

Running head: Bayes's Theorem

A Mathematical Proof and Example That Bayes's Theorem Is
Fundamental to Actuarial Estimates of Sexual Recidivism Risk

Theodore Donaldson, Ph.D.

Independent Practice

Morro Bay, California

Richard Wollert, Ph.D.

Independent Practice

Vancouver, Washington

Published in *Sexual Abuse: A Journal of Research and Treatment*, 2008, 12, 56-85.
(c) Association for the Treatment of Sexual Abusers (<http://sajrt.sagepub.com>)

This article may not exactly replicate the final version
published in the ATSA journal. It is not the copy of record.

Abstract

Expert witnesses in Sexually Violent Predator (SVP) cases often rely on actuarial instruments to make risk determinations. Many questions surround their use, however. Bayes's Theorem holds much promise for addressing these questions. Some experts nonetheless claim that Bayesian analyses are inadmissible in SVP cases because they are not accepted by the relevant scientific community (Doren, 2006). This position is illogical because Bayes's Theorem is simply a probabilistic restatement of the way that frequency data are combined to arrive at whatever recidivism rates are paired with each test score in an actuarial table. The paper at hand presents a mathematical proof and example validating this assertion. The advantages and implications of a logic model that combines Bayes's Theorem and the null hypothesis are also discussed.

A Mathematical Proof And Example That Bayes's Theorem Is
Fundamental To Actuarial Estimates of Sexual Recidivism Risk

Many states in the United States have enacted legislation allowing for the post-prison civil commitment of sex offenders as “sexually violent predators” (SVPs; Covington, 1997; Doren, 2002; Miller, Amenta, & Conroy, 2005). As the first stage of this process, offenders thought to meet the commitment standards are referred by review panels to prosecutors. Secondly, prosecutors decide if commitment petitions should be filed. Then, the courts determine if probable cause exists for evaluating the possibility that these offenders might be SVPs. Finally, the courts determine which respondents should be committed as SVPs. Regarding this step, Miller and her colleagues (2005, p. 31) have observed that

“The criteria commonly found in state SVP commitment laws closely resemble those set forth in *Kansas v. Hendricks* (1997). Such laws generally include four elements: (i) a history of sexual offenses, (ii) a mental abnormality, (iii) volitional impairment, and (iv) as a result of mental abnormality, the individual is likely to engage in acts of sexual violence. The reader is referred to Table 1 (pp. 32-36) for a review of statutory definitions across the 16 states with relevant civil commitment laws. Although variation exists among state definitions, the similarity of these definitions to those accepted by the Supreme Court as constitutional is clear.”

Decision making in SVP cases is a complex process because SVP laws do not define all the terms they invoke. Consequently, the judicial system relies heavily on the opinions of psychologists and other expert witnesses. Explicitly or implicitly, an expert in an SVP case renders two types of opinions (Wollert, 2006). First, he or she determines, typically after

surveying a large mass of potential evidence, the certainty with which it may be asserted that a given respondent is positive for a mental abnormality and is prone to recidivating. Then the expert tests the null hypothesis that the respondent is not an SVP based on the degree of certainty that these opinions command. If a "reasonable degree of certainty" is attained on all issues, the null hypothesis is rejected; if not, the null hypothesis stands.¹

Because the validity of the SVP construct was established by legislative action rather than a professional consensus based on the application of the scientific method, considerable uncertainty exists over basic questions that bear on SVP cases. Which diagnoses are reliable and constitute valid mental conditions? What causal linkages are demonstrable between a mental condition, lack of volitional control, and sexual dangerousness? What is the optimum time period for the reliable projection of recidivism estimates? To what extent do current actuarial instruments provide estimates that speak to the commitment standards specified in SVP legislation?

These issues need to be clarified. If they remain unresolved, an increasing amount of doubt may be cast on the extent to which experts can substantively assist the court in discharging its fact-finding role.

Although different methods may be used to advance the knowledge base in this area, and also to derive certainty opinions in individual SVP risk assessments, perhaps the most promising is associated with Bayes's Theorem. Bayes's Theorem, which combines information concerning the chances of encountering a given condition with the chances of encountering evidence associated with this condition, has been a cornerstone of probability theory since the dissemination of Bayes's (1764) essay on "The Doctrine Of Chances." Over the last 60 years,

many statisticians (e.g., de Finetti, 1964; Jaynes & Bretthorst, 2003; Jeffreys, 1939; Ramsey, 1931; Savage, 1954) have elaborated a mathematical foundation that supports its use as a method of scientific inquiry. In the process, they have called attention to the unique power of Bayes's Theorem for weighting evidence and, as a result, estimating the level of certainty with which any legal or scientific theory (hereafter symbolized as " R ") may be held on the basis of a specified set of evidence (hereafter symbolized as " S "). Because of these advantages, Bayesian analysis has elicited an "explosion of interest" from such varied scientific disciplines as biochemistry, ecology, medicine, oncology, pharmacology, public health, and statistics (Fienberg, 2006; Woodworth, 2004).

A few researchers, drawing on the promise of Bayesian analysis, have pioneered the application of Bayes's Theorem to the study of issues pertaining to the civil commitment of sexually violent predators. In one article, for example, Janus and Meehl (1997) reported a Bayesian analysis indicating that the statutorily-defined probability threshold for recidivism, referred to as point (iv) in the first paragraph, was attainable only under some rather questionable assumptions – that is, that the recidivism base rate for sex offenders was quite high and that the actuarial test evidence relied on to identify recidivists was quite accurate. In another, Wollert (2006) presented evidence that all current actuarials are largely useless for identifying future sexual recidivists because they incompletely specify the negative effects of age on recidivism; he also emphasized that even revised and more accurate actuarials might be useless for this purpose because the base rate of sexual recidivism has dropped dramatically in the United States in recent years (Langan, Schmitt, & Durose, 2003; Washington State Institute for Public Policy, April 26, 2005). In still another, Mossman (2006) argued that it is inappropriate for an expert to attribute a

tabled recidivism rate for a given actuarial score in one population to an SVP respondent in a different population unless the expert verifies that the Bayesian coordinates (i.e., actuarial accuracy and the recidivism base rate) are the same in both populations.

Each of these articles have important practice implications for experts who typically rely on actuarial tests as a foundation for opining that civil commitment respondents are more likely than not to recidivate. Accordingly, these articles have come under careful scrutiny and criticism. Some critics have argued that little weight should be attached to the conclusions of a specific Bayesian analysis. For example, Doren and Epperson (2001) argued that Janus & Meehl (1997) assumed that the highest plausible recidivism rate for prison releasees was 45% but that even higher estimates “may be thought of as representing a low-end estimate for the relevant base rate” and that “the more likely base rate estimate for the ‘commitment class’ population might be somewhat greater than 50%” (p. 48).

Other criticisms are linked to a much different type of claim, however – that is, that Bayesian analyses are inadmissible in civil commitment cases because they are not accepted by the relevant professional community. Doren (2006), for example, has alluded to this proposition by asserting that “predictions of recidivism ... use” probabilistic concepts “like ‘true positives’ and ‘false positives’ ” (p. 4) and further claiming that

“civil commitment risk evaluators do not make recidivism predictions ... hence the statistical concepts related to the accuracy of predictions do not apply ... the error being described is not specific to the statistical analysis that someone presents, but represents an erroneous fundamental assumption underlying the statistical analysis” (pp. 4-5).²

The authors believe it is fair and clarifying to question the assumptions underlying a particular Bayesian analyses – as long as a reasonable case can be made that other assumptions are actually justified.³ An anti-statistical position such as that articulated by Doren, however, threatens the integrity of science in the courtroom. For example, he filed at least one declaration in late 2004 in support of a Frye motion to suppress the admissibility of a colleague's research results, obtained through the application of Bayesian reasoning, on the grounds that the colleague's "theory that actuarially-ascertained risk estimates can be reduced by factoring in age through a specific numerical process is clearly not generally accepted in the field" (Doren, December 23, 2004). Furthermore, he was joined in this opinion by two other well-known psychologists (Hanson, January 21, 2005; Quinsey, January 14, 2005). Although a hearing did not take place, the fact that it was possible to elicit declarations that either intentionally or unwittingly attacked a method that is accepted by virtually all scientific disciplines underscores the legal and scientific problems that are associated with a stance that, at the very least, does not appreciate the importance of Bayesian analysis.

Overall, we believe that it is mathematically illogical for those who use actuarial tests in their civil commitment evaluations to claim that Bayes's Theorem is not accepted by the relevant scientific community. The reason this is the case is that Bayes's Theorem is simply a probabilistic restatement of the way that frequency data are combined to arrive at whatever recidivism rates are paired with each test score in an actuarial table. This assertion, if shown to be true, means that the large number of experts who currently accept actuarial tests must also necessarily accept Bayes's Theorem as a valid method for reaching certainty opinions in SVP cases.

In the next section of this paper a mathematical proof and example of the foregoing assertion will be presented. In the concluding section the implications of this proof will be discussed for setting boundaries on the claims that experts may advance as to the non-acceptance of Bayes's Theorem by those who undertake SVP evaluations. The advantages and implications of Bayes's Theorem for advancing research and practice will also be considered.

Proof and Example

Actuarial tables that are used to estimate the risk of sexual recidivism are almost always conceptualized in terms of cells that contain frequency data. Table 1, for example, shows that Q_j , the score-wise recidivism rate for any row in an actuarial table, is the quotient of dividing the number of sex offenders in a row who recidivate ($n_{j,R+}$) by the total number of sex offenders in that row ($N_j = n_{j,R+} + n_{j,R-}$). This calculation is also expressed in the following formula:

$$(1) \quad Q_j = \frac{n_{j,R+}}{(n_{j,R+} + n_{j,R-})}$$

Insert Table 1 about here

The design of Table 1 naturally focuses the attention of readers on frequency data. Conditional probabilities are compiled at the same time that frequency data are compiled, however. For example, the values in the left-most column of Table 1 refer to S_j , or the scores that subjects may obtain on an actuarial. Each S_j may be related to either frequency data or conditional probability data. When all scores are related to conditional probability data, the

expression on the right side of equation (1) is symbolized in the fifth column of Table 1 as $P(R^+|S_j)$ and read as the “estimated probability of recidivism on the condition of each score S_j .” Therefore, although Q_j refers to a frequency analysis and $P(R^+|S_j)$ refers to a probability analysis, the value of Q_j equals the value of $P(R^+|S_j)$ when the two variables are in the same row of an actuarial table.

It is also true that another set of conditional probabilities $P(S_j|R^+)$, read as the “estimated probability of each score S_j on the condition of recidivism,” is known whenever the frequency of recidivists with each S_j (i.e., the intersections of column R_j^+ and each row that is enumerated under column S_j in Table 1) and the total number of recidivists (the intersection of column R_j^+ and the “Total” row under column S_j) are known. The reason for this is that $P(S_j|R^+)$ is the quotient obtained by dividing the number of recidivists with S_j by the total number of recidivists. Similarly, a third set of conditional probabilities $P(S_j|R^-)$, read as the “probability of each score S_j on the condition of nonrecidivism,” is known if the number of nonrecidivists in each score group (i.e., the intersections of column R_j^- and each row that is enumerated under column S_j) and the total number of nonrecidivists (the intersection of column R_j^- and the “Total” row under column S_j) are known.

These considerations indicate that Table 1 may be expanded to call the attention of readers to estimated conditional probabilities as well as frequency data. When this is done, as in Table 2, it is apparent that each

$$(2) \quad P(S_j|R^+) = \frac{n_{j,R^+}}{N_{R^+}},$$

and that each

$$(3) \quad P(S_j|R_-) = \frac{n_{j,R_-}}{N_{R_-}}.$$

Multiplying each side of equation (2) by N_{R_+} , it follows that:

$$(2a) \quad n_{j,R_+} = P(S_j|R_+) \times N_{R_+}.$$

Similarly, multiplying each side of equation (3) by N_{R_-} , it follows that:

$$(3a) \quad n_{j,R_-} = P(S_j|R_-) \times N_{R_-}.$$

Inserting equations (2a) and (3a) into equation (1), it follows that:

$$(4) \quad Q_j = P(R_+|S_j) = \frac{P(S_j | R_+) \times N_{R_+}}{P(S_j | R_+) \times N_{R_+} + P(S_j | R_-) \times N_{R_-}}.$$

Insert Table 2 about here

Equation 4 is a version of Bayes's Theorem that is rarely encountered. The most frequently encountered version (cf., Fienberg, 2006; Iversen, 1984; Woodworth, 2004) of the theorem may be realized, however, by multiplying each side of equation (4) by $(1/N) / (1/N)$, so that

$$(5) P(R_+ | S_j) = \frac{P(S_j | R_+) \times \frac{N_{R_+}}{N}}{P(S_j | R_+) \times \frac{N_{R_+}}{N} + P(S_j | R_-) \times \frac{N_{R_-}}{N}},$$

or, since $\frac{N_{R_+}}{N} = P(R_+)$, which is the base rate for recidivism, and

$\frac{N_{R_-}}{N} = P(R_-)$, which is the base rate for nonrecidivism (and also the remainder of $1 - P(R_+)$),

$$(6) P(R_+ | S_j) = \frac{P(S_j | R_+) \times P(R_+)}{P(S_j | R_+) \times P(R_+) + P(S_j | R_-) \times P(R_-)}.$$

Table 3, compiled for the purpose of illustration from data reported by Hanson (2005)⁴, exemplifies an actuarial table that includes both frequency data and Bayesian coordinates. Inserting the tabled information into equations (1) and (6), it may be seen that the two equations yield the same risk estimates in the right-most column (with any slight differences being due to rounding error). Solving equation (6) for the values in each row points up the equally significant fact that each $P(R_+ | S_j)$ in Table 3, and thus each Q_j (the result of dividing R_{j+} by N_j), depend on the base rate of recidivism ($P(R_+)$; see the intersection of the "Symbol" row and the last column of Table 3), the distribution of scores among recidivists (each $P(S_j | R_+)$ in column 3), and the

distribution of scores among nonrecidivists (each $P(S_j|R-)$ in column 5). As pointed out by Mossman (2006), the result of dividing each $P(S_j|R+)$ by its corresponding $P(S_j|R-)$ is equal to the likelihood ratio for each score (LR_j , presented as the seventh column of Table 3). These likelihood ratios are the slope coefficients for a risk instrument's ROC curve, which may be derived by plotting the value of each $P(S_j|R+)$ and its associated $P(S_j|R-)$ on the y and x axes, respectively, of a two-dimensional graph where the vectors range from zero to 100 percent.

Insert Table 3 about here

Discussion

This paper has presented a mathematical proof and example that the score-wise estimates of sexual recidivism for an actuarial table based on Bayesian coordinates are equal to the estimates for an actuarial table based on frequency data. An actuarial table, in other words, is a data base that is specified in terms of both frequency information (e.g., $n_{j,R+}$, $n_{j,R-}$ and Q_j) and probability coordinates (e.g., $P(R+)$, $P(R+|S_j)$, and $P(R-|S_j)$). Whenever a new subject is added to this table each of the data categories it contains are updated simultaneously. Score-wise recidivism rates may be compiled simultaneously using frequency and Bayesian formulas at any time during the development of a table on a new sample and, when this is done, the obtained results will always be precisely identical.

These considerations point to the conclusion that those who use actuarial tables for the purpose of risk assessment are actually effecting an application of Bayes's Theorem. It is therefore self-contradictory and misleading for any experts who use actuarial instruments to testify that Bayes's Theorem is a novel and idiosyncratic approach to risk assessment. To avoid

this situation, experts who are asked to opine in court as to the novelty and idiosyncrasy of an evaluation method should first determine whether it can be translated into the operations presented in equation (6) or an equivalent set of operations such as those included in the odds ratio version of Bayes's Theorem (Wollert, in press). If so, the method is neither novel nor idiosyncratic.

Conversely, by acknowledging the scientific legitimacy and value of Bayes's Theorem, researchers and expert evaluators gain access to a method of analysis that may be quite advantageous in some situations for reaching certainty opinions about the risk of sexual recidivism in SVP cases. For example, suppose that the five-year base rate for sexual recidivism among prison releasees in California, Iowa, Minnesota, and Washington were found, as seems realistic from documents recently disseminated by state and federal agencies (Adkins, Huff, & Stageberg, December 2000; Padilla, October 10, 2006; Langan, et al., 2005; Minnesota Department of Corrections, April 2007; Washington State Institute for Public Policy, April 26, 2005), to approximate 5% rather than the 12% reported in Table 3. On the assumption that the values of $P(S_j|R_+)$ and $P(S_j|R_-)$ from Table 3 are generalizable to offenders from these states⁵, solving equation (6) shows that the estimated recidivism rate for those with high Static-99 scores does not exceed 15% (i.e. $P(R_+|S_j) = (.219 \times .05) / ((.219 \times .05) + (.066 \times .95)) = .15$). This result, in turn, underscores the extreme caution that experts in these states need to exercise before claiming that they are reasonably certain that a civil commitment candidate is likely to sexually recidivate. More generally, it points to the conclusion that, in the absence of adequate disclosures and qualifications, applying a risk instrument to populations with different recidivism rates than the sample on which it was developed is a very poor professional practice (Amenta,

Guy, & Edens, 2003). It is so poor, in fact, that it potentially violates section 9.02.b of the *Ethical Principles of Psychologists* (American Psychological Association, December 2002), section VII.D of the *Specialty Guidelines of Forensic Psychologists* (American Psychological Association, March 9, 1991), and various standards (e.g., 11.15, 11.16, 11.18, 12.13, 12.16, 12.17) cited in the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999).

Although the foregoing example highlights the power of Bayesian analysis for clarifying recidivism determinations, this approach may also be applied to other conditions of interest in SVP proceedings. In particular, a Bayesian framework may be relied upon to derive certainty opinions concerning whether or not a respondent is positive for a mental disorder, suffers from a volitional impairment, or is predisposed to the commission of sex crimes to such a degree that he represents a menace to the community (Wollert, in press).

At the same time that it conveys numerous advantages, Bayes's Theorem encourages the adoption of certain practices. For researchers, it emphasizes the importance of providing fully-specified data sets, such as that presented in Table 3, and archiving data so that Bayesian coordinates are readily accessible. For experts who are assessing a condition of interest in an SVP evaluation, an integration of the null logic model with Bayes's Theorem underscores the central importance of specifying the base rate of that condition in the most well-defined reference group including the respondent. Furthermore, it focuses their attention on the following questions that have also been alluded to at other points in this paper:

- What is the appropriate level of certainty to use in a particular analysis?

- What evidence suggests that the certainty standard for rejecting the null hypothesis - that is, that the respondent is not an SVP - might be unattainable?
- What is the evidence supporting the affirmative hypothesis that the respondent is an SVP?
- What is the power of the evidence in support of the affirmative hypothesis for detecting the condition of interest being evaluated (i.e., the likelihood ratio, defined in Table 3 as $LR_j = P(S_j|R_+) / P(S_j|R_-)$)?

Surveying the procedures that are commonly used to reach certainty opinions, the British statistician Dawid (2002) observed that “the current state of legal analysis of evidence seems to me similar to that of science before Galileo, in thrall to the authority of Aristotle and loth to concede the need to break away from old habits and thoughts” (p. 71). The Null-Bayes Logic Model discussed in this paper represents a valuable tool for breaking through this impasse and holds great potential for improving the accuracy of certainty opinions as well as their credibility. Consequently, its study and application has much to offer to forensic psychology in general, and to the clarification of issues related to sex offender research and evaluation in particular.

References

- Adkins, G., Huff, D., & Stageberg, P. (December 2000). *The Iowa Sex Offender Registry and recidivism*. Iowa City, IO: Iowa Department of Human Rights.
- Amenta, A. E., Guy, L., & Edens, J. F. (2003). Sex offender risk assessment: A cautionary note regarding measures attempting to quantify violence risk. *Journal of Forensic Psychology Practice*, 3 (1), 39-50.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Assn.
- American Psychological Association (December 2002). *Ethical Principles of Psychologists and Code of Conduct*. Washington, D.C.: American Psychological Association.
- American Psychological Association (March 9, 1991). *Specialty Guidelines for Forensic Psychologists*. Washington D.C.: American Psychological Association.
- Bayes, T. (1764). An essay toward solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53, 370-418.
- Covington, J. R. (1997). Preventive detention for sex offenders. *Illinois Bar Journal*, 85, 493-498.
- Dawid, A. P. (2002). Bayes's Theorem and weighing evidence by juries. In R. Swinburne (Ed.), *Proceedings of the British Academy: Vol. 113. Bayes's Theorem* (pp. 71-90). London: Oxford University Press
- de Finetti, B. (1964). Foresight: Its logical laws, its subjective sources. In H.

- E. Kyburg and H. Smokler (Eds.), *Studies in Subjective Probability* (pp. 91-158). New York: Wiley.
- Doren, D. M. (2002). *Evaluating sex offenders*. Thousand Oaks, CA: Sage.
- Doren, D. M. & Epperson, D. (2001). Great analysis, but problematic assumptions: A critique of Janus & Meehl (1997). *Sexual Abuse, 13*, 45-51.
- Doren, D. M. (2006). Battling with Bayes: When statistical analyses just won't do. *Sex Offender Law Report, 7* (4), 49-50, 60-61.
- Doren, D. M. (December 23, 2004). Affidavit in re the detention of Mitchell Gaff, Cause No. 94-2-07248-8. Superior Court of Washington for Snohomish County.
- Fienberg, S. E. (2006). When did Bayesian inference become "Bayesian." *Bayesian Analysis, 1* (1), 1-40.
- Hanson, R.K. (2005). *The validity of Static-99 with older sexual offenders*. Unpublished manuscript. Ottawa, Canada: Public Safety and Emergency Preparedness Canada.
- Hanson, R. K. (January 21, 2005). Affidavit in re the detention of Mitchell Gaff, Cause No. 94-2-07248-8. Superior Court of Washington for Snohomish County.
- Hanson, R. K. & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24*, 119-136.
- Iversen, G. R. (1984). *Bayesian statistical inference*. Beverly Hills: Sage.
- Janus, E. S. & Meehl, P. E. (1997). Assessing the legal standard for predictions of dangerousness in sex offender commitment proceedings. *Psychology, Public Policy, and Law, 3*, 33-64.
- Jaynes, E. T & Bretthorst, G. (2003). *Probability theory: The logic of science*. New York: Cambridge.

- Jeffreys, H. (1939). *The theory of probability*. Oxford: Oxford University Press.
- Kansas v. Hendricks*, 521 U.S. 346 (1997).
- Langan, P. A., Schmitt, E., & Durose, M. (2003). *Recidivism of sex offenders released from prison in 1994*. Washington, DC: U.S. Department of Justice.
- Miller, H. A., Amenta, A., & Conroy, M. (2005). Sexually violent predator evaluations: Empirical evidence, strategies for professionals, and research directions. *Law and Human Behavior*, 29, 29-54.
- Minnesota Department of Corrections (April 2007). Sex offender recidivism in Minnesota. Retrieved on 7.13.07 from <http://www.corr.state.mn.us/>.
- Mossman, D. (2006). Another look at interpreting risk categories. *Sexual Abuse*, 18, 41-63.
- Padilla, J. (October 10, 2006). *Response to court order by the Honorable Robert A. Barclay*. California: Psychology Department, California Department of Mental Health.
- Quinsey, V. L. (January 14, 2005). Affidavit in re the detention of Mitchell Gaff, Cause No. 94-2-07248-8. Superior Court of Washington for Snohomish County.
- Ramsey, F. P. (1931). *The foundations of mathematics*. London: Routledge and Kegan Paul.
- Savage, L. J. (1954). *The foundations of statistics*. New York: John Wiley and Sons.
- Washington State Institute for Public Policy (April 26, 2005). *Sex offender sentencing in Washington State: Recidivism rates*. Document No. 05-08-1203. Olympia, WA.
- Wollert, R. (2006). Low base rates limit expert certainty when current actuarial tests are used to identify sexually violent predators: An application of Bayes's Theorem. *Psychology, Public Policy, and Law*, 12, 56-85.

- Wollert, R. (in press). Poor diagnostic reliability, the Null-Bayes Logic Model, and their implications for sexually violent predator evaluations. *Psychology, Public Policy, and Law*.
- Woodworth, G. (2004). *Biostatistics: A Bayesian introduction*. New York: Wiley.
- Woodworth, G. G. & Kadane, J. (2004). Expert testimony supporting post-sentence civil incarceration of violent sexual offenders. *Law, Probability, and Risk*, 3, 221-241.

Authors' Note

Portions of this paper were presented at the 2007 meeting of the Western Psychological Association. The authors are indebted to Dr. Jacqueline Waggoner for her comments on an earlier version of the manuscript. Correspondence should be sent to either author. Dr. Donaldson may be contacted by mail at 350 Arbutus Ave., Morro Bay, CA 93442, by telephone at (805) 772-5086, and by e-mail at theodoredonaldson@sbcglobal.net. Dr. Wollert may be contacted by mail at 602 E. 31st St., Vancouver, WA 98663, by telephone at (360) 737-7712, and by e-mail at rwwollert@aol.com.

Footnotes

¹ These comments indicate that the standard an evaluator adopts to define “reasonable certainty” is critical for determining the status of any given null hypothesis. With a low standard, the null hypothesis would be rejected for almost all SVP candidates. With a high standard, the null hypothesis would be rejected for a much smaller number. Regarding future risk of recidivism, many laws define the standard as “likely” or “more likely than not,” suggesting a certainty threshold in excess of 50% (Woodworth & Kadane, 2004). Statutes leave evaluators to set other certainty standards, however, for evaluating the last three concepts referenced in the excerpt from Miller et al. (2005) that we cited in the opening paragraph. With respect to this issue the authors believe that experts should generally entertain a certainty standard on the order of 95% to 99% for three reasons. One is that this level closely approximates the analogous legal standard of “beyond a reasonable doubt.” Another is that this threshold is customarily adopted in psychological research. The last is that we feel that indefinite deprivation of liberty, which is the practical result of post-prison civil commitment, warrants a high level of certainty because it ranks among the most serious of legal dispositions. Other forensic mental health professionals may disagree with us on this issue. If so, it would be clarifying if they were to specify an alternative set of certainty standards alongside arguments in their defense.

² The assertion that civil commitment risk evaluators do not make recidivism predictions seems predicated on the belief that they avoid unqualified endorsements of declarative sentences about the future (e.g., that the defendant will recidivate). Considering both qualified and unqualified endorsements, it is important to keep in mind that the behavior of SVP evaluators is

clearly an exercise in the application of probability theory. According to Woodworth (2004), probabilities invariably refer to the strength of belief that may be placed in specified declarative sentences (e.g., the defendant committed a sex offense in the past, the defendant is positive for a mental abnormality in the present, or the defendant will recidivate in the future). Certainty opinions of evaluators therefore always have some reference to a declarative sentence. However, expert testimony is not prophecy and all expert conclusions rest on notions of statistical prediction, which necessarily includes uncertainty. Contending that prediction falls in a separate and unique category, and that the concepts of predictive science are only applicable to endeavors directed towards making unqualified statements, is therefore unrealistic, restrictive, and inaccurate.

³ Bayesian philosophy holds that some assumptions are more reasonable than others because of evidence or experience and that analysis of the viability of a theory must be based only on reasonable assumptions. Analyses that proceed from questionable assumptions are, in other words, open to question. In their critique of Janus & Meehl (1997), for example, Doren and Epperson (2001) grounded their base rate assumptions in old data that were collected on non-representative samples and analyzed with techniques (survival analyses) that were incompatible with the recidivism concepts used by Janus & Meehl. They also did not verify that the level of test accuracy under their base rate assumptions was comparable to the level of test accuracy calculated by Janus & Meehl. Their analysis, as a result, did not address the content of Janus and Meehl's paper. It is therefore essential, whether one is conducting or rejoining a Bayesian analysis, to determine that the "foundational probabilities" that are relied upon are relevant to the issue at hand. This is a particularly important principle for those working in the SVP field to

keep in mind in that speculative criticisms of Bayesian results probably have the net effect of discouraging the application of these methods to a field where they have typically been underutilized. This, in turn, would almost certainly limit development of the field's knowledge base.

⁴The original experience table for Static-99 was based on several different samples for which most data were collected over 25 years ago. A five-year recidivism rate of 18% was obtained when these samples were first combined by Hanson and Thornton (2000). Hanson (2005) has recently disseminated an updated experience table for Static-99 based on data from a more contemporaneous set of samples. The decrease in recidivism from 18% for the original experience table to 12% for the updated version presented as Table 3 is one more indicator of the reduction in recidivism that has occurred with time. In line with our analysis of the extent to which a base rate reduction results in a reduction in recidivism probabilities, the recidivism probabilities associated with high scores for the updated Static-99 are considerably less than those reported in the original version. In fact, they are sufficiently reduced that it is doubtful that any long-term recidivism probabilities for the samples in the updated version of Static-99 will be extremely high.

⁵This is a reasonable assumption when ROC curves for the developmental sample are similar to those for the application sample.

Table 1

Formulas for the Derivation of an Actuarial Table on the Basis of Frequency Information

S_j	R_{j+}	R_{j-}	N_j	Recidivism Rate (Q_j or $P(R+ S_j)$)
4	$n_{4,R+}$	$n_{4,R-}$	$n_{4,R+} + n_{4,R-}$	$n_{4,R+} / (n_{4,R+} + n_{4,R-})$
3	$n_{3,R+}$	$n_{3,R-}$	$n_{3,R+} + n_{R-}$	$n_{3,R+} / (n_{3,R+} + n_{3,R-})$
2	$n_{2,R+}$	$n_{2,R-}$	$n_{2,R+} + n_{2,R-}$	$n_{2,R+} / (n_{3,R+} + n_{3,R-})$
1	$n_{1,R+}$	$n_{1,R-}$	$n_{1,R+} + n_{1,R-}$	$n_{1,R+} / (n_{1,R+} + n_{1,R-})$
Total	$\sum n_{j,R+}$	$\sum n_{j,R-}$	$\sum n_{j,R+} + \sum n_{j,R-}$	$\sum n_{j,R+} / \sum (n_{j,R+} + n_{j,R-})$
Symbol	N_{R+}	N_{R-}	N	Q

Note: The abbreviations in the header stand for the following variables: S_j = a score of j (j may range from 1 to 4 in the table above); R_{j+} = number of recidivists with a score of j ; R_{j-} = number of nonrecidivists with a score of j on test S ; N_j = number of offenders with a score of j ; Q_j = the recidivism rate for offenders with a score of j . The number of recidivists, nonrecidivists, and offenders are calculated in second, third, and fourth columns of the "Total" row, and the score-wise recidivism rate is calculated in the fifth column. Symbols that refer to each of these quantities are included in the "Symbol" row. Q is the sample base rate of recidivism.

Table 2

An Actuarial Table That is Fully Specified in the Sense of Presenting Both Frequency Information and Probability Coordinates

S_j	R_{j+}	$P(S_j R_+)$	R_{j-}	$P(S_j R_-)$	N_j	$Q_j = P(R_+ S_j)$
4	$N_{4,R+}$	$n_{4,R+} / N_{R+}$	$n_{4,R-}$	$n_{4,R-} / N_{R-}$	$n_{4,R+} + n_{4,R-}$	$n_{4,R+} / (n_{4,R+} + n_{4,R-})$
3	$N_{3,R+}$	$n_{3,R+} / N_{R+}$	$n_{3,R-}$	$n_{3,R-} / N_{R-}$	$n_{3,R+} + n_{3,R-}$	$n_{3,R+} / (n_{3,R+} + n_{3,R-})$
2	$N_{2,R+}$	$n_{2,R+} / N_{R+}$	$n_{2,R-}$	$n_{2,R-} / N_{R-}$	$n_{2,R+} + n_{2,R-}$	$n_{2,R+} / (n_{2,R+} + n_{2,R-})$
1	$N_{1,R+}$	$n_{1,R+} / N_{R+}$	$n_{1,R-}$	$n_{1,R-} / N_{R-}$	$n_{1,R+} + n_{1,R-}$	$n_{1,R+} / (n_{1,R+} + n_{1,R-})$
Total	$\sum n_{j,R+}$	100%	$\sum n_{j,R-}$	100%	$\sum n_{j,R+} + n_{j,R-}$	$\sum n_{j,R+} / \sum (n_{j,R+} + n_{j,R-})$
Symbol	N_{R+}		N_{R-}		N	Q

Note. S_j , R_{j+} , R_{j-} , N_j , Q_j , and $P(R_+|S_j)$ refer to the same concepts as they did in Table 1. $P(S_j|R_{j+})$ stands for the probability of S_j on the condition of R_{j+} . $P(S_j|R_{j-})$ stands for the probability of S_j on the condition of R_{j-} .

Table 3

An Example Of A Fully-Specified Actuarial Table Showing That Score-wise Recidivism Rates Are Calculated Either By Solving Bayes's Theorem or By Adding and Dividing Frequency Information (Source: Hanson, 2005)

S_j	R_{j+}	$P(S_j R+)$	R_{j-}	$P(S_j R-)$	N_j	LR_j	$Q_j = R_{j+}/N_j = P(R+ S_j)$
4	92	.2190	199	.0660	291	3.3182	.316
3	157	.3738	575	.1908	732	1.9591	.214
2	114	.2714	1,198	.3975	1,312	.6828	.087
1	57	.1357	1,042	.3457	1,099	.3925	.052
Total	420	100%	3,014	100%	3,434		.120
Symbol	N_{R+}		N_{R-}		N		$Q = P(R+)$

Note: The abbreviations in the first 7 header columns stand for the following variables: S_j = a score of j on Static-99 (4 = high, 3 = moderately high, 2 = moderately low, 1 = low); R_{j+} = number of recidivists with a score of j on Static-99; $P(S_j|R+)$ = the probability of S_j on the condition of $R+$; R_{j-} = number of nonrecidivists with a score of j on Static-99; $P(S_j|R-)$ = the probability of S_j on the condition of $R-$; N_j = the number of offenders with a score of j on test S ; LR_j = the "likelihood ratio" of $P(S_j|R+) / P(S_j|R-)$. The far right column of each row reports the recidivism rates for each score, which can be obtained either from the frequency data or Bayes's theorem, i.e., the probability estimates. The formula for calculating risk percentages (Q_j) therefore equates to the formula for calculating $P(R+|S_j)$. The number of recidivists, nonrecidivists, and offenders are calculated in the "Total" row, and the sample-wise recidivism rate (Q or $P(R+)$) is calculated in the last column of the row. Symbols that refer to each of these quantities are included in the "Symbol" row. The base rate for nonrecidivism ($P(R-)$) is equal to either $1-Q$ or $1-P(R+)$.