

Validities and Abilities in Criminal Profiling: A Critique of the Studies Conducted by Richard Kocsis and His Colleagues

Craig Bennell

Natalie J. Jones

Carleton University, Canada

Paul J. Taylor

University of Liverpool, UK

Brent Snook

Memorial University of Newfoundland, Canada

In a recent issue of this journal, Kocsis reviewed the criminal profiling research that he and his colleagues have conducted during the past 4 years. Their research examines the correlates of profile accuracy with respect to the skills of the individual constructing the profile, and it has led Kocsis to draw conclusions that are important to the profiling field. In this article, the authors review the contributions of the Kocsis studies and critique their methodological and conceptual foundations. The authors raise a number of concerns and argue that data from the Kocsis studies fail to support many of the conclusions presented in his recent review. The authors present evidence in support of their assertions and provide recommendations that will allow future research in the area to generate data that are more meaningful and generalizable.

Keywords: *criminal profiling; accuracy; skills*

In a recent issue of this journal, Kocsis (2003a) provided a comprehensive review of the profiling research that he and his colleagues have conducted during the past 4 years (Kocsis, 2003b, 2004; Kocsis, Hayes, & Irwin, 2002; Kocsis, Irwin, Hayes, & Nunn, 2000). Their research attempts to identify the correlates of profile accuracy

Authors' Note: A reply from Dr. Richard Kocsis will appear in the August 2006 issue. Research for this article was supported by a grant awarded to the first author from the Office of the Dean, Faculty of Arts and Social Sciences, Carleton University. Correspondence concerning this article should be addressed to Craig Bennell, Department of Psychology, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada; phone: (613) 520-2600; fax: (613) 520-3667; e-mail: cbennell@connect.carleton.ca. Please address correspondence to Craig Bennell, Department of Psychology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada.

with respect to the skills of the individual constructing the profile. In addition to exploring the accuracy of professional profilers, Kocsis typically examined police officers, psychologists, students, and psychics.¹ These groups have been chosen because Kocsis believes they possess particular skills that professional profilers view as important to their discipline, namely investigative experience, an appreciation of the criminal mind, the capacity for logical reasoning, and a high degree of intuition, respectively (Hazelwood, Ressler, Depue, & Douglas, 1995).

Kocsis (2003a) draws four major conclusions that are potentially important to the profiling field. First, Kocsis claims that his data demonstrate that profilers “surpassed all of the compared groups in the total number of correct [profiling] predictions” (p. 134). Second, Kocsis argues that the poor performance of police personnel “fail[s] to support the asserted importance of investigative experience as the key skill necessary for proficient profiling” (p. 135). Third, Kocsis states that the capacity for logical reasoning appears crucial to success in profiling because his student sample “actually surpassed the sampled psychologists, making them arguably the most proficient group after the profilers” (p. 137). Fourth, Kocsis contends, “The combined data suggest little support for the use of psychics [and the intuition they supposedly represent] in accurately generating the characteristics of an unknown offender” (p. 138).

The studies of Kocsis and his colleagues have received considerable attention (e.g., Winerman, 2004). We also believe their research is valuable and applaud them for tackling some important questions. However, there are a number of issues that must be addressed before the conclusions of Kocsis (2003a) can be accepted as conclusive. This review highlights these concerns. We first present an overview of the method employed by Kocsis and his colleagues to examine profile accuracy because this procedure forms the basis of many of our concerns. We then discuss how conceptual and methodological flaws in Kocsis’ research undermine the arguments presented by Kocsis (2003a). We conclude by proposing recommendations as to how the various issues discussed can be addressed in subsequent investigations.

An Overview of the Kocsis Research Paradigm

To investigate the profiling abilities of their participants, Kocsis and his colleagues have adopted a paradigm similar to that used by Pinizzotto and Finkel (1990). The Kocsis paradigm involves collecting profiling data from groups of participants believed to possess one of the skills that professional profilers view as integral to their vocation. Group performances are then compared to determine whether a specific type of participant (e.g., police officers) and therefore a particular type of skill (e.g., investigative experience) produces a more accurate profile.

Profiling data are obtained by getting each participant to review case material surrounding a solved crime or crime series. Participants are then presented with a

multiple-choice questionnaire that surveys various offender characteristics including physical features, cognitive processes, offense behaviours, and social dynamics (Kocsis et al., 2000, also examined predictions of personality characteristics). Their task is to complete the questionnaire by selecting the options they believe match the characteristics of the offender known to have committed the crime. In their initial study, Kocsis et al. (2000) also requested that participants construct a written profile.

Profile accuracy in the Kocsis studies is measured by comparing the profile constructed by each participant to a correct profile constructed by the police officer who was originally responsible for investigating the crime in question. The higher the degree of correspondence between the two profiles, the greater the participant's accuracy. In each of their investigations, Kocsis and his colleagues divide these accuracy measures into subcomponents representing physical features, cognitive processes, offense behaviours, and social dynamics (plus personality characteristics on one occasion). Two omnibus accuracy scores are also computed. The first is based on the overall number of correctly answered questions across all subcomponents (excluding personality characteristics). The second is based on participants' responses to questions similar to those employed by Pinizzotto and Finkel (1990).

Problems With Accuracy Arguments

A number of our criticisms relate to the way that Kocsis and his associates have measured and compared profiling accuracy across the sampled groups and the conclusions they have drawn from their analyses.

Are There Problems With Kocsis' Accuracy Measure?

If Kocsis' measure of profile accuracy (for the multiple-choice profiles) is unsound, then all of the arguments that follow from this analysis are open to criticism. Although we do not claim that his approach to measuring profile accuracy is fundamentally flawed, we do have concerns about this measure.

Our primary concern relates to the degree to which certain items on the multiple-choice questionnaires are open to interpretation by both the participant and the investigator and, thus, the extent to which disagreements between a participant's profile and the investigator's profile reflect differences in interpretation versus genuine inaccuracies. In our opinion, answers to numerous items on these questionnaires appear to be highly dependent on the interpretation of the respondent.² This issue is salient when answering questions about the offender's height (short, average, tall, very tall), the offender's general build (thin, average, solid, fat), the offender's hair color (red, brown, black, gray, bald), the offender's hairstyle (bald, very short, short straight, short curly, long straight, long curly), the offender's alcohol consumption (nil, low, medium, in binges, high), and so on. Answers to such questions are based on the perspective of the people answering them. Even if the offender were known to all in question, responses

to such items would be subjective and would likely vary from person to person (e.g., the first author of this article, who is 6 feet 4 inches, would view a person who is 5 feet 11 inches as quite short, whereas the second author, who is 5 feet 7 inches, would view the same person as quite tall).

A related concern is how an investigator (even one with detailed knowledge of the case) could accurately determine the answers to other questions included on the multiple-choice questionnaires. For example, responding to questions about the offender's primary motive for the offense, whether the offender had fantasies about killing someone before the crime, whether the offender felt comfortable in the area where the offense took place, and whether the offender experienced remorse about the offense, all require a subjective interpretation. Kocsis does not clarify how the investigators in his studies answered these questions (presumably they were able to interview the offender in some depth following the arrest). However, some researchers maintain that it can be difficult to uncover answers to these sorts of questions with any confidence or accuracy, even after an extensive interview with the offender (e.g., Canter, 2000). Thus, individuals' interpretations of these questions will yield varying answers, which will also bias Kocsis' accuracy measure.

How Do Professional Profilers Compare (in Relative Terms) to the Other Groups?

We also believe there are difficulties in how Kocsis and his colleagues have analyzed and interpreted their data. For example, we disagree that Kocsis' data support his conclusion that professional profilers outperformed all other groups and that students are the next most proficient group after profilers.

Does it make sense to collapse groups? Based on the results presented by Kocsis et al. (2000) and Kocsis (2004), it seems erroneous to conclude that professional profilers in these studies outperformed other groups. In Kocsis et al. (2000), for example, there were no statistically significant differences observed between the mean accuracy scores of the profilers and the scores of any other group. In fact, the only significant differences consisted of psychologists outperforming police officers on predictions of physical features and offense behaviours. Even Kocsis and his colleagues (2000) recognized this when they stated, "In spite of their training, knowledge, and experience, profilers did no better than anyone else in the correct identification of features of the offender or offense" (p. 321).

However, Kocsis et al. (2000) proceed to argue:

Profilers did *descriptively* outperform all other groups on the two omnibus measures of accuracy and on two of the submeasures On the other two submeasures, the profilers were the second most accurate group, *with the difference between them and the most accurate group (psychologists) negligible and easily attributable to sampling error.* (p. 321, italics added)

This argument is strange because the greater performance of profilers (on the two omnibus measures and the two submeasures) was also negligible and, therefore, might also be attributable to sampling error.

What is even more irreconcilable is the decision by Kocsis et al. (2000) to undertake “a more sensitive analysis” (p. 321) to determine whether the profilers produce more accurate profiles than the other groups when those groups were combined to form an aggregate nonprofiler group. This procedure does not permit a more sensitive analysis of participant accuracy because the aggregation of raw data is known to reduce the sensitivity of data to individual variation (i.e., regression to the mean). This approach likely explains why the resulting analysis found that, “On every measure of accuracy, the profilers answered more questions correctly than the non-profilers” (p. 321; though it should be stressed that profilers were significantly more accurate than the nonprofilers on only 3 of the 7 accuracy measures).³

Collapsing the nonprofiler groups into one aggregate group tips the scales in the direction of finding significant results in favor of the profilers because the nonprofiler group contains the very low accuracy scores of the police and psychic groups. This suggests that, if the aggregation procedure were used to examine the performance of other isolated groups, it is possible that the relative accuracy results would be comparable to those obtained when comparing profilers to the combined nonprofiler group. If this occurred, the conclusions of Kocsis et al. (2000) would have to be reconsidered. For example, if the psychologists outperformed a nonpsychologist group that included professional profilers, it could be argued that mental health professionals are the superior profilers.

Although it is difficult to reanalyze the data provided by Kocsis et al. (2000), primarily because the standard deviations of accuracy scores are not specified, we explored what might have occurred if the aggregation procedure had been applied to different group combinations. For each profile component (e.g., physical features), we assigned every participant the mean accuracy rating for his or her corresponding group. Although our findings suggest that profilers did tend to outperform nonprofilers ($p < .05$, except when accuracy scores were compared for personality features), psychologists also tended to outperform nonpsychologists ($p < .05$, except when accuracy scores were compared for social dynamics). Based on our analysis, students performed about as well as nonstudents, whereas police officers and psychics performed worse than non-police officers and nonpsychics, respectively.

What do the meta-analysis results mean? 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). To compare results from these investigations, “The accuracy scores were first standardized within crime type by converting the scores to z scores” (Kocsis, 2003a, p. 132). Kocsis (2003a) argues that this procedure allows him to draw conclusions about the relative accuracy of the

groups that he and his colleagues have previously tested. Based on the results of this analysis, it appears as though profilers have produced the most accurate profiles, followed by students and then psychologists.

Standardized scores might theoretically allow for meaningful comparisons between groups across studies. However, it is our contention that the particular form of meta-analysis conducted by Kocsis (2003a) creates a bias in favor of certain groups, namely profilers and students. This bias is because of two related reasons. First, the same participant groups were not sampled in each study. For example, whereas Kocsis et al. (2000) and Kocsis (2004) both tested profilers and students, only Kocsis et al. (2000) tested psychologists. Second, the most recent investigation (Kocsis, 2004) clearly involved a profiling task of lesser difficulty than the task required of participants in previous research, as indicated by the generally superior performance of participant groups in this latest study (see Table 1). Therefore, the use of these z scores to draw firm conclusions about relative group performance is misleading (e.g., the z scores calculated for the psychologist group might have been higher if they had participated in the most recent, less difficult profiling task).⁴

How Accurate Are Professional Profilers Really?

Across all of their studies, Kocsis and his colleagues have focused on the relative performance of participant groups. Although they provide some data on absolute accuracy, they do not draw any conclusions about the accuracy of professional profilers (or any other group) in absolute terms. This is unfortunate because absolute accuracy is more important than relative accuracy in applied settings where police require effective profiles. Demonstrating that professional profilers are more accurate than others in producing profiles does not rule out the possibility that they (or the other groups) can be very inaccurate in their predictions. In fact, this does appear to be the case in some of Kocsis' studies given that the absolute accuracy of the participating profilers is often unimpressive.

Table 1 contains the absolute accuracy scores for Kocsis' participant groups across the various components included in the multiple-choice profile (accuracy scores for personality characteristics could not be calculated because we are not provided with information on the total possible score that could be achieved for this component). These data show that profilers, relative to other groups, are more accurate on certain components. For example, compared to the psychologists taking part in the study by Kocsis et al. (2000), profilers are more successful when it comes to making predictions about the offender's cognitive processes and social dynamics. However, with respect to these predictions, profilers still only reach accuracy levels of 46% and 43%, respectively.

Although some may view this degree of accuracy as investigatively useful as long as the profile is not relied on exclusively, we disagree for the following two reasons. First, it seems to us that a minimum standard must be set for profiling advice to be

TABLE 1
Absolute Accuracy Scores Across the Kocsis Studies
for the Various Groups Sampled

Participant Group (<i>n</i>)	Total Accuracy (%)	Cognitive Processes (%)	Physical Characteristics (%)	Offense Behaviours (%)	Social Dynamics (%)
Kocsis, Irwin, Hayes, and Nunn (2000)					
Profilers (5)	46	46	60	57	43
Psychologists (30)	42	32	61	58	26
Students (31)	40	29	57	52	29
Police (35)	39	36	57	44	26
Psychics (20)	38	37	47	52	23
Kocsis, Hayes, and Irwin (2002)					
Homicide detectives (12)	33	27	53	45	17
Senior detectives (31)	39	34	51	53	24
Trainee detectives (19)	39	35	54	50	26
Police recruits (50)	39	36	53	45	29
Students (31)	41	31	57	53	31
Kocsis (2004)					
Profilers (3)	70	81	79	67	58
Fire investigators (12)	54	55	70	61	38
Detectives (13)	49	56	64	62	30
Students (21)	59	65	72	71	41

considered useful (e.g., at least half of the information provided by the profiler must be accurate) or we run the danger of considering any level of profile accuracy investigatively useful (e.g., 30%, 20%, 10%, etc.). Second, if the level of accuracy attained by the profilers in the Kocsis studies is to be considered investigatively useful, then by extension the information provided by any of Kocsis' groups that achieve similar levels of accuracy could inappropriately be considered a potential asset to a serial crime investigation (e.g., using profiles constructed by college students).

The absolute accuracy scores obtained by the professional profilers and other groups sampled in Kocsis (2004) are higher than are those achieved in previous studies. The improved performance of participants in this study is perhaps attributable to the fact that Kocsis (2004) presented participants with information pertaining to a crime series (rather than a single crime, as used in previous studies). Using a series of crimes presumably provides participants with more relevant information on which to base a profile. Moreover, presenting participants with a crime series is a more realistic depiction of the circumstances in which profilers are typically employed as consultants. The problem with drawing any meaningful conclusions from the absolute accuracy scores in this particular study is that the profiler sample only consisted of three participants.

What Do the Written Profiles Tell Us About Accuracy?

In Kocsis et al. (2000), each participant was asked to construct a written profile of the unknown offender prior to completing the multiple-choice profile. Although these written profiles were not analyzed in the original study, they were evaluated in a subsequent investigation (Kocsis, 2003b). One of the reasons for studying these written profiles was to circumvent the problems associated with the multiple-choice profile, which “imposed various recognition and response parameters that are not typically encountered in the actual investigative application of the profiling technique” (p. 38). As discussed below, we agree that the multiple-choice questionnaires do not provide an appropriate test of profile accuracy. Therefore, an assessment of written profiles is potentially useful. Unfortunately, the accuracy of the written profiles has never been examined. Rather, Kocsis (2003b) adopted Pinizzotto and Finkel’s (1990) strategy of analyzing the length and types of descriptions included in the profiles.

Consistent with the Pinizzotto and Finkel (1990) study, Kocsis (2003b) found that “professional profilers wrote more lengthy profiles than did any of the other groups examined. In addition, these longer profiles contained more predictions about the offender” (p. 42). However, the results were less conclusive when the profiles were broken down into smaller components. For example, Kocsis found no significant group differences in the amount of information provided about the unknown offender’s physical features. Thus, even in terms of the length and detail of criminal profiles, the results of Kocsis provide only modest support for the superior performance of professional profilers.

We are not absolutely certain what the length of a profile or its level of detail actually conveys. We do believe, however, that it is a mistake to assume that longer, more detailed profiles are more accurate until this is demonstrated empirically (and these findings are replicated). This is what Kocsis (2003b) seems to be implying when he states that the results from Kocsis et al. (2000) and Kocsis (2003b) indicate that the quality (i.e., accuracy) and quantity (i.e., length) of a profile are related. Support for our argument comes from empirical studies that have demonstrated that profile quality and quantity are not necessarily related. For example, O’Keeffe and Alison’s (2000) study demonstrated that psychic detectives wrote longer, more detailed profiles compared to a control group, but the profiles written by psychics were not more accurate.

Problems With Skill Arguments

The purpose of the Kocsis research program is not simply to determine which participant group is superior in terms of profiling accuracy but also to elucidate why they are superior. Their approach to generating such explanations is to gather profiling data from participant groups that supposedly possess different skills. It is assumed that the performance of a particular group reflects the importance of that group’s particular skill to successful profiling. We have several concerns with this approach.

Have Kocsis and His Colleagues Even Been Comparing Skills?

Kocsis and his colleagues suggest that their participant groups possess representative skills (e.g., psychologists possess an appreciation of the criminal mind). However, on only one occasion has Kocsis actually assessed his participant groups to determine whether they possess the skills ascribed to them. No assessment of skills was conducted in Kocsis et al. (2000) or Kocsis (2004). On the other hand, Kocsis et al. (2002) appear 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 recruited in subsequent investigations. Moreover, this manipulation check should have been applied to test whether the other groups of participants possessed the skills they were said to possess.

The assessment of skills is central to Kocsis' research program. In fact, we believe that no valid conclusions about important profiling skills can be drawn from the accompanying data without this assessment. In addition, we believe that the results of this assessment warrant detailed presentation and discussion, thereby enabling the reader to judge for himself or herself whether the sampled groups possess their purported skills. As we argue below, one difficulty in conducting this assessment is that some of the skills Kocsis discusses may prove difficult to define and evaluate (e.g., intuition). Despite these obstacles, an attempt at skills assessment is foundational to the arguments in Kocsis' studies.

How Likely Is It That the Groups Possess the Skills They Are Said to Possess?

Without empirical data to support the notion that participant groups possess different skills, the reader is left to make this determination. Hence arises another serious problem for Kocsis (2003a), as we consider it very unlikely that certain groups actually possess the skills attributed to them. The self-declared psychic group is the best case in point. This group was "selected to demonstrate the capacity for intuition" (Kocsis, 2003a, p. 131). The fact that professional psychics claim to possess a high degree of intuition does not necessarily indicate that they do. Indeed, there is evidence to suggest that psychic-like predictions can be generated by techniques such as cold reading that have nothing to do with psychic powers or intuition (e.g., Hyman, 1981). Likewise, existing research, some of which is cited by Kocsis, indicates that it is doubtful psychics possess any more intuition than the average person (O'Keeffe & Alison, 2000; Wiseman, West, & Stemman, 1996). Thus, it is inappropriate to recruit a group of psychics to represent the skill of intuition when research fails to provide evidence of this skill in psychics.

Psychics are not the only group whose attributed skills are questionable. For example, it is arguable that many of the psychologists recruited by Kocsis do not have

a good appreciation of the criminal mind. Indeed, the psychologists examined by Kocsis et al. (2000) "had not formally studied forensic or criminal psychology" (p. 316), which is the area of psychology that seeks to understand the complexities of the criminal mind. Kocsis (2003b) reiterates this point when he indicates that the "psychologists, as far as we are aware, did not have any special knowledge or expertise in criminal behaviour per se" (p. 43). Given the confusion within the forensic psychology community about what drives the criminal mind, it is unlikely that general clinical psychologists have much insight regarding criminals or the way in which their minds operate.

We know of no research that directly supports our argument that psychologists lack an appreciation for the criminal mind, but there is indirect evidence. For example, it is common for psychologists with no specialized training in forensic psychology to conduct assessments within judiciary contexts (Bartholomew, 1981; Skeem & Golding, 1998). Basing their evaluations on traditional clinical skills, psychologists often demonstrate significant disagreement with respect to a defendant's specific abilities and impairments (Skeem & Golding, 1998). Furthermore, Skeem and Golding (1998) found that training failed to improve the quality of these psychologists' forensic reports. Although these investigations pertain to assessment, it is reasonable to assume that similar issues pervade other areas of forensic psychology such as criminal profiling.

What Do the Skills Even Mean and Do They Coexist?

We have two additional concerns about the specific skills that Kocsis and his colleagues set out to examine. The first relates to how the skills are defined. With the exception of investigative experience, Kocsis and his associates do not provide operational definitions of the various skills discussed. Consequently, even if one assumes that the participants possess the skills they were attributed, it is still unclear what abilities these participants actually possess. For example, the so-called skill of intuition is not defined by Kocsis and his colleagues, despite the fact that numerous definitions are available (e.g., intuition as immediate insight; Metcalfe & Weibe, 1987). If the skills that Kocsis has emphasized cannot be operationalized in such a way that they can be reliably and accurately measured, then this calls into question the very issue that drives the Kocsis research program (i.e., identifying skills that relate to success in profiling).

Our second issue is the implicit assumption that the skills highlighted in the Kocsis studies do not coexist within participants. Kocsis (2004) does acknowledge this fact in statements such as: "It is recognized that science students are not exclusively representative of skills of logical and objective reasoning" (p. 348). However, the experimental approach taken by Kocsis in terms of linking individuals' performance to a particular skill is based on the supposition that skills do not coexist. For example, how can one directly equate student performance in a profiling task with the importance of logical reasoning without assuming that this is the only skill the students possess? It is plausible that every participant in the Kocsis studies possess a degree of each skill (with the exception of investigative experience). If our argument is valid, generating simple cause-effect statements (i.e., students perform well, therefore logical reasoning

must be important) would be inappropriate. It may be that other skills possessed by students, or unique combinations thereof, underlie their profiling ability.

Recognizing the coexistence of skills is important for the Kocsis research program. Without this underlying assumption, Kocsis cannot claim to be testing the argument posited by Hazelwood et al. (1995) that successful profilers are characterized by a collection of skills. For his results to be meaningful, Kocsis should have assessed participants for representative skills and for every other skill that has been defined as important for profiling.

Problems With Validity Arguments

In addition to our concerns about the accuracy and skill arguments put forward by Kocsis (2003a), we also have several general concerns regarding the validity of his findings. In particular, we question the degree to which the results presented by Kocsis are externally valid given the fact that the research procedure he uses in his studies is, in his own words, "artificial" (Kocsis, 2003a, p. 139). In all fairness to Kocsis and his associates, throughout their articles they have discussed a number of the issues to be presented in this section. However, although aware of these problems, Kocsis (2003a) reaches conclusions that are often not appropriately tempered in light of the validity issues.

Were Enough Professional Profilers Sampled?

The number of profilers that have taken part in the Kocsis research program is small compared to other areas of research. Kocsis (2003a) recognizes this fact, but states that "the current sample of 11 individual professional profilers does . . . represent the largest empirical sample currently available" (p. 134; as indicated in Note 4, we only count 8 profilers across Kocsis' studies). We appreciate the difficulty in recruiting consenting profilers for such research studies and agree that the combined Kocsis sample represents the largest profiler sample to date. However, the recognition of small sample sizes does not circumvent the problems associated with this issue. It is simply not possible to generalize the results reported by Kocsis (2003a) to the wider profiling community based on a sample of 11 (or 8) profilers. In addition, as indicated above, the only study in which profilers have performed satisfactorily is Kocsis (2004), and the profiler sample in that study consisted of only three participants.

The inability to generalize beyond the Kocsis research program is compounded further by the fact that the recruited profilers are heterogeneous in their composition. According to Kocsis (2003a), the sampled profilers employ very diverse approaches (e.g., ranging from a reliance on psychodynamic ideas to the use of behavioural typologies) and come from different countries (e.g., the United Kingdom, the United States, Australia). 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.

Why Was More Weight Placed on the Multiple-Choice Profile Exercise?

The majority of participants across the Kocsis studies have been evaluated based solely on their performance on a multiple-choice profiling task. We believe this multiple-choice task poorly reflects how profiling is executed in the real world, a fact that was recognized by Kocsis (2003a) when he stated, "This work only examines profiling performance within an artificial context of a multiple-choice questionnaire" (p. 139). On account of this artificial context, Kocsis should limit the weight he places on multiple-choice profile results in terms of not overstating the generalizability of his findings. Requiring participants to construct written profiles, rather than complete multiple-choice profiles, more closely resembles the task of actual profilers. Curiously, after having participants complete the written profiling task in their first study (Kocsis et al., 2000), Kocsis abandoned this task completely in his subsequent investigations. Even when Kocsis and his colleagues collected written profiles, they did not analyze them for accuracy. 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. Moreover, it would allow researchers to determine empirically whether a relationship exists between written and multiple-choice profiles.

Were Different Time Limits Imposed on the Different Groups?

Kocsis has stated that no time limit was imposed on his participants for completing the profiling task. However, we would argue that, as a result of their data-collection procedure, Kocsis and his colleagues might have inadvertently imposed different time limits on the sampled participant groups. Our concern has to do with the fact that "the survey was completed by mail by the profiler group . . . [whereas] the inventory was administered in person to the members of all other groups" (Kocsis et al., 2000, p. 318).⁵ This difference would likely impose a shorter time limit on participants being group tested. At the very least, this could explain why profilers in the study by Kocsis (2003b) constructed slightly longer, more detailed profiles compared to other participants. The time difference, if it did exist, may also have had an effect on the accuracy of the multiple-choice profiles generated by the various groups because the profilers could theoretically process the material during a longer period and access useful resource material (even though they were explicitly instructed not to do so). Even if this time issue is not viewed as important, the time taken by every group to complete their profiling task was probably far shorter than the actual time it takes a professional profiler to construct a profile (Boon, 1997, indicates that a single

profile can take upwards of 40 hours to construct). This fact might also have an effect on the degree to which the results reported by Kocsis (2003a) can be generalized.

Why Was the Degree of Interaction Kept to a Minimum?

Profiling is often viewed as an interactive process that involves brainstorming sessions (Douglas & Burgess, 1986). The fact that Kocsis and his colleagues set up their experiments so that no interaction could take place among participants, or between participants and the researchers, again reduces the generalizability of their findings to naturalistic settings. We assume that Kocsis implemented these conditions to maintain a degree of control over the experimental situation. However, if participants are unable to construct their profiles under realistic conditions (e.g., by asking questions for clarification), how are we to know how well they might perform on genuine profiling tasks? This issue might be particularly relevant to professional profilers. It may be the case that profilers are able to pose the right questions to investigators (or to other profilers), and this is what permits them to piece crimes together and construct accurate profiles. Given this possibility, it might have been preferable to forgo a degree of experimental control by allowing participants to interact so that more valid results could be obtained.

Recommendations for Future Studies

Based on our critique of the Kocsis research program, a number of recommendations can be made that will hopefully allow future studies in this area to yield results that are more meaningful and generalizable.

Recommendations Related to Accuracy

We have concerns as to what Kocsis' accuracy measure actually represents because numerous items contained on the multiple-choice questionnaires are either open to interpretation by respondents (including the investigator) or are unverifiable. One easy way to address this problem in the future is to increase the degree of objectivity associated with the multiple-choice questions. For example, rather than using subjective multiple-choice options for offender height, such as short, specific height ranges could be used, such as 5 feet 0 inches to 5 feet 4 inches.

We also recommend that aggregate comparisons be abandoned as a method for comparing sampled groups or, at the very least, that the method be exhausted so that valid conclusions about the relative accuracy of the various groups can be drawn. We suggest that future research abandon aggregate comparisons because real differences in profiling accuracy should be evident without aggregation. The fact that this method of analysis has been necessary potentially indicates that profilers do not possess any particular set of skills that make them more effective than other sampled groups (at least in these artificial testing conditions). We are encouraged by Kocsis'

use of meta-analytic procedures and are confident that more formal meta-analytic techniques could be used to uncover patterns in the data once a sufficient number of studies have been conducted. However, to reach this stage, researchers must ensure the standardization of their data-collection procedure across studies and make every attempt to include the same types of participant groups (e.g., psychologists).

Lastly, we believe it is important to focus on the absolute performance of the various groups in addition to their relative performance. It is valuable to examine relative group performance to determine whether the individuals responsible for profiling offenders in our society are in fact superior profilers. However, it is just as important, if not more so, to establish the success of professional profilers in absolute terms so that police agencies can make informed decisions about the role of criminal profiling in their investigations.

Recommendations Related to Skills

Perhaps our most important recommendation is that participants be assessed for the various skills discussed by Kocsis. This is crucial if any convincing conclusions are to be drawn from this body of research. To address this issue, several matters require attention. First, the various skills that Kocsis attempted to examine need to be operationalized in such a way that leaves no doubts as to the construct being measured. Only once these definitions have been developed can we proceed to assess whether participants possess those skills and, if so, to what extent.

Existing assessment procedures could be used to assess some of the skills that Kocsis and his colleagues view as important. For example, in the future, researchers might want to use the Watson-Glaser Critical Thinking Appraisal Form to assess participants for their capacity for logical reasoning (Inlow & Chovan, 1993). For other skills, such as an appreciation of the criminal mind, new assessment procedures will have to be developed.

The results that emerge from the use of these assessment procedures will allow researchers to be more confident that they are measuring the skills they intended to measure. Using a full battery of assessment procedures on every participant will also enable researchers to deal with the coexistence of skills in the same participant. In all likelihood, multivariate statistical procedures will then be required to accurately determine whether a relationship exists between these various skill sets and profiling ability.

Recommendations Related to Validity

If future research in this area is to be translated into usable policy and practice, the validity of the research findings must improve. In this vein, our first recommendation in this section relates to sample size. As stated previously, the number of profilers sampled by Kocsis across investigations is small and not representative of profilers in general. Hence, every attempt must be made in the future to convince more profilers to participate in such research.⁶ To this end, profilers will have to perceive the value of partaking

in these studies and view the experiments as a fair reflection of their profession. Many of the changes that we recommend in the following paragraph may assist in this capacity.

There are a number of changes that can be made to the data-collection procedure used by Kocsis and his colleagues to increase the validity of future results. First, participants should be required to construct written profiles, the accuracy of which should be given equal weight to the multiple-choice profiles that are now commonly used by Kocsis. Second, every attempt should be made to ensure that all participants are tested under the same conditions, with particular attention paid to the amount of time participants are allotted to complete the task. If it is not possible to test the profilers in person because of logistical constraints, consideration should be given to testing every participant by mail. Third, researchers should reconsider the value of allowing participants to ask questions during the testing period. If profiling is typically conducted in a collaborative fashion, it is important to ensure that these conditions are approximated in research.

Conclusion

Although we applaud Kocsis and his colleagues for asking some pertinent questions about the abilities of professional profilers and for developing a long-term research program to explore these issues, we also have several methodological and conceptual concerns with the research. Some of the more serious problems relate to the soundness of their accuracy measure, the biases in their methods of analysis, the lack of any skills assessment, and the validity of their data-collection procedure. Based on these and other related issues, we argue that it is premature to accept the interpretation of current data offered by Kocsis (2003a). We expect that future studies will be conducted in this area and hope that some of our recommendations will allow for the collection of data that are more meaningful and generalizable.

Notes

1. Kocsis and his colleagues (Kocsis, 2003a, 2003b, 2004; Kocsis, Hayes, & Irwin, 2002; Kocsis, Irwin, Hayes, & Nunn, 2000) have also examined other groups, but the samples listed here are among the most commonly used.

2. Interestingly, Kocsis, Irwin, Hayes, and Nunn (2000) did recognize this problem when it came to their personality checklist but obviously felt the same problem did not apply to their multiple-choice questionnaire. When first introducing their results from the personality checklist, they state, "Our approach to this question is only *exploratory* and *impeded partially by ambiguities* as to how to measure whether a respondent is correct in his or her psychological interpretation" (p. 322, italics added). 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.

3. In our opinion, Kocsis and his colleagues have overstated the significance of these findings in subsequent studies as well. As we have stated here, in Kocsis et al. (2000), no significant results were found when profilers were compared (individually) to any other group. Furthermore, only a few significant

differences emerged when Kocsis et al. used the aggregation procedure. However, when discussing these results in a later study, Kocsis (2004) states that, "Kocsis et al. (2000) found that professional profilers outperformed *all* [italics added] the nonprofiler groups with regard to the overall accuracy of their predictions" (p. 343). This statement does not correspond to the available data.

4. Another problem with this meta-analysis relates to the sample sizes reported by Kocsis (2003a; Table 1, p. 133). Kocsis maintains that, across the three studies on which the analysis was based, he sampled 11 profilers, 36 psychologists, 85 science students, 25 specialist detectives, 12 nonpolice specialists, 85 general police officers, 50 police recruits, and 20 psychics. However, many of these numbers do not appear to add up. For example, as indicated in Table 1 of the current article, we calculate that Kocsis has sampled only 8 profilers, not 11, across his studies. Instead of 36 psychologists, we calculate that Kocsis has sampled just 30.

5. Kocsis et al. (2000) claim that most participants were tested in person for security reasons. It seems strange to us that the same security issues did not apply to the profilers.

6. Using psychologists or students as proxy groups in future studies to combat low profiler participation, as recommended by Kocsis (2003a), is not a sensible option in our opinion. His rationale is that these participants performed only slightly worse than professional profilers in previous studies and, moreover, that they are much more accessible and willing to participate in this research compared to professional profilers. Given the problems we have raised in this article, it appears presumptuous to suggest the use of proxy groups based on existing accuracy data. In addition, it seems to us that the recommendation to use proxy groups defeats the sole purpose of the Kocsis studies. If psychologists or students perform as well as professional profilers, why do we even need professional profilers?

Acknowledgments

We would like to thank an anonymous reviewer for his or her many helpful comments on an earlier draft of this article.

References

- Bartholomew, A. A. (1981). Criminal intent and the psychologist. *Australian Psychologist, 16*, 413-421.
- Boon, J. C. W. (1997). The contribution of personality theories to psychological profiling. In J. L. Jackson & D. B. Bekerian (Eds.), *Offender profiling: Theory, research, and practice* (pp. 43-59). Chichester, UK: Wiley.
- Canter, D. V. (2000). Offender profiling and criminal differentiation. *Legal and Criminological Psychology, 5*, 23-46.
- Douglas, J. E., & Burgess, A. E. (1986). Criminal profiling: A viable investigative tool against violent crime. *FBI Law Enforcement Bulletin, 55*(12), 9-13.
- Hazelwood, R. R., Ressler, R. K., Depue, R. L., & Douglas, J. E. (1995). Criminal personality profiling: An overview. In R. R. Hazelwood & A. W. Burgess (Eds.), *Practical aspects of rape investigation: A multidisciplinary approach* (2nd ed., pp. 115-126). Boca Raton, FL: CRC Press.
- Hyman, R. (1981). The psychic reading. In T. A. Sebeok & R. Rosenthal (Eds.), *The clever Hans phenomenon* (pp. 169-181). New York: New York Academy of Sciences.
- Inlow, F. H., & Chovan, W. (1993). Another search for the effects of teaching thinking and problem solving skills on college students' performance. *Journal of Instructional Psychology, 20*, 215-223.
- Kocsis, R. N. (2003a). Criminal psychological profiling: Validities and abilities. *International Journal of Offender Therapy and Comparative Criminology, 47*, 126-144.
- Kocsis, R. N. (2003b). An empirical assessment of content in criminal psychological profiles. *International Journal of Offender Therapy and Comparative Criminology, 47*, 38-47.

360 International Journal of Offender Therapy and Comparative Criminology

- Kocsis, R. N. (2004). Psychological profiling of serial arson offenses: An assessment of skills and accuracy. *Criminal Justice and Behavior, 31*, 341-361.
- Kocsis, R. N., Hayes, A. F., & Irwin, H. J. (2002). Investigative experience and accuracy in psychological profiling of a violent crime. *Journal of Interpersonal Crime, 17*, 811-823.
- Kocsis, R. N., Irwin, H. J., Hayes, A. F., & Nunn, R. (2000). Expertise in psychological profiling: A comparative assessment. *Journal of Interpersonal Violence, 15*, 311-331.
- Metcalf, J., & Weibe, D. (1987). Intuition in insight and non-insight problem-solving. *Memory and Cognition, 15*, 238-246.
- O'Keeffe, C., & Alison, L. J. (2000). Psychic rhetoric in psychic detection. *Journal of the Society for Psychical Research, 64*, 26-38.
- Pinizzotto, A. J., & Finkel, N. J. (1990). Criminal personality profiling: An outcome and process study. *Law and Human Behavior, 14*, 215-234.
- Skeem, J. L., & Golding, S. L. (1998). Community examiners' evaluations of competence to stand trial: Common problems and suggestions for improvement. *Professional Psychology: Research and Practice, 29*, 357-367.
- Winerman, L. (2004). Does profiling really work? *APA's Monitor on Psychology, 35*, 67.
- Wiseman, R., West, D., & Stemman, R. (1996). An experimental test of psychic detection. *Journal of the Society of Psychical Research, 61*, 34-45.