

DRO

Deakin University's Research Repository

Coyle, Ian R., Halon, Robert L., Campbell, Terence W., Thomson, Donald M. and Woskett, Jessica
2016, Alice in recidivism land: The Queens logic and child protection workers' assessment of sexual
dangerousness, *American journal of forensic psychology*, vol. 34, no. 1, pp. 5-36.

This is the published version.

©2016, American Journal of Forensic Psychology

Reproduced with the kind permission of the copyright owner.

Available from Deakin Research Online:

<http://hdl.handle.net/10536/DRO/DU:30090640>

**ALICE IN RECIDIVISM LAND: THE QUEENS LOGIC
AND CHILD PROTECTION WORKERS' ASSESSMENT
OF SEXUAL DANGEROUSNESS**

**Ian R. Coyle, Ph.D., Robert L. Halon, Ph.D., Terence W. Campbell, Ph.D.,
Donald M. Thomson, Ph.D., LL.B., Jessica Woskett**

This article is based on a case that highlights the dearth of cogent and pertinent risk assessment information in the reports and testimonies of many involved in assessing risks to children in real life problematic circumstances. In the case, the decision to exclude an intrafamilial sexual offender from the family home was made in large measure on the basis that the offender's wife accepted the offender's "denial of guilt" to accusations previously made against him. Keeping families apart should not be entertained without reliable and valid evidence pointing to that decision. However, as will be seen, the evidence often relied upon by child protection workers, albeit in good faith, is neither. The assessment processes described in the case point directly to what appears to be a wider lack of knowledge specific to assessment of recidivism, to misrepresentation of risk assessment information, and to overarching epistemological issues that appear to be widely ignored and/or misunderstood within the overall field of risk assessment and threat management. The purpose of this article is to describe how the information used to decide upon exclusion in the case cannot be considered probative and to iterate the methodological processes that must be considered in such cases if miscarriages of justice are to be avoided.

THE QUEEN'S LOGIC

"...There's the King's Messenger. He's in prison now, being punished: and the trial doesn't even begin till next Wednesday: and of course the crime comes last of all." "Suppose he never commits the crime?" said Alice. "That would be all the better. Wouldn't it?" the Queen said... Alice felt there was no denying that. "Of course it would be all the better," she said, "but it

wouldn't be all the better his being punished." "You're wrong there at any rate." said the Queen. "Were you ever punished?" "Only for faults" said Alice. "And you were all the better for it, I know!" the Queen said triumphantly. "Yes, but then I had done the things I was punished for." said Alice. "That makes all the difference." "But if you hadn't done them," the Queen said, "that would have been better still: better, and better, and better!" (Lewis Carroll, *Alice's Adventures in Wonderland*)

The parole period of a father was set to expire. He had been convicted of having "consensual" sexual intercourse with his then-15-year-old stepdaughter who had subsequently returned to live with her biological father. The offender had been on parole for two years, during which he had met all of his parole conditions and developed a most successful business. Three professionals, who were senior in terms of qualifications and experience, including the first author of this article, had independently assessed the father for release on parole and were in agreement that his risk of recidivism was low. Representatives of Child Safety Services (CSS), Department of Community Services, Queensland, objected to the father returning to reside with his family and, instead, applied for a Court order requiring him to have only third party supervised access to his biological children until they attained the age of sixteen. That the offender's wife had recently given birth to their latest child made the potential consequences of the CSS application profoundly more negative for the offender and his family.

A few months before the term of father's full sentence was to expire, his parole officer—also an employee of the Department of Community Services—applied for an order prohibiting the father from attending the birth of his latest child on the basis that attending the hospital would pose an unacceptable risk to the newborn. The application failed before the Supreme Court of Queensland. Undeterred, the parole officer then filed a formal complaint to Child Safety Services (CSS) stating that the father still posed a risk of recidivism, in particular against his biological offspring.

A social worker at CSS who was junior in terms of qualifications and experience became the Applicant to The Court in order to prosecute the parole officer's "high risk"; the contention being that there was an unacceptable risk of recidivism on the basis that the offender's wife accepted his denial of guilt. In an Affidavit filed in the Children's Court of Queensland, the social worker initially argued, on an *ex parte* basis, that because the father's wife believed him to be innocent she would not demonstrate adequate "protective factors" and, further, that there was considerable scientific justification for her position. Pursuant to a court order arising from the application, another psychologist, junior in terms of qualifications though not in experience, then assessed the offender, his wife and children. In a Kafkaesque twist to what was already a bizarre case, the psychologist started his assessment session with the offender's wife by showing her a photograph of two copulating kangaroos. The psychologist based his overall argument to a significant degree around the false idea that denial of risk does predict recidivism, yet concluded that the offender posed a low risk of recidivism; the "denial raises risk" issue will be discussed in more depth later in this article

Counsel for the respondent father argued, in essence, that the applicant's opinions of the father's risk of recidivism and dangerousness to his biological children were ill founded, because they were not supported by science or other empirical evidence. The court case largely involved expert evidence that is of significant moment in proceedings of this specific type and generally insofar as the court's reception of expert evidence sought to be adduced by Child Protection Service (CPS) workers and other professionals is concerned.

THE CSS CASE FOR EXCLUDING SEXUAL OFFENDING PARENTS FROM THE FAMILY HOME: APPLYING THE QUEEN'S LOGIC

Despite the lack of empirical evidence to support the risk proposition or justify "no contact" or "supervised contact" orders, CSS maintained there are a number of propositions that underlie their conclusions that such limitations are the best option. That is, even though the

decision to exclude an intrafamilial sexual offender from the family home, or impose other draconian sanctions such as requiring only supervised contact, may be empirically unfounded, it is, nonetheless, the best option. For example, Wilson (1), in harmony with the Queen's logic, noted that such decisions would often be undertaken on the basis of inadequate information and great urgency:

“...the fact that an allegation may later prove unfounded should not, by itself, dissuade us from using this remedy. These error costs are no greater than those that occur when the state removes a child who is later found not to have been abused.”

The basis for this best option argument is, so it is asserted, found by considering a number of factors commonly cited by Child Protective Services workers to justify their risk assessment decisions in such cases. The similarity of this basis to circular argument is hard to deny.

Although the factors said to be commonly cited are not independent, for convenience they are dealt with in what follows in the order that they typically appear in CPS workers' applications for such banishment. Sexual Abuse Counselling Services, an organization within the Department of Communities, which provides, so it is asserted, independent advice to CSS in Queensland, routinely cites these factors as justification for assigning severity of risk

CROSS-OVER AND RECIDIVISM IN INTRAFAMILIAL SEX OFFENDERS

CPS workers often note that the risk of recidivism is “almost guaranteed” and cite older publications on convicted offenders who have, usually anonymously, admitted to other offenses or have been convicted of additional offenses while incarcerated. Indeed, the authors have seen exactly this “almost guaranteed” phrase employed on numerous occasions even in the complete absence of any interview or assessment of the alleged or convicted offender. The argument advanced by CPS workers to support this contention is based, *inter alia*, on the concept of “cross-over”; that is, the propensity of intrafamilial sexual offenders

to offend against different genders, children of different ages and non-biological/biological children.

While it is true that some intrafamilial sexual offenders are serial offenders, this is by no means the norm and such conclusions are drawn from studies of offenders with particular characteristics. For example, CPS workers frequently quote Studer et al. (2) in which it is reported that of 328 convicted sex offenders who participated in the study (150 were intrafamilial offenders) 22%—72—of the intrafamilial offenders claimed to have had other intrafamilial victims. The fact is, therefore, 78% were not convicted of intrafamilial child sexual abuse. Compared to the intrafamilial offenders, the inter-familial offenders' rate of self-disclosure and/or proven convictions for other victims was 12.9%. These results have, broadly speaking, been replicated in more recent studies (3-5). We want to point out that, in the above discussion, the Queen's logic exists in comparing apples (claiming to have) to oranges (not convicted of).

Another significant contamination in the risk assessment literature concerned with calculating and yielding percentages of risk is the choice of terms such as "claims to (or known to) have done such and such" with which to determine the number of recidivists in the lot, versus counting the number in the developmental sample who were "not convicted of such and such" to represent non-recidivists. That lack of comparison alone ("claims to" versus "not convicted") makes actual existing base rates the overwhelmingly most reliable basis on which to estimate the probability of recidivism. Said another way, the language of actuarial risk assessment itself is its own demise; the base rates of re-offense, whether defined by arrests, charges, convictions, or diversionary dispositions, is the most valid and reliable factor for estimating the risk posed by an individual.

Simply put, with the caveat that using disparate definitions makes findings of much of the research unreliable, depending on the study cited some 20-24% of offenders have reported after being convicted that they offended against children who were: of the opposite gender;

different in terms of their biological relationships to the offender; or materially different in age range from their victim in the index offense for which they were convicted (3-5). It is critical to note, however, that all of these studies are based on the worst of the worst samples of sexual offenders; typically, those being processed for civil commitment at the conclusion of their custodial sentences. That is, offenders who are being assessed under more or less parallel legislation in overseas jurisdictions to the Queensland Dangerous Prisoners (Sexual Offenders) Act 2003 (Sexually Violent Predator involuntary commitment legislation in the U.S.).

Further, because incest is usually defined in the scientific literature as involving the extended family, including stepchildren, it is simply not possible to say with certainty what the percentage of cross-over offending from biological to non-biological children and vice versa really is. In fact, alternative terms used in the scientific literature, often seemingly coined *ad hoc*, to define intrafamilial offending as involving “related victims” can include the most minimal relationship by consanguinity, such as non-blood related cousins. It seems obvious that sexual offending against a biological child and then progressing to offending against a non-biological family member, including, say, a cousin, is vastly different than the obverse. This is a lacuna in the scientific literature that does not appear to be addressed by most commentators on this subject. Yet, because of differential degrees of inhibition, the cultural if not genetic imperative against incest, in the biological sense of the word, is strong (6). Hence, it would be expected that progression from sexual offending against biological children to sexual offending against stepchildren/extended family members would be easier and occur much more often than the obverse.

In another example of the Queen’s logic, on the basis of studies of demonstrably dangerous and/or serial offenders, the authors have seen CPS workers assert that the pattern of cross-over offenses so identified (i.e., in that population sample) proves that all intrafamilial offenders are at greater risk of reoffending when compared to other sexual of-

fenders. This claim is not justified. Conclusions drawn from a set of data that is based on one type of offender (e.g., demonstrably dangerous) cannot be considered, *ipso facto*, to apply to any other type of offender. Studies based on such retrospective data analysis also prove nothing: for those convicted of intrafamilial sexual abuse, the base rate of recidivism (the probability that any randomly selected intrafamilial sex offender will recidivate either generally or with respect to a specific offense within a defined time) is overwhelmingly more important in estimating future risk than is admission of prior offenses. Ignoring the effect of the base rate of the occurrence of any phenomena, as seems to occur routinely in these CPS cases, bedevils all areas of science (7-14).

Ignoring the effect of the base rate on the probability of any event occurring is referred to as base rate neglect. This is a ubiquitous cognitive heuristic that is inimical to proper assessment of the most reliable probability estimate of an event occurring. Ignoring or not mentioning base rate information has long been recognized as a fundamental problem leading to massive unreliability in the proffered estimate of risk, usually overestimation of the true probability of the occurrence of any event, such as recidivism in sexual offenders (11, 15). Neal and Grisso provide an example that neatly demonstrates base rate neglect:

“For example, Casscells, Schoenberger, and Graboys (1978) asked Harvard Medical School faculty, staff, and 4th-year medical students the following question:

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming that you know nothing about the person’s symptoms or signs?
_____%

The correct Bayesian answer under the most plausible interpretation of the problem is about 2%. Specifically, 51 people out of 1000 would test positive (1 true positive and 50 false positives). Of the 51 people with positive tests, 1 would actually have the

disease. Expressed as a proportion, this is $1/51 = 0.019$ or 1.9%. But only 18% of the Harvard-affiliated participants had an answer close to 2%. Forty-five percent of this distinguished group said the answer was 95%, thereby completely neglecting the base rate information” (11, page 203).

Base rates of recidivism as measured by longitudinal studies after offenders have been convicted are not affected by admissions of prior offenses. It proceeds from this that the focus of CPS workers on the post-conviction findings of other sexual offenses by intrafamilial sexual offenders has absolutely no relevance to the base rate of future offenses. Base rates will remain unperturbed by *ex post facto* academic studies. However, when assessing the risk of recidivism, great mischief arises from conflating the results of studies measuring offenses occurring prior to conviction with offenses reported and/or occurring post-conviction.

A further problem arises when relying on data obtained from convicted offenders while they are in a treatment programs. In such programs there is compelling pressure on offenders to acknowledge guilt because unless guilt is acknowledged they simply cannot progress to release. It has been demonstrated that, under such circumstances simply to get on with “treatment,” offenders will often admit to behavior that is remote to the point of fantasy (16). Accordingly, considerable skepticism is justified and necessary when the rate of crossover from one type of sex offense to other types is purported to have been identified in studies aimed at teasing out the relationships between variables related to intrafamilial offending. There is no reason to believe that recidivism rates based on that kind of research are reliable—usually representing significant overestimates—and every reason to believe that they are unreliable. Conversely, offenders deny their offenses with authority figures they do not trust but acknowledge them to those they do trust. Put this way, denial is not necessarily dichotomous but may represent a continuum, such that exactly when minimi-

zation of the severity of an offense becomes denial is a matter of conjecture.

The fact that some 20-24% of the “worst of the worst” of intra-familial sex offenders may demonstrate crossover has no bearing on the base rate of crossover occurring in the group of intrafamilial sexual offenders as a whole. One reasonable estimate of crossover offending is in the order of 5% over 5 years and 10% over 19 years (17). Yet, based on more recent data, those recidivism rates will significantly overestimate the actual recidivism rate because, for one thing, they are based on outdated (1982-1992) longitudinal studies. It is well established that the base rate for sexual recidivism of all types has been declining precipitously (18, 19). Although the reasons are not fully understood, the base rate decline is up to 59% when comparing current samples on commonly used actuarial measures that are based on data that is more than 10 years old, such as the Static-99 in its various iterations, and other nomothetic tests employed to predict recidivism, such as the Risk Matrix-2000 (RSM-2000) (9, p. 274).

Generally, the base rate for recidivism of all sexual offenders of the same age range as the offender in this case (he was 39 at the time of the application that is the subject of this article) is in the order of 7.6% according to the age-stratified data provided by Wollert and colleagues (20). Their data are based on an eight-year follow-up sample of 9,305 sexual offenders and are the most comprehensive and thoroughly analyzed, and include current data in which age stratification is included. However since the data on which this test (MATS-1) was developed involved longitudinal studies conducted over the past 17 years, it will also overestimate recidivism due to the falling rate of recidivism, at least among sexual offenders in first-world countries. Absolutely no mention was made of this seminal fact in the material contained in the affidavit filed by the applicant.

THE ACCURACY OF CHILD PROTECTION WORKERS' JUDGMENTS

Early studies demonstrated that “chance” was the mean accuracy of CPS workers in predicting future harm of all types. For example, Ruscio (21) in a review of studies evaluating the accuracy of CPS workers in determining whether to place children into foster care found that 48% of placements were unnecessary and 45% of children who should have been placed in foster were not. The criteria for making these decisions were the identified presence, or absence, of maltreatment of one type or another. This finding has since been repeatedly demonstrated in numerous studies, most particularly with respect to CPS workers making judgments about abuse of whatever type. For example, Gambrill and Shlonsky (22) have argued that 22% to 42% percent of child protection cases are incorrectly classified vis-à-vis abuse. More recent studies involving assessment by CPS workers of putative sexual abuse (23, 24) have confirmed this pessimistic assessment.

In addition to base rate neglect, various biases affect the objective assessment of risk by CPS workers just as such biases affect all non-empirically derived assessments (22, 25, 26). Other important biases in the forensic context are confirmatory bias (paying attention to things that confirm preexisting bias and/or an initial preconception while ignoring those things which disconfirm the bias and preconception) and overestimation bias, which is closely related to base rate neglect. The latter is routinely observed when the base rate of the occurrence of an event, in the cited case sexual recidivism, is low. In this situation CPS workers and mental health professionals involved in risk assessment of all disciplines routinely overestimate the likelihood of the event occurring (27). In addition, numerous studies have demonstrated that, despite the accuracy of decisions made by CPS workers being at chance level or, at best and only sometimes, accurate about 75% of the time, CPS workers express the utmost confidence in their judgments. Unequivocal evidence from empirical studies on CPS practice shows that such confidence is not related to performance (e.g., Regehr, et al. [24]).

A series of studies published in the social work domain have demonstrated that CPS workers routinely conflate their perceptions of accuracy of judgment with their personal convictions as to the accuracy of their judgments (22, 26, 28). Nowhere is this inflated and unfounded confidence in their judgments more apparent than in intrafamilial child sexual abuse cases where blame is attributed to the non-offending mother. In the Australian forensic context, Breckenridge and Baldry (29) reported that 61% of CPS workers *felt* that some mothers were aware of the abuse while 10% of CPS workers *believed* that the majority of mothers were aware of the abuse and did nothing (felt and believed being the operative concepts). There is absolutely no evidence to support these well-entrenched perceptions about what the mothers knew and to what degree what they knew or did not know contributes to reliable assessment of risk. Nonetheless, such contentions continue, in one shape or form, to this day. In fact, it is well established that up to 95% of mothers were not aware of the abuse prior to it being discovered by others (See, generally, Wilson, [1, 30])

The unwarranted confidence of CPS workers in the accuracy of their judgments vis-à-vis recidivism of intrafamilial sexual offenders is particularly inimical to objective risk assessment when CPS workers consider the involvement of the non-offending parent, overwhelmingly the mother. Seminally, Regehr et al. noted, “The perception that the mother denied or minimized the abuse increased the worker’s confidence that the child was at risk” (24, p. 626).

Despite this confidence, we could find not one single, reputable study in the entire scientific lexicon that supports this well-entrenched belief among CPS workers. What might explain this demonstrated lacuna in the knowledge of doubtless well-intentioned professionals? One obvious explanation for this ill-founded belief among CPS workers and other professionals—and the public at large for that matter—seems to be, first and foremost, the entrenched desire to protect children. With that over-riding goal (i.e., always err on the side of caution), conflating the results of crossover offending among intrafamilial

child sex offenders with the base rate of recidivism in this group of offenders follows almost organically. Another possible explanation involves ignorance of the low base rate of recidivism of child sexual offenders generally, leading to overestimation bias. Another explanation involves gross ignorance of fundamental statistical concepts that seem to bedevil professionals of all disciplines who make judgments about all types of recidivism. Nowhere is this lack of understanding of basic statistical concepts more important than in attempting to extrapolate from individual indicia that have been identified as being correlated to one degree or another with sex offense recidivism.

DENIAL OF GUILT AND RECIDIVISM IN SEXUAL OFFENDERS

Research into the importance of denial of guilt as a risk factor for sexual offense recidivism informs that it is not a significant predictor of sex offense recidivism and may even be negatively correlated with recidivism (31-33)—perhaps for the prosaic reason that those who most strongly deny their guilt are, in fact, innocent. Leaving putative innocence aside, it is critical to recognize that of all so-called risk factors, none, including base rate, which is the most reliable single predictor, is so strongly correlated with recidivism that it can be considered a reliable or valid predictive factor in isolation from all other factors. In a seminal meta-analysis based on 28,972 sexual offenders, Hanson and Bussière (31) identified some 57 factors that correlate with recidivism in sexual offenders, denial of guilt being one of them. Save phallometric measurement, which is impractical to employ in the overwhelming majority of cases, none of the factors emerging from their study correlated more than 0.19 with recidivism. This means that the highest amount of the total variability accounted for by any one factor—routinely considered to be the number of prior offenses—was only 3.6%.¹ Hanson and Bussière (31) obtained correlations between denial

¹ Technically this is referred to as the coefficient of determination. This is defined as 1-minus the correlation coefficient squared, where +/-1 is defined as perfect correlation.

and sex offender recidivism that ranged from -0.05 to 0.09 for a 95% confidence interval. In other words, denial may ever-so-slightly increase or decrease recidivism and risk but we don't know which.

More recent studies have broadly confirmed these findings (32, 33). With the exception of age at 65 years or older, the period in which the rate of sexual recidivism approaches zero, it is absurd to base a risk assessment overwhelmingly on any one factor, whatever that factor is. Acceptance of the offending partner's denial of guilt was not identified in the meta-analyses of Hanson and Bussière (31) or Hanson and Morton-Bourgon (32, 33) as having any relationship to recidivism. No subsequent comprehensive study has seriously challenged these findings.

There is, to the authors' knowledge, only one study that claims to show that denial predicts recidivism in any way (34). Nunes et al., claim to have demonstrated that with a total sample of 489 sex offenders, of whom 244 were intrafamilial offenders, those who were assessed as low risk compared statistically to the group more likely to recidivate than those who were high risk. There are a number of compelling problems with this study that involve idiosyncratic and general epistemological issues that create a virtual checklist of factors appearing in risk assessment research that must be carefully analyzed before any degree of reliability or validity can be assigned to both its methodology and results. Even were the methodology and definitions acceptable, the effect reported in terms of relative difference amounts to only a 5% to 10% absolute difference in recidivism rates of comparable populations of offenders.

A TEMPLATE FOR ASSESSING RELIABILITY AND VALIDITY IN RISK ASSESSMENT

We devote considerable space to analysis of the Nunes et al. (34) study, concluding that, whether taken *in toto* or each factor independently, there are compelling grounds to be skeptical about this study if not reject it altogether. Hence, this study serves as a veritable template for assessing the reliability and validity of a study in risk assessment.

The study's authors state that, to their knowledge, the statistical analysis they employed has never, before or since, been employed in such a study. Indeed, those authors noted (p. 100) that the procedures they employed were so unusual that they were verified through "consultation with an expert in medical statistics" who was anonymous, which further weakens the scientific underpinnings of the research. There is no evidence presented to show that the statistical methods employed comply with the mathematical assumptions necessary for applying interval level measurements to such small samples vis-à-vis some of the subgroups sampled.

To compound the methodological errors, assessments of "low risk" and "high risk" were made using a "modified RRASOR", which was not adequately described by the authors. In fact, it has never been adequately described or subjected to peer review to the best of our knowledge. In addition, the authors relied upon psychometric tests for psychopathy (PCL-R) that have been completely discredited in predicting recidivism in any particular individual (9, 35, 36). Finally, in this study, confirmatory bias would be difficult, if possible, to control (37, 38).

The Odds Ratio identified in this study is, at best, 2.9:1. Nunes and colleagues' definition of the "odds ratio" is:

"The odds ratio can be interpreted as an increase or decrease in the predicted odds of recidivism, or in the case of a dichotomous predictor (e.g., denial) as the odds of recidivism in one group as compared to the other (e.g., denier vs. admitter). An odds ratio of 1.00 would reflect no relationship between the predictor and the outcome." (34, page 96).

It might seem that an Odds Ratio of 2.91 indicates that the deniers are 2.91 times more likely to recidivate,² which would be wrong since it fails to take account of the underlying base rate of recidivism. Consid-

² The Odds Ratio is not the same as the Risk Ratio. With low base rates the Risk Ratio will overestimate, often significantly, the Odds Ratio.

ering a base rate of 8% (at the upper end of current base rate tallies) and applying Bayes' Theorem,³ it can be calculated that the probability of making a correct decision with a base rate of recidivism taking into account a test which has a demonstrated Odds Ratio of 2.91 is only 0.15 (14). In other words, on the basis of their reported results, the probability of being correct in arguing the offender was likely to recidivate because he denied guilt is only 15%. Generally a risk factor must have a very strong relative risk to be useful as a screening instrument at the individual level. To attain the minimum default scientific default standard of 95% accuracy (i.e. $p=0.95$), an Odds Ratio of 200:1 is needed with such a low base rate according to some authors, although others set the bar significantly lower (35, p. 261).

The definition used for related victims is overly broad: "biological children, step-children, adopted children, nieces, nephews, cousins, grandchildren or siblings in their index offense" (34, p. 94). On this basis alone, the generalizability of this study vis-à-vis the probability of an offender convicted for sexually offending against a stepchild then recidivating against a biological child is called into serious question. The definition of "denial" includes those who partially denied their guilt. Since "partial deniers" are often simply minimizing the severity of their offense, the issue of the import of "denial" is hopelessly confounded (39, 40).

The reported rate of recidivism in the group identified as "low risk" was, depending on which base rate data is relied upon, some four to six times that which would be expected of intrafamilial sex offenders as a group. Nunes et al. state specifically that:

"Readers should be cautioned that the study *intentionally* over-sampled recidivists and the proportion of sexual recidivists

³ This is mathematically unassailable. As Wollert (14, p. 176) notes: "In general, Bayes' Theorem (Bayes, 1764) is a tool for assessing the probability that a theory—for example, that a person with heart disease will die in 5 years—is true when considered in light of the diagnostic accuracy...of some piece of evidence, such as a disease criterion or a test score, and what is known about the overall, or base rate, probability of the focal outcome."

(32.9%; 24/73) and would not be representative of any particular population [emphasis added]” (34, p. 99).⁴

On this basis alone, this study should not be relied upon to draw conclusions about the recidivism risk of sexual offenders in the instant case or for the overwhelming majority of sexual offenders as the base rate of sex offense recidivism is four times lower than the rate reported in this study.

THE (LACK OF) REPLICABILITY CRISIS IN SCIENCE

The abject failure to systematically analyze the manifest failings of this research by any of those who have cited the Nunes et al. (34) study illustrates a number of insidious, poorly understood and ubiquitous problems in science generally and forensic science in particular. It may fairly be regarded as an outlier in statistical terms; that is, a result so unusual that it should not be relied upon without replication; an issue directly addressed by the authors’ explicit caution as to the limitations of their study. It is therefore alarming that recent studies in related areas by Australian forensic psychologists with considerable experience are silent on this point (41). They are not alone in this. At the time of writing this article, we found no evidence that any of the 64 studies that have cited Nunes et al. (34) has properly considered the profound methodological limitations of this study that proceed as an inevitable consequence of the *intentional* oversampling, nor its other methodological failings. It is tempting to write this off as inadequate attention to methodological and professional detail, as was evidenced by the psychologist in the instant case opening his interview/examination of the offender’s spouse with the picture of copulating kangaroos. However, writing off inadequate analysis as “inadequate attention” would be a mistake as not one of the studies citing this article with approba-

⁴ Thus, the sample size of recidivists classified as “low risk” that recidivated was only 23, according to the authors. This is an abysmally low sample size on which to seek to overturn empirical evidence based on over thirty thousand sexual offenders who have been the subject of major, long-term, studies. See, generally, Hanson and Bussière (31) and Hanson and Morton-Bourgon (32, 33).

tion has addressed the more fundamental problems with replication. This is a pan-global problem for science (42, 43). Amazingly few journals printing psychological and/or risk assessment issues publish replication studies; most want only “original” work. Without confirming replication studies, however, all such research and theory are merely exploratory and cannot be considered valid as courtroom evidence.

The value of direct replication has been at the forefront of recent scientific debate that contends that the credibility of research in the behavioral and social sciences is compromised by a culture of enquiry that promotes inflated effects and false-positive errors (44, 45). Systematic reforms in scientific practice are advocated on the basis that science’s hallmark function of self-correction is undermined by deviations from the truth-seeking principle (45, 46). Ioannidis’ summation best describes this predicament:

“In the absence of replication efforts, one is left with unconfirmed (genuine) discoveries and unchallenged fallacies, where in many areas of psychological science, perpetuated and unchallenged fallacies may compromise the majority of the circulating evidence” (45, p. 645).

Pashler and Harris (46), for one source, contend that psychological science is facing a “replicability crisis” such that it won’t necessarily correct itself unless direct (not just conceptual) replications are performed. Ioannidis (45) contends that the focus of scientific research needs to shift from discovering something “novel and significant to getting to the truth” (p. 648). Ioannidis (45), reported that, in a series of papers published in the well-regarded journal *Perspectives on Psychological Science*, the prevalence of replication studies performed following “positive” findings was very low—in the order of 1-5%. Makel and colleagues (47) noted that replications in psychological science are “extremely uncommon.”

More alarmingly, Pashler and Harris (46) have demonstrated that with an underlying base rate of 10%, which is a reasonable maximum

to apply for all sex offenders, and statistical power of 80%, which grossly overestimates the power of all techniques used to determine recidivism in sexual offenders, the proportion of studies yielding true positives will still be a mere 8%, while the proportion yielding false positives will be 36%. It is not until the underlying base rate, or prior probability, reaches 75% that, with a statistical power of 35%, false positives drop to the default position in science of 5%. An underlying base rate of 75% is never achieved even with the “worst of the worst” of sexual offenders. It is hard to think of any phenomena in forensic psychology specifically or psychology in general where the underlying base rate is 75% except for blinding obvious phenomena such as right-handedness (technically referred to as right-laterality). The consequences of ignoring base rates described by Neal and Grisso (11), is useful to keep in mind when considering these findings.

The obverse side of this problem is subtler: the prospect of false negative findings occurring as a result of poorly framed hypotheses and poor experimental design (44). While this is important, given the demonstrated bias against publication of negative findings, this problem is clearly the lesser of two evils.

OBJECTIVE METHODS OF RISK ASSESSMENT OF SEXUAL OFFENDERS

Levenson and Morin (48, p. 76), noting that “perpetrator factors related to the risk of repeat sex offending have been virtually ignored in the development of CPS risk assessment protocols,” proposed that risk assessment of future child sexual abuse can be improved using “empirically derived risk models.” The difficulty here is that the epistemological and statistical principles underlying risk assessment/management of sexual offenders are exceedingly complex but CPS workers with a social work background who encompass a significant proportion if not the overwhelming majority of such individuals simply do not have the training in epistemology and statistics to understand the limitations of many empirically derived risk models. Unfortunately, it is not just social workers that lack this training. It goes

not too far to say that there is no evidence that the majority of forensic psychologists in Australia or those in the U.S. serving as expert witnesses in sexually violent predator commitment comprehend all the issues that must be considered to form reasoned judgment in such matters. Cooke and Michie (15) contend that ignorance of Bayes Theory and the limitations of relying on the omnibus statistic the Area Under the Curve (AUC) derived from Receiver Operating Characteristic (ROC) analysis is rife among forensic psychologists. They are not alone in making these observations (9, 10).

ESTIMATING RISK: TRAPPED IN A CORNER WITHOUT OUR “BASE RATES”

The continuing decades-long line of research into the issue of how well the clinical judgment in a variety of professions accurately predict recidivism repeatedly found it to be no better than chance (9, 49, 50). An equally bald fact is that empirically derived risk models have poor predictive accuracy when applied to individuals (35, 51). As far as nomothetic or actuarial risk assessment instruments (ARAI) are concerned, it is absolutely unequivocal that they almost never meet the legal test of the Balance of Probabilities let alone Beyond Reasonable Doubt. Coyle and Halon (10) provide a fuller explanation of scientific and legal standards of proof in the context of ARAIs.

As Cooke and Michie (15) pithily note in discussing the predictive accuracy of the Risk Management 2000 (RM2000), an ARAI designed to predict violent and sexual recidivism:

“The prediction based on the RM2000s that a Very High Risk offender will reoffend will be wrong 93% of the time; predictions that an offender in either the Very High Risk or the High Risk groups will reoffend will be wrong 96% of the time. The numbers necessary to detain in order to prevent one instance of recidivism are large.”

Relatedly, with respect to the most frequently employed ARAI, the Static-99 in its various iterations, Campbell and De Clue (8) have demonstrated that at the usual cut-off score employed on this test

(>6/13), it is necessary to detain 6 individuals to prevent one instance of recidivism. Conversely, some 18 individuals judged to be at below this cut-off score (<6/13) could be safely released without one instance of recidivism. Considering that one point scoring variations between different assessors on the Static-99 occur some 45% of the time (52), the jurisprudence implications of these statistics are clear and profound.

Cooke and Michie (15) draw a bleak conclusion when evaluating the overall utility of ARAIs in predicting sexual recidivism:

“Given the perceived crisis of replication in our field, and the frequent failure of science to be “self-correcting” (Ioannidis, 2012), their prescription for ethical practice should, in our view, be applied—and soon—to the area of violence risk. Menken famously remarked ‘There is always an easy solution to every human problem—neat, plausible and wrong.’ Sexual violence is a complex human problem, ARAIs may, to some, appear to be both neat and plausible, but unfortunately, they are wrong—and dangerously so.”

There is an important caveat to the conclusions of Cooke and Michie as to the utility of ARAIs in meeting legal standards of proof although not in a way intended by the developers and proponents of ARIAs. If we ask “How useful are ARIAs for ruling *out* recidivism” the answer is illuminating. For example, assuming a Static-99R score of 6 and using the Routine group of offenders for comparison purposes, ruling *out* recidivism risk will result in being *correct* 95% of the time; 19 offenders scoring 6 or below can be released and only one of them will commit a new sexual offense. Therefore the accuracy of the Static-99R used correctly, that is, for “ruling out” recidivism is rather impressive. In a similar vein, Cooke and Michie noted with respect to the RM-2000:

“It should perhaps be noted in passing that 98% accuracy in the prediction of proven sexual reoffending can be easily achieved

with their data—*by predicting everyone will fail to reoffend* (i.e., 977/1,000) [emphasis added]” (15, p. 50).

What then is to be made of the adamant opinions of CPS workers that recidivism is almost guaranteed in cases such as the matter discussed here? Outright skepticism if not immediate rejection of all assertions that are completely unsupported by objective, scientific evidence springs to mind. More fundamentally, the problem that judicial officers face in making a decision inevitably involves consideration of the gravity of harm that might proceed if a false negative error is made; conflation of risk with potential gravity of harm is an ever-present reality in sex abuse cases. In theory, predicting the future with respect to recidivism of any type requires an assessment to be made independently of the restrictions that might be applied to offenders on the basis that they might reoffend. However, in practice, these decisions are not independent.

THE CALCULUS OF RISK ASSESSMENT

In a study with 153 real jurors who were post-polled with regard to the interaction between the probability of offending and an assessment of risk of recidivism the conflation of these two factors was most clearly demonstrated:

“Most jurors (81.7%) considered a 15% estimated chance of recidivism to mean that the respondent was “likely” to reoffend, and many (53.6%) even considered a 1% chance to indicate likely reoffense. Jurors who heard lower risk estimates in trials were more likely to report that a low chance of recidivism (as low as 1%) indicated an offender was likely to reoffend” (53, p. 293).

Simply put, jurors view risk more in terms of the severity of potential harm than in terms of statistical probability; which, of course, they cannot be expected to understand well enough. Given discretion to define tolerable risk, jurors consider even a statistically low degree of

risk intolerable. The difficulty here is patent. The threshold for making decisions in child protection matters is not clear. Hence, CPS workers are free to set whatever thresholds they choose. For example, if the base rate of recidivism over five years is 8% and CPS workers regard a risk of 1% as being unacceptable then the laws designed to protect children from sexual offenders no longer serve their original purpose of identifying “likely” offenders.

Are judges also likely to conflate the risk of recidivism with the gravity of harm? There is only one study that has addressed this point. Monahan and Silver (54) found that a 26% perceived likelihood by judicial officers that a patient would commit a violent crime was, on average, sufficient to justify civil commitment; a terminology employed in the U.S. to mean preventive detention. Eleven percent of judges would order commitment for a patient who presented a 1% risk of committing a violent act; 38.5% of judges would order commitment with an 8% risk; 23.1% would order commitment with a 26% risk; and 26.9% of judges would order commitment with a 56% risk. These figures bear comparison to those reported for jurors. Clearly, the judges in the singular study reported are less likely to conflate risk and the gravity of harm, but they do integrate the two factors into their calculus of risk. Whether these findings hold true for other jurisdictions is a live issue but it is simply foolish to think that judicial officers are immune to the cognitive biases and cognitive heuristics that influence decision making in other people, no matter how diligently they seek to apply the law; especially since, like laypersons, they typically do not understand the statistics or the methodology of risk assessment. In any event, the law is hopelessly muddled as to what is the standard of proof in the various states and territories in Australia and in other countries that have enacted preventive detention legislation and with respect to parole decisions and associated child protection legislation (9). The interrelationship between the risk of harm and the gravity of harm in assessing risk is rarely discussed yet it is of seminal import.

As Coyle and Halon note:

“Either wittingly or unwittingly, the various state legislation in Australia encompasses this issue by reference to the concept of an unacceptable risk to the community should SVP offenders who are considered to pose a serious risk of dangerousness be released. It seems that exactly the same thing happens in individual cases in the USA as a result of the tendency of the triers of fact to err on the side of ‘caution,’ although erring on the side of not making a False Negative decision is a better way of putting it. Certainly, the experience in Minnesota in the USA where only one SVP offender preventively detained has been released in the last 23 years speaks to this” (10, p. 221).

Exactly the same conflation between assessing the probability of an event occurring and the consequences of that event occurring obtains when considering matters such as the instant case, child protection matters and parole matters generally.

ALLEGIANCE EFFECTS

The problem of forensic evaluators scoring risk assessment instruments differently in the field (as opposed to laboratories) encompasses different types of error. One error is due to evaluator differences, which may be attributable to training effects and inherent difficulties in interpreting risk assessment instruments. Boccaccini et al. summarize this situation thusly:

“When forensic evaluators assign scores to offenders on a risk assessment instrument, we assume that the offenders receive different scores because they differ on the trait measured by the instrument. At the same time, we know that it is impossible to measure any trait or behavior perfectly and that any score an evaluator assigns to any offender is a product of both the offender’s true level of the trait and at least some measurement error” (55, p. 337).

A second type of error is allegiance bias, which is not always due to a systematic and deliberate attempt to score an instrument or push a clinical interview in one direction or the other. The potential for allegiance bias has long been recognized and has led to a not unreasonable skepticism of the role of expert witnesses within the legal system. This was most famously commented on by Sir George Jessel, the Master of the Rolls, in *Lord Arbinger v. Ashton*, in 1873:

“In matters of opinion I very much distrust expert evidence, for several reasons...An expert is not like an ordinary witness, who hopes to get his expenses, but he is employed by the person who calls him. Now it is natural that his mind, however honest he may be, should be biased in favour of the person employing him and accordingly we do find such bias...*undoubtedly there is a natural bias to do something serviceable to those who employ you and adequately remunerate you* [emphasis added]” (56, p. 374).

There is no evidence that we know of to suggest that human nature has changed so much (if at all) that we are no longer affected by this bias. There is no evidence that allegiance bias is not as alive and well in the 21st century as it was in the 19th. A series of studies conducted by Murrie and colleagues (57-59) has confirmed this. Most of this effect is clearly not conscious and systematic bias, just as Sir George Jessel noted.

A third source of error is systematic bias: more correctly called scientific fraud. In the U.S., this has been unequivocally proven with respect to a number of ostensibly highly qualified experts retained by the organs of the state to independently assess sex offenders for civil commitment. Some consistently find, in the absence of cogent evidence to support their conclusions, that 98-99% of those they assess should be preventively detained (60). Discombobulating though this may be for judicial officers, it is a problem that infects the curial process most clearly when it comes to risk assessment of future events.

Whatever the motivation of these psychologists, one thing is clear: this sort of behavior brings the problem of allegiance bias, if not fraud, to the front and center stage of risk assessment. Ultimately, in the case we have described, the weight of expert opinion overcame the preconceptions of the psychologist retained by CSS. It is rare for respondents to be able to afford the expert assistance required in this matter and the result could easily have been different.

CONCLUSIONS

According to standard concepts of jurisprudence in those jurisdictions that have embraced the common law, the law is not concerned with cost-benefit analysis. It is not even directly concerned with justice. The law is concerned with the application of legal principals and concepts that, in theory, will result in justice more often than not. It is hoped that by adhering to legal principles and concepts justice will be delivered with a high degree of confidence. However, nowhere in any common law jurisdiction is this quantified for any particular case. Nor can it be: the myriad facts that each case presents preclude that prospect.

The law guarantees that a decision will be made but it does not guarantee outcomes. Yet an outcome is precisely what the law seeks to require of those engaging in the task of risk-analysis of dangerousness. Estimates derived from any currently available empirical measures cannot predict the future behavior of individuals with anything approaching that implicit in even the minimum legal standards of proof. Estimates of absolute and relative accuracy in risk assessments derived from ARAIs give the impression of scientific accuracy that is simply not warranted. It is arrant nonsense to assert otherwise, unless, like Humpty Dumpty, when we use a word it means what we want it to mean.

In the absence of compelling evidence to the contrary, such as in those very rare cases where volitional control is absent or other compelling clinical considerations such as antidepressant-induced hypo-

mania obtain (see, for example: Coyle [9], p. 274-275), it is simply silly for anyone to assert that the risk of recidivism in sex offenders is all but guaranteed. Parenthetically it might be noted that CPS workers trained as social workers cannot make clinical judgments in Australia. While clinical social workers in the U.S. are trained to make diagnoses and form clinical opinions, most social workers we have encountered do not have the training to enable them to form clinical opinions and make diagnoses, although there is a danger that they may think they do. Although they are no doubt tempted by the siren song that "it is better to be safe than sorry," recommending putting or keeping a father out of the home tears apart the supportive fabric of the family. In turn, the level of social support available to family members is severely compromised with the result that mood disorders and other psychological ills multiply (61). Keeping families apart should not be entertained unless cogent evidence is available. In a related context, it is sobering to note that when the *Dangerous Prisoners Sexual Offender Act* was mooted in Queensland it was estimated that only 10 offenders would be likely to be subject to its provisions and incarcerated at any one time. There are now some 220 individuals preventively detained under this legislation.

POSTSCRIPT

Ultimately the applicant (CPS) withdrew from the case after further expert reports were served. The offender was allowed to reside with his wife and children. He has not recidivated and his children and his business have prospered. At the discretion of the Queensland Police he no longer has to report his whereabouts and the like pursuant to the relevant legislation.

REFERENCES

1. Wilson RF: Removing violent parents from the home: a test case for the public health approach. *Virginia Journal of Social Policy and the Law* 2005; 12:3:1-48

2. Studer LH, Clelland SR, Aylwin AS, Reddon JR, Monro A: Re-thinking risk assessment for incest offenders. *International Journal of Law and Psychiatry* 2000; 23:1:15-22
3. Cann J, Friendship C, Gozna L: Assessing crossover in a sample of sexual offenders. *Legal and Criminological Psychology* 2007; 12:149-163
4. Levenson JS, Becker J, Morin JW: The relationship between victim age and gender crossover among sex offenders. *Sexual Abuse: A Journal of Research and Treatment* 2008; 20:1:43-60
5. Studer LH, Aylwin SA: Sexual offender subtyping: the incest offender question. 2006. www.sexual-offender-treatment.org
6. Bittles AH, Neel JV: The cost of human inbreeding and their implications for variations at the DNA level. *Nature Genetics* 1994; 8:2:117-121
7. Campbell TW: Predictive accuracy of Static-99R and Static 2002R. *Open Access Journal of Forensic Psychology* 2011; 3:82-106
8. Campbell TW, De Clue, G: Calibration performance indicators for the Static 99R: 2013 update. *Open Access Journal of Forensic Psychology* 2013; 5:82-88
9. Coyle IR: The cogency of risk assessments. *Psychiatry, Psychology and Law* 2011; 18:2:270-296
10. Coyle IR, Halon RL: Humpty Dumpty and risk assessment: a Reply to Slobogin, in *Preventive Detention: Asking the Fundamental Questions*. Edited by Keyzer P. Cambridge, Intersentia, 2013; 193-222
11. Neal TS, Grisso T: The cognitive underpinnings of bias in forensic mental health evaluations. *Psychology, Public Policy and Law* 2014; 20:2:200-211
12. Tversky A, Khaneman D: Judgment under uncertainty: heuristics and bias. *Science* 1974; 185:1124-1131
13. Tversky A, Khaneman D: The framing of decisions and the psychology of choice. *Science* 1981; 211:453-458
14. Wollert R: Poor diagnostic reliability, the null-Bayes Logic model, and their implications for sexually violent predator evaluations. *Psychology, Public Policy and the Law* 2007; 13:3:167-203

15. Cooke D, Michie C: The generalizability of the Risk Matrix 2000: on model shrinkage and the misinterpretation of the Area Under the Curve. *Journal of Threat Assessment and Risk Management* 2014; 1:1:42-55
16. Lacombe D: Consumed with sex: the treatment of sex offenders in risk society. *British Journal of Criminology* 2008; 48:55-74
17. Kingston DA, Firestone P, Wexler A, Bradford JM: Factors associated with recidivism among intrafamilial child molesters. *Journal of Sexual Aggression* 2008; 14:1:3-18
18. Helmus L, Hanson RK, Thornton D: Reporting Static-99 in light of new research on recidivism norms. *The Forum* 2009; 21:1:38-45
19. Tatar JR II, Streveler A: Sex Offender Recidivism after Release from Prison. State of Wisconsin Department of Corrections, 2015
20. Wollert R, Cramer E, Waggoner J, Skelton A, Vess J: Recent research (n=9,305) underscores the importance of using age-stratified actuarial tables in sex-offender risk assessments. *Sexual Abuse: A Journal of Research and Treatment* 2010; 22:4:471-490
21. Ruscio J: Information integration in child welfare cases: an introduction to statistical decision making. *Child Maltreatment* 1998; 3:2:143-156
22. Gambrill E, Shlonsky A: Risk assessment in context. *Children and Youth Services Review* 2000; 22:11-12:813-837
23. Herman S: Improving decision making in forensic child sexual abuse evaluations. *Law and Human Behavior* 2005; 29:1:87-120
24. Regehr C, Bogo M, Shlonsky A, Le Blanc V: Confidence and professional judgment in children's risk of abuse. *Research on Social Work Practice* 2010; 20:6:621-628
25. De Roma BM, Kessler ML, McDaniel R, Soto CM: Important risk factors in home-removal decisions: social caseworker perceptions. *Child and Adolescent Social Work Journal* 2006; 23:3:263-277
26. Munro E: Common errors of reasoning in child protection work. *Child Abuse and Neglect* 1999; 23:8:745-758
27. Koehler DJ, Brenner L, Griffin D: The calibration of expert judgment: heuristics and biases beyond the laboratory, in *Heuristics and Biases: The Psychology of Intuitive Judgment*. Edited by Gil-

- vich T, Griffin D, Kahneman D. New York, Cambridge University Press, 2002; 686-715
28. Harris MS, Hackett W: Decision points in child welfare: an action research model to address disproportionality. *Children and Youth Services Review* 2008; 30:2:199
 29. Breckenridge J, Baldry E: Workers dealing with mother blame in child sexual assault cases. *Journal of Child Sexual Abuse* 1997; 6:1:65-80
 30. Wilson RF: Recognizing the threat posed by an incestuous parent to the victim's siblings: Part I: appraising the risk. *Journal of Child and Family Studies* 2004; 13:143-162
 31. Hanson RK, Bussière MT: Predicting relapse: a meta-analysis of sexual offender recidivism studies. *Journal of Clinical and Consulting Psychology* 1998; 66:20:348-362
 32. Hanson RK, Morton-Bourgon KE: The characteristics of persistent sexual offenders: a meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology* 2005; 73:6:1154-1163
 33. Hanson RK, Morton-Bourgon KE: The accuracy of recidivism risk assessments for sexual offenders: a meta-analysis of 118 prediction studies. *Psychological Assessment* 2009; 21:1:1-21
 34. Nunes KL, Hanson RK, Firestone P, Moulden HM, Greenberg DM, Bradford JM: Denial predicts recidivism for some sexual offenders. *Sexual Abuse: A Journal of Research and Treatment* 2007; 19:2:91-105
 35. Cooke DJ, Michie C: Limitations of diagnostic precision and predictive utility in the individual case: a challenge for forensic practice. *Law and Human Behavior* 2010; 33:541-556
 36. Donaldson TS, Abbott BR, Michie C: Problems with the Static-99R prediction estimates and confidence intervals. *Open Access Journal of Forensic Psychology* 2012; 4:1-23
 37. Abbott BR: Applicability of the new Static-99 experience tables in sexually violent predator risk assessments. *Sexual Offender Treatment* 2009; 4:1:1-28
 38. Wollert R, Waggoner, J: Bayesian computations protect sexually violent predator evaluation from the degrading effects of confirma-

- tory bias and illusions of certainty: a reply to Doreen and Levenson. *Sexual Offender Treatment* 2009; 4:1:1-23
39. Yates PM: Is sexual offender denial related to sex offense risk and recidivism? A review and treatment implications. *Psychology Crime and Law* 2009; 15:2:183-199
 40. Harkins L, Beech AR, Goodwill AM: Examining the influence of denial, motivation and risk on sexual recidivism. *Sexual Abuse: A Journal of Research and Treatment* 2010; 22:1:78-94
 41. Freeman JE, Palk GR, Davey JD: Sex offenders in denial: a study into a group of forensic psychologists' attitudes regarding the corresponding impact upon risk assessment calculations and parole eligibility. *Journal of Forensic Psychiatry and Psychology* 2010; 21:1:39-51
 42. Ioannidis JPA: Contradicted and initially stronger effects in highly cited clinical research. *JAMA: Journal of the American Medical Association* 2005; 294:218-228
 43. Ioannidis JPA: Why most published research findings are false. *PLOS Medicine* 2005; 2:696-701
 44. Fiedler K, Kutzner F, Krueger JI: The long way from α -error control to validity proper: problems with a short-sighted false-positive debate. *Perspectives on Psychological Science* 2012; 7:6:661-669
 45. Ioannidis JPA: Why science is not necessarily self-correcting. *Perspectives on Psychological Science* 2012; 7:6:645-654
 46. Pashler H, Harris CR: Is the replicability crisis overblown? *Perspectives on Psychological Science* 2012; 7:6:531-536.
 47. Makel M, Plucker J, Hegarty B: Replications in psychology research: How often do they really occur? *Perspectives in Psychological Science* 2012; 7:537-542
 48. Levenson JS, Morin JW: Risk assessment in child sexual abuse cases. *Child Welfare* 2006; 85:1:59-82
 49. Garb HN: *Studying the Clinician: Judgment Research and Psychological Assessment*. Washington DC, American Psychological Association, 1998

50. Meehl PE: Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence. Minneapolis, University of Minnesota Press, 1954
51. Donaldson TS, Abbott B: Prediction in the individual case: an explanation and application of its use with the Static-99R in sexually violent predator risk assessments. *American Journal of Forensic Psychology* 2011; 29:1:5-35
52. Boccaccini MT, Murrie DC, Mercado C, Quesada S, Hawes S, Rice AK, Jeglic EL: Implications of Static-99 field reliability findings for score use and report. *Criminal Justice and Behavior* 2012; 39:1:42-58
53. Knighton JC, Murrie DC, Boccaccini MT, Turner DB: How likely is “likely to reoffend” in sex offender civil commitment trials? *Law and Human Behavior* 2014; 38:3:293-304
54. Monahan J, Silver E: Judicial decision thresholds for violence risk management. *International Journal of Forensic Mental Health* 2003; 2:1-6
55. Boccaccini MT, Ruffino KA, Murrie DC, Gardner BO: Evaluator differences in Psychopathy Checklist Revised factor and facet scores. *Law and Human Behavior* 2014; 38:4:337-345
56. Lord Arbinger v. Ashton, 17 LR Eq 358 (1873)
57. Murrie DC, Boccaccini MT, Johnson JT, Janke C: Does interrater (dis)agreement on Psychopathy Checklist scores in sexually violent predator trials suggest partisan allegiance in forensic evaluations? *Law and Human Behavior* 2008; 32:352-362. doi:10.1007/s10979-007-9097-5
58. Murrie DC, Boccaccini MT, Turner DB, Meeks M, Woods C, Tussey C: Rater (dis)agreement on risk assessment measures in sexually violent predator proceedings: evidence of adversarial allegiance in forensic evaluation? *Psychology, Public Policy, and Law* 2009; 15:19-53. doi:10.1037/a0014897
59. Rufino KA, Boccaccini MT, Hawes SW, Murrie DC: When experts disagreed, who was correct? A comparison of PCL-R scores from independent raters and opposing forensic experts. *Law and Human Behavior* 2012; 36:527–537. doi:10.1037/h0093988

60. Kriegman D: Evaluating bias in expert witness conclusions of “sexual dangerousness” Part 1. *Sex Offender Law Report* 2014; 15:4: 49-63
61. Campbell T, Lorandos D: *Cross Examining Experts in the Behavioral Sciences*, Chapter 8. Eagan, Minnesota, Thomson Reuters, 2013

ABOUT THE AUTHORS

Ian Coyle, Ph.D., Australian Forensic and Personal Injury Consultants, Queensland, Australia. Professor Coyle is on the Law Faculty at Bond University, Queensland, Australia; School of Psychology, Faculty of Health, Deakin University, Victoria, Australia; Department of Psychology, Faculty of Sciences, University of Southern Queensland, Australia. He maintains a private practice in Forensic Psychology and Human Factors Engineering in Southport in Queensland. A widely published author, he is a research consultant at La Trobe University.

Robert Halon, Ph.D. maintains a private practice in clinical and forensic psychology in San Luis Obispo, California. He is a Fellow and Advisory Board Member of the American College of Forensic Psychology, and author on issues related to DSM diagnoses and diagnostic criteria for forensic purposes, dangerousness risk assessment, child custody, interrelationship of clinical and forensic terms and diagnostic information, and critical book reviews.

Terence Campbell, Ph.D., forensic psychologist and author, died before this manuscript could be submitted. As a scientist and a man he is greatly missed.

Donald Thomson, Ph.D., LL.B is a professor in the School of Psychology at Deakin University, Victoria, Australia. Professor Thomson is also a barrister with extensive experience in forensic matters and has authored numerous works in forensic psychology. He is Fellow of the Australian Psychological Society.

Jessica Woskett is a doctor of psychology candidate in forensic psychology at Deakin University.