results and final reports was invaluable.

TABLE OF CONTENTS

Introduction	3
The Purpose of Risk Assessment in Child Welfare	3
Different Approaches to Risk Assessment	4
Important Qualities of a Risk Assessment Instrument	7
Review of Studies	10
WRAM	11
CFAFA (the “Fresno model”)	13
CARF	15
CERAP	16
CRC Actuarial Instruments for Risk Assessment	18
Risk Assessment Model of Child Protection (Ontario)	20
Utah Risk Assessment Scales	21
Risk Assessment Implementation Issues	22
Support for Implementation	22
Use Across the Life of a Case	23
Liability Issues	24
Risk and Safety Assessment in California	25
Discussion and Implications	26
Implications for Practice	29
Implications for Research	30
Summary... ..	31
References	34
Appendices	
A. BASSC Search Protocol	40
B. WRAM Instrument	41
C. CFAFA Instrument	43
D. CARF Instrument	47
E. CERAP Instrument	48
F. CRC Instrument	52

RISK AND SAFETY ASSESSMENT IN CHILD WELFARE: INSTRUMENT COMPARISONS

Introduction

Before child welfare agencies intervene with families, they are generally required to identify maltreatment or the risk of maltreatment.¹ Therefore, the assessment of risk is a critical part of child welfare agency work.² Most states in the US formalize the process of assessing risk by using some type of structured decision-making process or tool. Risk assessment instruments generally include broad categories of areas related to abuse and neglect, behavioral descriptions, procedures to determine levels of risk, and standardized forms to record this information.³

This report, commissioned by the Bay Area Social Services Consortium, reviews the research literature regarding different instruments of risk and safety assessment in child welfare. The first section describes current approaches to risk and safety assessment. The second section describes instruments for risk assessment, and summarizes the research findings regarding instrument reliability, validity, outcomes, and use with children and families of color. The third section raises the following issues related to implementation of risk assessment models in child welfare: supports for implementation, use across the life of a case, and liability concerns. It also describes California's efforts to institute a standardized process of risk and safety assessment across the state. The report concludes with implications for practice and research.

The Purpose Of Risk Assessment In Child Welfare

One goal of risk assessment is to focus limited resources on the children who are at greatest risk of maltreatment. Many have pointed out the reduced resources of social services agencies; guided risk assessment serves as a strategy for targeting scarce resources and services toward those who have the greatest need.⁴ A second purpose of a structured risk assessment process is to facilitate an accurate, less biased decision process for determining which cases should receive services.⁵

Researchers from a variety of academic fields, particularly psychology, have studied human decision-making and identified a number of common errors people tend to make in their predictions and decisions. For example, researchers have found that people tend to: 1) ignore the

probability of an event in making predictions about the likelihood of its occurrence;⁶ 2) be overconfident of their ability to predict an event;⁷ and 3) have difficulty weighing factors related to a decision.⁸ Part of the difficulty for human decision-makers may be that the availability of too much information, some of which is likely unrelated to the outcomes, increases the chance that irrelevant information will be used in the decision.⁹

Several studies consider the issue within the context of the child welfare field. In a study by Schuerman et al.,¹⁰ case vignettes were presented to practicing social workers and to expert practitioners and theoreticians in the field. In considering the same vignette, workers make different choices from one another in deciding whether or not a child should be removed from the home. While this variability was less when cases were of obviously low and very high risk, "...similar cases in the midrange of severity of family problems are treated quite differently by different experts and workers."¹¹ In a second study, researchers conducted a content analysis of 45 "public inquiries" completed in England on two types of cases: cases of child deaths by parental actions, and sexual abuse cases in which professionals' actions were suspected of being overzealous. The formal inquiries identified three general types of errors made by social workers. First, "the most striking and persistent criticism was that professionals were slow to revise their judgments...the current risk assessment of a family had a major influence on responses to new evidence."¹² Secondly, professionals were skeptical of new information when it conflicted with their initial view of a family, yet uncritical of new information when it supported their initial view. And third, evidence from some sources was more highly valued (doctor's statements regarding abuse and social worker's witnessing of injury for example) than other sources (neighbors, concerns of the public).¹³ These findings suggest that social workers are prone to the same difficulties in decision-making as has been identified in other fields, and have contributed to the notion that risk assessment processes are needed to aid social workers in decision-making.

Different Approaches To Risk Assessment

Currently, there are two major approaches to risk assessment in child welfare decision-making: a *consensus-based* model, and an *actuarial* model. Both involve a list of family or case characteristics believed to be associated with risk of maltreatment (a risk assessment

“instrument”). However, the two approaches differ in the processes used to identify factors for inclusion in the instrument, and how the instruments are utilized in practice.

Consensus-based instruments emphasize a comprehensive assessment of risk.¹⁴ Instrument items have been identified based upon various theories of child maltreatment, the research literature on maltreatment, and/or the opinions of with expert practitioners. Items on one instrument are often combined with items from another instrument, creating hybrid instruments that vary according to the needs or beliefs of the user. Sometimes factors are assessed numerically, and families categorized by their total score, while other instruments simply describe areas that are to be assessed by the worker, without necessarily providing direction in terms of that assessment; the worker considers the area identified and codes it high, moderate or low risk based upon his or her judgment. Consensus-based instruments tend to use the same instrument to predict all forms of maltreatment.¹⁵

Actuarial instruments use statistical procedures to identify and weigh factors that predict future maltreatment.¹⁶ Often the statistical analysis is done in the state or county in which the instrument will be applied. Actuarial instruments tend to use fewer factors than do consensus-based models and generally use different factors to predict the likelihood of physical abuse and of neglect. Each factor is scored and scores are summed into overall risk scores. Families are categorized into low, moderate or high risk groups, and receive correspondingly different service responses.

In the field of child welfare as well as in other professional fields, there is some debate regarding the best approach to risk assessment. A consensus-based approach to risk assessment utilizes the underlying theoretical assumption that the causes of child maltreatment are multi-dimensional and complex, and all – or as many as possible – of the related domains should be considered.¹⁷ These instruments can also help structure a worker’s process of information gathering for clinical assessments of risk,¹⁸ and provide documentation of the reasoning underlying their decision-making.¹⁹ Some argue that the more comprehensive approach of consensus-based instruments provides better information for casework decisions.²⁰

However, consensus-based models are criticized in social work research literature for the following reasons: 1) poor conceptualization (according to Rycus and Hughes, “measures are often poorly defined, nebulous and ambiguous, overly global, illogical, and very subjective”);²¹ 2) inconsistency in the type and number of variables included (about 40% of factors used in consensus-based instruments are unique to a single instrument, and none of the factors are shared by all the instruments. Some have as few as 6 elements, some have as many as 50²²); 3) use of the same variables to predict physical abuse, neglect, and sexual abuse, even though contributors and dynamics are often different for these types of maltreatment;²³ and 4) reliance upon characteristics associated with maltreatment, rather than recurrence of maltreatment,²⁴ or upon characteristics for which there is no research support.²⁵

Some researchers assert actuarial instruments are more reliable and valid than clinical judgment or consensus-based instruments since they help practitioners focus their risk assessments on a small set of case characteristics that have demonstrated a strong statistical relationship to future maltreatment.²⁶ In a meta-analysis of over 100 studies of various health and behaviorally oriented predictions that compared actuarial methods to clinical judgment, about 95% of the studies found actuarial processes to be equal to or superior to clinical judgment.²⁷ According to proponents, even a small set of factors generally does a better job predicting outcomes than does simply the use of clinical judgment.²⁸

Others feel actuarial instruments do not facilitate the clinical judgment of skilled practitioners. Additionally, since the basis for including a factor on an actuarial instrument is its statistical association with a poor outcome – generally recurrence of maltreatment – factors may not appear to be causally related to the outcome. While some assert this is entirely appropriate since the intent is to predict rather than explain, others claim that this lack of a logical, there are concerns regarding whether actuarial instruments provide enough information for front end workers to effectively implement new practices like differential response, an intervention which requires workers to refer families at risk of abuse or neglect to appropriate community services.

It is not clear which approach is more commonly used. There is no national database regarding risk assessment approaches used by states.²⁹ As of 1996, Lyons et al., report that 15 states used

the Illinois CANTS 17B instrument or some derivation of it, 4 used the CARF system, and 4 used WARM or some derivation, all consensus-based instruments. However, an increasing number of states are using an actuarial instrument as part of a “Structured Decision Making (SDM)” case management system developed by the Children’s Research Center.

Important Qualities of a Risk Assessment Instrument

Two psychometric qualities are of particular importance when considering the value of any risk assessment approach: predictive validity, which refers to an instrument’s accuracy in identifying families’ degree of risk; and inter-rater reliability, or the degree to which an instrument results in consistent decisions by social workers for similar cases.³⁰

Predictive validity refers to the ability of an instrument to predict a particular outcome. For example, SAT scores might be used to predict the likelihood of college completion; if high SAT scores are associated with college completion, SAT scores have predictive validity for this purpose. In child welfare, most studies consider whether risk assessment instruments accurately predict the occurrence of subsequent maltreatment.

Any kind of diagnostic or predictive instrument will produce some errors; that is, some cases will be identified as ‘high risk’ even though they are truly low risk (resulting in false positives); and some cases will be identified as low risk although they are truly high risk (resulting in false negatives). Maltreatment subsequent to initial referral is relatively rare, and rare phenomena are difficult to predict accurately.³¹ In addition, human behavior, with its myriad of influencing factors, is also notoriously difficult to predict.³² In child welfare, these classification errors are of concern because false negatives can be dangerous to the child, while false positives can result in poor targeting of agency resources.³³

Some researchers assess the validity of an instrument based upon the degree to which false negatives or false positives are minimized (or the *sensitivity* and *specificity* of the instrument).³⁴ However, when considering low base rate phenomena, a low false negative and false positive rate can be achieved by simply predicting the event never happens.³⁵ Such a strategy, of course, is not useful for protecting children and providing services to families that are in need of them.

An alternative strategy to assess the validity of a prognostic instrument often used in the medical field is to classify individuals into risk categories; that is, based upon personal characteristics, an individual falls into a group that tends to have a low, moderate, or a high risk of some adverse event. In medicine, this framework provides important information for a patient and doctor to decide what kind of preventive approach to take.³⁶ Proponents of this approach to validation assert classification provides important information for services and case decisions, even if it produces more false results than does predicting the event never happens, and should be the basis for determining the validity of a risk assessment instrument.³⁷

Most validity studies focus on predictive validity, but some have argued other forms of validity are important to consider as well. English & Graham assert convergent validity, or the degree to which a measure corresponds to other measures of the same or similar constructs is also relevant.³⁸

In addition to considering the validity of the instruments overall, it is also important to consider the validity of the measures and or outcomes and factors assessed by an instrument. For example, there is some question regarding the appropriate proxy to use for maltreatment recurrence. If *substantiated maltreatment* is used, it could be the case that maltreatment occurs, but is not detected.³⁹ If *subsequent referral* is used, it could overestimate occurrence, as some referrals will be unfounded. Others worry that if substantiation decisions themselves are biased, they would be a poor outcome upon which to assess validity.⁴⁰ And some believe the *severity* of the maltreatment should be incorporated into the criterion.⁴¹ Regarding the other factors, some authors have criticized “anchors” or indicators as being poor proxies for the underlying construct of interest.⁴² An additional criticism of current risk assessment instruments is that both consensus-based and actuarial instruments focus almost exclusively on interpersonal factors of parents and rarely upon neighborhood, community or societal factors that have been determined to be associated with maltreatment in etiological studies.⁴³

Inter-rater reliability refers to the degree to which an instrument results in similar decisions on similar cases, when those cases are assessed by different workers.⁴⁴ Assessing reliability also involves challenges. The goal is to determine whether multiple coders using the same instrument

would reach the same decision for the same situation. However, the practice environment within which one person is making decisions cannot be replicated for another person, to determine if he or she would make the same decision. Two alternative strategies have been used: in the first, different workers read the same case file and their assessments or predictions of risk are compared; in the second, different workers read a case vignette, and their assessments or predictions of risk are compared. Both of these strategies are problematic for different reasons. A vignette is consistent, but artificial and limited in terms of the information that is provided to the decision-maker. Case files are richer and messier, and thus more like the real world, but could be missing critical information that was present in the real world situation.⁴⁵ Additionally, the statistical estimate (“r”) used to assess reliability can be problematic; there could be a high correlation between two scores, even if the two coders have consistently different scores.⁴⁶

Child welfare agencies have an expectation that “risk assessment will have some effect on services provided.”⁴⁷ Therefore, another way to evaluate risk assessment strategies is to consider whether they have resulted in intended effects for children. That is, if use of a particular strategy is improving safety and risk decisions, and high risk children receive protective services such as foster care as a result, then changes in certain outcomes would be expected. For example, one might expect to see fewer recurrences of maltreatment on cases that had been investigated. Similarly, if workers were doing a better job assessing risk and safety in reunification decisions, one would expect to see a reduction in foster re-entry rates, and/or recidivism rates (maltreatment after reunification). Along similar lines, some studies assess the use of various instruments with children and families of color. Because of the well-known disparities in case processes and outcomes for families and children of color, it is important to determine whether a particular instrument is equally valid for different racial/ethnic groups.

Comparing the performance of various instruments is difficult because not all instruments have been assessed for their reliability, validity, or effects. Comparisons are further complicated by the variety of events used by different researchers to measure maltreatment recurrence. For example, many researchers use substantiated maltreatment subsequent to the initial investigation as the criterion against which to measure predictive validity, but others use subsequent referral, a measure of chronicity of subsequent referrals or substantiation, or placement of the child outside

the home. Furthermore, the observation time frame varies, from 30 days, 60 days, 6 months, or 12 months or longer. Finally, for both reliability and validity studies, different strategies are employed to assess these psychometric qualities, and different kinds of statistics are used to estimate them.

Review Of Studies

This review of the research literature examined studies of risk assessment instruments that assessed reliability and validity, as well as those considering the effects of the implementation of a risk assessment system on child and case outcomes, and use with families of various different racial/ethnic groups. This review uses specific search terms and search sources to identify research literature within the topic. This method of searching can reduce the potential for bias in the selection of materials. Using specific search terms, we searched social science and academic databases available through the University of California library. In addition, we searched websites specializing in systematic reviews, as well as research institutes (please see Appendix A for a description of the search strategy). Studies were excluded if they did not assess a particular identified instrument of risk assessment appropriate for use in a CPS investigation situation, or, in the case of predictive validity studies, if the outcome assessed was other than maltreatment recurrence or re-referral of the case.

In this section, research on seven instruments of risk and safety assessment is summarized. These are the Washington Risk Assessment Matrix (WRAM), the California Family Assessment Factor Analysis (?) (the “Fresno” instrument), the Child At Risk Field System (CARF), the CERAP (Child Emergency Response Assessment Protocol), the actuarial Risk Assessment instrument developed by the Children’s Research Center, the Risk Assessment Model of Child Protection from Ontario, and the Utah Risk Assessment Scale. For each instrument, findings of available studies on predictive validity, convergent validity, inter-rater reliability, outcomes, and racial/ethnic group differences are summarized.

WRAM

A consensus-based instrument, the WRAM was developed by Washington State social service agency in 1986. Its composition is constantly evolving based on new research evidence. The instrument is used at the initial investigation. Currently the tool has 37 items that fit within seven theoretical domains: 1) child characteristics, 2) severity of abuse/neglect, 3) chronicity of abuse/neglect, 4) caretaker characteristics, 5) caretaker/child relationship, 6) socio-economic factors, and 7) perpetrator access. To use the instrument, social workers assess and rate the level of risk that they perceive for each item on a five point scale. Based on these ratings, families are categorized into risk levels. The instrument assesses risk of maltreatment in general, rather than considering risk for different kinds of abuse separately (e.g. neglect, abuse, sexual abuse, etc). (See Appendix B for a copy of the instrument; note that the instrument may have been revised since this version was printed).

Predictive validity

In tests of predictive validity, performance of the WRAM was mixed. In one study on 1400 cases from four sites across the county found that while rates of subsequent investigation were higher for moderate or high risk families than for low risk families, rates of *substantiated* maltreatment for families in low, moderate or high risk groups were not significantly different.⁴⁸ When a slightly different outcome was used – the number of subsequent substantiated reports received by a family within two years – a similar difficulty accurately classifying families was seen.⁴⁹

Another study using different analytic techniques and outcomes resulted in similar findings. In this study from 10 child welfare agency offices in New Jersey (n=239), five of the seven theoretical domains of the instrument were adapted to be used as scales measuring the domains, and two items were assessed individually. When these variables were entered into a regression model predicting subsequent maltreatment, the variable measuring abuse severity was found to be *negatively* correlated with maltreatment; i.e., the more severe the abuse the less likely the parent to maltreat again. Variables measuring child characteristics, caretaker characteristics, the parent-child relationship, and socio-economic status were not associated with subsequent maltreatment, while the variable measuring child behavior problems was positively associated with subsequent maltreatment. The model had poor predictive power overall, explaining

approximately 6% of the variability in the outcome. In addition, based on a plot of the sensitivity and specificity of the instrument overall, authors concluded that the performance of the instrument “...might be characterized as generally poor.”⁵⁰

In a different study in the state of Washington (n=12,329), cases coded low or no risk by the instrument were less likely to have a subsequent referral within 18 months than were cases coded as moderate and high risk; moderate risk cases however were not less likely to be re-referred than were high risk cases. Eleven items from the instrument were positively associated with re-referral, and eight items with recurrence of substantiated maltreatment. However, the same study also showed that there was relatively little difference in average risk ratings between cases re-referred, and those not re-referred.⁵¹

Convergent validity

One study considered the convergent validity of the WRAM. In this study (n=261), instrument items were used as scales and correlated with other scales or items from well-known measures of the same constructs. A strong correlation would suggest the WRAM item did a reasonably good job of measuring that area. Alternative measures of constructs were available for only nine of the 37 items on the WRAM. Of these, items assessing child development and behavior problems were not associated with measures of similar constructs. Four of the 5 items related to the caregiver were found to be associated with measures of similar constructs, but the item related to stress and social support was not associated with related measures.⁵² Findings then are mixed; some items do appear to have a degree of construct validity, while others do not.

Inter-rater reliability

Two studies were identified that assessed the inter-rater reliability of the instrument. In the first study, four raters were asked to rate the same 80 cases using the instrument. Less than 14% of the time, all four workers classified families in the same way; just over half the time, three out of four workers did so. Because some portion of these agreements could be due to chance, a statistical correction was done which produces a “kappa” score. Kappa varies from -1 to +1; a kappa of 0 would mean the performance of the instrument was no better than chance. According

to the authors of the study, kappa's in the range of .5-.6 are generally considered acceptable; the score for the WRAM was 0.18.⁵³

Outcomes

No studies were found that considered the effects of implementation of the WRAM on case outcomes.

Racial/Ethnic group differences

Several studies have considered use of the WRAM with different racial/ethnic groups. One study of 8785 cases in Washington state found that African American and Native American families were more likely to be assigned to the highest risk level than their numbers in the referral population would suggest; Asian American families were under-assigned to the highest risk level.⁵⁴ However, it should also be noted that in a subsequent multivariate analysis of 12,329 cases in the same state, Native American families were in fact more likely to be re-referred, and Asian families less likely to be re-referred.⁵⁵ However, another study of 1400 cases in four sites found that approximately equal percentages for African American and White families were classified into each risk level by the WRAM.⁵⁶

CFAFA (the “Fresno model”)

A consensus-based instrument, the CFAFA, or the “Fresno Model,” is derived from an instrument originally developed by the state of Illinois (the Child Abuse and Neglect Tracking System or CANTS 17B). The instrument is no longer used in Illinois, but has been used most recently in California. The instrument is to be used throughout the life of a case. The CFAFA has 23 items that fit within five theoretical domains: 1) precipitating incident, 2) child assessment, 3) caregiver assessment, 4) family assessment, and 5) family-agency interaction. All types of maltreatment are considered together. A social worker rates each item as low, moderate or high risk; sums the number of items coded at each risk level and decides the overall level of risk. (See Appendix C for a copy of the instrument; note that the instrument may have been revised since this version was printed).

Predictive validity

Similarly to the WRAM, in tests of predictive validity, the CFAFA did not perform well. In a study of 1400 cases, while rates of subsequent investigation were higher for moderate or high risk families than for low risk families, rates of substantiated maltreatment for families in low, moderate or high risk groups were not significantly different.⁵⁷ This was also true when the number of subsequent substantiated reports received by a family within two years was used as the outcome instead of the presence of any subsequent substantiated report.⁵⁸

Another study examined the Illinois CANTS 17B, upon which the CFAFA is based. When four items from this instrument were used as variables in a multivariate model predicting subsequent maltreatment (n=239), none were associated with the outcome, and the model had poor predictive power overall, explaining only 1% of the variability in the outcome. In addition, when authors plotted the sensitivity and specificity of the instrument overall in “receiver operating characteristic curves,” the authors concluded that the “performance...might be characterized as generally poor.”⁵⁹

Convergent validity

No studies assessing the convergent validity of the CFAFA were found.

Inter-rater reliability

One study attempted to determine the inter-rater reliability of the instrument. Four raters were asked to rate the same 80 cases using the instrument. Just over 16% of the time, all four workers classified families into the same risk groups; about 45% of the time, three out of four workers did so. The kappa for the CFAFA was 0.184.⁶⁰

Outcomes

No studies were found that considered the effects of implementation of the CFAFA on case outcomes.

Racial/ethnic group differences

One study of 1400 cases considered use of the CFABA with different racial/ethnic groups. In this study, the CFABA classified approximately equal percentages of African American and White families into each risk level.⁶¹

CARF

A consensus-based instrument, CARF was one of the first risk assessment instruments to focus on safety as distinct from risk. It was developed by ACTION for Child Protection. The instrument is to be used throughout the life of a case. Fourteen items within the following 5 domains: 1) child; 2) parent; 3) family; 4) maltreatment; and 5) intervention. Four “qualifiers” are also to be considered: 1) duration of a negative influence; 2) pervasiveness of a negative influence; 3) acknowledgement by parents of a negative influence; and 4) control of the negative influence. All types of maltreatment are considered together. Each item or qualifier is rated on a four point scale; the average of the 14 items plus the average of the four qualifiers is summed and divided by 2 to arrive at a final risk score. The family is then categorized into no risk, low risk, moderate risk, significant risk, or high risk groups. (see Appendix D for a copy of the instrument; note that the instrument may have been revised since this version was printed).

Predictive validity

The performance of CARF on tests of predictive validity was mixed. In one study of 207 cases in New York state, families assigned to the highest risk group were more likely to have a subsequent referral than families assigned the lowest risk group, though the relationship only “approached” statistical significance. Particular items were not found to be associated with subsequent maltreatment.⁶²

Convergent validity

One study assessed the convergent validity of CARF (n=90). Social worker ratings of the “parent risk field” and the “family risk field” were not found to be associated with any of eight clinical measures of related constructs against which they were each tested. Social worker ratings of the “child risk field” was found to be associated with one clinical measures of parent-reported “child to parent violence;” however, the rating of child risk field was also found to be negatively

associated with level of child PTSD – the opposite direction than expected – and no relationship was found between the ratings and five other clinical measures of related constructs.⁶³

Inter-rater reliability

One study assessing the inter-rater reliability of CARF was found. In this study, 25-50 workers from several counties in Pennsylvania were asked to read case vignettes and assess the level of risk using the risk assessment instrument. Scores of all pairs of coders of a vignette were correlated and those correlations averaged. Alpha coefficients for the CARF instrument overall risk scores for three different vignettes varied widely, ranging between .067 (very low reliability) to .952 (very high reliability).⁶⁴

Outcomes

In a study comparing substantiation rates before and after implementation of CARF in one New York county (n=207), no differences were found. When maltreatment type was considered separately, physical neglect was found to be somewhat more likely to be substantiated subsequent to implementation than prior to implementation. No difference in before and after substantiation rates were found for physical maltreatment sexual maltreatment, medical neglect, emotional maltreatment, or educational neglect.⁶⁵

Racial/ethnic group differences

No studies were found that considered the use of CARF with different racial/ethnic groups.

CERAP

A consensus-based instrument, the CERAP was developed by Illinois Department of Child and Family Services, the American Humane Association, the University of Illinois, and field experts as a “safety assessment.” The instrument is to be used throughout the life of the case. The CERAP is a single list of 14 items. All types of maltreatment are considered together. The social worker notes the presence or absence of each item; if any of the items are present, the social worker decides whether the child is “safe” or “unsafe.” If the worker decides the child is unsafe, a safety plan is developed. Training includes a rigorous testing and certification process. (See

Appendix E for a copy of the instrument; note that the instrument may have been revised since this version was printed.)

Predictive validity

One study attempted to assess the predictive validity of the CERAP. Since the CERAP focuses on safety assessment, the outcome used was subsequent referral within 60 days. Use of the CERAP at two different case time points -- initial investigation (n=380), and within 5 days of case opening (n=350) -- was assessed. At initial investigation, neither overall safety assessment nor number of safety factors identified associated with subsequent referral within 60 days, either in bivariate tests or multivariate instruments controlling for CERAP completion, prior reports, total number caregiver problems, and service receipt.⁶⁶ Within 5 days of case opening, both safety assessment and number of safety factors identified for services associated with subsequent referral within 60 days in bivariate tests, but these relationships did not remain in multivariate instruments once other factors listed had been controlled for. Completion of the instrument, regardless of safety rating, was negatively associated with subsequent referral.⁶⁷

Convergent validity

No studies were found that considered the convergent validity of the CERAP.

Inter-rater reliability

No studies were found that considered the inter-rater reliability of the CERAP.

Outcomes

A series of studies considering changes in the 60 day maltreatment recurrence rates in one state after implementation of CERAP compared to the year before implementation have been conducted. These studies have found a reduction in the maltreatment recurrence rates that has been maintained for six years following implementation.⁶⁸ Several alternative explanations for the reductions -- increased use of out of home placement, another policy, nationwide trend -- were considered and ruled out in a follow-up study.⁶⁹

Racial/ethnic group differences

No studies were found that considered the use of the CERAP with different racial/ethnic groups.

CRC Actuarial Instruments for Risk Assessment

Slightly different versions of this actuarial risk assessment instrument have been developed for various jurisdictions by the Children's Research Center (CRC). These instruments are based upon the statistical association of variables with substantiated maltreatment injury, foster care placement, and reinvestigation within two years in each location. The instrument is used at the initial investigation, and includes two subscales of ten items each; one subscale assesses risk of neglect and the other risk of physical or sexual abuse. Each item is scored with a 0, 1, or 2 as indicated on the instrument and each subscale is summed. Based on the highest subscale score, a family is classified into a low, moderate, high, or very high risk category. In most jurisdictions, workers can override the risk classification and increase the risk rating by one level. (See Appendix F for a copy of the instrument; note that the instrument may have been revised since this version was printed).

Predictive validity

A number of studies have been conducted attempting to assess the predictive validity of the CRC risk assessment instruments. These instruments are able to distinguish between low/medium/high levels of risk of subsequent maltreatment -- that is, families categorized as high risk by the instrument have a distinctly higher rate of subsequent maltreatment than do families categorized as moderate risk; and moderate risk families have a distinctly higher rate of subsequent maltreatment than do families categorized as low risk. This was found to be true for subsequent maltreatment within 6 months,⁷⁰ 18 months,⁷¹ and 2 years.⁷² When the number of subsequent substantiated reports received by a family was used as the outcome instead of the presence of any subsequent substantiated report, a similarly good ability to accurately classify families was seen at both 18 months⁷³ and 24 months.⁷⁴ Additionally, in a multivariate study, families coded as higher risk showed a stronger association with subsequent maltreatment (controlling for ethnicity, county size, service receipt, and safety finding) than did families coded as moderate or low risk.⁷⁵

Convergent validity

No studies were found that considered the convergent validity of the CRC actuarial instruments.

Inter-rater reliability

The CRC risk assessment instruments performed fairly well in reliability studies. In one study, four raters were asked to rate the same 80 cases using the instrument. Over half of the time, all four workers classified families in the same way; 85% of the time, three out of four workers did so. The kappa score for the WRAM CRC Risk Assessment instrument was .562.⁷⁶ In a second study that involved coding vignettes, most workers scored subscales with a possible range from 0-20 within 4 points of each other; “ranges indicate some disagreement but general agreement because they are small and not excessively skewed”; somewhat lower consistency was realized when scores were combined for an overall risk score.⁷⁷

Outcomes

No study was found that assessed the outcomes of use of the CRC actuarial risk assessment specifically. One study assessed outcomes in one state following the implementation of an *array* of CRC instruments that included the actuarial risk assessment. Compared to a demographically matched set of counties not implementing the array of instruments, counties implementing the CRC array had lower rates of re-referral or substantiation for cases closed without services; families received more services, particularly if they were high risk; and referral rates, substantiation rates, removal rates, and injuries were lower.⁷⁸

Racial/ethnic group differences

Findings from studies that assess the use of the instruments with different racial/ethnic groups are mixed. Some studies found that the instrument classifies approximately equal percentages of all ethnic groups into each risk level;⁷⁹ that rates of recurrence for different risk categories are consistent across ethnic groups;⁸⁰ and that the association of scores with subsequent maltreatment does not differ by ethnic group.⁸¹ However, other studies have found that White families were somewhat more likely to be coded higher on more items than families of color;⁸² African American families scored slightly higher on neglect scale, and white families scored slightly higher on physical abuse/sexual abuse scale;⁸³ in Minnesota, Southeast Asian and

Hispanic families were slightly more likely to be coded as low risk than other ethnic groups, and Native American families were slightly more likely to be coded High and Intensive Risk (differences were greatest on the neglect scale).⁸⁴ Additionally, in the Minnesota study, when predictive validity was considered separately by ethnic group, distinctions between maltreatment rates of risk groups were smaller for Native American families due to the high referral rate in the Low Risk group.⁸⁵

Risk Assessment Model of Child Protection (Ontario)

Based upon scales originally developed by Magura and Moses (1986), this consensus-based instrument was modified by a research team from the University of Toronto in consultation with the Ontario Association of Children's Aid Societies. Twenty-two items within five domains (caregiver, child, family, intervention, and abuse/neglect) are assessed via a 4-point Likert-type scale scoring system. The social worker determines a total overall assessment of risk (low to high) and the cumulative risk score (the total of the ratings from the five domains). The instrument is to be used when deciding whether a child should be removed and placed into foster care. All types of maltreatment are considered together.

Predictive validity

No studies were found that assessed the predictive validity of the Ontario Risk Assessment Model using outcomes of subsequent referral or maltreatment.

Convergent validity

No studies were found that considered the convergent validity of the Ontario Risk Assessment Model.

Inter-rater reliability

One study was found that assessed the inter-rater reliability of the Ontario Risk Assessment Model. A reliability score of $r=.92$ is reported for the cumulative risk score, and of $r=.96$ for the overall risk. However, it is not clear how many coders were used in the reliability assessment, how many cases were assessed for reliability, or by what means reliability was determined.⁸⁶

Outcomes

No studies were found that considered outcomes of the Ontario Risk Assessment Model.

Racial/ethnic group differences

No studies were found that considered the use of the Ontario Risk Assessment Model with different racial/ethnic groups.

Utah Risk Assessment Scales

The Utah Risk Assessment Scales are based upon Family Risk Scales and Child Well-being scales originally developed by Magura and Moses (1986), and incorporate additional scale items developed by the a steering committee composed of members of the Utah Department of Social Services and the Utah Child Welfare Training Project. According the study authors, these additional items are an advantage of the Utah Risk Assessment Scales, as these steering committee members were “CPS experts whose practice wisdom and experiences in child protection enabled them to add additional scales useful for making judgments about child well-being such as the degree of emotional maltreatment.”⁸⁷ This consensus-based instrument is composed of 32 items within five domains: parent, child, family, maltreatment, and intervention, assessed via a Likert-type scale scoring system. The instrument is to be used by CPS intake and investigative workers. All types of maltreatment are considered together.

Predictive validity

No studies were found that considered the predictive validity of the Utah Risk Assessment Scales.

Convergent validity

No studies were found that considered the convergent validity of the Utah Risk Assessment Scales.

Inter-rater reliability

One study (reported in two articles) assessed the inter-rater reliability of the Utah Risk Assessment Scales. To assess inter-rater reliability, eight vignettes were developed describing

cases of varying severity. Social workers, supervisors and CPS experts were asked to review each vignette and provide a risk rating using the Utah Risk Assessment Scales. These risk scores were then correlated to determine reliability scores. Pearson's r coefficients ranged from .568 to .855 for the eight vignettes, each of which was assessed by $n=22-28$ raters. A "Spearman-Brown prophecy formula" was applied to these scores to provide a "stepped-up" reliability estimate; according to authors, this analytical technique allows for a more accurate estimate of inter-rater reliability. Stepped up estimates were all above .970.⁸⁸

Outcomes

No studies were found that considered outcomes of the Utah Risk Assessment Scales.

Racial/ethnic group differences

No studies were found that considered the use of the Utah Risk Assessment Scales with different racial/ethnic groups.

Risk Assessment Implementation Issues

This section describes three implementation issues: 1) supports for implementation, 2) use at decision points across the life of a case, and 3) liability issues. The effort to institute a standardized risk assessment system across the state of California is also described.

Support for Implementation

The implementation of a risk assessment system can be problematic. One study of the CARF system reported that less experienced workers found the instrument more useful than did more experienced workers, who felt it limited their ability to utilize their expertise. In addition, some workers said they adjusted CARF ratings to match their clinical impressions, rather than using the risk assessment instrument to guide their decisions.⁸⁹ Similarly, in another study examining the Illinois CANTS-17B, workers admitted that they inflated risk levels at intake in order to guarantee that a family would receive services.⁹⁰ And several studies of risk assessment implementation found that risk assessment forms were often missing from case files or incomplete.⁹¹

A number of risk assessment researchers suggest that the success or failure of any instrument is based less upon whether it is actuarial or consensus-based and more upon the quality of implementation.⁹² There is some consensus that the successful implementation of risk assessment instruments requires certain supports.⁹³ High quality and comprehensive training is important, and should facilitate a conceptual understanding of the instrument⁹⁴ as well as involve an experiential component.⁹⁵ In addition, for optimal implementation, staff may need other skills training besides training in the risk assessment instrument such as interviewing, time management, assessment, and treatment planning skills.⁹⁶ Secondly, the important role supervisors and managers play in the implementation process has also been emphasized.⁹⁷ Involving supervisors and direct line staff in the planning and implementation process as much as possible has been recommended; the greater awareness that staff have early on, the less resistance may be encountered long-term.⁹⁸ Agency administration can also show support by using implementation of a new risk assessment instrument to look at and adjust workload.⁹⁹ In addition, agencies may need to plan for ongoing monitoring and accountability.¹⁰⁰

Use Across the Life of a Case

Since a family's risk may change over time, it is important that risk be periodically reassessed.¹⁰¹ For example, few states have explicit guidelines for making screening decisions, and fewer still have formal instruments to guide this decision.¹⁰² The percentage of referrals that are screened out varies dramatically by state, from 5% in New Jersey to 78% in Vermont, yet states with higher investigation rates were just as likely to substantiate a referral as states with lower investigation rates.¹⁰³ In addition, caseworkers have difficulty revising their assessments of families once they have been made.¹⁰⁴ Therefore, a structured instrument could help workers attend to critical factors that would indicate changes in risk across the life of a case. However, using the same instrument to assess risk at different case points may be unwise.¹⁰⁵ Factors that predict maltreatment at one point, such as at investigation and prior to services, may not be the same as those that predict subsequent maltreatment at another time point, such as at reunification after service provision. For example, one study examining two time points (within 24 hours of the initial investigation of a case, and within 5 days of a case opening for in-home services) found that factors that predicted maltreatment recurrence at each time point were not always the same.¹⁰⁶

Unfortunately, there is very little research regarding the reliability or validity of instruments used at other points in the case.¹⁰⁷ A partial explanation for this is related to the nature of some of the outcomes, which don't lend themselves well to explorations of validity or reliability. For example, the safety assessment focuses on assessing imminent or current risk to a child, and is usually concerned about maltreatment "...of a moderate to severe nature."¹⁰⁸ It refers to the immediate situation. The likelihood of subsequent maltreatment within 60 days is fairly rare; the likelihood of subsequent maltreatment occurring within days is infinitesimal. Workers must attempt to predict this anyhow, but determinants of events so rare essentially cannot be determined via any kind of statistical instrument using the kinds of factors available to social workers.¹⁰⁹

Liability Issues

While legal liability in CPS work has long been a concern,¹¹⁰ little has been written specifically about liability as it relates to the use of risk assessment instruments. As more states and counties mandate the use of a risk assessment instrument, concerns have been raised about the potential increases in risk of legal liability for workers and agencies.

Opinions about the issue differ. The U.S. Supreme Court's 1989 decision in *DeShaney v. Winnebago County Department of Social Services* asserted that there existed no "special relationship" requiring the state to protect a child, if the child had not been taken into custody. Thus, the state would not be held liable for an incorrect assessment of risk for a child known to CPS.¹¹¹ In addition, while Section 1983 of the Civil Rights Act allows for individuals to bring suit against government employees, "...the duties of most government child welfare workers are discretionary in nature and, as such, provide them with either absolute, or...qualified immunity for their actions".¹¹² CPS workers are assumed to be using their best judgment and working in good faith when making decisions about children and families, and are therefore provided immunity for their work and safeguarded from lawsuits.

Introduction of a formal risk assessment strategy could affect this in several ways. One school of thought is that mandated use of risk assessment instruments could increase agencies' potential liability: by making the use of a risk assessment instrument a part of agency policy (rather than

leaving its use to the worker's discretion) a worker's right to claim immunity may be reduced.¹¹³ Secondly, some argue liability could depend on the extent to which a risk assessment instrument is utilized.¹¹⁴ If a county has required the use of a risk assessment instrument and a worker fails to use instrument correctly or does not use the instrument at all, these actions could be interpreted as negligence or failure to follow agency policies. And third, liability may depend upon the validity of the instrument. The use of un-tested instruments and/or inconsistent worker training on risk assessment instruments may increase an agency's vulnerability.¹¹⁵ "Agencies place themselves in a precarious legal position by claiming their decision-making is based on standardized, validated risk assessment protocols when it is not. A state or provincial child protective service system that endorses or mandates a formal risk assessment instrument 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 instruments, agencies may be subject to legal remedies."¹¹⁶

A contrasting opinion is that worker discretion and the good-faith conduct of job duties will continue to protect workers from potential lawsuits; risk assessment instruments would neither improve nor worsen a worker's chance of being sued. Nohejl et al.¹¹⁷ reason that questions of worker liability center on competence and decision-making. Poorly filled out or missing risk assessment instruments would only be evidence of "simple negligence", which does not incur liability.¹¹⁸ In addition, any inaccuracy in prediction would not make the worker liable; the worker's efforts would be covered under good faith immunity. Only if there was evidence of "gross negligence" could liability exist.¹¹⁹ Nohejl et al.¹²⁰ also assert that the selection of a risk assessment instrument and worker training are considered within the realm of discretionary decisions. While implementation of a poorly researched risk assessment instrument or inadequately trained workers may cause embarrassment in court, it is unlikely that any agency or individual worker could be sued. Based on this view, an agency would have to show signs of "deliberate indifference" for the safety of others for liability to become an issue.¹²¹

Risk and Safety Assessment in California

California's 58 counties have adopted a variety of strategies for conducting risk assessments. In the past, most counties used either the consensus-based CFAFA (the "Fresno" model), which

was adopted by the state in 1991, or the CFRA (California Family Risk Assessment), the actuarial instrument developed by the Children's Research Center (CRC). Some counties do not require the use of any instrument. A major three-year systems reform effort in California (2001-2004) resulted in a recommendation that a standardized risk assessment procedure be implemented throughout the state.¹²² A workgroup comprised of county representatives and state personnel was convened to outline the risk and safety assessment procedure to be implemented by the state. Rather than identifying a particular instrument for state-wide implementation, the workgroup developed a framework of the various domains and items that should be assessed for risk and safety across the life of a case (see Figure 1). The decision points throughout the life of a case at which risk and safety should be assessed include: 1) determining a response to a referral, 2) making the initial safety decision, 3) placement, 4) the referral disposition, 5) ongoing case planning decisions, 6) reunification, and 7) case closure. A county can use either an actuarial or a consensus-based instrument as long as the areas relevant for each decision point are assessed.

Discussion and Implications

Overall, the available research suggests that the actuarial risk assessment instruments appear to have greater predictive validity and inter-rater reliability than the consensus-based instruments. The better predictive validity of the actuarial instruments may be related to their development via the statistical identification of the strongest predictors of a particular outcome in that state or county. Unless the sample used to develop that model was different from the typical population referred to child welfare agencies in that jurisdiction (or there were major changes in the local context or population demographics), it would be reasonable to assume that those variables in the model would continue to be predictive on later cohorts. The processes for identifying factors for consensus-based instruments, as outlined earlier, may simply be less accurate at identifying the strongest predictors of maltreatment.

Convergent validity was not assessed for the actuarial instruments and the performance of consensus-based instruments in this area was generally poor. These instruments may be unreliable, they may not be good measures of underlying concepts, or both problems may exist.

Figure 1: Recommended Statewide Safety Assessment System for California

	DETERMINE RESPONSE	INITIAL SAFETY DECISION	PLACEMENT	REFERRAL DISPOSITION	INITIAL/CHANGES IN CASE PLANNING	REUNIFICATION	CASE CLOSURE
Purpose	<p>Determine the appropriateness of the Child Abuse Report for a response.</p> <p>Determines the urgency of the response needed & the Response Team (who will respond).</p>	<p>Ensuring the safety of the child.</p> <p>Begin to identify risk factors of concern in the family.</p>	<p>Ensure the safety of the child by providing the least-restrictive alternative to their home environment.</p>	<p>Determine if a family needs continued public child welfare services, community services or has no current service needs.</p> <p>Examine why problems may be present, and what may be causing the symptoms and behaviors that may result in maltreatment and begins the process of helping the family.</p>	<p>To identify the best possible strategies for changing the conditions/behaviors that lead to harm or the risk of harm to the child.</p>	<p>To determine if it is safe to return a child to his or her home.</p>	<p>To safely terminate public CWS involvement.</p>
Standard Areas For Review	<ol style="list-style-type: none"> Current and Prior Maltreatment. Current and Prior CWS history. Child’s strengths and vulnerability. Cultural and language considerations. Perpetrator access to child. Violence Propensity. Social Environment. Caregiver Protective Capacity. Home Environment. Ability to meet child’s needs. Caregiver / child interaction. Ability to locate. 	<ol style="list-style-type: none"> Current and prior maltreatment. Child’s strengths and vulnerability. Cultural and language considerations. Perpetrator access to child. Violence propensity. Social environment. Caregiver protective capacity. Home environment. Ability to meet child’s needs. Caregiver / child interaction. Safety interventions. Pre-placement preventive services. 	<ol style="list-style-type: none"> Current and prior CWS history. Child’s strengths and vulnerability. Cultural and language considerations. Child’s immediate and ongoing needs. Level of care to meet child’s needs. Substitute Care provider’s willingness and ability to provide care and ensure safety. Substitute care provider’s strengths and willingness to support child’s case plan. Sibling placement considerations. Child’s permanency needs. Visitation. 	<ol style="list-style-type: none"> Current and prior maltreatment. Current and prior CWS history. Child’s strengths and vulnerability. Cultural and language considerations. Perpetrator access to child. Violence Propensity. Social Environment. Caregiver Protective Capacity. Home Environment. Ability to meet child’s needs. Caregiver / child interaction. Ability to locate. Safety interventions. Caregiver willingness to change. 	<ol style="list-style-type: none"> Current and prior maltreatment. Child’s strengths and vulnerability. Cultural and language considerations. Social Environment. Child’s permanency needs. Visitation. Contributing factors requiring intervention. Current and previous social services. History of criminal behavior. Basic needs. Medical / Dental Care. Mental Health / Coping Skills. Child Development. Education needs. Parenting Skills and practices. Child’s relationships with peers and adults. Substance abuse. Domestic violence. Delinquent behavior. 	<ol style="list-style-type: none"> Current and prior maltreatment. Current and prior CWS history. Child strengths and vulnerability. Cultural and language considerations. Perpetrator access to child. Violence propensity. Social Environment. Caregiver protective capacity. Home Environment Ability to meet child’s needs. Caregiver / child interaction. Child’s permanency needs. Visitation. Subsequent referrals. Caregiver’s compliance / progress toward objectives within case plan. 	<ol style="list-style-type: none"> Current and prior maltreatment. Child’s strengths and vulnerability. Cultural and language considerations. Perpetrator access to child. Violence Propensity. Social Environment. Caregiver protective capacity. Home Environment. Ability to meet child’s needs. Caregiver / child interaction. Child’s permanency needs. Ability to locate. Child’s permanency needs. Subsequent referrals. Caregiver’s compliance / progress toward objectives within case plan.

Actuarial instruments may show better reliability because factors in these instruments are more often objective; the factors in consensus-based instruments tend to be more subjective and less precise. For example, in a question related to prior CPS history using the consensus-based WRAM, the social worker is required to determine whether past incidents were ‘isolated’ or ‘intermittent,’ and whether there is evidence of ‘minor’ abuse and neglect or ‘moderate’ abuse and neglect. In comparison, the CRC instrument asks ‘whether or not’ there was a prior injury to a child from abuse or neglect, or ‘whether or not’ there was a prior investigation. Generally, more precise, well-defined, objective, and clearly articulated measures are more likely to be reliable because differences of opinion about the meaning or coding of factors are minimized.¹²³ In support of this idea, English & Graham (2000) note that in one study of the CRC actuarial instrument,¹²⁴ more objective items on the instrument (such as the age of child) had higher reliability than more subjective items (such as, “was child inadequately supervised”). In addition, some consensus-based instruments often require coders to use their judgment regarding the level or risk related to an area. That is, rather than asking workers “whether or not” maltreatment previously occurred, a consensus-based instrument asks workers to assign a level of risk to the broad area of previous maltreatment. Different workers could weight the same category differently.

In terms of outcomes, studies of the CERAP and the CRC instruments suggested that implementation resulted in improved outcomes. These instruments may be improving the accuracy of worker assessments of risk, resulting in fewer high risk children being left at home to be re-abused. The findings regarding the use of the instruments with different racial/ethnic groups are mixed.

It is important to note that the available research is limited. For any particular instrument, there were only a few studies available, or sometimes only a single study was found. Therefore, conclusions about the risk assessment instruments should be considered preliminary and in need of further study.

Implications for Practice

Some of the continuing debate regarding the best approach to assessing risk and safety may be due to a lack of clarity regarding the purposes of “risk assessment.” Distinctions between *risk assessment* and *family assessment* can be somewhat unclear¹²⁵ and a number of researchers have argued that they have often been confused.¹²⁶ “While risk assessment is designed to accurately estimate the likelihood of future incidents of maltreatment, 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 their environment that may contribute to or mitigate maltreatment.”¹²⁷ Rycus & Hughes (2003) argue that these case activities are separate and distinct, and that neither is served by attempting to use a single instrument to do them both, or even by attempting to do them at the same time.

If the goal of an assessment is to *predict the likelihood of the recurrence of maltreatment* in order to provide services to the families at greatest risk, this is clearly a risk assessment. The research evidence strongly suggests that the actuarial instrument will produce a more accurate and reliable prediction than the consensus-based instruments. On the other hand, if the goal of an assessment activity is to *gain a comprehensive understanding of the service needs of a family or individual*, a family needs assessment instrument may be more appropriate. In this situation, a consensus-based instrument may do a better job because it incorporates more items and thus provides more information. However, consensus-based instruments did not have high convergent validity, suggesting they may not accurately measure the relevant characteristics and thus would not necessarily be helpful in family assessments.

Finally, most research on risk assessment acknowledges that the use of any kind of risk assessment instrument, actuarial or consensus-based, requires good clinical skills.¹²⁸ Even the CRC actuarial instrument contains numerous items that require clinical judgment to score, and allow for a clinical override based on family characteristics or dynamics that are likely to affect risk, but are not included on the actuarial instrument. As Ereth et al. (2003) have noted, “...A caseworker can sense things that an actuarial instrument would ignore or could not employ... Many characteristics of human subjects simply cannot be quantified empirically and actuarial models cannot easily account for rare events.”¹²⁹ Therefore, clinical judgment can never be

eliminated from any risk assessment process. In fact, many researchers in child welfare stress that instruments for risk and safety assessment should be understood as decision aids to enhance or expand upon clinical judgment, rather than as a competing approach.¹³⁰ As Munro (1999) observed, "...Errors can be reduced if people are aware of them and strive consciously to avoid them. The challenge is to devise aids to reasoning that recognize the central role of intuition and do not seek to ignore or parallel it but, by using our understanding of its known weakness, offer ways of testing and augmenting it."¹³¹

Implications for Research

Most of the available research on risk and safety assessment instruments is limited to a few fairly well-known instruments: the WRAM, CERAP, CARF, CFAFA (the "Fresno" instrument), and the CRC actuarial instrument. Clearly, further research in the area of risk assessment is needed, including a better understanding of the following: the instruments currently in use, predictors of future maltreatment at other decision points across the life of a case, the appropriate definition of maltreatment recurrence, and implementation issues.

One useful endeavor would be an on-going survey of states' use of risk assessment instruments. Currently, there is no process by which such information is gathered, updated and made available to the practice and research community. As a result, it is not clear how many jurisdictions use actuarial versus consensus based instruments, nor which instruments they use. In addition, of research that has been done, much has been conducted by researchers associated with the instruments. While this research is of high quality and has been published in respected peer-reviewed outlets, studies conducted by independent researchers that confirm the findings of those studies might add credence to the conclusions.

Along the same lines, the current research focuses primarily on one decision point in the case; namely, the initial investigation. There is growing interest in instruments to assist in decision-making at other points in the case. Therefore, there is a clear need for the development of research-based instruments that can be validated for other decision points such as placement and reunification.

Certainly, predictive validity studies are appropriate for any instrument, actuarial or consensus-based, that attempts to assess the likelihood of future maltreatment. Some suggest that other relevant outcomes besides recurrence, such as severity of abuse should be considered as well.¹³² Whether tests of construct validity are appropriate for predictive instruments is less clear. However, if a distinction is clarified between risk assessment and family assessment, construct validity clearly is relevant to considerations of family assessment. On-going validity and reliability studies are necessary so that the instruments can be continually refined and improved over time.

Many have also suggested that the effects of services provided must also be taken into consideration.¹³³ Wald and Woolverton (1990) assert that a risk assessment instrument is “truly useful only if it identifies the likelihood of re-abuse *given specific interventions*.” The provision of services may mitigate risk; and therefore prediction that fails to take services to be provided into account may possibly overestimate risk. Others have pointed out the lack of consideration of neighborhood factors or child-caregiver interactions in current instruments,¹³⁴ which should also be taken into account in future research. Lastly, the nature and quality of implementation have been neglected in the research literature. It is important to consider the quality and nature of the implementation process, and worker acceptance or resistance to the instrument.¹³⁵

Summary

This structured review of the available research literature on instruments of risk and safety assessment in child welfare suggests that actuarial models have stronger predictive validity and inter-rater reliability than do consensus-based models. Neither approach emerged as superior in terms of affecting outcomes or treating racial/ethnic groups equitably. This review was limited by: 1) the lack of studies on decision points other than initial investigation, 2) the variability in definitions and measures across studies, and 3) the relatively small number of studies examining risk assessment instruments. Nonetheless, the findings of this structured literature review should be useful to agencies and practitioners when considering the various approaches to risk and safety assessment in child welfare.

Endnotes

- ¹ Wald & Woolverton (1990)
- ² Wald & Woolverton (1990)
- ³ Rycus & Hughes (2003)
- ⁴ Camasso & Jagannathan (2000); Leschied, Chiodo, Whitehead, Hurley & Marshall (2003); Rycus & Hughes (2003); Wald & Woolverton (1990)
- ⁵ Rycus & Hughes (2003)
- ⁶ Kahneman & Tversky (1982)
- ⁷ Kahneman & Tversky (1982)
- ⁸ Grove & Meehl (1996)
- ⁹ Dawes (1994)
- ¹⁰ Scheurman, Rossi & Budde (1999)
- ¹¹ Scheurman, Rossi & Budde (1999), p. 609
- ¹² Munro (1999) p.748
- ¹³ Munro (1999)
- ¹⁴ English (1999)
- ¹⁵ English (1999)
- ¹⁶ Rycus & Hughes (2003)
- ¹⁷ English & Graham (2000)
- ¹⁸ English (1999)
- ¹⁹ Doueck, English, DePanfalis & Moote (1993); English (1999)
- ²⁰ Nasuti & Pecora (1993)
- ²¹ Rycus & Hughes (2003) p.13
- ²² Rycus & Hughes (2003)
- ²³ Rycus & Hughes (2003)
- ²⁴ Wald & Woolverton (1990)
- ²⁵ McDonald & Marks (1991)
- ²⁶ Ereth, Johnson & Wagner (2003)
- ²⁷ Grove & Meehl (1996)
- ²⁸ Dawes (1979)
- ²⁹ Lyons, Doueck & Wodarski (1996)
- ³⁰ Rycus & Hughes (2003)
- ³¹ Munro & Rungay (2000)
- ³² Baird, Ereth & Wagner (1999)
- ³³ Camasso & Jagannathan (2000)
- ³⁴ Lyons et al.(1996)
- ³⁵ Baird, Ereth & Wagner (1999); Dawes (1994)
- ³⁶ Altman & Royston (2000); Baird & Wagner (2000)
- ³⁷ Baird, Ereth & Wagner (1999); Rycus & Hughes (2003)
- ³⁸ English & Graham (2000)
- ³⁹ English & Graham (2000)
- ⁴⁰ Morton (1999)
- ⁴¹ Morton & Salovitz (n.d.)
- ⁴² Morton (2004a)
- ⁴³ Galasso (2001)
- ⁴⁴ Rycus & Hughes (2003); Schuerman et al. (1999)
- ⁴⁵ Baird, Ereth & Wagner (1999)
- ⁴⁶ Baird, Ereth & Wagner (1999)
- ⁴⁷ Fluke, Wells, England, Walsh, English, Johnson, Gamble & Woods (1993), p. 118
- ⁴⁸ Baird & Wagner (2000)
- ⁴⁹ Baird & Wagner (2000)
- ⁵⁰ Camasso & Jagannathan (1995)
- ⁵¹ English, Marshall, Brummel & Orme (1999)
- ⁵² English & Graham (2000)
- ⁵³ Baird, Wagner, Healy & Johnson (1999)

-
- ⁵⁴ English, Marshall, Brummell & Orme (1995)
⁵⁵ English et al. (1999)
⁵⁶ Baird & Wagner (2000)
⁵⁷ Baird & Wagner (2000)
⁵⁸ Baird & Wagner (2000)
⁵⁹ Camasso & Jagannathan (1995)
⁶⁰ Baird, Wagner, Healy & Johnson (1999)
⁶¹ Baird & Wagner (2000)
⁶² Doueck, Levine & Bronson (1993)
⁶³ Kolko (1998)
⁶⁴ Fluke et al. (1993)
⁶⁵ Doueck, Levine & Bronson (1993)
⁶⁶ Fuller, Wells & Cotton (2001)
⁶⁷ Fuller et al. (2001)
⁶⁸ Nieto & Garnier 2001; Garnier & Nieto (2002)
⁶⁹ Fluke, Edwards, Bussey, Wells & Johnson (2001)
⁷⁰ Johnson (2004)
⁷¹ Baird & Wagner (2000)
⁷² Johnson (2004); Loman & Siegel (2004)
⁷³ Baird & Wagner (2000)
⁷⁴ Loman & Siegel (2004)
⁷⁵ Johnson (2004)
⁷⁶ Baird, Wagner, Healy & Johnson (1999)
⁷⁷ Loman & Siegel (2004)
⁷⁸ Wagner, Hull & Luttrell (1995)
⁷⁹ Baird & Wagner (2000); Baird, Ereth & Wagner (1999); Johnson (2004)
⁸⁰ Baird, Ereth & Wagner (1999)
⁸¹ Johnson (2005); Johnson (2004)
⁸² Johnson (2005); Johnson (2004)
⁸³ Baird, Ereth & Wagner. (1999)
⁸⁴ Loman & Siegal (2004)
⁸⁵ Loman & Siegal (2004)
⁸⁶ Lescheid et al. (2003)
⁸⁷ Nasuti (1998) p. 57
⁸⁸ Nasuti (1998); Nasuti & Pecora (1993)
⁸⁹ Doueck, Levine & Bronson (1993)
⁹⁰ Lyle & Graham (2000)
⁹¹ Kolko (1998); Lyle & Graham (2000)
⁹² Hollinshead & Fluke (2000)
⁹³ DePanfilis (1996); Lyle & Graham (2000)
⁹⁴ Cicchinelli (1990); Repetosky & Bailey (1988)
⁹⁵ Costello (1995); Depanfilis (1996)
⁹⁶ Costello (1995); DePanfilis (1996); Thompson, Brown, & Pecora (1989)
⁹⁷ Albers & Roditti (2004); Depanfilis (1996); Rycus & Hughes (2003); Sheets (1992)
⁹⁸ Sheets (1992)
⁹⁹ DePanfilis (1996); Squadrito, Neuenfeldt, & Fluke (1994)
¹⁰⁰ Costello (1995); Repetosky & Bailey (1988)
¹⁰¹ Munro (2004)
¹⁰² Downing, Wells, & Fluke (1990)
¹⁰³ Tumlin & Geen (2000)
¹⁰⁴ Munro (1999)
¹⁰⁵ Wald & Woolverton (1990)
¹⁰⁶ Fuller et al. (2001)
¹⁰⁷ Rycus & Hughes (2003); Zuravin, Orme & Hegar (1995)
¹⁰⁸ Fuller et al. (2001)

-
- ¹⁰⁹ Johnson (2004)
¹¹⁰ American Federation of State, County and Municipal Employees (1999); Kanani, Regehr & Bernstein (2002)
¹¹¹ Nohejl, Doueck & Levine (1992)
¹¹² Curran (1995), p. 17
¹¹³ Curran (1995); Davidson (1991)
¹¹⁴ Curran (1995); Davidson (1991)
¹¹⁵ Curran (1995); Davidson (1991); Rycus & Hughes (2003); Wald & Woolverton (1990)
¹¹⁶ Rycus & Hughes (2003), p.29
¹¹⁷ Nohejl et al. (1992)
¹¹⁸ Nohejl et al. (1992), p.195
¹¹⁹ Nohejl et al. (1992), p.196
¹²⁰ Nohejl et al. (1992)
¹²¹ Nohejl et al. (1992), p. 196
¹²² Child Welfare Services Stakeholders Group (2003)
¹²³ Rycus & Hughes (2003)
¹²⁴ Wood (1997)
¹²⁵ Lyons et al. (1996)
¹²⁶ Rycus & Hughes (2003); Wald & Woolverton (1990)
¹²⁷ Rycus & Hughes (2003), p.11
¹²⁸ Doueck, English, DePanfilis & Moote (1993); Johnson (2004)
¹²⁹ Ereth et al. (2003) p.3
¹³⁰ Ereth et al. (2003); Fuller et al. (2001); Munro (1999)
¹³¹ Munro (1999) p.756
¹³² Wald & Woolverton (1990)
¹³³ English & Aubin (1990); Milner (1994); Nasuti (1998)
¹³⁴ Morton (2004b)
¹³⁵ English & Aubin (1990)

References

- Albers, M. & Roditti, M. (2004). *Management values and decision-making instruments: Nuts and bolts of risk assessment in California*. Bay Area Academy, Berkeley, CA.
- Altman, D.G. & Royston, P.R. (2000). What do we mean by validating a prognostic model? *Statistics in Medicine*, 19, 453-473.
- American Federation of State County and Municipal Employees (1999). *Liability and child welfare workers*. Available online at: <http://www.afscne.org/publications/child/cww99205.htm>.
- 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). Risk assessment in child protective services: Consensus and actuarial instrument reliability. *Child Welfare*, 78(6), 723-748.
- Camasso, M. J. & Jagannathan, R. (2000). Instrumenting the reliability and predictive validity of risk assessment in child protective services. *Children and Youth Services Review*, 22(11/12), 873-896.

-
- Camasso, M. J. & Jagannathan, R. (1995). 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.
- Child Welfare Services Stakeholders Group (2003). *CWS Redesign: The future of California's child welfare services - final report*. State of California Health and Human Services Agency, Department of Social Services.
- Cicchinelli, L. (1990). Risk assessment: Expectations, benefits and realities. *Fourth National Roundtable on CPS Risk Assessment*. San Francisco, CA: American Public Welfare Association.
- Costello, T. (1995). Why is it so hard to implement risk assessment? *Ninth National Roundtable on CPS Risk Assessment*. San Francisco, CA: American Public Welfare Association.
- Curran, T. F. (1995). Legal issues in the use of CPS risk assessment instruments. *From the APSAC Advisor, 8*(4), 1-4.
- Davidson, H. (1991). Risk assessment and the law: Unsettled issues. *Fifth National Roundtable on CPS Risk Assessment*. San Francisco, CA: American Public Welfare Association.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision-making. *American Psychologist, 34*, 571-582.
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York, NY: The Free Press.
- DePanfilis, D. (1996). Implementing child mistreatment risk assessment systems: Lessons from theory. *Administration in Social Work, 20*(2), 41-59.
- Doueck, H. J., English, D. J., DePanfalis, D. & Moote, G. T. (1993). Decision-making in child protective services: A comparison of selected risk assessment systems. *Child Welfare, 72*(5), 441-452.
- Doueck, H. J., Levine, M. & Bronson, D. E. (1993). Risk assessment in child welfare services: An evaluation of the Child at Risk Field System. *Journal of Interpersonal Violence, 8*(4), 446-467.
- Downing J. D. Wells S. J. & Fluke J. (1990). Gatekeeping in child protective services: A survey of screening policies. *Child Welfare, 69* (4), 357-369.
- English, D. J. (1999). Evaluation and risk assessment of child neglect in public child protection services. In H.Dubowitz, (Ed.), *Neglected children: Research, practice, and policy* (pp. 191-210). Thousand Oaks, CA: Sage Publications, Inc.
- English, D. & Aubin, S. W. (1990). Outcomes for screened-out and low-risk cases within

-
- a child protective services risk assessment system. *Fourth National Roundtable on CPS Risk Assessment* (pp.41-58). San Francisco, CA: American Humane Association.
- English, D. J. & Graham, J. C. (2000). An examination of relationship between children's protective services social worker assessment of risk and independent LONGSCAN measures of risk constructs. *Children and Youth Services Review*, 22(11-12), 897-933.
- English, D. J., Marshall, D., Brummel, S. & Orme, M. (1995). A preliminary examination of similarities and differences in the assessment of risk for different ethnic groups. *Ninth National Roundtable on CPS Risk Assessment*, pp.195-218. San Francisco, CA: American Human Association.
- English, D. J., Marshall, D., Brummel, S. & Orme, M. (1999). Characteristics of repeated referrals to child protective services in Washington state. *Child Maltreatment* 4(4), 297-307.
- Ereth, J., Johnson, K. & Wagner, D. (2003). *New Mexico Children, Youth and Families Department Foster Provider Risk Assessment Study*. Madison, WI: Children's Research Center.
- Fluke, J., Edwards, M., Bussey, M., Wells, S. & Johnson, W. (2001). Reducing recurrence in child protective services: Impact of a targeted safety protocol. *Child Maltreatment*, 6(3), 207-218.
- Fluke, J., Wells, S., England, P., Walsh, W., English, D., Gamble, T. & Woods, L. (1993). Evaluation of the Pennsylvania approach to risk assessment. *Seventh National Roundtable on CPS Risk Assessment* (pp.116-170). San Francisco, CA: American Humane Association.
- Fuller, T. L., Wells, S. J. & Cotton, E. E. (2001). Predictors of maltreatment recurrence at two milestones in the life of a case. *Children and Youth Services Review*, 23(1), 49-78.
- Galasso, L. B. (2001). *Toward the prevention of child maltreatment through risk assessment: Evaluation of an ecological, prospective instrument of risk for child abuse potential*. Dissertation, Michigan State University Department of Psychology.
- Garnier, P. & Nieto, M. (2002). *Illinois Child Endangerment Risk Assessment Protocol evaluation: Impact on short-term recurrence rates – year six+*. Children and Family Research Center, University of Illinois at Urbana-Champaign.
- 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.
- Hollinshead D. & Fluke J. (2000). What works in safety and risk assessment for child protective services. In M. Kluger G. Alexander and P. Curtis (Eds.), *What Works in Child Welfare* (pp. 67-74). Washington D.C.: CWLA Press.
- Johnson, W. (2004). *Effectiveness of California's child welfare Structured Decision-making*

-
- Instrument (SDM): A prospective study of the validity of the California Risk Assessment.*
Unpublished report.
- Johnson, W. (2005). Effects of a research-based risk assessment on racial/ethnic disproportionality in service provision decisions. In D. Derezotes, J. Poertner & M. F. Testa (Eds.), *Race Matters in Child Welfare: The Overrepresentation of African American Children in the System*. Washington, D.C.: CWLA.
- Kahneman, D. & Tversky, A. (1982). On the psychology of prediction. in D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgment under Uncertainty: Heuristics and Biases* (pp. 50-98). New York, NY: Cambridge University Press.
- Kanani, K., Regehr, C. & Bernstein, M.M. (2002). Liability considerations in child welfare: Lessons from Canada. *Child Abuse and Neglect*, 26, 1029-1043.
- Kolko, D. J. (1998). CPS operations and risk assessment in child abuse cases receiving services: Initial findings from the Pittsburgh service delivery study. *Child Maltreatment*, 3(3), 262-275.
- Leschied, A. W., Choido, D., Whitehead, P. C., Hurley, D., & Marshall, L. (2003). The empirical basis of risk assessment in child welfare: The accuracy of risk assessment and clinical judgment. *Child Welfare*, 82(5), 527-540.
- Loman L. A. & Siegel G. L. (2004). *An evaluation of the Minnesota SDM Family Risk Assessment*. St. Louis, MO: Institute for Applied Research. Available from: <http://www.iarstl.org>.
- Lyle, C. G. & Graham, E. (2000). Looks can be deceiving: Using a risk assessment instrument to evaluate the outcomes of child protection services. *Children and Youth Services Review*, 22(11/12), 935-949.
- 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.
- McDonald, T.P. & Marks, J. (1991). A review of risk factors assessed in child protective services. *Social Services Review*, 65, 112-132.
- Milner, J. S. (1994). Assessing physical child abuse risk: The Child Abuse Potential Inventory. *Clinical Psychology Review*, 14(6), 547-583.
- Morton, T.D. (1999). The increasing colorization of America's child welfare system: The overrepresentation of African American children. *Policy & Practice*, 57(4), 23-30.
- Morton, T. D. (April, 2004a). Where assessment fails. *Commentary*. Duluth, GA: Child Welfare Institute.
- Morton, T. D. (Sept, 2004b) *Caretaker and child interactions in child maltreatment*. Duluth, GA: Child Welfare Institute.

-
- Morton, T. D. & Salovitz, B. (n.d.) *Evolving a theoretical instrument of child safety in maltreating families*. Unpublished manuscript.
- Munro, E. (1999). Common errors of reasoning in child protection work. *Child Abuse and Neglect*, 23(8), 745-758.
- Munro, E. (2004). A simpler way to understand the results of risk assessment instruments. *Children and Youth Services Review*, 26(9), 873-883.
- Munro, E. & Runggay, J. (2000). Role of risk assessment in reducing homicides by people with mental illness. *British Journal of Psychiatry*, 176, 116-120.
- Nasuti, J. P. (1998). Risk assessment in child protective services: Challenges in measuring child well-being. *Journal of Family Social Work*, 3(1), 55-70.
- Nasuti, J. P & Pecora, P. J. (1993). Risk assessment scales in child protection: A test of the internal consistency and interrater reliability of one stateside system. *Social Work Research and Abstracts*, 29(2), 28-33.
- Nieto, M. & Garneir, P. (2001). *Illinois Child Endangerment Risk Assessment Protocol evaluation: Impact on short-term recurrence rates – year five*. Children and Family Research Center, University of Illinois at Urbana-Champaign.
- Nohejl C. Doueck H. J. & Levine M. (1992). Risk assessment implementation and legal liability in CPS practice. *Law & Policy*, 14(2 & 3), 185-203.
- Repetosky C. & Bailey D. (1988). The risk "fit": Integrating risk assessment with case management and practice decisions in child protection and child welfare services. *Second National Roundtable on CPS Risk Assessment*. San Francisco, CA: American Public Welfare Association.
- Rycus, J. S. & Hughes, R. C. (2003). *Issues in risk assessment in child protective services: Policy white paper*. Columbus, OGH: North American Resource Center for Child Welfare, Center for Child Welfare Policy.
- Scheurman, J., Rossi, P. H. & Budde, S. (1999). Decisions on placement and family preservation. *Evaluation Review*, 23(6), 599-618.
- Schwalbe, C. (2004). Re-visioning risk assessment for human service decision making. *Children and Youth Services Review*, 26(6), 561-576.
- Sheets D. (1992). How Texas Got S.M.A.R.T.: A description of the rapid application design and development process used by Texas to design and implement a statewide risk assessment system. *Sixth National Roundtable on CPS Risk Assessment*. San Francisco, CA: American Public Welfare Association.

-
- Squadrito, E., Neuenfeldt, D., & Fluke, J. (1994). Findings of Rhode Island's two-year research and development efforts. *Eighth National Roundtable on CPS Risk Assessment*. San Francisco, CA: American Public Welfare Association.
- Thompson L., Brown J. & Pecora P. J. (1989). Implementation of risk assessment systems: Training and quality assurance issues. *Third National Roundtable on CPS Risk Assessment*. San Francisco, CA: American Public Welfare Association.
- Tumlin, K. C. & Geen, R. (2000). The decision to investigate: understanding state child welfare screening policies and practices. *No.A-38 in New Federalism Issues and Options for States, Series A*. Washington, D.C.: The Urban Institute.
- Wagner, D., Hull, S. & Luttrell, J. (1995). Structured decision making in Michigan. *Ninth National Roundtable on CPS Risk Assessment*, pp. 167-191. San Francisco, CA: American Humane Association.
- Wald, M.S. & Woolverton, M. (1990). Risk assessment: The emperor's new clothes? *Child Welfare*, 69(6), 483-511.
- Wood, J.M. (1997). Risk predictors for re-abuse or re-neglect in a predominantly Hispanic population. *Child Abuse and Neglect*, 21(4), 379-389.
- Zuravin, S. J., Orme, J. G. & Hegar, R. L. (1995). Disposition of child physical abuse reports: Review of the literature and test of a predictive instrument. *Children and Youth Services Review* 17(4), 547-566.

APPENDIX A: BASSC SEARCH PROTOCOL

Search Terms

(Child maltreatment or child abuse or child neglect) AND (risk assessment or safety assessment)

Academic databases for books and articles

Pathfinder or Melvyl

ArticlesFirst

ERIC

Expanded Academic ASAP

Family and Society Studies Worldwide

PAIS International

PsychInfo

Social Science Citation Index

Social Services Abstracts

Social Work Abstracts

Sociological Abstracts

Systematic Reviews

Campbell Collaboration

Cochrane Library

Research Institutes

Brookings Institution

CASRC (San Diego)

Chapin Hall

GAO

Manpower Demonstration Research Corporation

Mathmatica Policy Research, Inc.

National Academy of Sciences

RAND

Urban Institute