Journal of Investigative Psychology and Offender Profiling
J. Investig. Psych. Offender Profil. **2**: 23–41 (2005)
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI**: 10.1002/jip.21

Between a ROC and a Hard Place: A Method for Linking Serial Burglaries by *Modus Operandi*

CRAIG BENNELL* and NATALIE J. JONES
Department of Psychology, Carleton University, Ottawa, Ontario, Canada

Abstract
The purpose of this study is to determine if readily available information about commercial and residential serial burglaries, in the form of the offender's modus operandi, *provides a statistically significant basis for accurately linking crimes committed by the same offender. Logistic regression analysis is applied to examine the degree to which various linking features can be used to discriminate between linked and unlinked burglaries. Receiver operating characteristic (ROC) analysis is then performed to calibrate the validity of these features and to identify optimal decision thresholds for linking purposes. Contrary to crime scene behaviours traditionally examined to link serial burglaries, the distance between crime site locations demonstrated significantly greater effectiveness as a linking feature for both commercial and residential burglaries. Specifically, shorter distances between crimes signalled an increased likelihood that burglaries were linked. Thus, these results indicate that, if one examines suitable behavioural domains, high levels of stability and distinctiveness exist in the actions of serial burglars, and these actions can be used to accurately link crimes committed by the same offender. Copyright © 2005 John Wiley & Sons, Ltd.*

INTRODUCTION

The task of linking unsolved serial crimes is particularly relevant in police investigations. However, it is also a considerable challenge, especially in the absence of hard forensic evidence such as fibres, fingerprints, or DNA (Grubin, Kelly, & Brunsdon, 2001). Without such physical evidence, linking crimes may hinge solely upon behavioural information revealed from a thorough examination of crime scene characteristics and offence locations. These aspects of the criminal event are popularly regarded as the offender's *modus operandi* (MO) and they have been the subject of limited empirical study.

Historically, the MO concept has been associated with the assumption that a given offender will exhibit similar behaviours across their crimes and, furthermore, that these

*Correspondence to: Craig Bennell, Department of Psychology, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada, K1S 5B6. E-mail: cbennell@connect.carleton.ca

behaviours will be relatively distinct from those of other offenders committing similar types of crimes (Green, Booth, & Biderman, 1976). More recently, however, it has been argued that these assumptions are incorrect (e.g. Douglas & Munn, 1992). Most law enforcement personnel now assume that serial offenders will modify aspects of their criminal activity over the course of their crime series due to a range of situational and learning factors. As a result, the use of MO as a tool to link serial crimes is presently applied sparingly and with extreme caution.

However, it is possible that law enforcement officers have dismissed the MO concept prematurely due to having searched for signs of stability and distinctiveness in inappropriate behavioural domains. Indeed, results of recent research suggest that this may be the case. For example, Bennell and Canter (2002) demonstrated that certain behaviours exhibited by serial commercial burglars in the UK, principally their selection of crime site locations, demonstrated potential as an accurate crime-linking variable. Such spatial behaviours are not typically considered in attempts to link serial burglaries.

The objective of the present study is to replicate and extend the research conducted by Bennell and Canter (2002). In contrast to that original investigation, the data collected for this study pertains to both commercial and residential burglaries. Additionally, all of the analyzed crimes were committed in a different police jurisdiction and over a different time frame than the ones examined previously. Thus, a thorough analysis of the current data will determine the extent to which the results presented by Bennell and Canter (2002) generalize across different crime types, geographic areas, and temporal periods.

From a theoretical perspective, these results will be consequential in their indication of whether similar underlying psychological processes influence how different types of burglars commit their crimes. From a practical perspective, these findings will signify whether investigative strategies employed to link serial burglaries should be tailored to specific geo-demographic situations and crime types, or whether global strategies suffice. Thus far, no research has examined this issue as it relates to serial burglary. However, it seems to be the case that global linking strategies, which draw on a range of crime scene actions, are the norm for interpersonal serial crimes, such as rape and murder (e.g. Collins, Johnson, Choy, Davidson, & Mackay, 1998).

We begin by discussing a novel approach to behavioural linking that was originally proposed by Bennell and Canter (2002), whereby the problem is viewed as a diagnostic decision-making task. We then present a detailed description of an analytical technique, termed receiver operating characteristic (ROC) analysis, which can be applied to empirically validate and calibrate this proposed linking approach.

The behavioural linking problem as a diagnostic task

Bennell and Canter (2002) have argued that the behavioural linking problem can be fruitfully conceptualized as a diagnostic task '. . . similar, for example, to diagnosing cancer in radiology, assessing risk in psychiatry, predicting storms in meteorology, etc.' (p. 154). This analogy can be drawn because, fundamentally, the problem for the diagnostician in all of these tasks is identical: He or she must accurately and effectively identify and apply relevant diagnostic information in order to render the best possible decision. The decisions in each of these tasks are also basically equivalent. For example, given a pair of burglaries, the two possible predictions are that the crimes are linked (i.e. committed by the same person) or unlinked. Similarly, the two possible realties are that the burglaries are actually linked or unlinked. As a result of combining these two statements, four possible decision outcomes emerge, two of which are correct and two of which are incorrect. Such is the case for any two-alternative, yes–no type diagnostic task (Swets, 1988).

These four decision outcomes are generically referred to as hits, correct rejections, false alarms, and misses. In the current context, a *hit* occurs when the prediction that two crimes are linked is correct. A *correct rejection* occurs when unlinked crimes are accurately identified. A *false alarm* refers to the prediction that two crimes are linked when they are not. Finally, a *miss* denotes the prediction that two crimes are unlinked but they are not. The goal for the diagnostician, then, is to seek ways to increase the probability of rendering a correct decision and/or to decrease the probability of making an incorrect decision.

Conceivably, the best way to arrive at this objective is through the identification of features that reliably correspond with one diagnostic alternative but not the other. For example, many forms of cancer can now be reliably diagnosed from features observed in individuals with cancer that are not present in those people without the disease (e.g. Getty *et al*., 1997). In theory, this technique is also applicable to behavioural linking. For example, as stated by Bennell and Canter (2002), this approach '. . . would require the identification of some linking feature, or set of features, reliably associated with crimes committed by the same offender(s) that are not associated with crimes committed by different offenders' (p. 154).

In reality, however, it has in the past been proven difficult to identify such features. Criminal behaviour is far too complex and ambiguous to permit the identification of a subset of actions that perfectly discriminate between linked and unlinked crimes. In addition, there are strong empirical grounds for expecting that such discriminators are unlikely to exist for crimes such as burglary (Canter, 2000). The sheer frequency with which burglary is committed decreases the probability of deciphering unique behavioural elements that can be used to discriminate between offenders. So, is a traditional MO approach to behavioural linking in cases of serial burglary doomed to fail?

The results reported by Bennell and Canter (2002) suggest that, while perfect discriminators are unlikely to be found, the aforementioned approach can still be effective by identifying features that *tend* to occur at different rates for linked versus unlinked crimes. They proposed that across-crime similarity measures relating to various aspects of burglary behaviour might be useful in this capacity. More specifically, Bennell and Canter (2002) examined the degree of across-crime similarity exhibited by their sample of serial burglars in relation to crime site selection choices, entry behaviours, characteristics of the targeted properties, and items stolen. They showed that the inter-crime distance (i.e. the distance between two crime site locations) was by far the most effective linking feature, in that shorter distances reliably signalled an increased likelihood that burglaries were linked. In contrast, the behaviours that have traditionally been used to link serial burglaries, such as an offender's method of entry, proved to be much less useful. In fact, these conventional MO indicators resulted in burglaries being linked at a level that was only slightly greater than chance.

Bennell and Canter (2002) have also argued that, if categorical criteria are not available for the purpose of behavioural linking, such that their presence or absence indicates the correct decision, then an appropriate decision threshold must be established. This threshold refers to a cut-off point along a continuum of evidence whereby any value obtained above that point (or below it, depending on the evidence) results in a positive diagnostic decision (Swets, 1992). In our case, this threshold refers to a particular across-crime similarity score (relating to inter-crime distances, entry behaviours, items stolen, etc.) that defines the degree of similarity which two burglaries must exhibit before they are predicted to be linked. According to Swets (1996), the general goal is to set this threshold in order to '. . . produce the best balance among the four possible decision outcomes for the situation at hand' (p. 3).

Focusing on inter-crime distances for 86 solved commercial burglaries committed by 43 offenders in one area of a major UK city, Bennell and Canter (2002) demonstrated the importance of setting an appropriate threshold when linking serial burglaries. For example, when an inter-crime distance of 0.70 km was defined as a threshold, whereby any two burglaries committed within 0.70 km of one another would be linked, they were able to accurately identify 52.4% of linked burglaries and 93.2% of unlinked burglaries. Using a more lenient threshold of 2.50 km, accurate classification was possible for 61.9% of linked burglaries but for only 67.7% of unlinked burglaries. Thus, relatively small changes to the decision threshold can potentially have an enormous impact on the effectiveness of linking decisions.

Examining diagnostic tasks using ROC analysis

Bennell and Canter (2002) demonstrated the utility of a technique known as ROC analysis for setting optimal decision thresholds. This is a procedure widely accepted across many diagnostic fields for evaluating decision-making performance (Swets, Dawes, & Monahan, 2000). The technique was originally developed for the radar field to examine the effectiveness of radar technicians in distinguishing meaningful blips on their radar screens from background noise. Since that time, the technique has been applied to fields such as radiology, engineering, and psychology to measure discrimination accuracy and to identify appropriate decision thresholds (Swets *et al*., 2000).

In the current context, ROC analysis illustrates how the probability of making certain types of linking decisions is subject to change for a particular linking feature, or set of features, as decision thresholds are varied from lenient to strict (Swets, 1996). These probabilities are calculated using the frequencies of the four decision outcomes previously discussed (i.e. hits, correct rejections, false alarms, and misses). More specifically, for each of the decision outcomes represented in Table 1, conditional probabilities are calculated using the formulae provided. These obtained values represent the estimated likelihood of rendering a specific prediction given a certain reality (Swets *et al*., 2000).

Since the probabilities of each column in Table 1 add up to 1, only a single cell in each column is required to measure accuracy (Swets, 1988). The probability of making a hit

Table 1. Possible decision outcomes in the behavioural linking task

		Reality: Actually linked	Reality: Actually unlinked	
Prediction:	Linked	a hit pH[a] = $a/(a + c)$	b false alarm pFA[b] = $b/(b + d)$	$a + b$
	Unlinked	c miss pM[c] = $c/(a + c)$	d correct rejection pCR[d] = $d/(b + d)$	$c + d$
		$a + c$	$b + d$	$a + b + c + d = N$

[a]pH = probability of a hit; [b]pFA = probability of a false alarm; [c]pM = probability of a miss; [d]pCR = probability of a correct rejection.

(pH) and the probability of making a false alarm (pFA) are those most frequently considered. These probabilities also have the most serious implications in the current criminal justice context. It is important to recognize that these probabilities vary as a function of the decision threshold in a directly proportional manner (Swets, 1992). For example, when using inter-crime distance as a linking feature, adopting a relatively strict decision threshold (e.g. linking burglaries within 0.01 km of each other) would result in a low pFA but it would also yield a low pH. To increase pH, a more lenient threshold could be adopted (e.g. linking burglaries within 100.00 km of each other), but this would consequently create an increase in pFA.

An optimal decision threshold can be selected through any number of procedures (Swets, 1992). Most of these procedures would draw upon knowledge of the prior probability of both linked and unlinked crimes that exist in the police jurisdiction under consideration, as well as the costs and benefits associated with incorrect and correct decision outcomes, respectively. Other procedures would ignore this information totally and simply set a decision threshold that will allow the diagnostician to achieve a predetermined value of pH or pFA that is desirable for the situation at hand (Swets, 1992). For example, a police force may decide that it does not have the resources to exceed $p\text{FA} = 0.05$ for a particular task and, therefore, this rate will determine the appropriate decision threshold.

Once the values of pH and pFA have been calculated across various decision thresholds (this is typically done via a computer program), these probabilities are plotted on a ROC graph (pH on the y-axis and pFA on the x-axis). As illustrated in Figure 1, when the coordinate points are connected, the result is typically a concave downward curve, known as a ROC curve, which starts at the lower left corner of the graph (where the decision thresholds are strict) and ends in the upper right corner (where the decision thresholds are lenient) (Swets, 1988). The *area under* the ROC *curve*, referred to as the AUC, is a measure of discrimination accuracy for the particular linking feature(s) that gave rise to that curve.

AUCs can range from 1.00 (indicating perfect accuracy) to 0.00 (indicating perfect inaccuracy). An AUC of 1.00 represents a ROC curve that follows the left and upper axes of the graph, whereas an AUC of 0.00 represents a ROC curve that follows the bottom and right axes of the graph. An AUC of 0.50 (indicating chance accuracy) corresponds to a ROC curve that follows the positive diagonal on the graph, going from the bottom left corner to the upper right corner.

Using ROC analysis to examine the accuracy of behavioural linking decisions is advantageous for at least two reasons. First, the AUC is a measure of discrimination accuracy that is independent of the particular decision threshold adopted. This occurs because the AUC represents the position of the entire ROC curve in its graph rather than any single point along it (Swets, 1988). Alternative measures of discrimination accuracy (e.g. percentage correct) are biased by threshold placement. When employing such biased techniques, it would be impossible to determine whether the level of accuracy achieved when using a particular linking feature is due to the inherent discriminatory power of that feature or is simply attributable to the decision threshold adopted.

Second, because the AUC is based on the proportions of the various decision outcomes, rather than their raw frequencies, it does not depend on the relative frequencies of the diagnostic alternatives in any given sample (Swets, 1988). Thus, in our case, AUCs can be compared across samples that vary in terms of their frequencies of linked and unlinked crimes. This advantage is particularly important in the current study because we are proposing to directly compare levels of discrimination accuracy associated with various features used for linking across both commercial and residential burglary, as well

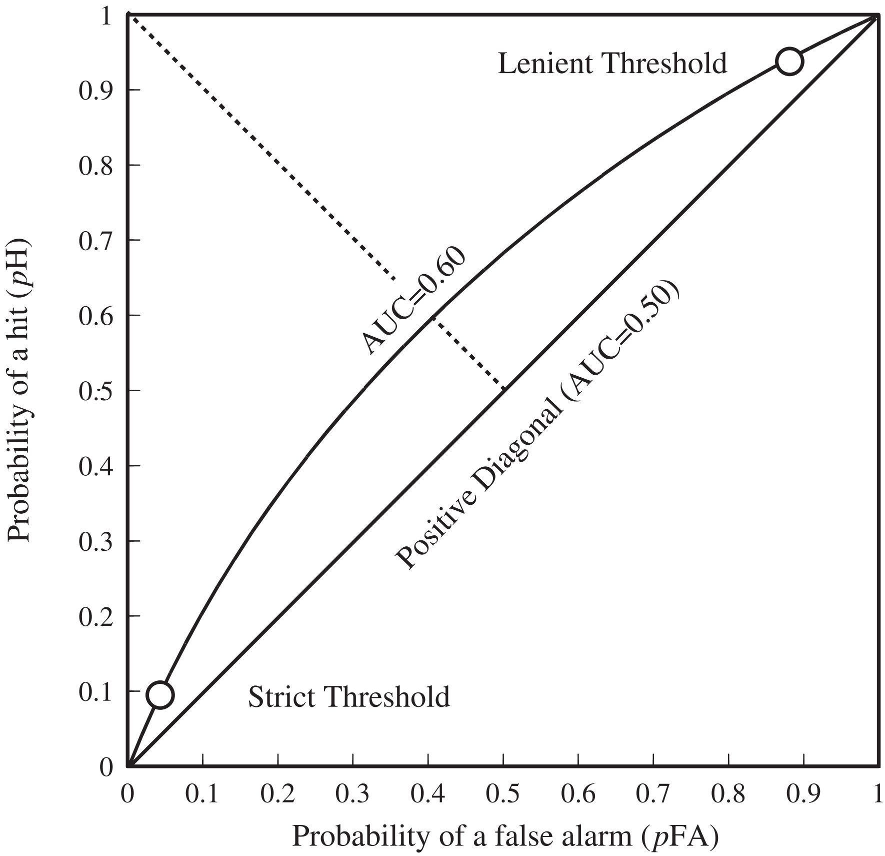

Figure 1. A ROC curve, for a specific level of discrimination accuracy, indicating a lenient and a strict decision threshold.

as across the current samples of burglary and the sample examined by Bennell and Canter (2002).

The aims of the present study

The present investigation attempts to determine if readily available (i.e. observable) information about burglaries can provide a statistically significant basis for linking these crimes to a common offender. ROC analyses will be used to calibrate the validity of the various linking features for classification purposes, applied on their own, and in combination. The present study will also consider the issue of how to identify optimal decision thresholds.

Our primary goal is to determine whether the results reported by Bennell and Canter (2002) generalize across different crime types, geographic areas, and temporal periods. If similar underlying psychological processes influence how different types of burglars (i.e. commercial versus residential) commit their crimes, comparable results would be expected. However, it is also plausible that the findings presented by Bennell and Canter (2002) are specific to the commercial burglaries they examined. In particular, the spatial behaviour of burglars may vary across crimes, space, and time, depending as it does on environmental factors such as population density, target distribution, and travel routes. These are features that Brantingham and Brantingham (1993) argued were subject to significant fluctuation across crime type, space, and time.

METHOD

The sample

The present sample of commercial and residential serial burglaries was randomly extracted from a database of burglaries housed in a large metropolitan UK police force that consists of three police districts. In accordance with the definition used by Bennell and Canter (2002), commercial and residential burglaries featured in the present study were defined, respectively, as burglaries committed against a property that functions as a business or a dwelling. A serial burglar was defined as any offender convicted of two or more burglaries.

The three police districts focused on in this study differ from one another on several features (they also differ from the police district originally examined by Bennell and Canter, 2002). District 1 covers 153 km^2, has a population of 286 000 persons, and a population density of 1875 persons/km^2. District 2 only covers 112 km^2, but it has a population of 457 000 persons, and a population density of 4053 persons/km^2. District 3 covers 230 km^2, has a population of 439 000 persons, and a population density of 1467 persons/km^2.

The burglaries examined in this investigation were all committed between January 1995 and December 1999 (which is slightly earlier than the burglaries examined by Bennell & Canter, 2002). The entire commercial burglary sample consisted of 57 serial burglars, altogether responsible for 634 crimes, with offence series ranging in length from three to 24 crimes. The entire residential burglary sample consisted of 51 serial burglars responsible for 660 crimes, with offence series ranging from three to 23 crimes. For the purpose of this study, however, a smaller subset of burglaries was selected for analysis.

More specifically, burglaries were grouped on the basis of the police districts where they were committed and three burglaries were randomly selected from each offender for analysis. As discussed by Bennell and Canter (2002), this procedure was applied to maintain a constant distribution of offences across burglars. Accordingly, there is increased confidence that the results of this study are not biased by undue weighting being assigned to very prolific burglars who may have displayed particularly high or low levels of behavioural stability or distinctiveness across their crimes.

Potential biases in the data

Trained crime analysts coded all of the offence information pertaining to these burglaries. However, because the information was entered directly into a computer database immediately following the commission of each crime, an assessment of coding reliability was precluded. There are several additional limitations associated with the data, which were also applicable to the data collected by Bennell and Canter (2002). First, only solved burglaries are included in this investigation. It is possible that solved burglaries are characterized by higher levels of behavioural stability and distinctiveness relative to unsolved cases, potentially explaining why the former were solved in the first place. Hence, the generalizability of the findings to unsolved burglaries occurring in the same police jurisdiction may be limited. Second, the sole reliance on police records as a source of data can be problematic because this data was collected for investigative rather than research purposes (Canter & Alison, 2003). Different police officers would have coded the data using a variety of coding procedures, thus introducing unknown biases into the data.

While these issues potentially weaken the quality of the data, there is no method by which they can be remedied. In addition, it should be stressed that these problems will likely add 'noise' to the data, reducing the probability of significant results. Thus, should the current research findings demonstrate significance, it will be despite the quality of the data rather than because of it. Regardless of the aforementioned disadvantages, there is an important benefit associated with the use of data extracted from genuine police records. Namely, findings derived from such data will likely demonstrate ecological validity. In other words, they will have practical relevance because they are based on the forms of information upon which the police actually operate.

The linking features

Previous research indicates that three broad domains cover the majority of behaviours typically exhibited by commercial and residential serial burglars (e.g. Maguire & Bennett, 1982; Walsh, 1986). These domains include (1) entry behaviours (e.g. whether the offender entered through the front door of the property), (2) target characteristics (e.g. whether the offender burgled a property with an external alarm), and (3) items stolen (e.g. whether the offender stole cash from the property). Within the police database utilized for this study, information pertaining to each of these behavioural domains was coded in dichotomous fashion across all of the offences, indicating the presence (1) or absence (0) of particular behaviours.

In addition to these three domains, a fourth aspect of the burglaries was examined. This information pertained to the offender's spatial behaviour, expressed as the distance in kilometres between every pair of burglaries. Spatial behaviour was included in this study primarily because of its effectiveness as a linking feature in the investigation conducted by Bennell and Canter (2002). In addition, the examination of spatial behaviour is consistent with a growing body of literature indicating that burglars exhibit predictable spatial mobility patterns (Wiles & Costello, 2000) and the broader research that has shown the interpretability of a variety of criminal geographic patterns (Canter, 2003). Within the police database used for this study, information relevant to spatial behaviour was available in geo-coded x–y coordinates.

Computational procedures

The dependent variable in this investigation is the dichotomous classification of whether the same offender committed a pair of burglaries (i.e. linked versus unlinked). The independent variables are continuous across-crime similarity measures relating to each of the behavioural domains and the spatial domain. As discussed previously, the hypothesis behind these independent variables is that a higher degree of similarity will be exhibited across crimes committed by the same offender. Thus, it is expected that, compared to unlinked crimes, linked offences will be characterized by shorter inter-crime distances and higher across-crime similarity scores for entry behaviours, target characteristics, and items stolen.

Two specially designed computer programs were used to calculate these similarity measures. The first program accepts as input a series of dichotomously coded variables pertaining to each of the three behavioural domains. These variables indicate the presence or absence of specific behavioural features of which these domains are comprised. For each behavioural domain, the program then provides as output a similarity measure between every pair of burglaries, calculated using Jaccard's coefficient (Jaccard, 1908). This similarity measure ranges from 0 (indicating no similarity across the burglaries) to 1 (indicating total similarity across the burglaries). The second program served to analyze the spatial domain, accepting as input the geo-coded x–y coordinates of burglary locations and providing as output the distance in kilometres between every pair of burglaries.

Statistical procedures

Logistic regression analysis was performed to examine the extent to which the various linking features can be used to accurately link burglaries committed by the same offender. The use of logistic regression analysis is appropriate in this study given the dichotomous nature of the dependent variable (Tabachnik & Fidell, 1996). Logistic regression analyses were first run on the linking features separately to determine the extent to which these regression models could be used to accurately predict linked burglary pairs. Logistic regression models that included the optimal combination of linking features were then developed using a forward stepwise regression procedure in order to determine if these models exhibited higher levels of predictive accuracy and fit with the data relative to the other logistic regression models.

In order to reduce the potential for bias that stems from developing and testing the regression models on the same sample of burglaries, the commercial and residential burglary samples were each randomly split in half to form experimental samples (upon which the models were developed) and test samples (upon which the models were validated). Results from the test samples should be indicative of how the logistic regression models might perform on burglaries that have yet to be observed (Efron, 1982).

ROC analyses were then carried out to evaluate the degree to which the various logistic regression models could be used to accurately classify burglary pairs in the test samples (as linked or unlinked) and to identify optimal decision thresholds. This procedure allows us to obtain measures of classification accuracy that are not biased by the placement of the decision threshold, as would be the case if we simply classified the burglary pairs using the default threshold of $p > 0.5$, which is adopted when using the SPSS logistic regression function. The precise data entered into the ROC analyses were the estimated linkage probabilities for every burglary pair comprising the test sample, along with the data representing whether the burglary pairs were actually linked or unlinked. All ROC analyses were performed using ROCKIT©, a computer package designed by the Department of Radiology at the University of Chicago (Metz, Herman, & Roe, 1998).

RESULTS

Logistic regression analyses

Logistic regression models were first developed to examine the predictive accuracy of each linking feature separately (see Tables 2 and 3). Consistent with our hypothesis, the signs

Table 2. Summary of the commercial burglary single feature regression results

Model	District	Logit (SE[a])	Wald's (df[b])	R^2
Inter-crime distances	1	−0.17 (0.04)	15.56 (1)***	0.08
	2	−0.91 (0.09)	104.76 (1)***	0.23
	3	−0.85 (0.11)	61.88 (1)***	0.51
Entry behaviours	1	0.76 (0.51)	2.19 (1)	0.01
	2	1.35 (0.34)	15.68 (1)***	0.01
	3	1.52 (0.56)	7.30 (1)**	0.02
Target characteristics	1	1.18 (0.31)	14.34 (1)***	0.03
	2	1.79 (0.20)	76.76 (1)***	0.06
	3	1.46 (0.33)	19.95 (1)***	0.05
Items stolen	1	1.45 (0.42)	11.95 (1)***	0.02
	2	0.97 (0.35)	7.74 (1)**	0.01
	3	1.79 (0.35)	26.15 (1)***	0.06

[a]SE = Standard error; [b]df = Degrees of freedom.
**p* < 0.05; ***p* < 0.01; ****p* < 0.001.

Table 3. Summary of the residential burglary single feature regression results

Model	District	Logit (SE[a])	Wald's (df[b])	R^2
Inter-crime distances	1	−0.77 (0.13)	37.39 (1)***	0.44
	2	−1.14 (0.09)	163.96 (1)***	0.31
	3	−0.46 (0.11)	16.01 (1)***	0.41
Entry behaviours	1	1.77 (0.64)	7.69 (1)**	0.03
	2	2.18 (0.30)	54.46 (1)***	0.08
	3	1.44 (0.80)	3.20 (1)	0.03
Target characteristics	1	0.28 (0.44)	0.41 (1)	0.00
	2	0.94 (0.26)	12.69 (1)***	0.01
	3	1.42 (0.66)	4.60 (1)*	0.05
Items stolen	1	1.60 (0.69)	5.39 (1)*	0.02
	2	1.54 (0.37)	16.94 (1)***	0.01
	3	1.91 (1.56)	1.49 (1)	0.02

[a]SE = Standard error; [b]df = Degrees of freedom.
**p* < 0.05; ***p* < 0.01; ****p* < 0.001.

of the logit coefficients in these models indicate that, relative to unlinked burglaries, linked burglaries are consistently characterized by shorter inter-crime distances (i.e. negative logit coefficients) and higher levels of across-crime similarity with respect to entry behaviours, target characteristics, and items stolen (i.e. positive logit coefficients). These findings hold irrespective of the police district where the burglaries were committed and regardless of whether the burglaries were commercial or residential.

In addition, many of the regression models were found to have a high degree of predictive accuracy (as measured by Wald's statistic) and a satisfactory fit with the data (as measured by R^2). As indicated in Tables 2 and 3, the regression models consisting of inter-crime distances are particularly effective at distinguishing between linked and unlinked burglaries. In contrast, the models consisting of across-crime similarity scores for entry behaviours, target characteristics, and items stolen are far less effective in this capacity.

Logistic regression models were then developed to identify the optimal combination of linking features across the various sub-samples of burglary data. As indicated in Tables 4

Table 4. Summary of the commercial burglary optimal features regression results

Optimal model	District	Logit (SE[a])	Wald's (df[b])	R^2
Inter-crime distances	1	–0.16 (0.04)	14.01 (1)***	0.11
Items stolen		1.40 (0.44)	10.28 (1)***	
Target characteristics		0.87 (0.33)	7.18 (1)***	
Inter-crime distances	2	–0.81 (0.09)	89.61 (1)***	0.25
Target characteristics		0.87 (0.22)	15.26 (1)***	
Entry behaviours		1.06 (0.36)	8.71 (1)**	
Items stolen		0.93 (0.38)	5.96 (1)**	
Inter-crime distances	3	–0.86 (0.11)	61.22 (1)***	0.53
Items stolen		1.41 (0.47)	8.95 (1)**	

[a]SE = Standard error; [b]df = Degrees of freedom.
**p* < 0.05; ***p* < 0.01; ****p* < 0.001.

Table 5. Summary of the residential burglary optimal features regression results

Optimal model	District	Logit (SE[a])	Wald's (df[b])	R^2
Inter-crime distances	1	–0.87 (0.14)	39.31 (1)***	0.49
Entry behaviours		3.63 (0.91)	16.05 (1)***	
Inter-crime distances	2	–1.09 (0.09)	155.13 (1)***	0.34
Entry behaviours		1.65 (0.34)	23.95 (1)***	
Items stolen		1.68 (0.46)	13.17 (1)***	
Inter-crime distances	3	–0.45 (0.11)	16.01 (1)***	0.41

[a]SE = Standard error; [b]df = Degrees of freedom.
**p* < 0.05; ***p* < 0.01; ****p* < 0.001.

and 5, apart from the fact that inter-crime distances are included in every optimal regression model, there are no other obvious similarities in the models in terms of the independent variables that were included. That is, the optimal regression models are, to some extent, crime and district specific. The only other common features of the optimal regression models are their good degree of fit with the data and their consistent levels of predictive accuracy, which are equal to or slightly higher than any of the single feature regression models. Importantly, however, the levels of predictive accuracy and fit associated with the optimal models are never notably higher than those associated with the regression models that consist of inter-crime distances in isolation.

ROC analyses

The regression models constructed from data in the experimental samples indicate that many commercial and residential burglary behaviours exhibited by offenders in this police jurisdiction can be used to accurately predict which burglary pairs are linked. However, the predictive accuracy of the models varies across crime type and police district. ROC analyses were conducted to evaluate the degree to which the various logistic regression models could be used to accurately classify burglary pairs in the test samples (as linked or unlinked).

To perform these analyses, the regression models presented in Tables 2 to 5 were used to calculate estimated probabilities for every burglary pair in the commercial and resi-

Table 6. Summary of the commercial burglary ROC results

Variable	District	AUC[a]	SE[b]	95% CI[c]
Optimal	1	0.76***	0.04	0.68–0.82
	2	0.89***	0.01	0.86–0.91
	3	0.89***	0.01	0.86–0.91
Inter-crime distances	1	0.76***	0.04	0.69–0.83
	2	0.88***	0.01	0.85–0.91
	3	0.88***	0.01	0.85–0.91
Entry behaviours	1	0.57	0.04	0.49–0.64
	2	0.57**	0.03	0.52–0.62
	3	0.57**	0.03	0.52–0.62
Target characteristics	1	0.60**	0.04	0.52–0.68
	2	0.62***	0.03	0.56–0.67
	3	0.62***	0.03	0.56–0.67
Items stolen	1	0.52	0.06	0.45–0.60
	2	0.52	0.03	0.47–0.58
	3	0.62	0.03	0.47–0.58

[a]AUC = Area under the curve; [b]SE = Standard error; [c]CI = Confidence interval.
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table 7. Summary of the residential burglary ROC results

Variable	District	AUC[a]	SE[b]	95% CI[c]
Optimal	1	0.94***	0.01	0.92–0.96
	2	0.91***	0.01	0.89–0.94
	3	0.85***	0.03	0.79–0.91
Inter-crime distances	1	0.94***	0.01	0.92–0.97
	2	0.91***	0.01	0.88–0.93
	3	0.85***	0.03	0.79–0.91
Entry behaviours	1	0.57	0.05	0.48–0.66
	2	0.62***	0.03	0.57–0.67
	3	0.59	0.07	0.45–0.73
Target characteristics	1	0.53	0.04	0.45–0.61
	2	0.57**	0.03	0.52–0.62
	3	0.64*	0.06	0.52–0.75
Items stolen	1	0.56	0.05	0.47–0.65
	2	0.59***	0.03	0.54–0.64
	3	0.63*	0.07	0.50–0.76

[a]AUC = Area under the curve; [b]SE = Standard error; [c]CI = Confidence interval.
$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

dential test samples. These probabilities were then utilized to construct separate, cross-validated ROC graphs for each logistic regression model, including the optimal models. Due to space limitations, we have not presented the actual ROC graphs that were constructed. Instead, tables are provided that summarize the relevant results (the actual ROC graphs are available upon request from the first author).

As illustrated in Tables 6 and 7, the results from the ROC analyses are in accordance with the results yielded from the regression analyses, suggesting that the regression models do possess a degree of validity. For example, the use of inter-crime distances consistently results in ROC curves that indicate a high degree of discrimination accuracy regardless of crime type or police district. In addition, while the three individual behavioural domains often produce levels of discrimination accuracy beyond what would be expected by chance, their AUCs are significantly lower than those for inter-crime distances. Finally, ROC curves associated with the optimal regression models reveal AUCs that are equal to or slightly greater than the AUCs for inter-crime distances.

Swets (1988) argues that AUCs close to 0.50 are non-informative, AUCs between 0.50 and 0.70 indicate low levels of accuracy, AUCs between 0.70 and 0.90 indicate good levels of accuracy, and AUCs between 0.90 and 1.00 reflect high levels of accuracy. Thus, according to these guidelines, the ROC curves resulting from the optimal regression models and the models consisting of inter-crime distances convey good to high levels of discrimination accuracy. On the other hand, the ROC curves resulting from the remaining regression models (i.e. those including across-crime similarity scores for one of the three behavioural domains) represent low levels of accuracy.

Given the results of our ROC analyses, we focused on inter-crime distances for the purpose of identifying optimal decision thresholds. To establish these thresholds, a measure known as Youden's index was calculated for each of the regression models (Hilden, 1991). Youden's index is calculated using the formula, $J = pH + pCR - 1$, where the subtraction of 1 from $pH + pCR$ ensures that J always lies between 0 and 1. The goal is to select the decision threshold that results in the highest possible J value, where both pH and pCR are equal to 1 (i.e. no incorrect linking decisions are made). This can be accomplished by calculating Youden's index for all of the decision thresholds used to construct the ROC curve (decision thresholds in this case are set along the range of p-values obtained when using the logistic regression model). The values of Youden's index can then be plotted on the y-axis against the various decision thresholds on the x-axis. The highest point on this graph indicates the optimal decision threshold. Note that for the practical purpose of determining the inter-crime distance to which a decision threshold relates, decision thresholds set along the range of p-values produced by the regression models can be related back to the original distance data. For example, instead of advising the police to use a decision threshold of, say, $p > 0.02$ to link burglaries, one can recommend the use of a threshold that is easier to understand and implement, say an inter-crime distance of <1.40 km.

As indicated in Table 8, the optimal decision thresholds for inter-crime distances vary slightly across crime type and police district, as do the pH and pFA values that are achieved using these thresholds. For example, in terms of actual inter-crime distances, the optimal decision thresholds are slightly larger for commercial burglaries and for District 1.

To appreciate the practical significance of using inter-crime distances over other linking features, one can calculate exactly how many more hits (or how many less false alarms) would be made when using each particular decision threshold (see Swets *et al.*, 2000). For example, using the optimal threshold calculated for inter-crime distances with respect to commercial burglaries in District 1, $pH = 0.76$ and $pFA = 0.28$. Using the optimal thresholds for the other linking features in this district, the pH and pFA values achieved are $pH = 0.52$ and $pFA = 0.41$ (for entry behaviours), $pH = 0.52$ and $pFA = 0.42$ (for target characteristics), and $pH = 0.51$ and $pFA = 0.51$ (for property stolen). Thus, if a police investigator were primarily interested in maximizing pH, he or she could make $76 - 52 = 24$ additional hits for every 100 burglary pairs encountered by relying on inter-crime distances instead of entry behaviours. On the other hand, if the investigator were primarily interested in minimizing pFA, he or she could make $41 - 28 = 13$ fewer false alarms for every 100 burglary pairs encountered by considering inter-crime distances over entry behaviours.

Table 8. Decision thresholds for regression models consisting of inter-crime distance and the resulting values of pH and pFA

District	Commercial burglary	Residential burglary
1	$p > 0.01$ (<3.00 km)	$p > 0.15$ (<2.60 km)
	pH[a] = 0.76, pFA[b] = 0.28	pH = 0.89, pFA = 0.17
2	$p > 0.01$ (<2.30 km)	$p > 0.04$ (<2.10 km)
	pH = 0.86, pFA = 0.28	pH = 0.85, pFA = 0.19
3	$p > 0.17$ (<2.30 km)	$p > 0.24$ (<2.20 km)
	pH = 0.93, pFA = 0.17	pH = 0.77, pFA = 0.26

[a]pH = Probability of a hit; [b]pFA = Probability of a false alarm.

DISCUSSION

Logistic regression analysis and ROC analysis were applied in this study to determine if the results obtained by Bennell and Canter (2002) in relation to commercial burglary could be replicated using a different commercial burglary data set and extended to the crime of residential burglary. As in the previous investigation, both forms of analysis clearly support the possibility of utilizing certain objectively available aspects of a commercial or residential burglar's actions to systematically perform the behavioural linking task, despite potential problems with police data.

This finding stands in stark contrast to recent suggestions that an offender's MO is too dynamic to be of practical value in linking serial crimes (e.g. Davies, 1992; Douglas & Munn, 1992; Turvey, 2002). Although the lower levels of discrimination accuracy achieved when using traditional MO indicators are consistent with these arguments, it is nonetheless possible to identify certain behavioural elements of burglary that are relatively stable and distinct across a crime series.

Maximizing discrimination accuracy through the use of inter-crime distance

As in the study conducted by Bennell and Canter (2002), relative to traditionally considered linking behaviours, inter-crime distance was the feature most effective in linking serial burglaries. In fact, the level of discrimination accuracy attained when using inter-crime distances in the present study surpassed the level achieved in that previous investigation (AUC = 0.80), particularly for residential serial burglary. More specifically, in the present study, the average AUCs found for inter-crime distances in the case of commercial and residential burglary were 0.84 and 0.90, respectively.

The general consistency of this finding across commercial and residential burglaries is important. It suggests that similar psychological processes may underlie crime site selection choices for the two forms of burglary. For example, both types of burglars appear to be choosing relatively distinct geographic areas in which to commit their crimes. That is, to a large extent, the burglars examined in this study, and the previous one, appear to commit their crimes in non-overlapping offending territories. If this were not the case, it would not have been possible to achieve the degrees of predictive accuracy yielded in these studies.

The only case in which inter-crime distances were not particularly effective for linking purposes was for the offenders committing commercial burglaries in District 1. It is possible that, compared to the other districts, there are fewer and more spatially dispersed commercial properties in this particular area. If this were the case, commercial burglars in District 1 would have to target the same geographic areas being pursued by other burglars, thus decreasing the extent to which inter-crime distances could be used to distinguish linked from unlinked crimes (i.e. the offending territories exploited by these burglars would overlap to a greater extent than for burglars committing crimes in other districts). However, a more detailed analysis is required to verify this assumption, requiring more extensive data and the development of analyses that take account of areas with a high-density of crime.

From a practical perspective, the results of the current study suggest that similar linking strategies may be productive across both types of burglaries examined and across the three different police districts, at least in terms of the linking features that should be focused on. More specifically, as Bennell and Canter (2002) argued previously, inter-crime distances may prove to be effective as an initial filter in the behavioural linking task as a means of reducing the number of potential links that originally require consideration.

This filter idea accords well with the findings reported in several recent investigations. For example, Grubin *et al*. (2001) provided evidence that linking accuracy could be enhanced in cases of serial sexual assault by taking into account the offender's spatial behaviour. However, Grubin and his colleagues collected data from right across the UK and, therefore, it is not surprising that the offenders in their study overlapped minimally in terms of their areas of criminal activity. Given that burglaries are not typically investigated at the national level, it is of great practical significance that we have been able to demonstrate the linking potential of inter-crime distances as related to serial burglaries committed within much smaller police districts.

Should law enforcement officers adopt linking strategies based on traditional MO indicators (e.g. the offender's method of entry), the results reported here suggest they will likely experience great difficulty. While very slight improvements to linking accuracy can be achieved by considering various MO indicators, in addition to inter-crime distances, police forces will have to consider whether these increases are worth the time and effort required to collect, record, and analyze data beyond inter-crime distances. It is possible that the low levels of predictive accuracy associated with traditional MO indicators in the present study are due to the limited information available in UK police records on the details of burglary-related actions. If this is the case, a more refined approach to recording what does or does not happen in these crimes may yield more promising results. It is also plausible that the information collected on burglaries needs to be interpreted more intelligently in relation to the opportunities which particular properties provide for the crime. However, all such possibilities require considerable increases in the resources put into collecting information on these crimes.

Why are inter-crime distances so effective at linking serial burglaries?

A variety of factors probably contribute to the linking effectiveness of inter-crime distances over other, more traditional behavioural domains. The first potentially important factor is that the location of crime sites can be readily recorded in a very reliable and accurate fashion by the police, which may allow for the emergence and detection of consistent patterns of spatial behaviour. The same does not apply to other behaviours examined in this study. For example, the recording of items stolen in a burglary may be particularly unreliable and inaccurate. Not only will the coding of this information depend on the items the police choose to record as stolen and on the items the property owner chooses to declare were stolen, it will also depend on the property that was available to steal when the burglar entered the building.

A second factor relates to situational influences on burglary behaviour. Of all the various decisions made by a serial burglar during the commission of his crimes, his choice of burglary locations is possibly the most crucial, being the one decision over which the burglar has considerable control. Therefore, it reasonably follows that this aspect of burglary behaviour will be the most consistently exhibited while the other, more context-dependent behaviours will be less consistent. For example, the items an offender steals from a given property is very situation-specific, and therefore less stable, depending as it does on item availability at the particular crime scene.

Third, borrowing from an argument presented by Funder and Colvin (1991) in the non-criminal domain, the degree of variance associated with the linking features in question may explain their relative capacity to discriminate between linked and unlinked crimes. The idea is that those features associated with greater variance within a sample more readily permit the emergence and potential observation of behavioural differences between offenders. A descriptive analysis of the linking features used in this study certainly indicates that the degree of variance associated with inter-crime distance is much larger than the variances associated with the other behavioural domains. This difference is probably due to the nature of the various linking features (i.e. a greater range of actions on the part of burglars is possible in the spatial domain compared to the other behavioural domains) and the specific measures used to quantify similarity in this study (i.e. inter-crime distances allow for a greater number and range of similarity scores than do Jaccard coefficients).

Finally, data bias cannot be discounted as a possible explanation for the effectiveness of inter-crime distances in linking serial burglaries within the context of this investigation. In particular, the fact that we have focused solely on solved crimes has probably resulted in an over-estimation of AUCs for inter-crime distances. Indeed, one of the reasons why certain burglaries may be linked and solved in the first place is that they are committed in close proximity to one another. On the other hand, linked crimes that do not fit this pattern may remain unlinked and unsolved. In addition, it is important to stress the fact that, in this study, only relatively small samples of burglaries were collected from each police district (i.e. we did not examine the total population of burglaries in the three districts). Accordingly, our results may be insensitive to areas within each police district that had a very high density of burglaries, known as burglary hotspots, where the task of using inter-crime distances to discriminate between burglaries committed by different offenders would be very difficult. In other words, our sampling process would tend to draw out burglaries that are spread over the area of study, which would inflate our AUCs for inter-crime distances. This study can therefore only be regarded as a first step towards an operational system that would ultimately consider burglary hotspots and other local factors that are not accounted for when sampling from a particular area. However, the present findings do suggest that such larger scale studies would be well worth undertaking.

The importance of selecting appropriate decision thresholds

One of the primary advantages of using the ROC approach to examine the behavioural linking task is that discrimination accuracy and the impact of setting different decision thresholds can be considered simultaneously. Surprisingly, few researchers involved in behavioural linking research have discussed the issue of setting decision thresholds (e.g. Green *et al.*, 1976; Grubin *et al.*, 2001; Santtila, Korpela, & Häkkänen, 2004) and it was only very recently that this issue was formally considered (by Bennell & Canter, 2002). The present investigation, as well as the previous one, illustrates that determining appropriate decision thresholds is as important for effective behavioural linking as the identification of accurate linking features. Indeed, it is futile to recognize the general utility of inter-crime distances in linking a pair of burglaries in ignorance of the proximity that must be met before two crimes should be considered linked.

The results reported here clearly demonstrate that linking performance depends not only on the inherent discriminatory power of a particular linking feature, or features, but also on the exact position of the decision threshold. In general, *p*H and *p*FA will both increase as the decision threshold becomes more lenient and both decrease as the decision threshold becomes more stringent. This highlights the need to identify appropriate decision thresholds for behavioural linking that produce the desired balance between the various decision outcomes.

In this investigation, some potentially important findings emerged from the identification of optimal decision thresholds for inter-crime distances. Perhaps of greatest significance is the apparent crime and district specificity of the optimal decision thresholds. For example, the optimal thresholds were slightly larger for commercial burglaries compared to residential burglaries and slightly larger for burglaries occurring in District 1 compared to Districts 2 and 3.

From a theoretical perspective, this indicates that certain characteristics of commercial burglary, especially as they occur in District 1, produce slight differences in offender mobility patterns. As discussed, we suspect that the reasons for these differences may be related to the distribution of commercial properties in that district (e.g. being more spatially dispersed). However, further research must be conducted to confirm this hypothesis. From a practical perspective, this finding suggests that linking strategies might have to be tailored to fit each police district. Inter-crime distances could still be considered uniformly across districts (since they are the most effective linking feature) but it may be advisable for the police to adjust their thresholds for determining the conditions under which two burglaries are to be linked.

Notwithstanding the above, we should stress that the procedure we employed for defining optimal decision thresholds in this study may not be the most appropriate. Our procedure of calculating Youden's index maximized the probability of making correct linking decisions and minimized the probability of making incorrect linking decisions. However, while this approach is clearly rational and preferable to many alternatives, it is based on certain assumptions that are unlikely to be valid in the current context. For example, with this approach, the prior probabilities of linked and unlinked burglaries are assumed to be equal, and this is unlikely to be the case in any jurisdiction (unlinked crimes will typically be much more common). Furthermore, costs and benefits associated with incorrect and correct linking decisions are assumed to be equivalent, which again is unlikely in operational settings. Unfortunately, estimating optimal decision thresholds can be difficult, though not impossible, and will require a thorough cost-benefit analysis to be conducted by individual police forces.

FUTURE RESEARCH

Little research has been conducted that directly examines the issue of behavioural linking. Therefore, it is premature to offer any fervent recommendations at this juncture as to the specific manner in which such a task should be executed in police investigations. However, with additional research, this goal should be attainable and it is our conviction that ROC analysis will prove to be an invaluable tool in this regard.

In our opinion, four issues in particular warrant immediate consideration. First, different sets of linking features should be examined to determine their level of discrimination accuracy compared to the features we have examined here. With the exception of inter-crime distances, all of our linking features possess relatively low levels of discriminatory power and there is a distinct possibility that other features will be more effective. This emphasises the importance of collecting much richer data on crimes that are to be linked.

Second, as Canter and Youngs (2003) have argued, it will be important in the future for researchers in this area to identify the most salient features of a crime, taking into account the overall frequency with which various behaviours are exhibited by offenders. This process is worthwhile since very low or very high frequency behaviours will likely lack the discriminatory power that is needed to successfully link crimes.

Third, different types of crimes should be analysed to determine if they can also be accurately linked using the proposed approach. Shifting the focus from serial burglary onto crimes such as serial rape appears logical considering the fact that police forces rarely have sufficient time or manpower to thoroughly investigate the high number of property crimes committed (Vancouver Board of Trade, 2003). Arguably, it would be much more advantageous for the police to have an effective linking system in place for the types of crimes that warrant the greatest amount of their investigative attention. The results here accord well with those reported by Grubin *et al*. (2001) so the potential for successfully applying ROC analysis to inter-crime distances in cases of unsolved serial rape is very great. This technique also has applied potential in court proceedings, where the defence or prosecution may wish to draw on ROC analyses to support an argument that a given set of rapes are either linked or unlikely to have been committed by the same offender.

Fourth and finally, linking procedures like the one developed in the present investigation should be compared to other linking approaches in existence to determine their relative efficacy. On a related note, it may be worthwhile to compare the performance of actuarial linking procedures to the performance of human judges (e.g. crime analysts) to determine if these statistical methods can provide a useful decision support system. Statistical methods like those used in the current study have proven effective in this regard in other diagnostic fields (e.g. Getty, Pickett, D'Orsi, & Swets, 1988; Seltzer *et al*., 1997).

REFERENCES

Bennell, C., & Canter, D. V. (2002). Linking commercial burglaries by *modus operandi*: tests using regression and ROC analysis. *Science & Justice*, *42*, 153–164.

Brantingham, P. L., & Brantingham, P. J. (1993). Environment, routine and situation: toward a pattern theory of crime. In R. V. Clarke & M. Felson (Eds.), *Routine activity and rational choice* (pp. 259–294). New Brunswick, NJ: Transaction.

Canter, D. V. (2000). Offender profiling and criminal differentiation. *Legal and Criminological Psychology*, *5*, 23–46.

Canter, D. V. (2003). *Mapping murder: the secrets of geographical profiling*. London: Virgin Books.

Canter, D. V., & Alison, L. J. (2003). Converting evidence into data: the use of law enforcement archives as unobtrusive measurement. *The Qualitative Report*, *8*, 151–176.

Canter. D. V., & Youngs, D. (2003). Beyond 'offender profiling': the need for an investigative psychology. In R. Bull, & D. Carson (Eds.), *Handbook of psychology in legal contexts* (2nd ed.) (pp. 171–205). Chichester, UK: Wiley.

Collins, P. I., Johnson, G. F., Choy, A., Davidson, K. T., & Mackay, R. E. (1998). Advances in violent crime analysis: the Canadian violent crime linkage analysis system. *Journal of Government Information*, *25*, 277–284.

Davies, A. (1992). Rapist's behaviour: a three-aspect model as a basis for analysis and the identification of serial crime. *Forensic Science International*, *55*, 173–194.

Douglas, J. E., & Munn, C. (1992). Violent crime scene analysis: *modus operandi*, signature, and staging. *FBI Law Enforcement Bulletin*, *6*, 1–10.

Efron, B. (1982). *The jackknife, the bootstrap and other re-sampling plans.* Philadelphia, PA: Society for Industrial and Applied Mathematics.

Funder, D. C., & Colvin, C. R. (1991). Explorations of behavioral consistency: properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, *60*, 773–794.

Getty, D. J., Pickett, R. M., D'Orsi, C. J., & Swets, J. A. (1988). Enhanced interpretation of diagnostic images. *Investigative Radiology*, *23*, 240–252.

Getty, D. J., Seltzer, S. E., Tempany, C. M. C., Pickett, R. M., Swets, J. A., & McNeil, B. J. (1997). Prostate cancer: relative effects of demographic, clinical, histologic, and MR imaging variables on the accuracy of staging. *Radiology*, *204*, 471–479.

Green E. J., Booth, C. E., & Biderman, M. D. (1976). Cluster analysis of burglary M/Os. *Journal of Police Science and Administration*, *4*, 382–287.

Grubin, D., Kelly, P., & Brunsdon, C. (2001). *Linking serious sexual assaults through behaviour*. London, UK: Home Office.

Hilden, J. (1991). The area under the ROC curve and its competitors. *Medical Decision Making*, *11*, 95–101.

Jaccard, P. (1908) Nouvelle recherches sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat*. 44, 223–270.

Maguire, M., & Bennett, T. (1982). *Burglary in a dwelling: the offence, the offender and the victim*. London, UK: Home Office.

Metz, C. E., Herman, B. A., & Roe, C. A. (1998). Statistical comparison of two ROC curve estimates obtained from partially paired data sets. *Medical Decision Making*, *18*, 110–121.

Santtila, P., Korpela, S., & Häkkänen, H. (2004). Expertise and decision making in the linking of car crime series. *Psychology, Crime, and Law*, *10*, 97–112.

Seltzer, S. E., Getty, D. J., Tempany, C. M. C., Pickett, R. M., Schnall, M. D., McNeil, B. J., & Swets, J. A. (1997). Staging prostate cancer with MR imaging: a combined radiologist-computer system. *Radiology*, *202*, 219–226.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*, 1285–1293.

Swets, J. A. (1992). The science of choosing the right decision threshold in high-stake diagnostics. *American Psychologist*, *47*, 522–532.

Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics*. Mahwah, NJ: Lawrence Erlbaum Associates.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*, 1–26.

Tabachnik, B. B., & Fidell, L. S. (1996). *Using multivariate statistics*. New York, NY: Harper Collins Publishers.

Vancouver Board of Trade. (2003, October 24). *Report on property crime in Vancouver*. Retrieved on 23 July 2004 from http://www.boardoftrade.com/policy/ PROPERTYCRIME-FINAL-OCT24-03-web.pdf

Turvey, B. (2002) *Criminal profiling: an introduction to behavioral evidence analysis* (2nd ed.). London, UK: Academic Press.

Walsh, D. (1986). *Heavy business: commercial burglary and robbery*. London, UK: Routledge.

Wiles, P., & Costello, A. (2000). *The road to nowhere: the evidence for travelling offenders*. London, UK: Home Office.

ACKNOWLEDGEMENT

The authors would like to thank David Canter for his many helpful comments on an earlier draft of this paper.

