

The feasibility of using crime scene behaviour to detect versatile serial offenders: An empirical test of behavioural consistency, distinctiveness and discrimination accuracy

Tonkin, Matthew; Woodhams, Jessica

DOI:
[10.1111/lcrp.12085](https://doi.org/10.1111/lcrp.12085)

Document Version

Early version, also known as pre-print

Citation for published version (Harvard):

Tonkin, M & Woodhams, J 2015, 'The feasibility of using crime scene behaviour to detect versatile serial offenders: An empirical test of behavioural consistency, distinctiveness and discrimination accuracy', *Legal and Criminological Psychology*. <https://doi.org/10.1111/lcrp.12085>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

The feasibility of using crime scene behaviour to detect versatile serial offenders: An empirical test of behavioural consistency, distinctiveness and discrimination accuracy

Abstract

Purpose: To test whether geographical, temporal and Modus Operandi (MO) crime scene behaviours can be used to support behavioural case linkage (BCL) with crime series that contain several different types of offence.

Methods: Crime scene data relating to 749 solved commercial burglaries and robberies were extracted from the databases of the Metropolitan Police Service, London, England. From these data, 2,231 linked crime pairs (containing two crimes committed by the same offender) and 273,422 unlinked crime pairs were created (two crimes committed by different offenders). Three measures of similarity were calculated for each crime pair: 1) the kilometre-distance between crimes (inter-crime distance); 2) the number of days between crimes (temporal proximity); and 3) a statistical measure of similarity in MO behaviour (Jaccard's coefficient). Statistical tests of difference, binary leave-one-out logistic regression, and Receiver Operating Characteristic analysis were used to determine whether the three measures of similarity could be used to distinguish between linked and unlinked crime pairs, some containing only burglaries (burglary pairs), some containing only robberies (robbery pairs) and some containing both burglaries and robberies (cross-crime pairs).

Results: Linked and unlinked crime pairs could be distinguished with a high level of accuracy (AUCs > .90), with the highest accuracy when combining inter-crime distance, temporal proximity and Jaccard's coefficient. These findings were replicated with the burglary pairs, robbery pairs and cross-crime pairs.

This is the prepublication version of a paper now accepted for publication by Legal and Criminological Psychology for which the DOI is 10.1111/lcrp.12085

Conclusions: Offender behaviour is sufficiently consistent and distinctive to support the use of BCL with versatile crime series, as well as with burglary crime series and robbery crime series.

Introduction

Behavioural case linkage (BCL) is often used in the absence of physical trace material as a method for identifying linked crime series, thereby helping the police apprehend prolific serial offenders who are responsible for a disproportionate amount of crime and who impose significant costs on society (Piquero, Farrington, & Blumstein, 2007; Woodhams, Hollin, & Bull, 2007).

The ability to link crimes behaviourally relies upon offenders repeating certain elements of their crime scene behaviour from one offence to the next (behavioural consistency) and their behaviour being different from that of other offenders (behavioural distinctiveness) (Woodhams et al., 2007). A number of studies have sought to test these theoretical assumptions of BCL, but the literature is limited by the lack of replication research, the fact that the data tested are not always an accurate reflection of the real-world context in which BCL would be used and that most of the research has only examined crime series that contain one type of crime (e.g., series that contain solely residential burglaries). The research reported in this paper aimed to address each of these limitations, thereby strengthening the potential contribution that this literature can make to BCL theory and practice.

Behavioural Consistency and Distinctiveness in Offending Behaviour

As mentioned above, researchers of BCL have suggested that the reliable and accurate linking of crime relies upon two theoretical assumptions: behavioural consistency and distinctiveness (e.g., Woodhams et al., 2007). These assumptions were originally proposed to explain non-criminal human behaviour; in particular, they have been applied to the study of personality (e.g., Mischel & Shoda, 1995; Mischel, Shoda, & Smith, 2004). There is, however, a range of evidence to suggest that these assumptions might also apply to offending behaviour. For example, there is a large body of research that has reported script-like

behaviour and consistency amongst burglars when selecting targets, breaking in to a property and searching for items to steal (see Bennett & Wright, 1984; Nee & Taylor, 2000; Taylor & Nee, 1988; Wright & Decker, 1994). Furthermore, the existence of offending scripts that have the potential to generate behavioural consistency has been discussed in relation to a variety of other offending behaviours, including firesetting (Butler & Gannon, 2015), sexual offending (Ward & Hudson, 1998, 2000), robbery (Cornish, 1994) and carjacking (Topalli, Jacques, & Wright, 2015), to name but a few.

In terms of behavioural distinctiveness, Bouhana, Johnson, and Porter (2014) highlight research that indicates a range of individual differences between offenders in terms of their perceptions of risk when selecting targets, their sensitivity to situational factors, the level of pre-offence planning they engage in and the target characteristics that attract them to offend (Bennett & Wright, 1984; Nee & Meenaghan, 2006; Wright & Decker, 1994, 1997). These individual differences might allow for the emergence of distinctive offending behaviour that can be used to differentiate the crimes of one offender from those of another. Thus, there is a range of evidence to support the notion that some degree of consistency and distinctiveness might be expected in the behaviour of serial offenders.

It is important, however, to recognise that within this literature many offenders *do not* report behaving in a consistent way from one crime to the next, with variation observed in a range of offence behaviours, including target selection, search behaviour and items stolen (Nee & Meenaghan, 2006). Likewise, many of the offenders have reported identical search patterns and stole very similar items during their offences (typically cash, jewellery and documents) (e.g., Bennett & Wright, 1984; Nee & Meenaghan, 2006; Nee & Taylor, 2000; Taylor & Nee, 1988). Consequently, there appears to be a degree of homogeneity in offender behaviour, which would make it difficult to accurately distinguish the crimes of one offender from those of a different offender (i.e., there may be a lack of behavioural distinctiveness).

This brief review indicates that a degree of consistency and distinctiveness may exist in offending behaviour, but we should also expect behavioural variation across a series of crimes and a certain amount of homogeneity amongst offenders in their offending behaviour. This raises the following question: do serial offenders display enough behavioural consistency and distinctiveness to support the potential use of BCL in practice? A growing body of research has sought to address this question, and it is to this literature that we now turn.

Empirical Tests of Behavioural Consistency, Distinctiveness and Discrimination

Accuracy

Approximately 30 empirical papers have been published in the last 15-20 years that seek to test whether sufficient offender behavioural consistency and distinctiveness exist to allow linked crimes to be distinguished from unlinked crimes (referred to as discrimination accuracy hereafter). These studies support the existence of offender behavioural consistency and distinctiveness for some offenders, some of the time, in a range of person- and property-oriented crimes, including commercial and residential burglary (e.g., Bennell & Jones, 2005; Bouhana et al., 2014), commercial and personal robbery (Burrell, Bull, & Bond, 2012; Woodhams & Toye, 2007), arson (e.g., Santtila, Fritzon, & Tamelander, 2004), rape/sexual assault (e.g., Yokota, Fujita, Watanabe, Yoshimoto, & Wachi, 2007), homicide (e.g., Melnyk, Bennell, Gauthier, & Gauthier, 2011) and auto theft (e.g., Tonkin, Grant, & Bond, 2007).

Within this literature a common metric that has been used to assess discrimination accuracy is the Area Under the Curve (AUC), which is produced by Receiver Operating Characteristic (ROC) analysis. It is beyond the scope of this article to provide a comprehensive review of ROC analysis and its relative merits as a method for quantifying discrimination accuracy, but the interested reader is referred to Bennell, Jones, and Melnyk (2009) for further information. For the purposes of the current article, though, it is worth

noting that the AUC typically ranges from 0.50 (indicating a chance level of discrimination accuracy) to 1.00 (which indicates that every time a linked crime pair is randomly selected it is more similar in terms of offender behaviour than a randomly selected unlinked crime pair). Consequently, large AUC values that statistically exceed chance ($AUC = 0.50$) indicate that offender behavioural consistency and distinctiveness exist at a level that is sufficient to allow linked crimes to be accurately distinguished from unlinked crimes.

A recent review of the BCL literature using ROC analysis (Bennell, Mugford, Ellingwood, & Woodhams, 2014) suggested that 2% of the AUCs reported fall in the non-informative range ($AUC < .50$), 29% fall in the low range ($AUC = 0.50 - 0.70$), 54% fall in the moderate range ($AUC = 0.70 - 0.90$), and 15% fall in the high range ($AUC > 0.90$) (Swets, 1988). These findings, therefore, provide mixed support for the existence of consistency and distinctiveness in offender crime scene behaviour, which further underscores the above discussion suggesting that consistency/distinctiveness do not necessarily apply to all offenders all of the time. Nevertheless, provided appropriate behaviours are relied upon there is evidence demonstrating that linked and unlinked crimes can be distinguished at a level that far exceeds chance, thereby suggesting that sufficient levels of consistency and distinctiveness exist to support BCL. Researchers and practitioners should, however, exercise caution when interpreting these findings for a number of reasons, which will now be discussed.

The Limitations of Previous BCL Research

While it is beyond the scope of this paper to provide a comprehensive review of limitations, a few key points relevant to the current study will be noted (see Tonkin, 2014, for a more detailed review). The first limitation of note is that there are few replication studies, which limits the extent to which the findings can be used to draw robust and generalisable conclusions that can guide the development of theory and practice. For example, there is only

one study testing consistency, distinctiveness and discrimination accuracy in commercial robbery (Woodhams & Toye, 2007).

Second, the majority of previous research has focused on testing consistency, distinctiveness and discrimination accuracy with samples that contain only one type of crime (e.g., series consisting solely of residential burglaries). This is despite the fact that many offenders (particularly the most prolific) are versatile in their offending (Farrington, Snyder, & Finnegan, 1988; Piquero et al., 2007). Existing research does not, therefore, provide guidance for conducting BCL with series that contain several different types of crime. Fortunately, however, recent research has started to address this issue (Tonkin, Woodhams, Bull, Bond, & Palmer, 2011; Tonkin, Woodhams, Bull, & Bond, 2012). These studies have demonstrated that simple measures of geographical and temporal behaviour (inter-crime distance and temporal proximity¹) are able to achieve moderate to high levels of discrimination accuracy when used to distinguish between linked and unlinked crime pairs that contain a range of violent, sexual and property-related offences (AUCs = 0.79 – 0.90). These findings have been demonstrated with samples containing both solved and unsolved crimes².

While these findings are promising, the literature on cross-crime linkage is still preliminary and a number of significant limitations exist. First, the previous research in this area has examined a very narrow range of offender behaviour (geographical and temporal behaviour only), which is problematic because there may be situations where this information may be either unavailable or unreliable. A second limitation is that both previous studies of cross-crime linkage (Tonkin et al., 2011, 2012) were conducted in the same geographical region, which limits the practical and theoretical value of this work. A third limitation of the research by Tonkin et al. (2011, 2012) was that their data only included two crimes per offender, rather than all of the crimes committed by each offender within the sampling

period. Consequently, the data used in these studies did not replicate the real life investigative context within which BCL would be used (i.e., with datasets that contain an uneven number of offences per offender). If the aim is to develop statistical linkage models that can be applied in practice, it is important that these models are developed on data that replicate reality as closely as possible.

The Current Study

Given the above discussion, the current study sought to extend the BCL literature by examining a large dataset of commercial burglaries and robberies. These crime types were chosen for several reasons: 1) they pose a considerable problem for police forces around the world (Cowen & Williams, 2012); 2) there are many offenders who commit both commercial burglary and commercial robbery offences (e.g., Wright & Decker, 1997), which means that methods for behaviourally linking across these crime types would be of value; and 3) commercial burglary and robbery share a number of offender behaviours, such as property stolen and target selection behaviour, that make it possible to examine cross-crime linkage using Modus Operandi (MO) behaviours (which has never been done before).

The current study, therefore, contributes to the literature in three important ways. 1) It presents the first empirical test of whether MO behaviours can be used to support cross-crime linkage. 2) It presents the first empirical test of cross-crime linkage using data that contain an uneven number of offences per offender (which is closer to the real life investigative context within which BCL would be used). 3) It replicates key findings relating to cross-crime linkage (Tonkin et al., 2011, 2012), commercial burglary (Bennell & Jones, 2005) and commercial robbery (Woodhams & Toye, 2007) in a geographical location not previously tested, using one of the largest datasets yet compiled for the purposes of BCL research.

Method

Data

All detected commercial burglary and commercial robbery crimes³ committed between 01/01/2010 and 31/03/2013 were extracted from the crime databases of the Metropolitan Police Service, London, England. From these data a sub-section was selected for analysis, which consisted of 749 commercial burglaries and robberies committed by 214 serial offenders⁴. These data contained all commercial robbery series committed over the study period ($n = 84$ series, 237 crimes, average series length = 2.82 crimes), all series containing both commercial burglaries and robberies ($n = 46$ series, 151 crimes, average series length = 3.28 crimes) and 18.96% of the burglary series ($n = 84$ series, 361 crimes, average series length = 4.30 crimes).

For each crime in the dataset, information pertaining to 67 dichotomised behavioural variables was used to examine behavioural consistency, distinctiveness and discrimination accuracy (see the Appendix). Low frequency variables that occurred in < 10% of crimes were excluded because these variables are unlikely to be of use when linking the majority of crimes (Santtila et al., 2008). In addition to the behavioural variables, the geographical location of the crime (x, y coordinates) and the estimated time of the offence (committed from dates/times and committed to dates/times) were extracted from the crime databases.

Analytic Strategy

A specially designed piece of software that has been utilised in numerous studies of BCL (e.g., Bennell & Canter, 2002; Bennell & Jones, 2005) was used to create all possible linked and unlinked crime pairs from the above data. These crime pairs were then split into those pairs that contained one commercial burglary and one commercial robbery (the cross-crime pairs, $n = 132,160$, 183 linked, 131,977 unlinked), those pairs that contained two burglaries (the burglary pairs, $n = 100,128$, 1,732 linked, 98,396 unlinked) and those pairs that contained two robberies (the robbery pairs, $n = 43,365$, 316 linked, 43,049 unlinked).

Three measures of behavioural similarity were calculated for each crime pair: 1) the straight-line kilometre distance between the two crimes in each pair (the inter-crime distance, ICD); 2) the number of days between the two crimes in each pair (based on the mid-point of the committed from/committed to dates; temporal proximity, TP); and 3) Jaccard's coefficient, which is a statistical measure of how similar two crimes are behaviourally (which was calculated based on the 67 dichotomous MO variables mentioned previously). The formula for calculating Jaccard's coefficient is: $a / (a + b + c)$, where 'a' refers to the number of behaviours present in both crimes in the pair (1/1) and 'b' and 'c' refer to the number of behaviours present in one crime but not in the other (1/0 and 0/1).

The assumption underpinning these three measures of similarity was that crimes committed by the same serial offender (i.e., linked crime pairs) would be more similar in terms of MO behaviour and closer together geographically and temporally than crimes committed by different serial offenders (i.e., unlinked crime pairs). This finding would suggest that behavioural consistency and distinctiveness exist to some degree in offender crime scene behaviour.

In order to test whether the assumptions of consistency and distinctiveness hold within these data and whether linked crimes could be accurately distinguished from unlinked crimes, three separate analyses were performed. First, Mann-Whitney U tests⁵ were conducted to statistically compare the linked crime pairs with the unlinked crime pairs in terms of ICD, TP and Jaccard's coefficient (separate analyses were performed to compare linked versus unlinked cross-crime pairs, linked versus unlinked burglary pairs and linked versus unlinked robbery pairs; Bonferroni corrected $\alpha = .006$). These analyses allowed us to examine how consistency and distinctiveness varied across different types of offender crime scene behaviour (geographical, temporal and MO) and how they varied across different types of offence series (burglary, robbery and cross-crime).

In the second phase of the analysis a series of binary logistic regression analyses were conducted using a leave-one-out (LOO) classification. In these analyses the dependent variable was linkage status (1 = linked crime pair; 0 = unlinked crime pair) and the independent variables were ICD, TP and Jaccard's values indicating similarity in MO behaviour. A separate logistic regression was conducted for each of the three independent variables at each level of analysis (cross-crime pairs, burglary pairs and robbery pairs), thereby producing a total of nine simple regressions. These analyses indicated how successfully ICD, TP and MO similarity could distinguish between linked and unlinked crime pairs (when used on their own, not in combination). In addition, three stepwise regression analyses were conducted using the forward likelihood ratio method (one regression for each level of analysis), which indicated whether superior discrimination accuracy could be achieved by combining the three measures of behavioural similarity.

The LOO classification procedure involved removing each crime pair from the sample one at a time and the remaining data were then used to develop a logistic regression model, which was subsequently applied to the extracted pair to produce a predicted probability value (ranging from 0, indicating a low predicted probability of the crime pair being linked, to 1.00, indicating a high predicted probability of the pair being linked). This pair was then returned to the dataset and the procedure repeated with the next pair until a probability value had been calculated for all linked and unlinked crime pairs in the sample (Woodhams & Labuschagne, 2012). The use of cross-validation procedures like this is particularly important in the current area of research, where the aim was to develop findings that can be used to guide future law enforcement investigations.

The third phase of the analysis involved using these predicted probability values to construct ROC curves. Twelve separate ROC curves were constructed, corresponding to the nine simple and three stepwise regression analyses described above. These ROC curves

provided an insight into how successfully the three measures of behavioural similarity (ICD, TP and MO similarity) were able to distinguish between linked and unlinked crime pairs.

RESULTS

Statistical Comparison of Linked and Unlinked Crime Pairs

The Mann-Whitney U tests reported in Table 1 indicate that linked crime pairs had statistically larger ($p < .001$) Jaccard's values and shorter ICD and TP values than unlinked crime pairs, which was a finding that existed across all three levels of analysis (i.e., with the cross-crime pairs, with the burglary pairs and with the robbery pairs). Overall, these findings suggest that consistency and distinctiveness exist at a level that exceeds chance in all three types of offender crime scene behaviour and at all three levels of analysis. Consequently, it should be possible to distinguish with some accuracy between linked and unlinked crime pairs using behavioural similarity.

[Table 1 about here]

Binary Logistic Regression Analysis

To further investigate behavioural consistency, distinctiveness and discrimination accuracy with these data, a series of simple and stepwise binary logistic regression analyses were conducted using a LOO classification method (see Table 2). All regression models were statistically significant ($p < .001$), which indicates a degree of success when attempting to distinguish between linked and unlinked burglary, robbery and cross-crime pairs using ICD, TP and similarity in MO behaviour. These findings further support the notion that consistency and distinctiveness exist in offender crime scene behaviour at a level that exceeds chance.

When the three measures of behavioural similarity are compared, it is clear that the greatest consistency, distinctiveness and discrimination accuracy was achieved by the ICD, followed by the TP and then similarity in MO behaviour (as indicated by the model χ^2 and R^2

values). However, the highest levels of consistency, distinctiveness and discrimination accuracy were achieved when combining the three measures of similarity into stepwise models, with each predictor making a statistically significant and unique contribution to predictive accuracy. Hence, the stepwise models achieved larger model χ^2 and R^2 values than the single-factor regression models, with these findings applying to cross-crime, burglary and robbery pairs.

[Table 2 about here]

ROC Analysis

To further test consistency, distinctiveness and discrimination accuracy, ROC curves were constructed using the predicted probability values produced as a result of the regression analyses (see Table 3). From Table 3 it is clear that all 12 regression models were able to distinguish between linked and unlinked crime pairs to a statistically significant degree ($p < .001$), which further suggests that relative consistency and distinctiveness exist in offender crime scene behaviour.

When comparing the different types of offender crime scene behaviour, ICD and TP demonstrate statistically higher levels of consistency, distinctiveness and discrimination accuracy than MO behaviours, as indicated by the non-overlapping AUC confidence intervals (Knezevic, 2008). However, the highest levels of consistency, distinctiveness and discrimination accuracy were achieved when combining the three measures of behavioural similarity, with the stepwise models achieving high AUC values (AUCs $> .90$; Swets, 1988) that were statistically larger than those for the single-factor regression models. These findings apply to cross-crime, burglary and robbery pairs.

[Table 3 about here]

Discussion

The aim of the current study was to address key limitations within the BCL literature, thereby helping to build a more robust and reliable evidence-base from which theoretical and practical conclusions can be drawn. The main contributions of this research are that: 1) it presents the first empirical test of whether MO behaviours can be used to support cross-crime linkage; 2) it presents the first empirical test of cross-crime linkage using data that contain an uneven number of offences per offender (which is closer to the real life investigative context within which BCL would be used) and 3) it replicates key findings relating to cross-crime linkage (Tonkin et al., 2011, 2012), commercial burglary (Bennell & Jones, 2005) and commercial robbery (Woodhams & Toye, 2007) in a geographical location not previously tested, using one of the largest datasets yet compiled for the purposes of BCL research.

Overall, the serial offenders in this sample displayed a sufficient degree of consistency and distinctiveness to allow linked and unlinked crimes to be distinguished at a level that far exceeds chance. Importantly, these findings applied to all three types of offender behaviour examined in this study (ICD, TP and MO similarity) and to all three levels of analysis (cross-crime, burglary and robbery). Consequently, there is evidence to support the use of BCL in practice and the current study provides statistical formulae that might be used to facilitate this process (as discussed below).

There are, however, important differences in the level of consistency, distinctiveness and discrimination accuracy observed in this study as a function of the type of offender behaviour and the type of offence examined.

In terms of the former, offender geographical and temporal behaviour (the ICD and TP, respectively) demonstrated greater consistency, distinctiveness and discrimination accuracy than MO behaviours. These findings are consistent with previous research on commercial burglary, residential burglary and auto theft (e.g., Bennell & Jones, 2005; Bouhana et al., 2014; Tonkin et al., 2008). A variety of explanations have been proposed for

this superior performance (see Bennell & Jones, 2005), including that an offender can exert greater control over decisions about where and when to commit a crime than s/he can over decisions about what to steal, whether violence is used etc., which depend to some extent on situational characteristics at the crime scene. As explained by personality researchers, behaviours under primary control of the actor tend to be more consistent than those that are heavily influenced by the situational context (Funder & Colvin, 1991). An alternative explanation for the findings is that geographical and temporal information are more easily and objectively recorded than some MO behaviours, such as whether a property was searched in a tidy or untidy manner (which is a subjective judgment) and what property was stolen (which depends on what a victim is willing and able to report as stolen). The ease of recording information would inevitably impact on data quality, with lower data quality making it more difficult to detect meaningful patterns of consistency, distinctiveness and discrimination accuracy. Thus, the larger AUC values for ICD and TP compared with MO behaviours may simply be a result of differences in data quality, rather than necessarily due to inherent differences in the consistency and/or distinctiveness of offender behaviour.

In addition to type of crime scene behaviour, there was also variation as a function of crime type. More specifically, the level of consistency, distinctiveness and discrimination accuracy for MO behaviours was greater amongst robbery crime pairs (AUC = .82) than either burglary (AUC = .66) or cross-crime pairs (AUC = .63). This finding replicates the previous work of Woodhams and Toye (2007), who reported high levels of discrimination accuracy using MO behaviours that were far greater than those achieved in other property-oriented crimes, such as burglary and car theft (e.g., Bennell & Canter, 2002; Bennell & Jones, 2005; Tonkin et al., 2008). Moreover, when the ICD, TP and MO similarity were combined in the stepwise model an AUC value of 0.97 was observed (the largest reported in

the current study), thereby suggesting high levels of consistency, distinctiveness and discrimination accuracy for robbery offences.

In contrast, the AUC values for ICD and TP were somewhat comparable across burglary, robbery and cross-crime pairs, which suggests that these measures offer similar potential for behavioural linking regardless of crime type (except TP with burglary pairs, which achieved a lower AUC compared to robbery and cross-crime pairs).

Having considered the main findings and some potential explanations for these findings, the theoretical and practical implications will now be briefly explored. Perhaps the most striking finding from this study is that statistically significant AUC values were observed for the cross-crime pairs, which indicates that offenders demonstrate a degree of consistency in their crime scene behaviour, even when engaging in two very different offending behaviours (e.g., the presence of at least one victim, and sometimes multiple victims, in robbery but not burglary creates a number of very different considerations for an offender). While surprising, this finding is logical when considered in light of the personality literature, which suggests that behavioural consistency should be expected, even across seemingly very different situations, provided the actor perceives these situations as psychologically similar (e.g., Furr & Funder, 2004; Sherman, Nave, & Funder, 2010). The literature on offender decision-making has demonstrated that burglary and robbery offenders are often motivated by the same need (i.e., to gain quick and easy money), and these crime types are often discussed and used interchangeably by offenders (e.g., Bennett & Wright, 1984; Wright & Decker, 1994, 1997). It is, therefore, logical to predict that burglary and robbery might be perceived in a similar way psychologically by offenders, which would help to explain the consistency observed in the current study for cross-crime pairs.

In this study the ICD achieved the highest discrimination accuracy of all three measures, and this accuracy was comparable across all three levels of analysis. These

findings lend support to several seminal theories of offender behaviour (such as rational choice theory, routine activities theory and crime pattern theory), which suggest that offenders seek to minimise the efforts and risks involved in offending (e.g., by returning to geographical locations that are familiar to them). Moreover, they suggest that similar psychological processes are involved in the production of criminal spatial behaviour, irrespective of crime type, which is exactly what one would predict from seminal theories such as crime pattern theory and rational choice theory (Brantingham & Brantingham, 1981, 1984; Clarke & Felson, 1993). These findings also lend support to the notion that the near-repeat phenomenon can be explained by the same offender returning to that geographical area in order to commit further crimes (Bernasco, 2008).

In terms of the practical implications of this study, the findings suggest that there is significant potential for BCL using geographical, temporal and MO behaviour. This is reassuring given the already extensive use of this procedure by law enforcement agencies around the world (e.g., Labuschagne, 2012; Snook, Luther, House, Bennell, & Taylor, 2012; Yokota et al., 2007). Importantly, this study suggests that BCL can function not just within a single crime type (burglary or robbery) but that it is possible to use offender crime scene behaviour to identify linked crime series containing multiple crime types. Given that the most prolific offenders are typically the most versatile (e.g., Piquero et al., 2007), this study provides an important step towards improved methods for investigating those offenders that commit a disproportionate amount of crime and impose considerable costs on society.

In the future it may be possible to develop a decision-support BCL tool based on these findings, which would analyse large crime databases in a quick and efficient manner, using a combination of geographical, temporal and MO information to create a prioritised list of potentially linked crimes for further investigation by an analyst. This would help to tackle one of the fundamental challenges faced when conducting BCL in practice, which is the vast

amount of information that must be processed that inevitably makes manual searching of large databases impractical and associated with a high risk of analytical error. Given staff cuts and the rapidly reducing resources available to the police, such tools may be a valuable asset to analysts involved in BCL.

Before such systems can be implemented in practice, however, a significant amount of testing would be required and the existing limitations of research must be addressed. The primary limitation of this study is that the analyses relied on detected crimes and did not include non-serial offences, which does not reflect the data with which BCL would be used in practice. While research has suggested that these issues may not impact on findings as much as anticipated (Tonkin, Santtila, & Bull, 2012; Tonkin et al., 2012), future research must endeavour to continue testing BCL using unsolved and non-serial offences.

A further potential limitation is that the data utilised in this study were taken directly from police crime databases. While this is a strength because it is important to test the principles of BCL with real-world data, there is necessarily a compromise in terms of a lack of experimental control (e.g., not being able to test the inter-rater reliability of the data). Thus, there may be a number of unidentified and uncontrollable inaccuracies within the data that impact on the degree of consistency, distinctiveness and discrimination accuracy observed in the current study.

An important avenue for future research is to develop BCL decision-support tools and to test them in experimental scenarios and in practice with ongoing criminal investigations. This is vital because, while it is important to test the principles of consistency and distinctiveness, there are many other- equally important- practical issues (such as the availability of resources, the usability of linkage tools etc.) that are not tested by studies using the methodology adopted in this study. Unless these issues are explored, we will never truly

This is the prepublication version of a paper now accepted for publication by *Legal and Criminological Psychology* for which the DOI is 10.1111/lcrp.12085

know whether the BCL literature can contribute to more reliable, accurate and cost-effective methods of linking crime.

References

- Bennell, C., & Canter, D. V. (2002). Linking commercial burglaries by modus operandi: Tests using regression and ROC analysis. *Science and Justice*, *42*, 153-164. doi: 10.1016/S1355-0306(02)71820-0
- Bennell, C., & Jones, N. J. (2005). Between a ROC and a hard place: A method for linking serial burglaries by modus operandi. *Journal of Investigative Psychology and Offender Profiling*, *2*, 23-41. doi: 10.1002/jip.21
- Bennell, C., Jones, N. J., & Melnyk, T. (2009). Addressing problems with traditional crime linking methods using receiver operating characteristic analysis. *Legal and Criminological Psychology*, *14*, 293-310. doi: 10.1348/135532508X349336
- Bennell, C., Mugford, R., Ellingwood, H., & Woodhams, J. (2014). Linking crimes using behavioural clues: Current levels of linking accuracy and strategies for moving forward. *Journal of Investigative Psychology and Offender Profiling*, *11*, 29-56. doi: 10.1002/jip.1395
- Bennett, T., & Wright, R. (1984). *Burglars on burglary: Prevention and the offender*. Aldershot, UK: Gower.
- Bernasco, W. (2008). Them again? Same-offender involvement in repeat and near repeat burglaries. *European Journal of Criminology*, *5*, 411-431. doi: 10.1177/1477370808095124
- Bouhana, N., Johnson, S. D., & Porter, M. (2014). Consistency and specificity in burglars who commit prolific residential burglary: Testing the core assumptions underpinning

This is the prepublication version of a paper now accepted for publication by Legal and Criminological Psychology for which the DOI is 10.1111/lcrp.12085

behavioural crime linkage. *Legal and Criminological Psychology*. Advance online publication. doi: 10.1111/lcrp.12050

Brantingham, P. J., & Brantingham, P. L. (1981). *Environmental criminology*. Beverly Hills, CA: Sage.

Brantingham, P. J., & Brantingham, P. (1984). *Patterns in crime*. New York, NY: Macmillan.

Burrell, A., Bull, R., & Bond, J. W. (2012). Linking personal robbery offences using offender behaviour. *Journal of Investigative Psychology and Offender Profiling*, 9, 201-222. doi: 10.1002/jip.1365

Butler, H., & Gannon, T. A. (2015). The scripts and expertise of firesetters: A preliminary conceptualization. *Aggression and Violent Behavior*, 20, 72-81. doi: 10.1016/j.avb.2014.12.011

Clarke, R. V., & Felson, M. (1993). *Routine activity and rational choice*. New Brunswick, NJ: Transaction.

Cornish, D. (1994). The procedural analysis of offending and its relevance for situational prevention. In R. V. Clarke (Ed.), *Crime prevention studies* (Vol. 3, pp. 151-196). Monsey, NY: Criminal Justice Press.

Cowen, N., & Williams, N. (2012). *Comparisons of crime in OECD countries*. Retrieved from: www.civitas.org.uk/crime/crime_stats_oecdjan2012.pdf

Farrington, D. P., Snyder, H. N., & Finnegan, T. A. (1988). Specialization in juvenile court careers. *Criminology*, *26*, 461-487. doi: 10.1111/j.1745-9125.1988.tb00851.x

Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, *60*, 773-794. doi: 10.1037/0022-3514.60.5.773

Furr, R. M., & Funder, D. C. (2004). Situational similarity and behavioral consistency: Subjective, objective, variable-centred, and person-centred approaches. *Journal of Research in Personality*, *38*, 421-447. doi: 10.1016/j.jrp.2003.10.001

Knezevic, A. (2008). *Overlapping confidence intervals and statistical significance*. StatNews No. 73. Retrieved from: www.cscu.cornell.edu/news/statnews/stnews73.pdf

Labuschagne, G. (2012). The use of a linkage analysis as an investigative tool and evidential material in serial offenses. In K. Borgeson & K. Kuehnle (Eds.), *Serial offenders: Theory and practice* (pp. 187-215). Sudbury, MA: Jones & Bartlett Learning.

Melnyk, T., Bennell, C., Gauthier, D. J., & Gauthier, D. (2011). Another look at across-crime similarity coefficients for use in behavioural linkage analysis: An attempt to replicate

Woodhams, Grant, and Price (2007). *Psychology, Crime & Law*, *17*, 359-380. doi: 10.1080/10683160903273188

This is the prepublication version of a paper now accepted for publication by *Legal and Criminological Psychology* for which the DOI is 10.1111/lcrp.12085

Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualising situations, dispositions, dynamics, and invariance in personality structure.

Psychological Review, *102*, 246-268. doi: 10.1037/0033-295X.102.2.246

Mischel, W., Shoda, Y., & Smith, R. E. (2004). *Introduction to personality: Toward an integration* (7th ed.). New York, NY: Wiley.

Nee, C., & Meenaghan, A. (2006). Expert decision making in burglars. *British Journal of Criminology*, *46*, 935-949. doi: 10.1093/bjc/az1013

Nee, C., & Taylor, M. (2000). Examining burglars' target selection: Interview, experiment or ethnomethodology. *Psychology, Crime & Law*, *6*, 45-59. doi: 10.1080/10683160008410831

Piquero, A. R., Farrington, D. P., & Blumstein, A. (2007). *Key issues in criminal career research: New analyses of the Cambridge study in delinquent development*. New York, NY: Cambridge University Press.

Santtila, P., Fritzon, K., & Tamelander, A. L. (2004). Linking serial arson incidents on the basis of crime scene behavior. *Journal of Police and Criminal Psychology*, *19*, 1-16. doi: 10.1007/BF02802570

Santtila, P., Pakkanen, T., Zappalà, A., Bosco, D., Valkama, M., & Mokros, A. (2008). Behavioural crime linking in serial homicide. *Psychology, Crime & Law*, *14*, 245-265. doi: 10.1080/10683160701739679

This is the prepublication version of a paper now accepted for publication by Legal and Criminological Psychology for which the DOI is 10.1111/lcrp.12085

Sherman, R. A., Nave, C. S., & Funder, D. C. (2010). Situational similarity and personality predict behavioral consistency. *Journal of Personality and Social Psychology*, *99*, 330-343. doi: 10.1037/a0019796

Snook, B., Luther, K., House, J. C., Bennell, C., & Taylor, P. J. (2012). The Violent Crime Linkage Analysis System: A test of interrater reliability. *Criminal Justice and Behavior*, *39*, 607-619. doi: 10.1177/0093854811435208

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*, 1285-1293. doi: 10.1126/science.3287615

Taylor, M., & Nee, C. (1988). The role of cues in simulated residential burglary: A preliminary investigation. *British Journal of Criminology*, *28*, 396-401.

Tonkin, M. (2014). Testing the theories underpinning crime linkage. In J. Woodhams & C. Bennell (Eds.), *Crime linkage: Theory, research and practice* (pp. 107-139). London, UK: CRC Press.

Tonkin, M., Grant, T., & Bond, J. W. (2008). To link or not to link: A test of the case linkage principles using serial car theft data. *Journal of Investigative Psychology and Offender Profiling*, *5*, 59-77. doi: 10.1002/jip.74

Tonkin, M., Santtila, P., & Bull, R. (2012). The linking of burglary crimes using offender behaviour: Testing research cross-nationally and exploring methodology. *Legal and Criminological Psychology*, *17*, 276-293. doi: 10.1111/j.2044-8333.2010.02007.x

Tonkin, M., Woodhams, J., Bull, R., & Bond, J. W. (2012). Linking solved and unsolved crimes using offender behaviour. *Forensic Science International*, 222, 146-153. doi: 10.1016/j.forsciint.2012.05.017

Tonkin, M., Woodhams, J., Bull, R., Bond, J. W., & Palmer, E. J. (2011). Linking different types of crime using geographical and temporal proximity. *Criminal Justice and Behavior*, 38, 1069-1088. doi: 10.1177/0093854811418599

Topalli, V., Jacques, S., & Wright, R. (2015). "It takes skills to take a car": Perceptual and procedural expertise in carjacking. *Aggression and Violent Behavior*, 20, 19-25. doi: 10.1016/j.avb.2014.12.001

Ward, T., & Hudson, S. M. (1998). A model of the relapse process in sexual offenders. *Journal of Interpersonal Violence*, 13, 700-725. doi: 10.1177/088626098013006003

Ward, T., & Hudson, S. M. (2000). A self-regulation model of relapse prevention. In D. R. Laws, S. M. Hudson, & T. Ward (Eds.), *Remaking relapse prevention with sex offenders: A sourcebook* (pp. 79-101). Newbury Park, CA: Sage.

Woodhams, J., Hollin, C. R., & Bull, R. (2007). The psychology of linking crimes: A review of the evidence. *Legal and Criminological Psychology*, 12, 233-249. doi: 10.1348/135532506X118631

This is the prepublication version of a paper now accepted for publication by Legal and Criminological Psychology for which the DOI is 10.1111/lcrp.12085

Woodhams, J., & Labuschagne, G. (2012). A test of case linkage principles with solved and unsolved serial rapes. *Journal of Police and Criminal Psychology*, 27, 85-98. doi: 10.1007/s11896-011-9091-1

Woodhams, J., & Toye, K. (2007). An empirical test of the assumptions of case linkage and offender profiling with serial commercial robberies. *Psychology, Public Policy, and Law*, 13, 59-85. doi: 10.1037/1076-8971.13.1.59

Wright, R. T., & Decker, S. H. (1994). *Burglars on the job*. Boston, MA: Northeastern University Press.

Wright, R. T., & Decker, S. H. (1997). *Armed robbers in action: Stickups and street culture*. Boston, MA: Northeastern University Press.

Yokota, K., Fujita, G., Watanabe, K., Yoshimoto, K., & Wachi, T. (2007). Application of the behavioral investigative support system for profiling perpetrators of serious sexual assaults. *Behavioral Sciences and the Law*, 25, 841-856. doi: 10.1002/bsl.793

Tables

Table 1

Statistical Comparisons of Linked and Unlinked Crime Pairs in terms of Similarity in Offender Crime Scene Behaviour

Level of Analysis	Inter-Crime Distance	Temporal Proximity	Modus Operandi Behavioural Similarity
Cross-Crime Pairs	$Z = -20.08, p < .001^*, r = -.06$ Median (KM): Linked pairs = 3.04 Unlinked pairs = 13.92	$Z = -14.88, p < .001^*, r = -.04$ Median (days): Linked pairs = 40.00 Unlinked pairs = 206.00	$Z = 9.08, p < .001^*, r = .02$ Median (Jaccard): Linked pairs = .14 Unlinked pairs = .08
Burglary Pairs	$Z = -57.87, p < .001^*, r = -.18$ Median (KM): Linked pairs = 1.99 Unlinked pairs = 11.33	$Z = -38.77, p < .001^*, r = -.12$ Median (days): Linked pairs = 43.00 Unlinked pairs = 223.00	$Z = 30.81, p < .001^*, r = .10$ Median (Jaccard): Linked pairs = .18 Unlinked pairs = .10
Robbery Pairs	$Z = -24.45, p < .001^*, r = -.12$ Median (KM): Linked pairs = 3.60	$Z = -25.11, p < .001^*, r = -.12$ Median (days): Linked pairs = 15.00	$Z = 20.53, p < .001^*, r = .10$ Median (Jaccard): Linked pairs = .33

	Unlinked pairs = 15.88	Unlinked pairs = 199.00	Unlinked pairs = .13
--	------------------------	-------------------------	----------------------

Note. KM = Kilometres.

* Significant at the Bonferroni corrected α level of .006

Table 2

Simple and Stepwise Binary Logistic Regression Analyses Predicting Linkage Status

Model	Constant (SE)	Logit (SE)	Model χ^2 (df)	Wald (df)	Model R^2 (Cox and Snell – Nagelkerke)
Cross-Crime Pairs					
Inter-Crime Distance	-3.13 (.13)	-.48 (.03)	619.50 (1)*	277.31 (1)*	.01 – .23
Temporal Proximity	-5.11 (.11)	-.01 (.00)	222.76 (1)*	123.85 (1)*	.00 – .08
Similarity in MO Behaviour (Jaccard)	-7.21 (.10)	5.35 (.44)	100.15 (1)*	148.36 (1)*	.00 – .04
Stepwise Model (ICD + TP + Jaccard)	-2.67 (.18)	ICD: -.47 (.03) TP: -.01 (.00) Jaccard: 4.95 (.49)	938.72 (3)*	ICD: 270.65 (1)* TP: 121.75 (1)* Jaccard: 100.47 (1)*	.01 – .34
Burglary Pairs					
Inter-Crime Distance	-1.36 (.04)	-.48 (.01)	4622.03 (1)*	1950.50 (1)*	.05 – .28
Temporal Proximity	-2.87 (.04)	-.01 (.00)	1388.15 (1)*	919.80 (1)*	.01 – .09
Similarity in MO	-4.75 (.04)	4.88 (.15)	927.72 (1)*	1122.79 (1)*	.01 – .06

Behaviour (Jaccard)					
Stepwise Model (ICD + TP + Jaccard)	-1.12 (.06)	ICD: -.43 (.01) TP: -.01 (.00) Jaccard: 4.45 (.18)	6357.19 (3)*	ICD: 1710.04 (1)* TP: 703.18 (1)* Jaccard: 630.20 (1)*	.06 – .38
Robbery Pairs					
Inter-Crime Distance	-1.79 (.10)	-.35 (.02)	895.15 (1)*	445.92 (1)*	.02 – .25
Temporal Proximity	-2.62 (.08)	-.03 (.00)	852.73 (1)*	283.37 (1)*	.02 – .24
Similarity in MO Behaviour (Jaccard)	-6.38 (.11)	5.96 (.24)	476.73 (1)*	594.47 (1)*	.01 – .13
Stepwise Model (ICD + TP + Jaccard)	-1.63 (.16)	ICD: -.29 (.02) TP: -.02 (.00) Jaccard: 5.94 (.35)	1947.27 (3)*	ICD: 323.39 (1)* TP: 215.46 (1)* Jaccard: 293.32 (1)*	.04 – .53

* $p < .001$

Table 3

ROC Analyses Testing the Discrimination Accuracy of Three Measures of Behavioural

Similarity

Model	AUC (SE)	95% Confidence Interval	Classification Category (Swets, 1988)
Cross-Crime Pairs			
Inter-Crime Distance	.93 (.01)*	[.91, .94]	High
Temporal Proximity	.82 (.02)*	[.78, .85]	Moderate
Similarity in MO Behaviour (Jaccard)	.63 (.03)*	[.58, .68]	Low
Stepwise Model (ICD + TP + Jaccard)	.95 (.01)*	[.93, .96]	High
Burglary Pairs			
Inter-Crime Distance	.91 (.00)*	[.90, .91]	High
Temporal Proximity	.77 (.01)*	[.76, .78]	Moderate
Similarity in MO Behaviour (Jaccard)	.66 (.01)*	[.65, .68]	Low
Stepwise Model (ICD + TP + Jaccard)	.93 (.00)*	[.93, .94]	High
Robbery Pairs			
Inter-Crime Distance	.90 (.01)*	[.88, .92]	High
Temporal Proximity	.91 (.01)*	[.89, .92]	High
Similarity in MO Behaviour (Jaccard)	.82 (.01)*	[.79, .84]	Moderate

This is the prepublication version of a paper now accepted for publication by Legal and Criminological Psychology for which the DOI is 10.1111/lcrp.12085

Stepwise Model (ICD + TP + Jaccard)	.97 (.01)*	[.96, .98]	High
--	------------	------------	------

* $p < .001$

Footnotes

¹ Inter-crime distance refers to the straight-line kilometre distance between offences and temporal proximity to the number of days between offences. The assumption is that linked crimes will occur closer in time and space (i.e., will have shorter inter-crime distance and temporal proximity values) than unlinked crimes.

² When conducting studies in this area, researchers must know which crimes have been committed by which offenders, otherwise it will not be possible to test the predictive accuracy of the linkage models developed in these studies. That is, if researchers do not know which crimes have been committed by which offenders in real life, they will have no way of knowing whether their predictions are correct. Typically, researchers have determined the linkage status of crime pairs (linked vs. unlinked) by sampling crimes that are detected/solved. However, this does not reflect the real-world scenario in which BCL is used (i.e., with unsolved crimes), which has led researchers to examine BCL using data that contain unsolved offences. In these studies linkage status is confirmed via the recovery of matching DNA material across several crime scenes. This provides the researcher with a way of determining which crimes were in reality committed by the same offender and which were not.

³ In England and Wales the Home Office defines commercial burglary as the theft of property from business premises and commercial robbery as the theft of property from business premises that involves the actual or implied use of force.

⁴ It was necessary to take a sub-section of the data due to limitations in the amount of data that Excel and the specialised package used during the analyses were able to process.

⁵ Ideally a dependent-measures statistic would have been used because the process of creating all pairwise linked and unlinked crime pairs from the data meant that each crime appeared in multiple crime pairs. Consequently, the linked and unlinked crime pairs cannot be considered statistically independent. However, the non-parametric dependent-measures statistic, Wilcoxon signed-rank test, does not allow for unequal sample sizes across the two comparison groups, thereby meaning that the independent-measures Mann-Whitney test had to be utilised. The violation of independence that this necessarily causes should be borne in mind when interpreting this section of the findings. Nevertheless, this violation of the statistical assumptions is not substantial given that the subsequent ROC analyses were uninfluenced by this issue. It should also be noted that a non-parametric statistic was appropriate given the non-normally distributed data (as indicated by statistically significant Kolmogorov-Smirnov tests, $p < .001$, large skewness values and large kurtosis values).