Clinical Versus Actuarial Judgments in Criminal Justice Decisions: Should One replace the Other?

Stephen D. Gottfredson
Laura J. Moriarty
Virginia Commonwealth University

IN VIRTUALLY ALL decision-making situations that have been studied, actuarially developed devices outperform human judgments. This is true with respect to psychiatric judgments (see, for example, Meehl, 1965; Gough, 1962; Ennis and Litwack, 1974); graduate school admissions (e.g., Dawes, 1979; Dawes and Corrigan, 1974); prognostic judgments made by sociologists and psychiatrists relative to a parole-violation criterion (Glaser, 1955, 1962); parole board decisions (Gottfredson, 1961; Gottfredson and Beverly, 1962; Carroll, Wiener, Coates, Galegher, & Alibrio, 1982); mental health and correctional case worker judgments of offender risk (Holland, Holt, Levi, & Beckett, 1983), spousal assault (Hilton and Harris, 2005) and in other areas (Goldberg, 1970), including the analysis of credit risk (Somerville and Taffler, 1995). Indeed, a recent review and meta-analysis of 56 years' accumulation of research on the "clinical vs. statistical" prediction "problem" conducted as part of a Festschrift for Paul E. Meehl, a pioneer in the field, again confirms that statistical models outperform clinical decision-makers (Ægisdóttir, White, Spengler, Maugherman, Anderson, Cook, Nichols, Lampropoulos, Walker, Cohen and Rush, 2006).

The relative superiority of statistical to intuitive methods of prediction is due to many factors. For example, human decision-makers often do not use information reliably (e.g., Ennis and Litwack, 1974), they often do not attend to base rates (Meehl and Rosen, 1955), and this has been specifically illustrated in criminal justice decision-making (Carroll, 1977); they may inappropriately weight items of information that are predictive, or they may assign weight to items that in fact are not predictive; and they may be overly influenced by causal attributions (e.g., Carroll, 1978) or spurious correlations (Monahan, 1981). In fairness, it should be pointed out that there may be advantages to intuitive judgments as well. For example, human decision-makers can make use of information that cannot be made available to a statistical device (at least readily). Demeanor during an interview may be one such example. Other factors in favor of intuitive judgments are reviewed in Dawes (1975; Dawes, Faust, and Meehl, 1989). [1]

Given these facts, is there reason to still consider clinical judgments when determining risk-assessment within a justice system population? Indeed, with the 1998 publication of *Violent Offenders: Appraising and Managing Risk* (Quinsey, Harris, Rice and Cormier), we find an argument that we should not. "What we are advising is not the addition of actuarial methods to

existing practice, but rather the complete replacement of existing practice with actuarial methods" (p. 171; see Litwack, 2001 for a strong rebuttal in the arena of the assessment of dangerousness). We argue that even though statistical prediction is superior to clinical judgment in almost all settings, this does not obviate the need for nor value of clinical judgment in a variety of arenas, including some criminal justice venues. We use the roles of probation officers and correctional treatment specialists to provide examples.

Probation Officers and Correctional Treatment Specialists

Among the largest group of criminal justice professionals is that working in corrections. And with the vast number of adults and juveniles on probation, parole or incarcerated, the workload of these individuals is quite high.

According to the U.S. Department of Labor, Bureau of Labor Statistics, there are about 90,600 probation officers and correctional treatment specialists nationally (Bureau of Labor Statistics, 2006), and in the federal system, there are approximately 5,000 officers throughout the United States and its territories (personal communication, Richard Gayler, May 31, 2006). The number of adults on probation in 2004 was about 4.1 million (Glaze and Palla, 2005), for an average caseload nationally of about 46. The top three states in terms of employment of probation officers and correctional treatment specialists are California (13,090), Texas (6,100) and Florida (5,760). When examining data from 2004 that reports the state's community corrections population, we find that Texas has the largest, with 428,773 adults under supervision, followed by California with 384,852, and Florida with 281,170 (Glaze and Palla, 2005). Using these figures, the average caseloads range from a high of 70 to a low of 29.

Qualifications for employment as a probation officer or correctional treatment specialist vary by state, but a Bachelor's degree in social work, criminal justice or some other related field typically is required (Bureau of Labor Statistics, 2006). Some states require a more advanced degree— Master of Science or Master of Arts in a related field (psychology, sociology, criminology, etc.), often with an additional experiential requirement as well. Many states require that probation officers and correctional treatment specialists receive training, upon completion of which the candidate must pass a certification test. Typically, new officers work as trainees or for a probationary period before they become permanent employees.

Probation officers supervise offenders who have been placed on probation while correctional treatment specialists counsel and create rehabilitation plans for offenders to follow when they are no longer incarcerated or on parole (Bureau of Labor Statistics, 2006). Probation officers also spend a great deal of time in court, investigate offender backgrounds, write pre-sentence investigation reports, and recommend sentences and treatment plans. Correctional treatment specialists may work in jails, prisons, or probation or parole agencies, where they might evaluate the progress of inmates, develop parole and release plans, write case reports for parole boards and other decision-makers, and/or develop and write treatment plans and summaries for clients.

What we have then, is a large number of highly qualified and trained professionals who routinely are required to make prognostic decisions about offenders. Elsewhere, we have described ways to improve the reliability of these decision-making processes (Gottfredson and Moriarty, 2006). We have argued that the use of actuarial devices invariably increases the reliability and prognostic validity of decisions made in these settings (Gottfredson and Moriarty, 2006), and, as noted above, some would argue that the surest way to do this is to rely solely on statistical prediction, such as risk-assessment tools, as a way to increase the accuracy and reliability of these decisions (Quinsey, Harris, Rice and Cromier, 1999). Is it indeed time to supplant human judgment in justice system settings with the cold calculus of the actuary?

Human Judgment

Judgments are made routinely in a host of fields including psychiatry and psychology (Kleinmuntz, Faust, Meehl, & Dawes, 1990; Dawes et al., 1989); mental health (Ægisdóttier et al., 2006); dangerousness (Litwack, 2001); economics (Dawes, 1999); forecasting (Bunn & Wright, 1991), medicine, engineering, finance, management (Kleinmuntz et al., 1990, p. 146); interpersonal violence (Hilton, Harris & Rice, 2006; Mills, 2005); and forensics (Hilton, Harris, Rawson & Beach, 2005; Harris, Rice & Cormier, 2002), in addition to those noted earlier in this paper. In most cases, the literature reveals strong support for the accuracy of actuarial prediction over human judgment. This is a longstanding finding, replicated in dozens of venues (Dawes et al., 1989; Kleinmuntz et al., 1990; Westen & Weinberger, 2005). As noted earlier, there are many reasons to expect that actuarial methods will outperform human judgments. In addition to those reasons cited above, these methods may be expected to provide other benefits:

> Even when actuarial methods merely equal the accuracy of clinical methods, they may save considerable time and expense. … When actuarial methods are not used as the sole basis for decisions, they can still serve to screen out candidates or options that would never be chosen after more prolonged consideration. When actuarial methods prove more accurate than clinical judgment the benefits to individuals and society are apparent (Dawes et al., 1989, page 1673).

Why, then, should we continue to allow (indeed, require) probation officers and case managers to exercise individual discretion, when an actuarially-derived tool may be expected to perform better? As Westen and Weinberger (2005) remind us in a discussion of the pioneering work of Paul E. Meehl on clinical and statistical prediction, even though statistical prediction will routinely outperform clinical prediction, we should not lose sight of the fact that "actuarial procedures are far from infallible, sometimes achieving only modest results" (Dawes et al., 1989, p. 1673). [For a discussion of the methodological and statistical problems associated with such applications and the resultant fallibilities of such procedures, see Gottfredson and Moriarty, 2006]. Still, as Dawes (2005) concludes, "whenever statistical prediction rules (SPR's) are available for making a relevant prediction, they should be used in preference to intuition" (p. 1245).

But does the superiority of actuarial procedures over clinical judgment mean that there is no place for clinical judgment in predicting behavior? The answer is no: "an enormous amount of prediction is made by human judgment" (Darlington, 1986, p. 362). Simply put, clinical methods of decision-making rest in the decision-maker's head, while statistical or actuarial methods eliminate the human judgment with the "conclusions rest(ing) solely on empirically established relations between the data and the condition or event of interest" (Dawes et al., 1989, p. 1668). There are instances when clinicians can make valid inferences (Westen and Weinberger, 2005), and there are times when it is preferable to use both clinical and statistical judgments to predict behavior. As Darlington (1986, 362) reports, "This research does not suggest human judgment is generally unnecessary; rather it indicates that the most accurate predictions generally result from a predictive system in which human judgment and statistical analysis are mixed according to prescribed rules." Moreover, Dawes et al. (1989) assert that there are instances when clinical judgment might improve the actuarial method. They cite specifically the following examples: "judgments mediated by theories and hence difficult or impossible to duplicate by statistical frequency alone; select reversal of actuarial conclusions based on the consideration of rare events or utility functions that are not incorporated into statistical methods; and complex configural relations between predictive variables and outcome" (p. 1670).

The Place for Personal Judgment

As some assert, the dilemma once posed by "using either the head or the formula is no longer the main focus of contemporary decision research. Rather, the focus has long ago shifted to evaluating the use of both modes of information combination in tandem" (Kleinmuntz et al., 1990, p. 146; and see generally, Litwack, 2001).

In some of the most recent research examining violence risk assessment (Hanson, 2005), we find that the question has shifted from whether violence can be predicted, to what is the best method of risk assessment. The validation research typically found that a series of measures reviewed in the article showed moderate accuracy in predicting violent recidivism. The question then is: Might the prediction have been improved if clinical judgments were included as well?

This question is partly answered by Douglas, Yeoman, and Boer (2005), who studied violence risk in a sample of criminal offenders. Douglas and colleagues looked at the predictive validity of multiple indices of violence risk. Although they conclude that "several indices were related to violent recidivism with large statistical effect sizes,…." they also found the findings to be "inconsistent with a position of strict actuarial superiority, as HCR-20 [2] structured risk judgments—an index of structured professional or clinical judgments—were as strongly related to violence" (479).

Conclusion

With very few established exceptions, statistical prediction clearly outperforms clinical judgment. Accordingly, we certainly would not advocate use of clinical judgment over statistical prediction. For as Dawes reports,

> The superiority of statistical prediction is crystal clear when clinical judgment is pitted against actuarial analysis in a situation where both are based on the same information—so that the problem is basically one of how to combine it. It also has been found that clinical judgment in psychology is inferior in situations where the important variables captured by the statistical model constitute a proper subset of the variables considered by the clinician. It is also true that the statistical models need not even be optimal. Nevertheless, clinical psychologists make a great deal of money by relying on their intuitions for combining information and for making predictions, and in courts they eschew statistical models, instead proudly proclaiming that "in my experience…" What happens here is that the "inside view" is preferred to the outside one, despite massive evidence that that outside one is superior (Dawes, 1999, pp. 37-38).

However, there are times when a combination of the two may better serve clientele. As Dawes et al. (1989) report, "Clinicians might be able to gain an advantage by recognizing rare events that are not included in the actuarial formula (due to their infrequency) and that countervail the actuarial conclusion (p. 1670).

And while such incidents might be infrequent, it is also true that the probation officers and correctional counseling specialists must have a role in decision-making that goes beyond the mere administering of the risk-assessment devices. There is a place for human judgment and experience in the decision-making process, and we must value their continued consideration.

However, as noted by Sir Frances Bacon, "We do ill to exalt the powers of the human mind, when we should seek out its proper helps" (as quoted in Hogarth (1980)). In light of the well-known tendency for justice system decision-makers to concentrate on information that is demonstrably not predictive of offender behavioral outcomes (Gottfredson and Gottfredson, 1986), and the potential consequences of this for affecting the validity of prognostication (Gottfredson and Moriarty, 2006), caution is the order of the day.

Walters, G. D. (1997). A confirmatory factor analysis of the Lifestyle Criminality Screening Form. *Criminal Justice and Behavior*, *24*, 294-308.

Webster, C. D., Douglas, K. S., Eaves, D., & Hart S. D. (1997). *HCR-20: Assessing risk for violence* (Version 2). Burnaby, Canada: Simon Fraser University, Mental Health, Law, and Policy Institute.

Webster, C. D., Hucker, S., J., & Bloom, H. (2002). Transcending the actuarial versus clinical polemic in assessing risk for violence. *Criminal Justice and Behavior, 29*, 659-665.

Wormith, S. J., & Goldstone, C. W. (1984). The clinical and statistical prediction of recidivism. *Criminal Justice and Behavior*, *11*, 3-34.

Clinical Versus Actuarial Judgments in Criminal Justice Decisions: Should One replace the Other?

Ægisdóttier, S., White, M.J., Spengler, P.M., Maugherman, A.S., Anderson, L.A., Cook, R.S., Nichols, C.N., Lampropoulos, G.K., Walker, B.S., Cohen, G., & Rush, J.D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *Counseling Psychologist*, 34 (3), 341-382.

Bunn, D., & Wright, G. (1991). Interaction of judgmental and statistical forecasting methods: Issues and analysis. *Management Science*, 37 (5), 501-518.

Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook, 2006-07 Edition, Probation Officers and Correctional Treatment Specialists, on the internet at http://www.bis.gov/oco/ocos265.htm (visited on May 26, 2006).

Carroll, J.S. (1977). Judgments of recidivism risk: Conflicts between clinical strategies and base-rate information. *Law and Human Behavior*, *1* (2), 191-198.

Carroll, J.S., Wiener, R.L., Coates, D., Galegher, J., & Alibrio, J.J. (1982). Evaluation, diagnosis, and prediction in parole decision making. *Law and Society Review, 17* (1), 199-228.

Darlington, R.B. (1986). Combining human judgment and multiple regression. *Journal of Personality Assessment*, 50 (3), 362-395.

Dawes, R.M. (1975). Case by case versus rule-generated procedures for the allocation of scarce resources. In M. Kaplan and S. Schwartz (eds.), *Human judgment and decision processes in applied setting*, pp. 83-94. New York: Academic Press.

Dawes, R.M. (1979). The robust beauty in decision making. *Psychological Bulletin, 34* (7), 571-582.

Dawes, R.M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81*, (2), 95-106.

Dawes, R.M. (2005). The ethical implications of Paul Meehl's work on comparing clinical versus actuarial prediction methods. *Journal of Clinical Psychology*, 61 (10), 1245-1255.

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, 29-40.

Dawes, R. M., Faust, D., & Meehl, P.E. (1989). Clinical versus actuarial judgment. *Science, New Series*, 243 (4899), 1668-1674.

Douglas, K.S., Yeoman, M., & Boer, D.P. (2005). Comparative validity analysis of multiple

measures of violence risk in a sample of criminal offenders. *Criminal Justice and Behavior*, 32 (5), 479-510.

Ennis, B.J., & Litwack, T.R. (1974). Psychiatry and the presumption of expertise: Flipping coins in the courtroom. *California Law Review, 62*, 693-752.

Glaze, L.E., & Palla, S. (2005). Probation and Parole in the United States, 2004. Bureau of Justice Statistics Bulletin (NCJ 210676). Washington, DC: U.S. Department of Justice.

Glaser, D. (1955). The efficacy of alternative approaches to parole prediction. *American Sociological Review, 20*, 283-287.

Glaser, D. (1962). Prediction tables as accounting devices for judges and parole boards. *Crime and Delinquency, 8* (3), 239-258.

Goldberg, L.R. (1970). Man versus model of man: A rationale, plus some evidence for a method of improving on clinical inference. *Psychological Bulletin, 73*, 422-432.

Gottfredson, D.M. (1961). Comparing and combining subjective and objective parole predictors. *Research Newsletter, 3* (Sept-Dec.). Vacaville: California Medical Facility.

Gottfredson, D.M., & Beverly, R.F. (1962). Development and operational use of prediction methods in correctional work. Proceedings of the Social Statistics Section. Washington, DC: American Statistical Association.

Gottfredson, S.D., & Gottfredson, D.M. (1986). Accuracy of prediction models. In Blumstein, A. et al., (eds.), *Criminal Careers and "Career Criminals,"* Vol. II, Washington, DC: National Academy of Sciences Press.

Gottfredson, S.D., & Moriarty, L.J. (2006). Statistical risk assessment: Old problems and new applications. *Crime & Delinquency*, 52 (1), 178-200.

Gough, H.G. (1962). Clinical versus statistical prediction in psychology. In L. Postman (ed), *Psychology in the Making*, pp. 526-584. New York: Knopf.

Hanson, R. K. (2005). Twenty years of progress in violence risk assessment. *Journal of Interpersonal Violence*, 20 (2), 212-217.

Harris, G.T., Rice, M.E., & Cormier, C.A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior*, 26 (4), 377-394.

Hilton, N. Z., Harris, G.T., Rawson, K., & Beach, G.A. (2005). Communicating violence risk information to forensic decision makers. *Criminal Justice and Behavior*, 32 (1), 97-116.

Hilton, N. Z. & Harris, G.T. (2005). Predicting wife assault: A critical review and implications for policy and practice. Trauma, Violence and Abuse, 6 (1), 3-23.

Hilton, N. Z., Harris, G.T., & Rice, M.E. (2006). Sixty-six years of research on the clinical versus actuarial prediction of violence. *Counseling Psychologist*, 34 (3), 400-409.

Hogarth, R. M. (1980). Judgement and choice: The psychology of decision. New York: John Wiley.

Holland, T.R., Holt, N., Levi, M., & Beckett, G.E. (1983). Comparison and combination of clinical and statistical predictions of recidivism among adult offenders. *Journal of Applied Psychology, 68* (2), 203-211.

Kleinmuntz, B., Faust, D., Meehl, P.E., & Dawes, R. M. (1990). Clinical versus actuarial judgment. *Science, New Series*, 247 (4939), 146-147.

Litwack, T.R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology Public Policy and Law*, 7 (2), 409-443.

Meehl, P.E. (1965). Seer over sign: The first good example. *Journal of Experimental Research in Personality, 1*, 27-32.

Meehl, P.E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52* (3), 194-216.

Mills, J. F. (2005). Advances in the assessment and prediction of interpersonal violence. *Journal of Interpersonal Violence*, 20 (2), 236-241.

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques* . Beverly Hills, CA: Sage.

Quinsey, V.L., Harris, G.T., Rice, M.E., and Cormier, C.A. (1998). Violent offenders: Appraising and managing risk. Washington, D.C.: American Psychological Association.

Somerville, R.A., & Taffler, R.J. (1995). Banker judgment versus formal forecasting models: The case of country risk assessment. *Journal of Banking & Finance,* 19, 281-297.

Westen, D., & Weinberger, J. (2005). In praise of clinical judgment: Meehl's forgotten legacy. *Journal of Clinical Psychology*, 61 (10), 1257-1276.

Overcoming Sisyphus: Effective Prediction of Mental Health Disorders and Recidivism Among Delinquents

Archer, R. P. (1997). *MMPI-A: Assessing adolescent psychopathology* (2 nd Ed.). Mahwah, NJ: Erlbaum.

Archer, R. P., & Baker, E. M. (2005). Minnesota Multiphasic Personality Inventory—Adolescent. In T. Grisso, G. Vincent, & D. Seagrave (Eds.) *Mental health screening and assessment in juvenile justice* (pp. 240-252) . New York : Guilford Press.

Archer, R. P., Bolinskey, P. K., Morton, T. L., & Farris, K. L. (2003). MMPI-A characteristics of male adolescents in juvenile justice and clinical treatment settings. *Assessment, 10*, 400-410.

Aronen, E. T., Noam, G. G., & Weinstein, S. R. (1993). Structured diagnostic interviews and clinicians' discharge diagnoses in hospitalized adolescents. *Journal of the American Academy of Child ad Adolescent Psychiatry, 32,* 674-681.

Bartel, P., Borum, R., & Forth, A. (2000). *Structured Assessment for Violence Risk in Youth (SAVRY): Consultation Edition.* Tampa : Louis de la Parte Florida Mental Health Institute, University of South Florida.

Borum, R., Bartel, P., Forth, A. (2005). Structural Assessment of Violence Risk in Youth. In T. Grisso, G. Vincent, & D. Seagrave (Eds.) *Mental health screening and assessment in juvenile justice* (pp. 311-323) . New York : Guilford Press.

Butcher, J. N., Williams, C. L., Graham, J. R., Archer, R. P., Tellegen, A., Ben-Porath, Y. S., et al. (1992). *Minnesota Multiphasic Personality Inventory—Adolescent (MMPIA): Manual for administration, scoring and interpretation.* Minneapolis : University of Minnesota Press.

Cambell, M. A., Porter, S., & Santor, D. (2004). Psychopathic traits in adolescent offenders: An evaluation of criminal history, clinical, and psychosocial correlates. *Behavioral Sciences and the Law, 22,* 23-47.

Catchpole, R., & Gretton, H. (2003). The predictive validity of risk assessment with violent

Endnotes

[Clinical Versus Actuarial Judgments in Criminal Justice Decisions: Should One replace the Other?](#)

[Validation of the Risk and Resiliency Assessment Tool for Juveniles in the Los Angeles County probation System](#)

[How Much Risk Can We Take? The Misuse of Risk Assessment in Corrections](#)

[From Theory to Practice: The Lifecycle Document for the Results-Based Management Framework for the Federal Probation and Pretrial Services System](#)

Clinical Versus Actuarial Judgments in Criminal Justice Decisions: Should One replace the Other?

[1] Portions of the preceding discussion adapted from Gottfredson and Gottfredson (1986).

[2] Historical-Clinical-Risk-Management-20 (HRC-20).

[back to top](#)

Validation of the Risk and Resiliency Assessment Tool for Juveniles in the Los Angeles County probation System

[1] These estimates are based on unpublished analyses from Turner & Fain (2003).

[2] Police refer cases to the District Attorney in Los Angeles for processing. Youths charged with offenses for which the District Attorney must file and those youths who are detained in juvenile hall are directed to the Court for arraignment. The SDRRC was administered at this pre-plea stage for these "court" cases. Youths not initially referred to court—those generally with more minor offenses—are referred to Probation to make a determination of how to handle the case. These "non-court" cases can received a number of possible outcomes, including having the case closed, the youth being placed on informal probation, or the case being referred to court. The SDRRC was administered to "noncourt" cases at this point.

[3] Due to logistical restrictions, we were not able to pilot test the SDRRC in all area Probation Department offices.

[4] For weighting purposes, ethnicity was divided into five categories: black, Hispanic, white, other race, and unknown. Age was categorized as less than 13, greater than 18, and single years of age for ages 13-18.

[5] We were not able to use age in calculating weights because there were too few