Available online at www.sciencedirect.com
SCIENCE DIRECT®
Children and Youth Services Review
26 (2004) 561–576
Children and Youth Services Review
www.elsevier.com/locate/childyouth

Re-visioning risk assessment for human service decision making

Craig Schwalbe

University of North Carolina, School of Social Work, 301 Pittsboro, Chapel Hill, NC 27599-3550, USA

Available Online 23 March 2004

Abstract

This article examines the usefulness of actuarial risk assessment for high-stakes decision making in child welfare, mental health, criminal justice and juvenile justice. Review of the literature affirms the potential benefits of risk assessment instruments for decision making by human service professionals. However, research also hints at the underutilization of risk assessment in practice. Although a number of explanations may account for this, the needs of decision makers in the real world of day to day practice has received little attention in the literature. This article identifies insights from the Recognition primed decision making theory (RPD) that promise to strengthen the utility of actuarial risk assessment instruments. It argues that an actuarial risk assessment instrument, based on appropriate causal theory, would have a greater likelihood of utilization as compared to the a-theoretical instruments that predominate in current structured decision making systems.

Keywords: Risk assessment; Structured decision making; Naturalistic decision making; Statistical decision making

1. Introduction

Human services professionals make a continuous stream of decisions that influence the lives of their clients. A small number of their decisions may be characterized as ‘high-stakes’. These are decisions that weigh the use of powerful interventions aimed at the prevention of physical harm or criminal behavior. Examples include foster care placement in the child welfare system (English & Pecora, 1994), detention in the juvenile justice system (Hoge, 2002), civil commitment in the

E-mail address: cschwalb@email.unc.edu (C. Schwalbe).

0190-7409/04/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.childyouth.2004.02.011

mental health system (Monahan et al., 2001) and parole decisions in the criminal justice system (Bonta, 2002). For these and other high-stakes decisions, human services professionals recommend, and in some cases decide, the level of intervention that balances self-determination against the public mandate to reduce the likelihood of violence and injury.

The consequences of error can be fatal. Homicides committed by mentally ill perpetrators and child fatalities due to abuse or neglect capture the attention of the media and public. Commissions and studies are convened to discover what when wrong in the decision-making process and how can it be improved (Munro, 1996; Firestein, 1999; Munro & Rumgay, 2000). The findings of these inquiries focus attention to a fundamental characteristic of high-stakes decisions in the human services: given the magnitude of the issues involved, the consequences of decision-making errors can be grave.

To ensure accurate decision making, public social service systems have developed structured decision-making systems for human service professionals. These approaches guide decision making contingent upon the results of a variety of assessments. One procedure often included is actuarial risk assessment. These survey-like instruments measure risk factors known to relate to an outcome of interest and are subjected to empirical tests to document predictive validity. Often they are scored and summarized using labels such as 'low-risk', 'medium-risk' and 'high-risk' to denote the likelihood of a future event. They take advantage of the well documented superior predictive ability of actuarial methods of prediction (Grove & Meehl, 1996). Actuarial risk assessment instruments are intended to predict negative outcomes for clients, thereby identifying those high risk individuals who are in need of exceptional interventions as described above.

The purpose of this article is to critique the current state of risk assessment technology from the perspective of its utilization in practice. While this subject has been previously addressed from the standpoint of policy and administration (De-Panfilis, 1996), it has not been evaluated in terms of the theories of decision making. The premise of this article is that actuarial risk assessment instruments will influence human services decision making when they target specific decision-making processes as they occur in the day to day practice of human service professionals.

The article begins with a brief review of empirical research identifying weaknesses in human service professional decision making and the scant literature measuring the utilization of risk assessment instruments. It continues with a review and critique of the current theoretical framework for risk assessment, the Rational Choice Model. It concludes by recommending insights from the Naturalistic decision making theories that might strengthen the influence of risk assessment on human services decision making.

2. Research on human services decision making

Clients and consumers of human services require accurate decisions to ensure fair and effective treatment. However, empirical research into the decisions made by human service professionals shows that decision making is often out of step with these ideals. Specifically, research identifies two problems for which risk assessment instruments have been designed to neutralize: inconsistency across decision makers and the weak ability of human service professionals to predict important outcomes of interest.

Research shows that human service professional decisions are frequently inconsistent across decision makers. Using case vignettes, studies of child welfare workers (Rossi, Schuerman & Budde, 1999; Schuerman, Rossi & Budde, 1999) and mental health professionals (Hendryx & Rohland, 1997), for example, find that there is little agreement among them regarding the need for placement in foster care or inpatient hospitalization. Possible explanations for the lack of consistency include findings that social workers are more likely to justify treatment decisions using appeals to values and agency policy than to empirical research (Rosen, 1994) and make high-stakes decisions based on perceptions of the social worker–client relationship instead of client characteristics (Holland, 2000). These studies provide evidence that the likelihood of highly invasive interventions is at least somewhat dependent upon the fortunes of chance, which determines case assignment.

Research also confirms that human services professionals are inadequate predictors of future behavior for their clients (Skeem, Mulvey & Lidz, 2000). Their predictions are compromised by contextual and individual circumstances that can change on day to day basis (Dawes, Faust & Meehl, 1989; Gambrill & Shlonsky, 2000). In contrast, the predictive validity of the actuarial method has growing empirical support. Reviewing 135 studies conducted over a 60-year time span, Grove and Meehl (1996) found only eight studies in which predictions using 'clinical judgment' were more accurate than actuarial instruments. A well-constructed actuarial instrument appears to predict future behavior with greater accuracy than is possible by clinical judgment alone.

Researchers and policy-makers designed structured risk assessment to reduce errors associated with both of these problems. It does this by standardizing information collected for decision making and by integrating case information into accurate classifications using actuarial methods (Ruscio, 1998; Baird, Wagner, Healy & Johnson, 1999; Baird & Wagner, 2000). When used properly, risk assessment promises to ensure that high-stakes decisions are made equitably based on the presenting risk profiles of clients served by human services professionals.

3. Research on risk assessment utilization

There are few empirical evaluations of risk assessments' actual influence on the decision making of human services professionals. The available literature reports findings confined to specific social service agencies leaving results vulnerable to the variability of local agency culture. This limitation restricts our ability to make broad generalizations. Nevertheless, studies identified to date shed light on three of the reasons that risk assessment instruments may be underutilized even in settings in which they are mandated by agency rule.

First, risk assessment instruments may be subverted for unintended purposes by line staff. A study of risk assessment in a child welfare agency illustrates this possibility. The setting for the study was an urban child welfare agency in Minnesota (Lyle & Graham, 2000). Statewide policy required that child welfare workers complete a risk assessment instrument, modeled after the Illinois CANTS 17B, when a case was transferred from the intake unit to the ongoing child protective services unit and again when a case was closed. The CANTS 17B measures risk in 13 factors. It requires that assessing workers rate each item on a five point likert type scale from no risk to high risk. Researchers reported outcome data showing non-sensical reductions in some risk item scores. For instance, the mean score for ‘previous history of abuse or neglect’ showed a statistically significant reduction from 2.04 to 1.48. Upon further investigation, researchers learned that some intake workers inflated risk scores to ensure continued services for select families. In this case, risk assessment became a bureaucratic tool of some workers to achieve goals they felt over-rode the stated purpose of the instrument, a possibility left open by a design feature not unusual among risk assessment instruments that permits assessing workers latitude in how they judge the severity of some individual risk items.

Second, risk assessment findings may be ignored due to doubt about the psychometric properties of the instrument. This occurred in the Arizona juvenile justice system (Krysik & LeCroy, 2002). Court counselors felt that their risk assessment instrument, designed to predict recidivism, was difficult to complete due to the lack of reliable information. Of probation officers surveyed ($N = 56$), 64% reported that they did not use risk assessment findings in their court recommendations. 84% felt that they made more accurate predictions than the risk assessment instrument. As a consequence, up to 68% of risk assessment instruments were incomplete in some counties. It appears that juvenile justice risk assessment instruments were disregarded while decisions were made using alternate sources of information.

Third, risk assessment instruments may simply be ignored in favor of clinical testimony. A study of the decision-making practices of a psychiatric hospital found this effect (Hilton & Simmons, 2001). The study examined the transfers and discharges of psychiatric patients from a maximum security hospital in Ontario, Canada. In this setting, decision makers had access to the opinions of clinical teams working with patients as well as results of the Violence Risk Appraisal Guide (VRAG), an instrument validated for use on adult psychiatric and criminal populations (Harris, Rice & Quinsey, 1993). Hilton and Simmons examined 187 transfer and discharge decisions. They found that decisions were associated with clinician testimony but were not related to the results of the VRAG. Moreover, they discovered that clinician testimony was also uncorrelated with VRAG scores. Instead, clinicians considered information that was not included on the VRAG and did not integrate VRAG findings into their reports. In this hospital, empirically valid measures of violence risk were essentially ignored by decision makers.

While these studies do not provide conclusive evidence of the wide-spread disregard of risk assessment, they do raise important issues. The finding that organizational factors influence the utilization of risk assessment (Lyle & Graham, 2000) points to the need for effective management practices and training to encourage their use. The finding that human services professionals are attuned to the psychometric quality of risk assessment instruments (Krysik & LeCroy, 2002) reinforces the need for instruments that meet a high empirical standard. Overcoming these issues is critical to the success of structured risk assessment. Fortunately, resources that guide administrators (DePanfilis, 1996) and that ensure empirically sound instruments (Douglas, Ogloff, Nicholls & Grant, 1999; Jung & Rawana, 1999; Belfrage, Fransson & Strand, 2000; Camasso & Jagannathan, 2000; Monahan et al., 2000; Risler, Sutphen & Shields, 2000; Jones, Harris, Fader & Grubstein, 2001) are increasingly available.

The finding that some human services professionals utilize their own judgment to the exclusion of structured risk assessment (Hilton & Simmons, 2001) may not be surprising. What is surprising, however, is how little attention has been devoted to this phenomenon. One source for understanding this problem may be found through a review of the theory of decision making that guides risk assessment design and development.

4. Rational choice model

Risk assessment is currently conceptualized within the rational choice model of decision making. This theory, also referred to here as 'Statistical decision theory', is rooted in economics and probability theory (Lindley, 1985; Pratt, Raiffa & Schlaifer, 1995). It is a prescriptive theory designed to offer procedures for rational decision making given an uncertain future. Conceived in this paradigm, risk assessment is intended to reduce the uncertainty involved in high stakes human services decisions.

One procedure for making rational decisions is called Maximizing Subjective Expectation of Utilities (SEU). This procedure for decision making is commonly taught in economics and business as decision analysis. Decision analysts identify four components to all decision problems regardless of setting: outcomes, options, utilities and probabilities. Outcomes are the possible results following the implementation of a decision. For example, child welfare decision makers might, following substantiation of child maltreatment, define outcomes in terms of child safety and repeated child maltreatment. Options would include the array of intervention choices ranging from case closure to termination of parental rights. 'Utilities' are the preferences, or values, that the child welfare system attaches to some anticipated consequences over others. Often, utilities are defined in financial terms, such as the cost of intervention, although Lindley (1985) demonstrated procedures for quantifying subjective preferences such as child safety and client self determination. Finally, probabilities are estimated for each outcome given each intervention choice. In formal decision analysis, utilities and probabilities are quantified and integrated mathematically to recommend an optimal decision; that is, a recommended decision that provides the greatest likelihood of achieving the most favorable result among an array of choices. Lindley (1985) outlined a process of SEU decision making illustrated below. Consider a simplified example wherein a child welfare worker maximizes SEU to determine an appropriate intervention level for a family following

Table 1
Determine outcomes, choices, and consequences

Options	Outcomes	
	Absence of child maltreatment	Repeat child maltreatment
Close case	Safe child Low intrusion $	Injured child Low intrusion $
Outpatient services	Safe child Moderate intrusion $$	Injured child Moderate intrusion $$
Foster care placement	Safe child Possible psychological injury Extreme intrusion $$$	Injured child Possible psychological injury Extreme intrusion $$$

Note: '$' reflects the relative cost of intervention.

substantiation of child maltreatment. The decision-making process will unfold in three stages:

1. Define options and outcomes for a given decision problem. For example, after substantiating child maltreatment, a child welfare worker may face choices outlined in Table 1. In this example, a child welfare worker has three choices-case closure, outpatient services, and out of home placement—and faces two possible outcomes—absence of child maltreatment or repeated child maltreatment. In addition, each outcome has associated consequences depending upon the intervention selected. For example, consequences associated with these outcomes include safety vs. injury, degree of intrusion into family life and cost of intervention.
2. Assign numerical values to consequences to indicate their relative preference to the decision maker. These values are called 'utilities'. Begin by arraying consequences shown above into a list from the most desirable to the least desirable as shown in Table 2. Next, select any consequence on the list and assume it is a 'sure thing'. Then pose the following hypothetical gamble: 'at what probability will you be willing to give up the 'sure thing' for the chance at obtaining the best consequence in the list'. In the following example, the child welfare worker will give up 'injured child, moderate intrusion', $$ for a 20% chance at achieving the most desirable consequence but required an 85% chance before risking 'safe child, moderate intrusion', $$.
3. Finally, estimate probabilities for each outcome given each option using the best information available. Calculate subjective utilities using the equation shown in Table 3 and select the option providing the largest utility. In this example, outpatient services are selected as this option gives the greatest probability of achieving the most favorable result given the probability and utility assumptions outlined above.

When used, SEU promises at least three benefits prized by administrators and policy-makers. First, decisions will be consistent. Because SEU uses mathematical formulas, inconsistency should disappear when decisions are based upon precise

Table 2
Establish numerical values for preferences

Consequence	Utility values
Safe child Low intrusion \$	1.0
Safe child Moderate intrusion \$\$	0.85
Safe child Possible psychological injury Extreme intrusion \$\$\$	0.60
Injured child Low intrusion \$	0.30
Injured child Moderate intrusion \$\$	0.20
Injured child Possible psychological injury Extreme intrusion \$\$\$	0.10

Note: '$' reflects the relative cost of intervention.

calculations using the same information (Ruscio, 1998). Second, decisions will be rational. That is, they will conform to the laws of probability (Lindley, 1985; Pratt et al., 1995). Third, the decision-making process will be transparent; it will be open to critique. Admittedly, the assignment of utilities, and to some extent probabilities, is subjective. However, the SEU process requires clear specification of all elements of the decision-making process resulting in greater transparency; a key feature of decision making in settings marked by high levels of administrative and legal accountability (Hammond, 1996, 2000; Payne & Bettman, 2001). Because of these benefits, Statistical decision making is seen as the standard against which the decisions of human services professionals should be judged.

Most observers of human decision making will recognize that SEU is seldom explicitly used in real world settings. In fact, researchers in the Heuristics and Biases tradition demonstrated that most decision making does not approach the standards of the Statistical decision theory. With convincing regularity, laboratory

Table 3
Calculate an optimal decision

Options	Absence of child maltreatment		Repeat child maltreatment		Subjective utility
	Probability	Utility	Probability	Utility	
Do nothing	0.50	1.0	0.50	0.30	0.650
Outpatient services	0.75	0.85	0.25	0.20	0.688
Foster care placement	0.85	0.60	0.15	0.10	0.525

Note: Subjective utility $= P(\text{absence}) \times \text{utility} + P(\text{repeat}) \times \text{utility}$.

researchers examining people's choice behavior find that decision making systematically violates the laws of probability (Tversky & Kahneman, 1974, 2002). This occurs because of biases associated with the mental shortcuts people use to process information. These mental shortcuts have been called 'heuristics'. Tversky and Kahneman identified three heuristics in 1974. To date, at least 30 have been described in the literature (Croskerry, 2002).

Heuristics contribute to decision-making error by leading people to ignore key information about decision problems. For instance, it is common for people to discount empirical information in favor of strong intuitive impressions, to resist changing first impressions despite evidence favoring alternative explanations and to dismiss historical patterns, which should temper prediction confidence (Griffin & Tversky, 1992; Buehler, Griffin & Ross, 2002; Croskerry, 2002). Heuristics direct people's attention to the particularities of unique cases and away from probabilistic patterns found in the reference classes to which cases belong. For instance, people commonly fear violence and crime from mentally ill patients when in fact violence from this group occurs at a low frequency compared to the larger population (Stuart & Arboleda-Florez, 2001). Misunderstandings as these, divorced from empirical data, contribute to systematic errors in prediction and their consequent deleterious effects on decision making.

The findings of the Hueristics and Biases tradition have dominated the field of decision-aiding since Tversky and Kahneman's (1974) article. Typical advice to human services decision makers is directed toward the goal of reducing cognitive biases. For example, human services decision makers are advised to de-bias decision making through the use of 'cognitive forcing strategies' (Croskerry, 2003), to actively seek out alternative explanations and disconfirming evidence (Gambrill, 1990; Sheppard, 1995), to specify more options for decision problems (Dawes, 2001) and to consider the base rates of potential outcomes given a set of population characteristics (Murdach, 1994; Gambrill, 2003). The purpose of this advice is to conform the decision-making process to the Rational Choice Model.

Structured decision making systems in fields like child welfare and juvenile justice, which incorporate actuarial risk assessment strategies, approximates the SEU procedure. The benefit of risk assessment for these systems rests with the scientific method's capacity to identify markers of future behaviors and to weight these in an optimal formula useful for prediction. Their utility is buoyed by research in actuarial methods of prediction that demonstrate the superiority of this method over strategies based in clinical judgment and intuition (Dawes et al., 1989; Grove & Meehl, 1996). Their reliance on statistical computations promises to increase the reliability and objectivity of human services decision making by standardizing the information upon which decision makers use to select optimal choices (Ruscio, 1998; Baird et al., 1999; Gambrill & Shlonsky, 2000). Accordingly, risk assessment instruments should become valuable tools for improving the performance of human services decision makers.

Yet research cited above raises some doubt that risk assessment instruments have been widely embraced by human services professionals. Some possible reasons for continued resistance to actuarial predictions has been described in the literature

(Meehl, 1986). They include a lack of knowledge of cognitive biases or statistical remedies for them, a persistent belief in the efficacy of clinical judgment based upon experience and training, overconfidence, the dehumanizing feel of statistical equations, fear of computers, and an ethical belief against the standardization that accompanies the statistical method.

Other issues likely interfere with risk assessment utilization as well. One of the vexing problems related to human services decision making relates to the difficulty inherent in the prediction of future behavior. Hammond (1996, 2000) argues in the Cognitive Continuum Theory that cognitive strategies for decision making lie on a continuum anchored by intuition on one pole and by rational analysis on the other. He further argues that characteristics of the decision-making context require cognitive strategies that are more or less intuitive or analytical. Specifically, the availability of highly discriminating cues, called 'multiple fallible indicators', determines the most suitable type of decision making strategy. When indicators exist whose discriminability is well documented, rational-choice strategies may be employed. However, when indicators have poor empirical discriminability, strategies resembling intuition must be employed. Hammond notes that decisions involving the physical sciences are often amenable to rational-choice strategies while decisions involving the prediction of human behavior require greater use of intuition.

Human services professionals will immediately recognize the implication of Hammond's (1996, 2000) argument. To the extent that risk assessment can increase the discriminability of the 'multiple fallible indicators' that confront human services professionals engaged in high stakes decision making, it should enable these decision makers to engage in more rational-choice strategies. However, the predictability of human behavior in problems which are of concern to human services professionals has not been well established. This is particularly true of phenomena which occur at low base rates (Fuller & Cowan, 1999; Jones et al., 2001) and for clients who occupy a not-so-obvious middle-risk category for which human service professionals may be less confident about their predictions (McNiel, Sandberg & Binder, 1998). Thus, there may be natural limits to the extent that human service decision making can be aligned with the Rational Choice Model.

But the chief reason that may interfere with risk assessment utilization lies in another of the key findings of the Heuristics and Biases tradition: that people are influenced to a greater extant by a 'good story' than by statistical information (Dawes, 1999, 2001). People use 'stories' to explain the past and to project the present into the future. External comparison to referent classes, the hallmark of scientific validation, is not vital to the persuasiveness of 'good stories'. Instead, story based hypotheses are judged according to the standard of plausibility (is it reasonable that the sequence of events encompassed in the explanation could occur as stated?) and internal coherence (is the story logical?) (Tversky & Kahneman, 2002). The weakness of 'good stories' is their vulnerability to biases through the influence of heuristics as outlined above (Dawes, 2001). This weakness notwithstanding, Dawes (1999) argued that decision makers often ignore probabilistic empirical information in the absence of a 'good story' to explain it. For risk assessment instruments, the persuasiveness of story based hypotheses implies that un-contextualized empirical risk assessment findings will receive little attention from decision makers as compared to vivid presentations of ideographic information.

5. Naturalistic decision making

Naturalistic theories of decision making are an alternative approach to understanding the decision-making process. This class of decision-making theory originated from observations of decision making by experts in a wide range of fields in real world settings. They were developed using a grounded theory approach in contrast to the deductive approach that undergirds the Rational Choice Model. These theories offer insights into decision making processes that can be strengthened when supplemented by sensitively designed risk assessment instruments.

The origin of Naturalistic decision making (NDM) as a subfield of decision research can be traced to a conference sponsored by the Army Research Institute held in 1989 (Klein, Orasanu, Calderwood & Zsambok, 1993; Lipshitz, Klein, Orasanu & Salas, 2001a). In this conference, the common elements of a new class of decision-theories were identified based upon the observation of decisions made in real world settings. Participants in this conference agreed that the assumptions of the Rational Choice Model of decision making, developed through laboratory study, are incompatible with the contexts in which real world decisions are made. Where the laboratory studies of Heuristics and Biases researchers tightly controlled problem definitions, choices and expertise, NDM researchers observed that decision makers in the real world cope with ill-structured problems in dynamic environments characterized by competing goals, feedback loops, time pressure and high stakes (Orasanu & Connolly, 1993). The common elements of NDM theories included: a focus on situation assessment instead of deliberation over choices, use of mental imagery instead of calculation, context dependence instead of laboratory generalizability, dynamic processes instead of discrete decision events, and description-based prescription instead of normative prescription (Lipshitz, 1993). These features continue to characterize the later developments of NDM although research since that first conference has tended to emphasize the way that expertise is used by decision makers (Lipshitz, Klein, Orasanu & Salas, 2001b).

The theory of Recognition primed decision making (RPD) is one NDM theory describing the strategies decision makers use to make decisions in real world settings (Klein, 1998). It has been developed through systematic observations of expert decision makers in diverse settings such as firefighting, nursing, and naval command centers. RPD was assembled based upon observations of difficult 'high-stakes' decisions.

The heart of RPD is the recognition of patterns or prototypes (Klein, 1998). Pattern recognition begins with an assessment of the decision problem and concludes with a comparison of the problem to known prototypes developed through training and experience. Once recognized, decision makers usually consider only one course of action that 'makes sense' given the recognized prototype. The process is cyclical in that information provided through continuous feedback loops confirm recognized prototypes or are occasion to revise initial impressions and subsequently alter plans.

RPD researchers found that decision makers spend most of their energies directed toward the pattern matching process. Conversely, these theorists also found that expert decision makers spend little to any time examining alternative courses of action. Often, experts simply know what to do once a pattern has been recognized.

Not all situations conform to known prototypes, however. In novel cases, a single course of action may not be obvious to even the most expert of decision makers. According to RPD, expert decision makers confronted by novel decision problems engage in mental simulation to evaluate options (Klein, 1998). Unlike the concurrent comparison of multiple options prescribed by the Rational Choice Model, RPD describes how experts evaluate a series of options one at a time. Mental simulations are used to imagine the consequences of a course of action given the patterns observed in the first stage of the process. Problems identified through the simulation are occasion to either modify the strategy or to abandon it for a new one. Typically, decision makers take the first option that passes the standards imposed by the goals of the decision problem; a process described in the decision-making literature as 'satisficing' (Gigerenzer & Todd, 1999; Gigerenzer & Selten, 2001).

Occasionally, the processes of pattern matching and mental simulation can break down or contribute to decision making errors. When experts err, it is most often because either the process of pattern matching leads to inaccurate conclusions, as when a pattern is incorrectly specified, or when the process of mental simulation fails to correct a poorly conceived plan, as when favored options are 'fixed beyond repair'. Recommended corrections include increased training for decision makers in prototypes to increase pattern recognition proficiency and the use of the 'pre-mortem strategy' whereby decision makers assume that a favored plan has failed and must explain why it has failed (Klein, 1997, 1998). The strength of these recommendations is that they require no substantial modifications to the decision-making process itself. Because of this, they are more likely to be adopted than are strategies that require decision makers to use new or novel approaches to decision making.

6. NDM and actuarial risk assessment

The NDM theories identify processes that expert decision makers use to make high stakes decisions. These theories expose weaknesses in the decision-making process that could be strengthened by the introduction of well designed actuarial risk assessment instruments. Indeed, actuarial instruments that simultaneously target the pattern matching and mental simulation processes might be best positioned to achieve the ultimate goal of risk assessment—to improve the high-stakes decision making of human service professionals.

Well designed actuarial risk assessment instruments contribute to the pattern matching process. They help novice and expert human services professionals alike to identify combinations of risk factors that portend future problems and risks. In human services settings, patterns of risk that predict poor outcomes include cumulative risk and the presence of keystone risks (Fraser, 2004). The advantages of actuarial instruments are that they weight risk patterns consistently and that they provide decision makers with objective probability estimates. Therefore, to ensure useful risk assessment instruments, continuous attention must be paid to their predictive validity. Indeed, the sheer volume of validation studies conducted to date reflect acceptance of this ideal (Douglas et al., 1999; Jung & Rawana, 1999; Belfrage et al., 2000; Camasso & Jagannathan, 2000; Monahan et al., 2000; Risler et al., 2000; Jones et al., 2001).

Continuing challenges for the empirical validity of risk assessment remain, however (Gambrill & Shlonsky, 2000). These include refining definitions of important outcome variables, attending to the reliability and validity of specific items incorporated into instruments so as to reduce measurement errors, and balancing the sensitivity (probability of identifying true positives; high sensitivity leads to increases in false positives) and specificity (probability to identifying true negatives; high specificity leads to increases in false negatives) of risk assessment scales. Attention to these issues will maximize the contribution of risk assessment instruments toward helping decision makers identify patterns with high levels of expertise.

Well designed actuarial risk assessment instruments will also contribute to the mental simulations of human services decision makers. For this to happen, risk assessment findings need to be presented in such a way as to be easily incorporated into the 'stories' that are told to explain predictions and to justify interventions. While some may argue the sole and limited purpose of risk assessment is for empirical prediction, the limitation of this view is that statistical information in the absence of a causal hypothesis is likely to be ignored by many decision makers (Dawes, 1999). The most useful of risk assessment instruments will overcome this tendency of human service professionals by contributing substantively to causal explanations.

The challenge for risk assessment is to maintain high levels of predictive validity necessary for accurate pattern matching while increasing the correspondence of risk assessment findings to explanatory theory. In many instances, for example, current risk assessment instruments are a-theoretical. They indicate risk level without explaining the dynamic processes that might explain their findings. Alternatively, actuarial risk assessment instruments could identify theoretically relevant constructs derived from causal theory. For instance, a risk assessment instrument for juvenile justice decision making could be informed by the Social Development Model, an empirically supported theory of delinquency (Catalano & Hawkins, 1996; Hawkins et al., 2003). The findings of such an instrument could show that a youth is high risk for future offending owing to a lack of bonding to pro-social others and a lack of skills for pro-social involvements, for example. Subject to empirical demonstration of predictive validity, such an instrument would meet the pattern matching needs of juvenile justice authorities for accurate classifications along with providing a theoretically relevant hypothesis to explain how observed patterns of risk factors explain risk estimates.

Ultimately, theoretically based actuarial risk assessment instruments would support implementation of the evidence based practice movement. They could do this because theoretically based actuarial risk assessment instruments would contribute to hypothesis formation about (1) what risk patterns contribute to high risk; and

(2) what risk patterns, if disrupted through intervention, might prevent negative outcomes. These hypotheses, informed by theory and research, would lead to an increased ability to match interventions to clients. Thus, incorporating theoretically based actuarial risk assessment instruments into the process of high stakes decision making would lead to increasingly more appropriate, and effective, interventions.

7. Conclusion

This paper argues that adherence to the Rational Choice Model contributes to the underutilization of risk assessment instruments in the day to day practice of human service professionals. Exploration into the naturalistic decision making theories suggests that the usefulness of many current risk assessment instruments may be hampered by an exclusive focus on empirical prediction and reliability. While reinforcing the importance of these qualities, this paper proposes that empirical validation will not on its own ensure accurate and equitable decision making if instruments are designed without consideration for their utility. Along with empirical validity, decision makers require that the findings of risk assessment contribute to explanation. With increased attention to substantive theoretical concerns, risk assessment instruments will be poised to provide significant support to human service professionals making high-stakes decisions. Needed now are comparative studies of the impact of various risk assessment designs on the high-stakes decisions of human service professionals.

References

Baird, C., & Wagner, D. (2000). The relative validity of actuarial- and consensus-based risk assessment systems. *Children and Youth Services Review, 22,* 839–871.

Baird, C., Wagner, D., Healy, T., & Johnson, K. (1999). Risk assessment in child protective services: consensus and actuarial model reliability. *Child Welfare, 78*(6), 723–748.

Belfrage, G., Fransson, G., & Strand, S. (2000). Prediction of violence using the HCR-20: a prospective study in two maximum-security correctional institutions. *The Journal of Forensic Psychiatry, 11,* 167–175.

Buehler, R., Griffin, D., & Ross, M. (2002). Inside the planning fallacy: the causes and consequences of optimistic time predictions. In T. Gilovich, D. Griffin, D. Kahneman, *Heuristics and biases: the psychology of intuitive judgment* (pp. 250–270). New York: Cambridge University Press.

Buehler, R., Griffin, D., & Ross, M. (2002). Inside the planning fallacy: the causes and consequences of optimistic time predictions. In T. Gilovich, D. Griffin, D. Kahneman, *Heuristics and biases: the psychology of intuitive judgment* (pp. 250–270). New York: Cambridge University Press.

Camasso, M. J., & 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.

Catalano, R. F., & Hawkins, J. D. (1996). The social development model: a theory of antisocial behavior. In J. D. Hawkins, *Delinquency and crime: current theories* (pp. 149–197). New York: Cambridge University Press.

Croskerry, P. (2002). Achieving quality in clinical decision making: cognitive strategies and detection of bias. *Academic Emergency Medicine, 9*(11), 1184–1204.

Croskerry, P. (2003). Cognitive forcing strategies in clinical decisionmaking. *Annals of Emergency Medicine, 41*(1), 110–120.

Dawes, R. M. (1999). A message from psychologists to economists: mere predictability doesn't matter like it should (without a good story appended to it). *Journal of Economic Behavior & Organization, 39*(1), 29–40.

Dawes, R. M. (2001). *Everyday irrationality*. Boulder, CO: Westview Press.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical vs. actuarial judgment. *Science, 243*, 1668–1674.

DePanfilis, D. (1996). Implementing child mistreatment risk assessment systems: lessons from theory. *Administration in Social Work, 20*(2), 41–59.

Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: the HCR-20 violence risk assessment scheme and the psychopathy checklist: screening version. *Journal of Consulting and Clinical Psychology, 67*(6), 917–930.

English, D. J., & Pecora, P. J. (1994). Risk assessment as a practice method in child protective services. *Child Welfare, 73*(5), 451–473.

Firestein, R. E. (1999). *Why children continue to die: a report on child fatalities in New York City*. Retrieved May 28, 2003, from www.childrensrights.org.

Fraser, M. W. (2004). *Risk and resilience in childhood: an ecological perspective* (2nd ed). Washington, DC: NASW Press.

Fuller, J., & Cowan, J. (1999). Risk assessment in a multi-disciplinary forensic setting: clinical judgement revisited. *Journal of Forensic Psychiatry, 10*(2), 276–289.

Gambrill, E. (1990). *Critical thinking in clinical practice: improving the accuracy of judgments and decisions about clients*. San Francisco: Jossey-Bass.

Gambrill, E. (2003). Ethics, science, and the helping professions: a conversation with Robyn Dawes. *Journal of Social Work Education, 39*, 27–40.

Gambrill, E., & Shlonsky, A. (2000). Risk assessment in context. *Children and Youth Services Review, 22*(11–12), 813–837.

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality: the adaptive toolbox*. Cambridge, MA: The MIT Press.

Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology, 24*, 411–435.

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*, 293–323.

Hammond, K. R. (1996). *Human judgment and social policy: irreducible uncertainty, inevitable error, unavoidable injustice*. New York: Oxford University Press.

Hammond, K. R. (2000). *Judgments under stress*. New York: Oxford University Press.

Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: the development of a statistical prediction instrument. *Criminal Justice and Behavior, 20*, 315–335.

Hawkins, D. F., Smith, B. H., Hill, K. G., Kosterman, R., Catalano, R. F., & Abbott, R. D. (2003). Understanding and preventing crime and violence: findings from the Seattle social development project. In T. P. Thornberry, M. D. Krohn, *Taking stock of delinquency: an overview of findings from contemporary longitudinal studies*. New York: Kluwer Academic/Plenum Publishers.

Hendryx, M. S., & Rohland, B. M. (1997). Psychiatric hospitalization decision making by CMHC staff. *Community Mental Health Journal, 33*, 63–73.

Hilton, N. Z., & Simmons, J. L. (2001). The influence of actuarial risk assessment in clinical judgments and tribunal decisions about mentally disordered offenders in maximum security. *Law and Human Behavior, 25*(4), 393–408.

Hoge, R. D. (2002). Standardized instruments for assessing risk and need in youthful offenders. *Criminal Justice and Behavior, 29*, 380–396.

Holland, S. (2000). The assessment relationship: interactions between social workers and parents in child protection assessments. *British Journal of Social Work, 30*, 149–163.

Jones, P. R., Harris, P. W., Fader, J., & Grubstein, L. (2001). Identifying chronic juvenile offenders. *Justice Quarterly, 18*(3), 479–507.

Jung, S., & Rawana, E. P. (1999). Risk and need assessment of juvenile offenders. *Criminal Justice and Behavior, 26*(1), 69–89.

Klein, G. A. (1997). Developing expertise in decision making. *Thinking and Reasoning, 3*, 337–352.

Klein, G. A. (1998). *Sources of power: how people make decisions*. Cambridge, MA: MIT Press.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (1993). *Decision making in action: models and methods*. Norwood, NJ: Ablex Publishing Corporation.

Krysik, J., & LeCroy, C. W. (2002). The empirical validation of an instrument to predict risk of recidivism among juvenile offenders. *Research on Social Work Practice, 12*, 71–81.

Lindley, D. V. (1985). *Making decisions* (2nd ed). New York: John Wiley & Sons.

Lipshitz, R. (1993). Converging themes in the study of decision making in realistic settings. In G. A. Klein, J. Orasanu, R. Calderwood, C. E. Zsambok, *Decision making in action: models and methods* (pp. 103–137). Norwood, NJ: Ablex Publishing Corporation.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001a). Focus article: taking stock of naturalistic decision making. *Journal of Behavioral Decision Making, 14*(5), 331–352.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001b). Rejoinder: a welcome dialogue—and the need to continue. *Journal of Behavioral Decision Making, 14*(5), 385–389.

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*, 935–949.

McNiel, D. E., Sandberg, D. A., & Binder, R. L. (1998). The relationship between confidence and accuracy in clinical assessment of psychiatric patients' potential for violence. *Law and Human Behavior, 22*(6), 655–669.

Meehl, P. E. (1986). Causes and effects of my disturbing little book. *Journal of Personality Assessment, 50*, 370–375.

Monahan, J., Steadman, H. J., Appelbaum, P. S., Robbins, P. C., Mulvey, E. P., & Silver, E. (2000). Developing a clinically useful actuarial tool for assessing violence risk. *British Journal of Psychiatry, 176*, 312–319.

Monahan, J., Steadman, H. J., Silver, E., Appelbaum, P. S., Robbins, P. C., & Mulvey, E. P. (2001). *Rethinking risk assessment: the MacArthur study of mental disorder and violence*. Oxford: University Press.

Munro, E. (1996). Avoidable and unavoidable mistakes in child protection work. *British Journal of Social Work, 26*, 793–808.

Munro, E., & Rumgay, J. (2000). Role of risk assessment in reducing homicides by people with mental illness. *British Journal of Psychiatry, 176*, 116–120.

Murdach, A. D. (1994). Avoiding errors in clinical-prediction. *Social Work, 39*(4), 381–386.

Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. Klein, J. Orasanu, R. Calderwood, C. E. Zsambok, *Decision making in action: models and methods* (pp. 3–20). Norwood, NJ: Ablex Publishing Corp.

Payne, J. W., & Bettman, J. R. (2001). Preferential choice and adaptive strategy use. In G. Gigerenzer, R. Selten, *Bounded rationality: the adaptive toolbox*. Cambridge, MA: The MIT Press.

Pratt, J. W., Raiffa, H., & Schlaifer, R. (1995). *Introduction to statistical decision theory*. Cambridge, MA: The MIT Press.

Risler, E. A., Sutphen, R., & Shields, J. (2000). Preliminary validation of the juvenile first offender risk assessment index. *Research on Social Work Practice, 10*(1), 111–126.

Rosen, A. (1994). Knowledge use in direct practice. *Social Service Review, 68*, 561–577.

Rossi, P. H., Schuerman, J., & Budde, S. (1999). Understanding decisions about child maltreatment. *Evaluation Review, 23*(6), 579–598.

Ruscio, J. (1998). Information integration in child welfare cases: an introduction to statistical decision making. *Child Maltreatment, 3*(2), 143–157.

Schuerman, J., Rossi, P. H., & Budde, S. (1999). Decisions on placement and family preservation. *Evaluation Review, 23*(6), 599–618.

Sheppard, M. (1995). Social-work, social-science and practice wisdom. *British Journal of Social Work, 25*(3), 265–293.

Skeem, J. L., Mulvey, E. P., & Lidz, C. W. (2000). Building mental health professionals' decisional models into tests of predictive validity: the accuracy of contextualized predictions of violence. *Law and Human Behavior, 24*(6), 607–628.

Stuart, H. L., & Arboleda-Florez, J. E. (2001). A public health perspective on violent offenses among persons with mental illness. *Psychiatric Services, 52*(5), 654–659.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science, 185*, 1124–1131.

Tversky, A., & Kahneman, D. (2002). Extensional vs. intuitive reasoning: the conjunction fallacy in probabilistic judgment. In T. Gilovich, D. Griffin, D. Kahneman, *Heuristics and biases: the psychology of intuitive judgment* (pp. 19–48). New York: Cambridge University Press.