Validities and Abilities in Criminal Profiling

The Dilemma for David Canter’s Investigative Psychology

Richard N. Kocsis
Sydney, Australia

This article discusses misconceptions in Bennell, Jones, Taylor, and Snook’s (2006) critique of Kocsis (2003a) concerning the validities and abilities surrounding criminal profiling. It also, in part, serves to canvass the approach of investigative psychology and what may be viewed as an imperative for some of its proponents to challenge evidence into the validity of criminal profiling, which is not easily reconcilable with that ideology.

Keywords: criminal profiling; accuracy; skills

Possibly the greatest mystery surrounding criminal profiling has been its growth despite an absence of robust scientific evidence to validate it. Many of my past publications have emphasized this. Central to my argument has been the need for original, quantitatively based experimentation that seeks to impartially test profilers, the results of which I have consistently argued require publication in independent, academically peer-reviewed mediums. These are not exceptional criteria but rather the common standard encountered in disciplines that claim scientific credibility. I have sought to make a contribution in this regard via a series of publications (e.g., Kocsis, 2003a, 2003b, 2004; Kocsis & Hayes, 2004; Kocsis, Hayes, & Irwin, 2002; Kocsis & Heller, 2004; Kocsis, Heller, & Try, 2003; Kocsis, Irwin, Hayes, & Nunn, 2000; Kocsis & Middledorp, 2004; Kocsis, Middledorp, & Try, 2005).1

When I first learnt of the forthcoming manuscript by Bennell, Jones, Taylor, and Snook (2006), I was delighted with the thought that others had undertaken some independent, data-driven experiment. A reading of the article, however, showed that this was not necessarily the purpose of the article. Underpinning the circumstance of the article was, it seems, an important ideological issue relevant to the field of profiling involving an accidental clash between the findings of the research I have conducted.

Author’s Note: This article is a reply to the article by B. Bennell, N. J. Jones, P. J. Taylor, and B. Snook, published in the June 2006 issue of the International Journal of Offender Therapy and Comparative Criminology. Please address correspondence to Dr. Richard N. Kocsis, PhD, Forensic Psychologist in Private Practice, P.O. Box 662, Dee Why, Sydney, NSW 2099, Australia; e-mail: richard_kocsis@hotmail.com.
and the investigative psychology perspective. Consequently, it is in this broader context that I discuss the claims of Bennell et al. and the misunderstandings that underpin their article.

**Ideology and Rigor**

**The Ideology of Investigative Psychology**

In the 1990s, Professor David Canter coined the term *investigative psychology* to describe his research activities into criminal profiling, the rationale for this term presumably being to identify psychological endeavors associated with the investigation of crimes and in particular criminal profiling (Canter, 1994). Whether the previously existing and venerated disciplinary label of *forensic psychology* did not already adequately encapsulate such a notion is highly debatable (Arrigo & Shipley, 2005). By adopting the title investigative psychology, Canter differentiated his activities from that of others and founded a center with himself as its head with this nomenclature (Palermo & Kocsis, 2005). To fully appreciate Bennell et al.’s article, it is important to recognize the challenge research such as Kocsis’s (2003a) poses to proponents of investigative psychology.

The predominant tenet underpinning the ideology of investigative psychology appears to be a stylized method for the ideographic interpretation of crime behavior from which principles have been developed for use in criminal profiling. In what appears to be the first publication characteristic of this approach, Canter and Heritage (1989) argue that other endeavors in the field are empirically flawed. Over the years, Professor Canter has collaborated on numerous occasions with his students, both past and present, to publish a succession of studies that adopt this approach. Bennell et al. appear to identify with the investigative psychology approach.

Whilst proponents of investigative psychology have been actively producing publications throughout the years, I am not aware of there being any independently conducted, empirically robust, and scientifically peer-reviewed study that demonstrates that a sample of suitably qualified experts in investigative psychology can construct an accurate criminal profile. Moreover, there does not appear to be any independent evidence to suggest that a suitably qualified sample of experts employing investigative psychology techniques can accurately profile the characteristics of an unknown offender any better than individuals employing some other rivaling method for profiling.

Indeed, the available evidence at this time in support of investigative psychology as a valid method for profiling appears to be largely analogous to the field of criminal profiling generally. These predominantly being anecdotal examples (Canter, 1994) and consumer satisfaction surveys aimed at gauging the perceived utility of profiles provided by consulting police officers (Britton, 1997). As outlined in Kocsis and Hayes (2004), as well as Kocsis and Palermo (2005), these offerings either singularly or cumulatively do not constitute reliably robust sources of evidence. Consequently,
whilst the research of investigative psychology may sound good in theory, compelling
evidence confirming this approach as a valid (i.e., accurate) method for practitioners
to employ in profiling crimes has yet, to my understanding, to emerge.

The Dilemma for Investigative Psychology

In all fairness, this lack of independent evidence in support of investigative psychology
as a valid method for criminal profiling is not unique but largely inherent to all of the rival-
ing approaches. However, to appreciate the specific challenge for investigative psychology,
clarification of my own research is necessary. In the mid-1990s, one issue that interested
me was the desire to determine with a degree of scientific rigor whether the concept of pro-
filing was valid; that is, whether “profilers” could in fact proficiently predict the character-
istics of an unknown offender based on the behaviors exhibited in the commission of a
crime. The most scientifically pragmatic method of investigating this question involved, to
my mind, controlled experimentation involving the construction of criminal profiles by
profilers in comparison with other groups of participants.

Over the years, and inclusive of my most recent study, a total of 16 individuals who
provide profiles on a professional basis have been tested. The findings of this research
suggest that the sampled profilers can predict the characteristics of an unknown
offender, to varying degrees, more proficiently than many of the compared partici-
pants. In essence, these studies provide some long overdue, albeit still tentative, empir-
ical evidence in support of criminal profiling. Through the course of this research,
however, I have also indirectly and quite inadvertently, it seems, created a dilemma for
the ideology of investigative psychology.

The findings of Kocsis (2003a) suggest that a group of profilers employing a range
of approaches can construct, comparatively speaking, a profile more accurately than
other groups engaged in the same exercise. Proponents of investigative psychology,
however, appear to maintain that their approach to criminal profiling is the only scientif-
ically valid method for effective profiling because it employs methods that they believe
are superior to other approaches (Canter, 1994; Canter & Heritage, 1989). Consequently,
research that tends to show that profilers employing a range of approaches can be
effective at profiling tends to detract from arguments aimed at emphasizing the supe-
riority of investigative psychology methods. In short, the findings of Kocsis (2003a) do
not lend support to the contention that effective and valid profiling is the exclusive
domain of those employing investigative psychology techniques or any others for that
matter, and herein lies the dilemma.

Matters of Rigor

Possibly the most disappointing aspect of the article by Bennell et al. is that it
features an assortment of conceptual errors and what appears to be the selective
presentation of issues and, at times, a questionable juxtaposition of arguments.
Sufficient space to discuss in detail all discrepancies is simply not available and would
result in an intellectually tedious article. Consequently, only a few examples are provided to give an illustration of their flavor and occurrence.

An assortment of what I would describe as selective presentations are sprinkled throughout Bennell et al.’s article. One example is their reinterpretation of data from Kocsis et al. (2000, 2002) and Kocsis (2004), as displayed in Table 1 of their article. This table is presented without the control groups from any of the aforementioned publications. Presentation of the data in this manner appears to accomplish two goals. First, it prevents comparisons between the relative performance of the skill-based and control groups (a significant design feature to these studies). Second, it creates a context that serves to diminish the effort expended in the studies by portraying smaller numbers of participants than were actually sampled in each of the studies.

These discrepancies, however, are surpassed by one profound error that tends to demonstrate that Bennell et al. have fundamentally failed to properly understand the Kocsis (2003a) study, which states first in the introduction on page 127:

The conclusions offered in this article draw on the combined research of Kocsis, Irwin, Hayes, and Nunn (2000); Kocsis, Hayes, and Irwin (2002); and Kocsis (in press) [i.e. Kocsis, 2004] with some additional original data included in the analysis [italics added].

And then reiterates on page 132:

As the purpose of this article is to present a holistic impression of the data from all three of the previous studies as well as to incorporate additional data not previously used [italics added].

It should be apparent from these statements that Kocsis (2003a) is not simply a compilation of the Kocsis et al. (2000, 2002) and Kocsis (2004) studies as Bennell et al. appear to understand it. Rather, the analysis reported in Kocsis (2003a) represents the integration of data from the previous studies as well as the incorporation of additional original data specific to the Kocsis (2003a) study.

Unfortunately, having misunderstood the fundamental structure of Kocsis (2003a), Bennell et al. present a series of criticisms and arguments that appear to be premised and/or reflective of this misconception. One simple example is their suggestion that the participant numbers from the previous studies do not concord with those in Kocsis (2003a; see Note 4, Bennell et al., 2006, p. 359). Although Bennell et al. present this as an uneasy mystery, these disparities (as perceived by them) are simply the differences in the additional data present in Kocsis (2003a). Another manifestation of this problem is evident in Bennell et al.’s reinterpretation of the data presented in Table 1 of their article. This table is an inaccurate and thus invalid representation of data and conclusions drawn in Kocsis (2003a) as it does not reflect the additional data component.

A far more problematic manifestation of Bennell et al.’s apparent failure to properly understand the analysis of Kocsis (2003a) is evident in arguments they juxtapose to substantiate their criticisms. It appears that the strategy to much of their
article is to deconstruct the findings of Kocsis (2003a) by dissecting previous studies, as they seem to believe that such studies culminate and equate with the data analyzed in Kocsis (2003a). As the conclusions drawn in Kocsis (2003a) relate to more than these previous studies, this approach is misconceived and thus erroneous. An illustration of this juxtaposition emerges with Bennell et al.’s challenge to the finding that the data of Kocsis (2003a) fail to support the finding that the profilers outperformed all other groups. As one component to substantiate this challenge to the findings of Kocsis (2003a), Bennell et al. present a full page of argument concerning the collapsing of participant groups in Kocsis et al. (2000). Certainly, Bennell et al. are entitled to express their opinions concerning this procedure within the specifically confined context of the findings in Kocsis et al. (2000). Their comments, however, are spurious and irrelevant in the context of Kocsis (2003a) as this method of analysis does not occur in Kocsis (2003a) and the data analyzed in Kocsis (2003a) are more than that of the previous studies, such as Kocsis et al. (2000).

The Accuracy Arguments in Criminal Profiling

The first of three broad subsections to Bennell et al.’s article presents a range of arguments aimed at challenging the findings of Kocsis (2003a) concerning the accuracy of profilers. My rebuttals to these arguments are hampered, as previously explained, as they are premised on Bennell et al.’s apparently flawed understanding of the parameters of the analyzed data in Kocsis (2003a). This error is significant. By way of simple illustration, in the context of the sampled profilers, Bennell et al.’s criticisms fail to account for 27% of the data present in Kocsis (2003a) and from which its conclusions were specifically drawn. For this reason, the following discussion does not focus on responding to Bennell et al.’s criticisms relating to data, as any response could potentially be misinterpreted as some indication of their validity. Instead, the discussion hereafter focuses on issues of a methodological and conceptual nature, matters that Bennell et al. indicate are the focus of their concern (see Bennell et al., 2006, p. 345).

Accuracy Measures

Although not asserting them to be flawed, Bennell et al. opine about whether the multiple-choice questions used to measure responses were suitably reliable as a measure. Specifically, they speculate about possible subjectivity in the interpretation of certain questions (although this is expressed to be “many” questions in the context of Note 2 of their article) and how responses may not be entirely reflective of profile predictions but rather differences in the interpretation of questions between participants. It seems that the force of their argument hinges on the observation that “answers to such questions are based on the perspective of the people answering them” (Bennell et al., 2006, p. 346). The authors also query the reliability of the model answers that were developed and used to score all questionnaires.
It should be unnecessary to note that the questionnaires contain questions that are open to interpretation and that answers to these questions are based on the perspective of the people answering them, as arguably all questions are open to interpretation as an individual will answer a question according to how he or she personally perceives it, not according to how someone else perceives it. The purpose of a multiple-choice questionnaire is to focus a participant’s attention on identifying the answer that he or she feels best accords with his or her own perceptions. By using multiple-choice questions, one can effectively quantify the predictions made via the responses given. It is simply naïve to believe that in a social science experiment such as this, one could eliminate an individual’s biases and prejudices in answering questions posed to them, as individuals will invariably answer from the perspective of their own life experience and draw on their own understanding of prevailing community standards. Consequently, research participants who answered the administered questionnaire were on an equal footing in the sense that they were all individually confronted with the task of answering each identical question posed as they perceived and interpreted it to be. With this in mind, therefore, it is difficult to understand how Bennell et al. could legitimately argue that the measure adopted in the form of the questionnaire is inherently deficient and the responses of the participants thereby suspect.

Bennell et al. also conclude with the following observation: “Thus, individuals’ interpretations of these questions will yield varying answers, which will also bias Kocsis’ accuracy measure” (p. 347). It is one thing to assert that individuals’ interpretations of certain questions will yield varying answers, but it is quite another to suggest that this will in turn compromise the validity of the responses for the entire experiment, as this does not take into account the fact that each participant is faced with an identical set of questions. Perhaps Bennell et al. themselves, it appears, are not too committed to their arguments concerning the profile accuracy measure, however, given their cautious opening observation: “Although we do not claim that this approach to measuring profile accuracy is fundamentally flawed, we do have concerns about this measure” (p. 346).

Bennell et al. also seem to argue that it is not possible for participants to determine the answers to some of the questions included on the multiple-choice questionnaires (see p. 347). By way of illustration, they point to the question relating to an offender’s primary motive for the offense and query how this might be discerned by an investigator (even one with detailed knowledge of the case). Arguably an investigator could accurately determine a factor such as motive from any number of sources in the same way in which a tribunal of fact commonly called a “court” would attempt to discover this for the purpose of resolving who committed a particular crime. Namely, by regard to sources of third-party information and any defendant’s self-report including any confessions or admissions of the accused, witness accounts and statements, and any other available evidence suggestive of motive. In raising this query, Bennell et al. appear to miss the point of the experiment, which is intended to simulate, as far as possible, a profiling exercise in the real world. Such a perceived difficulty as discerning a
perpetrator’s motive is arguably analogous to a profiling exercise conducted in the course of a real criminal investigation. Clearly, there is no universal truth awaiting to be discovered by even the most conscientious and precise inquirer, so it is nonsensical to believe that one can expect to achieve perfectly reliable answers to any question and to suggest that in the absence of perfectly reliable answers to questions an experiment will be compromised. For this reason also, it is simply not realistic to assert that answers to questions in this experiment could not be reliably determined by an investigator intimately familiar with the investigation and with the knowledge of the court’s independent findings in respect of the matter.

How Do Professional Profilers Compare?

Bennell et al. raise two criticisms of an analytical nature to challenge the accuracy scores of profilers found in Kocsis (2003a). The first is their argument concerning the analysis of sampled groups that have been combined, whereas the second relates to the possibility of some bias in the analysis due to the use of different test instruments in some of the studies.

As previously mentioned, Bennell et al.’s arguments regarding collapsed groups are without merit in the context of the findings reported in Kocsis (2003a). Nonetheless, two potentially misleading issues raised by these authors require discussion. First, in advancing their arguments, Bennell et al. (p. 347) cite the Kocsis (2004) publication under the heading “Does It Make Sense to Collapse Groups?” As the Kocsis (2004) study does not feature the analysis of any collapsed groups, citation of it under this heading has the potential to mislead a reader into erroneously believing that this study features such an analysis.

Second, Bennell et al. make the statement “Based on the results presented by Kocsis et al. (2000) and Kocsis (2004), it seems erroneous to conclude that professional profilers in these studies [italics added] outperformed other groups” (p. 347). To support this statement, the aforementioned criticism concerning the analysis of collapsed samples in Kocsis et al. (2000) is raised. However, Bennell et al. do not provide any specific exposition of the statistically significant findings of the Kocsis (2004) study in discussing the accuracy scores (pp. 346–350). This is a pertinent omission by Bennell et al., as the Kocsis (2004) study does not, as mentioned, feature the analysis of collapsed samples but does indeed find statistically significant differences demonstrating that the sampled profilers outperformed other groups in the Kocsis (2004, p. 352, Table 1) study. Therefore, in presenting their arguments to challenge accuracy scores, Bennell et al. seemingly fail to acknowledge the existence of the statistically significant results in Kocsis (2004) that serve to directly contradict their assertion, which, as has just been explained, is factually inaccurate.

Finally, Bennell et al.’s criticism concerning the possibility of bias through the meta-analysis of different samples is also suggestive of the authors lacking a sound understanding of statistical standardization procedures. That is, they do not appear to
appreciate that a function of the statistical standardization of scores is to compensate for the very contingency they seem to raise (see author’s acknowledgment and Devore & Peck, 2005; Wild, 2000).

The Naiveté of “Absolute” Accuracy Scores

Another criticism raised by Bennell et al. is the suggestion that the evaluation of profile accuracy should be undertaken in absolute terms via, for example, the calculation of percentages (referred to by the authors as “absolute accuracy”). This is in contrast to the comparative analyses presented in studies such as Kocsis’s (2003a) wherein accuracy is judged, relatively speaking, in terms of performance among differing participant groups.

Bennell et al.’s argument concerning absolute accuracy is articulated in the statement “absolute accuracy is more important than relative accuracy in applied settings where police require effective profiles” (p. 349), but it appears unsubstantiated. The expertise of Bennell et al. to support such a unilateral claim, particularly with respect to the requirements of applied settings such as policing, is unclear. Indeed, the authors do not offer any citation in support of this assertion, proceeding instead with their discussion as if it were accepted dogma.

This proposition concerning absolute accuracy is suggestive of naiveté with respect to the capriciousness of how crimes are actually solved in an applied, as opposed to academic, context. Accuracy does not necessarily automatically equate with utility (Kocsis & Palermo, 2005). Arguably, the pivotal issue in an applied setting is not whether information contains any specified percentage of accuracy, as Bennell et al. appear to advocate elsewhere in their article (see p. 349). Rather, it is whether the information is productive in terms of contributing to the identification of the offender.8 Accordingly, what is more important in a real-world investigation is whether information provided in a profile possesses a higher standard of accuracy than what police investigators, for example, themselves may achieve, thus contributing meaningfully to an investigation. It is this operationally orientated perspective that studies such as Kocsis’s (2003a) offer via the comparative analysis of accuracy scores.

Bennell et al.’s argument concerning absolute accuracy is also questionable on a more technical level in that it appears to be based on an assumption that the achievable scores from the questionnaires can be validly equated to percentages. It is highly debatable whether profiles will ever achieve a standard of perspicuity where they could reliably achieve a 100% standard of accuracy (Kocsis, 2006). Accordingly, the calculated percentages presented in Table 1 by Bennell et al. are unlikely to be a reliable reflection of an individual’s prediction capability as they attempt to present them.

Written Profiles and Accuracy

The final component of Bennell et al.’s critique concerning the accuracy scores of profilers involves articulating what they view as a number of concerns related
to Kocsis (2003b) and, in particular, examination by the study of the content of written profiles and its relevance to accuracy. In sum, Bennell et al. appear to infer that Kocsis (2003b) presents the notion that the quality (i.e., accuracy) and quantity (i.e., length) of a written criminal profile are related. That is, longer/larger profiles are more accurate. Bennell et al. would be justified in criticising this notion if it had, in fact, been expressed. It appears, however, that these authors have overlooked the discussion contained in Kocsis (2003b, p. 44), which specifically cautions against this very conclusion.9 Although Kocsis (2003b) notes the presence of an apparent relationship between length and accuracy of written profiles, this finding is clearly reported in the context of the available data and qualified by advocating the need for further empirical investigation of the issue (Kocsis, 2003b, p. 44).

The Skills Arguments in Criminal Profiling

The second major subsection of Bennell et al.’s article presents several pages of discussion criticising the skill basis of the participants in the various studies. That is, they primarily question whether the sampled participants are representative, and homogeneously representative, of their identified skills and whether, in turn, based on their relative performance, these data provide any reliable insight into the value of such skills in the effective construction of a criminal profile.

Specifically, Bennell et al. argue that no real assessment has been undertaken to determine whether the groups “possess the skills ascribed to them.” As a consequence, they say, “no valid conclusions about important profiling skills can be drawn.” To illustrate this view, the authors turn their attention first to the psychic group as “the best case in point” (352).

In discussing the psychics, Bennell et al. query whether the skill of intuition is one that is actually possessed by the psychics. They also criticize my colleagues and me for not defining “the so-called skill of intuition” (p. 353) or assessing, as I already mentioned, whether the groups possessed the skills attributed to them. Such criticism seems peculiar given that Bennell et al. themselves concede that such assessment “may prove difficult to define and evaluate (e.g., intuition)” (p. 352) and offer no suggestion as to how the skill of intuition might be either tested or assessed.

Bennell et al. even go so far as to state the following: “We consider it very unlikely that certain groups actually possess the skills attributed to them” (p. 352). This statement also appears peculiar given that it conflicts with what the Kocsis et al. (2002) study found, a study that Bennell et al. themselves acknowledge, this being that the Kocsis et al. (2002) study did include an assessment of skills, the findings of which suggested that the student group did possess the capacity for logical reasoning ascribed to them (see Bennell et al., 2006, p. 352).
Nevertheless, Bennell et al. maintain that “psychics are not the only group whose attributed skills are questionable” (p. 352) and proceed to query whether the psychologists recruited for the studies have a good appreciation of the criminal mind. In arguing their case Bennell et al. state the following:

“We know of no research that directly supports our argument that psychologists lack an appreciation of the criminal mind, but there is indirect evidence” (p. 353). It seems that evidence they rely on in support of this assertion is a study by Skeem and Golding (1998), which, they report, indicates that “psychologists often demonstrate significant disagreement with respect to a defendant’s specific abilities and impairments” (Bennell et al., 2006, p. 353). It is unclear to me why Bennell et al. would refer to this study by Skeem and Golding in support of their argument, as this study lends no support to the proposition that psychologists lack an appreciation of the criminal mind. Rather, the Skeem and Golding study was concerned with evaluating the ability of sampled clinicians composed of PhD-level psychologists, psychiatrists, and social workers to effectively evaluate a defendant’s competence to stand trial. That is, their study examined the nature and quality of competency to stand trial assessments completed by such clinicians; for example, whether they failed to describe the psycho-legal reasoning underlying their clinical conclusions and whether they consulted appropriate third-party sources of information in preparing their forensic reports.

Even accepting that psychologists have differences in opinion regarding a defendant’s specific abilities and impairment, it is difficult to see how such differences in opinion are demonstrative of their lack of appreciation of a criminal mind. Despite this, Bennell et al. conclude that “it is unlikely that general clinical psychologists have much insight regarding criminals or the way in which their minds operate” (p. 353). The authors do not explain why it is unlikely, and they also erroneously surmise that the psychologists used in the study were “general clinical psychologists.”

Bennell et al. also present a challenge to the skill basis of the sampled psychologists, which in my opinion is potentially misleading. Although the general label assigned and adopted by Hazelwood, Ressler, Depue, and Douglas (1995) is an “appreciation of the criminal mind” and psychologists are placed under the label of this conceptualized skill, the skill that is examined in the context of my studies is a fundamental understanding of human psychology and “demonstrable psychological skills” (Kocsis et al., 2000, p. 312). Bennell et al.’s arguments in challenging the validity of the sampled psychologists appear to omit this point. Perhaps this omission is understandable given that raising the issue would necessitate some argument that trained psychologists, irrespective of their specialization, are not representative of psychological skills.

Indeed, it seems that Bennell et al. have overlooked some important matters concerning the experiments conducted in each of the studies. First, the assembled groups of profilers, police detectives, psychologists, and students are intended to be representative of the skills so-called and identified by Hazelwood et al. (1995) as necessary for effective
criminal profiling. Second, these groups serve as a refinement on the groups adopted for examination by Pinizzotto and Finkel (1990) in their studies by additionally incorporating the testing of psychics and detectives with varying levels of experience. Furthermore, Bennell et al. seem to ignore the fact that police occasionally use psychics to assist them in criminal investigations requiring additional insight (e.g., Geberth, 1996). Psychics are typically used in this context because they are perceived and indeed regularly hold themselves out as having some psychic like faculty that enables them at some ethereal level to have insight without reasoning, a talent colloquially known as “intuition.” In other words, psychics claim to have intuitive abilities, and the aim of the experiments is to test whether the claims of these and other groups of individuals have any basis in fact in terms of being useful for effective criminal profiling.

What I also find peculiar to Bennell et al.’s reasoning is their attenuation on psychics and intuition to substantiate their arguments. Why among the various studies that I have conducted do they predominantly focus their discussion on the sample of 20 psychics from the earliest study (i.e., Kocsis et al., 2000) instead of the other skill-based groups numbering collectively into the hundreds? This choice seems especially pertinent as the skill that has been predominantly featured and examined throughout my studies is that of investigative experience. For this reason, I would expect the issue of investigative experience to dominate Bennell et al.’s discussion. Instead, however, they focus their attention on the relevance or otherwise of psychics and the element of intuition.

If Bennell et al.’s belief that “no valid conclusions about important profiling skills can be drawn” (p. 352) without an assessment demonstrating that each participant in the study possessed the skills ascribed to him or her is to be accepted, then their argument ostensibly implies that police officers are not a reliable representation of the skills/attributes of investigative experience. Although police officers undoubtedly possess other skills, it is difficult for Bennell et al. to plausibly suggest that in the context of these studies and in comparison to psychologists, students, and psychics, police would not have as their most distinctive and representative skill investigative experience. This is an issue that Bennell et al. inadvertently appear to concede when they state, “It is plausible that every participant in the Kocsis studies possess a degree of each skill (with the exception of investigative experience)” (p. 353).

Also, it needs to be asked how likely it is that university students undertaking science degrees involving the logical and objective analysis of data would not as a group demonstrate these skills more than others. Especially given that these participants were recruited on the basis of their stated lack of skills and/or experience in law enforcement, psychology, or the belief that they possessed paranormal abilities that could assist them in profiling a crime.

Although the sampled participants were not evaluated via some formal psychometric instrument, this factor does not, as Bennell et al. seem to suggest, obviate
reasonable logic that the groups used in the context of these studies are valid, albeit approximate, representations of the predominantly examined skills.

Moreover, Bennell et al.'s criticisms concerning skills assessment have, it seems, little regard for the logistical imperatives underpinning a social science experiment of this type and magnitude. In a hypothetically ideal world, the perfect experiment could be devised and any issue thoroughly tested. The real world, however, is beset with what are potentially endless logistical obstacles that must be negotiated if experimentation and progress are to occur. The incorporation of some formal measurement to evaluate all the skills of participants in the experiments would, if indeed capable of such assessment, entail a significant logistical encumbrance to the production of quantitative research in a field that already suffers from a dearth of original data. In this respect, it needs to be considered whether the benefits potentially derived from the administration of some measure(s) satisfying the concerns of Bennell et al. would outweigh the logistical encumbrances it may present for conducting such already difficult research. That is, would the outcome of some psychometric measure yield any significant improvements to the experiment? To my mind, it is more constructive to ask whether, from a practical perspective, the majority of participant groups present an approximate, yet nonetheless, valid representation of their assumed skills. I argue that they do.

External Validity

The final broad subsection of Bennell et al.'s article features a variety of criticisms designed to challenge the external validity of the findings and procedures featured in the various studies I have conducted. These criticisms are presented using four subheadings. Accordingly, four subheadings are discussed in response.

Sample Sizes: Who Are Profilers Anyway?

Bennell et al. posit two challenges concerning the sampled profilers in Kocsis (2003a): first, whether sufficient profilers were collected, and second, whether the recruited participants were reflective of what Bennell et al. conceptualize as “profilers.”

Criticisms concerning the adequacy of sample sizes are potentially endless. Nonetheless, by this criticism, Bennell et al. highlight their failure once again to correctly understand the parameters of the analysis in Kocsis (2003a) by querying what they perceive to be disparities in the sample sizes. Furthermore, some hypocrisy is, in my opinion, in operation given that a number of these authors have published studies using comparatively small samples from what are arguably generally accessible populations involving logistically easier experiments. Indeed, this is a curious issue given that Bennell et al. (p. 354) fully acknowledge the difficulties in recruiting profilers.

In considering the adequacy of samples, Bennell et al. fail to provide exposition of two issues in particular that I consider important. These are the serious deficit of
empirical data testing profilers and the expressed tentative nature of the conclusions drawn from studies such as Kocsis’s (2003a). First, the significance of the findings in Kocsis (2003a) should not be underestimated, as this study currently represents the largest existing data set, thus far, to directly and empirically test the capabilities of expert profilers and other skill-based groups engaged in constructing a criminal profile. Bennell et al. even concede this point, as they state that they “agree that the combined Kocsis sample represents the largest profiler sample to date” (p. 354).

Second, and importantly, at no point has the research I have undertaken ever been represented as an authoritative treatise on the topic of criminal profiling. Rather, my findings are clearly reported as tentative observations concerning a practice for which scientific inquiry is long overdue. Further replication and development in the future is clearly emphasized (see Kocsis, 2003a, pp. 139–140), as is the reluctance of profilers to participate in independent research that seeks to impartially assess their skills (e.g., Kocsis, 2003a, p. 135; Kocsis, 2004, p. 348; Kocsis et al., 2000, p. 234). Bennell et al., however, seem to underestimate the difficulties associated with securing the participation of profilers in such experiments, as they simply, and rather naively, I would assert, propose by way of solution that more “profilers will have to perceive the value of partaking in these studies and view the experiments as a fair reflection of their profession” (p. 357).

In their discussion, Bennell et al. seem to also suggest that the findings may be unrepresentative, as some sampled profilers did not possess law enforcement experience. Specifically, they make the following statement:

In addition, certain profilers sampled did not have any law enforcement experience (Kocsis, 2004), creating grounds to question the extent to which the results generalize to the majority of profilers (at least in North America) who do have law enforcement experience. (p. 354–355)

The inference here seems to be that “genuine” profilers are only those who possess law enforcement experience. Such an inference is disturbingly reminiscent of arguments advanced by those keen to monopolize profiling for reasons of their own career interests (Kocsis & Coleman, 2000).

Finally, Bennell et al. seem to acknowledge that there are no uniformly recognized criteria for determining who can legitimately be called a profiler. In the absence of a universally agreed definition then, it seems pointless to criticize the profilers used in the studies in terms of their not being validly representative of profilers.16

**Multiple-Choice Questionnaires Vs. the Vagaries of Written Profiles**

Arguably, the most curious argument Bennell et al. advances as a criticism of the findings of studies such as Kocsis’s (2003a) is that they predominantly evaluate participants’ performance via a multiple-choice profiling task. The basis to this contention
appears to be sourced in the following statement: “This work only examines profiling performance within an artificial context of a multiple-choice questionnaire (Kocsis, 2003a, pp. 139)” (p. 355). Indeed, Bennell et al. appear to seize on the use of the word artificial as some admission of the invalidity of the research.17

The circumstances in which a criminal profile is constructed do not typically involve, for example, a police officer presenting a profiler with a case package to a crime and a multiple-choice questionnaire that nominates possible offender/crime characteristics that the profiler must respond to in order to provide his or her predictions. Instead, the process typically involves a police investigator consulting a profiler and providing that profiler with material for examination for the purpose of furnishing a profile, either verbally or in writing (Kocsis, 2006). Whilst, the method of communicating predictions via a multiple-choice questionnaire is atypical and therefore arguably artificial, what is not artificial or atypical is the nature of the predictions made. This important nuance between the mode of communication and the information conveyed, it appears, has been missed by Bennell et al. In this context, what needs to be and is assessed via studies such as Kocsis’s (2003a) is the accuracy of the information contained in profiles. Bennell et al., it seems, have confused the two issues.

In contrast to progressing scientific understanding of profiling Bennell et al.’s arguments will, in my view, stifle genuine progress. Namely, because the reliability and validity of any measure that attempts to interpret the vagaries of written profiles will undoubtedly involve far greater methodological limitations and confounds than the methods they presently criticize. Indeed, Bennell et al.’s own statements are suggestive of this and belie what I view as the sophistry to their rationale:

Although it would undoubtedly require more effort to develop a valid procedure for measuring the accuracy of written profiles, the resulting data would probably be more interpretable and meaningful. (p. 355)

The use of the word probably seems to indicate that the authors are not themselves convinced by their arguments.

**Time Parameters in Criminal Profile Construction**

Bennell et al. (p. 355) also voice concerns surrounding possible time parameters that the research participants may have been facing when participating in the various studies. Two observations seem pertinent to this issue. First, Bennell et al. pose the question “Were different time limits imposed on the different groups?” (p. 355). As Bennell et al. themselves acknowledge, absolutely no time limits were imposed, and as indicated in the studies, all participants were instructed to take as much time as they required to complete the questionnaire to the best of their abilities. Despite knowing this, the authors proceed to discuss the issue by speculating about what might have been the case if time limits were present. For example, “the time difference, if it
did exist may also have . . .” (p. 355). Exactly what is served by such speculation is not clear. Nor is it clear why the following observation is offered: “Boon, 1997, indicates that a single profile can take upwards of 40 hours to construct” (pp. 355-356). There is no universally agreed time for the construction of a profile, and no evidence is offered by Bennell et al. to support this statement of Boon’s (1997) one way or the other.

It is precisely because there is no set time for constructing a profile that no time limit was imposed on the participants in the studies conducted. In designing these studies, the need for parity in the experimental procedures was carefully considered. Accordingly, administration influences were guarded against wherever possible. One simple strategy was to ensure that all participants had a comfortable environment where they could complete the questionnaire using as much time as they felt necessary.

In short, Bennell et al. fail to consider the context of the presented task and the experimental conditions that prevailed. No data were collected from any participant who indicated that his or her abilities were being impeded in respect of the presented task or experimental conditions. Furthermore, none of the profilers indicated that they were being compelled to engage in some unorthodox procedure that impeded or compromised their abilities by participating, for example, on an individualized noninteractive basis. Irrespective of the conception of how profiles should be constructed, the sampled participants (e.g., the expert profilers) were satisfied with the conditions within which they were tested, as confirmed by their consent to, and participation in, the experiments. Accordingly, I argue the obtained data are indeed representative of the various participants’ capabilities.

Interaction Among Participants — How Are Accurate Profiles Constructed?

Another criticism articulated by Bennell et al. is the lack of interaction that participants were permitted. Specifically, Bennell et al. suggest that the construction of a criminal profile is an interactive process involving “brainstorming sessions” (p. 356). Findings from Kocsis (2003a) may not, therefore, Bennell et al. contend, be a valid representation of profiling capabilities, as such studies did not include brainstorming. Although it is not clear what exactly is meant by brainstorming, as no attempt is made to describe the concept, its relevance to the profiling process remains somewhat of a mystery given the context in which it is raised. Specifically, Bennell et al. quote the following from Douglas and Burgess (1986): “Profiling is often viewed as an interactive process that involves brainstorming sessions (Douglas & Burgess, 1986)” (p. 356).

The inclusion of this statement by Bennell et al. appears curious given that first, there is no evidence that I am aware of to support the proposition that profiling is a process necessarily involving brainstorming. Second, it is not clear from this statement
whether Douglas and Burgess (1986) themselves agree or disagree with the view that profiling involves brainstorming sessions. Bennell et al. make no attempt to take this statement further by failing to offer evidence in support of the view that the construction of profiles necessarily involves brainstorming sessions or any other type of sessions for that matter.

Finally, even accepting that profile construction is an interactive process involving something called brainstorming, how does one cogently operationalize this notion into a genuinely robust scientific experiment? As previously indicated with reference to multiple-choice questionnaires, the methodological flaws and confounding elements likely to emerge from such conceptions are likely to far outweigh the benefits of attempting to explore an issue such as brainstorming.

**Conclusion**

It is difficult to write a rebuttal to an article that purports to critique one’s research when that critique is premised on an erroneous understanding. In failing to correctly understand the Kocsis (2003a) study, Bennell et al. have arguably undermined the validity of their arguments, which in turn raises questions concerning the extent to which the findings of Kocsis (2003a) have been effectively challenged by these authors. It is my contention that the conclusions drawn in the Kocsis (2003a) study are valid; thus, the dilemma these findings present to the ideology of investigative psychology remains.

Unfortunately, failing to correctly understand the parameters of the Kocsis (2003a) study is not, for me, the only disappointing aspect of Bennell et al.’s article. Although they seem to echo my sentiments for more empirical research into evaluating profilers, it seems that rather than conduct their own original data-driven experiment to test the findings of my studies or others, they have instead opted for seeking to reinterpret the data collected by myself and to offer what I see as suggestions for the hypothetical ideal experiment.

This circumstance is not unusual in the field of criminal profiling, for as I keep emphasizing, one of the biggest problems confronting the state of profiling research today is the absence of original experimentation. Researchers in the field seem to favor producing publications that are either not driven by original data or to make use of more readily available forms of data that are not particularly relevant to the practical application of profiling. Clearly, such approaches are less onerous than designing a new experiment and persuading even a modest number of individuals such as profilers to participate in it, which is why perhaps it has proven so popular. As a result, however, very little in the way of new and instructive information about profiling emerges.

In my view, the most disappointing aspect of the article by Bennell et al. is what I regard as its selective presentation of issues by the omission of pertinent material. I have endeavored to highlight such omissions throughout my article, as this to my mind seriously undermines the value of Bennell et al.’s assertions.
In conclusion, unless and until more researchers invest the time and effort to conduct more challenging, original, quantitative research, our understanding of criminal profiling will fail to advance as it should. For this reason, I look forward to research that pursues the more difficult path of conducting independent experiments with original data that serve to genuinely expand our understanding of profiling and how it can work.

Notes

1. It should be clarified that because of the schedules of different journals, Kocsis (2004) appeared after the publication of Kocsis (2003a). Despite the chronological sequence of these publications, the study of Kocsis (2004) preceded the study of Kocsis (2003a).
2. With the exception of Pinizzotto and Finkel (1990) and my own research endeavors.
3. Evidence suggesting Bennell Jones, Taylor, Snook’s (2006) incorrect understanding of the Kocsis (2003a) study can be found in their description of the study. For example,

The second piece of evidence that Kocsis (2003a) draws on to support his arguments about the relative performance of the groups is the outcome of a form of meta-analysis in which he combined results from Kocsis et al. (2000), Kocsis et al. (2002), and Kocsis (2004) [italics added]. (p. 348)

and

Another problem with this meta-analysis relates to the sample sizes reported in Kocsis (2003a; Table 1, p.133). Kocsis maintains that, across the three studies on which the analysis was based…. [italics added] (Note 4, p. 359)

4. For example, Bennell et al. count 8 profilers from the previous studies and thus infer that the data do not accord with the 11 profilers reported in Kocsis (2003a). Bennell et al., it appears, fail to realize that Kocsis (2003a) incorporated another 3 profilers in addition to the 8 sampled from the three previous studies.
5. Bennell et al. (Note 2) state the following:

Presumably, if the reader agrees with our argument that many of the items on the multiple-choice questionnaires are open to interpretation, then all of the analyses conducted by Kocsis and his colleagues should also be viewed as exploratory and impeded partially by ambiguities. (p. 358)

6. It should be clarified that Bennell et al. do refer to Kocsis (2004) but only in the context of the arguments concerning meta-analysis aimed at challenging the findings of Kocsis (2003a) and then in the context of the converted percentages with reference to their contenotions of “absolute accuracy.” Consequently, Bennell et al. do not address the statistically significant findings in Kocsis (2004), which demonstrate that professional profilers in that study did outperform other groups.
7. In my view, another highly contentious assertion made by Bennell et al. is their claim that I have overstated the findings of my studies (Note 3, p. 358). They take particular issue with the following statement: “Kocsis et al. (2000) found that professional profilers outperformed all the non-profiler groups with regard to the overall accuracy of their predictions (Kocsis, 2004, p. 343)” (p. 359). Also, they opine that the available data of Kocsis et al. (2000) do not support this assertion. Once again, it would appear that Bennell et al. have failed to properly understand the material. The profiler sample in Kocsis, Irwin, Hayes, & Nunn. (2000, Table 1, p. 320) achieved an overall mean accuracy score of 13.80, which surpasses the scores of all nonprofiler groups (individual scores and/or combined). In contrast to the assertions by
Bennell et al., the data demonstrate a clear descriptive pattern that the profilers outperformed all of the non-profiler groups. This pattern is not stated as being statistically significant in the passage cited by Bennell et al., so it would appear that these authors have, again, drawn an erroneous inference. The gravity of Bennell et al.’s assertions is highlighted by the fact that a key finding of the Kocsis (2004, p. 354) study is that profilers were statistically better than the police detectives on the total accuracy measure. It is ironic that it is these same statistically significant findings of Kocsis (2004) that Bennell et al. omit to discuss in their critique of accuracy scores.

8. A hypothetical example to highlight the problems with Bennell et al.’s argument is to assert that a profile containing only 10% of information that was accurate should be considered invalid. As a matter of logic, this proposition would at first appear to be reasonable. The difficulty arising from this proposition, especially when placed in an applied context, is how this position could be sustained if the 10% of information that is accurate contains a pivotal clue to the identity of the perpetrator, which could in turn lead to their immediate arrest. Conversely, assuming that a profile possesses a 90% degree of accuracy is no assurance that such information may be meaningful in an investigative context in terms of leading to the apprehension of the offender.

9. For example,

we should acknowledge that our finding that profilers construct profiles that are longer and that contain more predictions does not in itself mean that those profiles are necessarily more useful to an investigation. In other words, quantity does not mean quality. Indeed, the best profile would be the shortest, containing only the correct name and address of the offender. (Kocsis, 2003b, p. 44)

10. For example, Kocsis, Hayes, & Irwin (2002) were dedicated to examining investigative experience.

11. As one of the few individuals to have empirically tested profilers, I am sufficiently confident to assert that the construction of scientifically valid (i.e., accurate) criminal profiles will have little relationship with skills related to the somewhat ethereal element of intuition. It is my suspicion that most reputable scientists possessing a sound understanding of scientific principles would share my assertion.

12. The apparent sophistry to Bennell et al.’s arguments is best highlighted by their hypothetical acceptance. If we assumed that some reliable instrument existed and was administered that measured investigative experience, is there any genuine expectation that this instrument would not demonstrate that police officers possess investigative experience and to higher levels than psychologists, students, or psychics? Furthermore, using this same hypothetical instrument, is there any genuine expectation that it would generally show that, for example, senior detectives employed for more than 15 years in law enforcement would not have far higher levels of investigative experience than detectives with only 2 years employment or police recruits receiving basic training (see Kocsis et al., 2002)? Alternatively, how plausible is it to expect that this instrument would reveal that detectives who are trained and/or who specialize in crimes of arson would not have far higher levels of investigative experience specifically pertinent to arson investigation than detectives who do not possess this (i.e., a qualitative distinction; see Kocsis, 2004)? The weaknesses of Bennell et al.’s arguments are that it is these highly probable classifications of inherent skills that, although not psychometrically proven, seem nonetheless entirely plausible.

13. Another puzzling statement of Bennell et al. concerns the value they accord to the examination of skills in terms of science students:

On the other hand, Kocsis et al. (2002) appears to have made some attempt to do this but only to determine whether students possessed the capacity for logical reasoning. Although we applaud their attempt to assess student skills, we think the results of this effort (based on 31 students) should not be generalized to other student groups. (p. 352)

The question to arise from this assertion is Why? Given that Snook, Canter, and Bennell have, in the past, collaborated on a study involving a slightly larger sample and consider their study valid, it needs to be
questioned why the sample in this particular study is considered so unrepresentative by Bennell et al. and its findings quickly dismissed. This is particularly poignant as, in contrast to Bennell et al.’s speculations, the results of Kocsis et al. (2002) present empirical data that directly undermine their arguments.

14. Readers should appreciate the weakness of Bennell et al.’s arguments concerning the evaluation of skills, as they themselves acknowledge the difficulties in evaluating such skills via some psychometric instrument. Perhaps no better demonstration of this difficulty is that Bennell et al. have not, to my knowledge, produced research involving such an instrument.

15. For example, “we only count 8 profilers across Kocsis’ studies” (Bennell et al., 2006, p. 354).

16. It should be noted that I have adopted a definition of those who professionally engage in the construction of profiles for active criminal investigations as being profilers. It is this representative definition of a profiler that has consistently featured throughout my studies.

17. This confusion with what is meant by the word artificial also relates to Bennell et al.’s opening paragraph concerning validity arguments.

18. Indeed, it also needs to be asked what compelling evidence exists to suggest that Douglas and Burgess (1986) enjoy an authoritative position to indicate what is the most effectual method is for the construction of a criminal profile.

Acknowledgments

The author acknowledges with gratitude Associate Professor Gillian Heller and Lecturer Jenny Middledorp, Department of Statistics, Macquarie University, for their advice in responding to criticisms concerning meta-analysis and statistical standardization.

References


