

**The Linking of Burglary Crimes using Offender Behaviour: Testing Research
Cross-Nationally and Exploring Methodology**

Tonkin, M., Santtila, P., & Bull, R.

Abstract

Purpose. The current study tests whether existing behavioural case linkage findings from the United Kingdom (UK) will replicate abroad with a sample of residential burglaries committed in Finland. In addition, a previously discussed methodological issue is empirically explored.

Methods. Seven measures of behavioural similarity, geographical proximity and temporal proximity are calculated for pairs of burglary crimes committed by 117 serial burglars in Finland. The ability of these seven measures to distinguish between pairs of crimes committed by the same offender (linked pairs) and different offenders (unlinked pairs) is tested using logistic regression and Receiver Operating Characteristic (ROC) analysis. Two methodologies for forming the unlinked pairs are compared; one representing the ‘traditional’ approach used by research and the other a new approach that represents a potentially more realistic and statistically sound approach to testing case linkage.

Results. A wider range of offender behaviours were able to distinguish between linked and unlinked crime pairs in the current Finnish sample than in previous UK-based research. The most successful features were the kilometre-distance between crimes (the

intercrime distance), the number of days separating offences (temporal proximity), and a combination of target, entry, internal and property behaviours (the combined domain).

There were no statistically significant differences between the two methodological approaches.

Conclusions. The current findings demonstrate that a wider range of offender behaviours can be used to discriminate between linked and unlinked residential burglary crimes committed in Finland than in the UK. The use of a more realistic and statistically sound methodology does not lead to substantial changes in case linkage findings.

The investigation of crime can be enhanced if the police are able to identify crimes that have been committed by the same person or the same group of people (i.e. link crimes as an offence series). The primary benefit of this process is that it allows the evidence from multiple investigations to be pooled, thus potentially increasing the quantity and quality of evidence that is available to prosecute offenders (Grubin, Kelly, & Brunsdon, 2001).

One approach that the police use to link crimes is through offender behaviour, whereby crimes that are committed in a similar way behaviourally are judged to have been committed by the same offender (linked crimes), whereas crimes that are committed in different ways are judged to be the work of separate offenders (unlinked crimes) (Woodhams, Hollin, & Bull, 2007). Behavioural case linkage relies on two theoretical assumptions about offender behaviour: offenders must demonstrate a degree of similarity in the way they behave from one crime to the next (behavioural consistency) and their behaviour must be different from the way in which other offenders behave (behavioural distinctiveness) (Bennell, Jones, & Melnyk, 2009; Woodhams et al., 2007).

Over the last decade a number of empirical studies have begun to test the validity of behavioural case linkage with a range of crime types (e.g. Bennell & Canter, 2002; Salfati & Bateman, 2005; Santtila, Fritzon, & Tamelander, 2004; Tonkin, Grant, & Bond, 2008; Woodhams & Toye, 2007). This research has focused on a range of issues, including the exploration of methodology (e.g. Bennell, Gauthier, Gauthier, Melnyk, & Musolino, 2010; Woodhams, Grant, & Price, 2007), extending the evidence for case linkage to new crime types (e.g. Tonkin et al., 2008; Woodhams & Toye, 2007), and

attempting to replicate existing findings in new geographical locations and time periods (Bennell & Jones, 2005; Markson, Woodhams, & Bond, 2010).

But, despite this growing body of research, significant work is still required. For example, although the case linkage findings relating to residential burglary have been replicated across the UK, there has been no attempt to replicate these findings abroad. Also, the range of research into methodological issues is somewhat limited. The aims of the current study are, therefore, twofold: 1) to examine the replicability of existing case linkage findings abroad using a sample of residential burglaries committed in Finland; and 2) to explore a methodological issue that has been discussed in the literature but is yet to receive substantial attention (Bennell & Canter, 2002; Woodhams, 2008).

Research into Behavioural Case Linkage with Burglary

There now exist a number of studies that have examined behavioural case linkage using burglary data (e.g. Bennell & Canter, 2002; Bennell & Jones, 2005; Ewart, Oatley, & Burn, 2005; Goodwill & Alison, 2006; Markson et al., 2010).

This research has found that certain offender behaviours demonstrate sufficient consistency and distinctiveness to allow linked crimes to be reliably distinguished from unlinked crimes. The kilometre-distance between offence locations (the intercrime distance) has been particularly successful in this task, with the intercrime distance outperforming target characteristics, entry behaviours, internal behaviours (such as offender search behaviour), and property stolen when differentiating between linked and unlinked burglaries. These findings have been shown to replicate in various locations around the UK (Bennell, 2002; Bennell & Jones, 2005; Markson et al., 2010).

The number of days separating burglaries (temporal proximity) has also been shown to reliably differentiate between linked and unlinked crimes (Markson et al., 2010). In Markson and colleagues' study, the temporal proximity was second only to the intercrime distance in terms of discrimination accuracy (it outperformed target, entry, internal, and property behaviours), and the combination of intercrime distance and temporal proximity was able to facilitate the greatest level of linking success. These findings are corroborated by studies that have utilised different methodologies (Ewart et al., 2005; Goodwill & Alison, 2006).

The case linkage literature on burglary can be criticised, however, because it has only ever focused on samples from the UK. It is, therefore, unclear whether the same findings apply in other countries. Indeed, different countries can be expected to vary in terms of physical and social geography, the availability, type and distribution of potential targets, and approaches to policing and data recording/storage, which might lead to variation in case linkage performance. We briefly review some of these differences now, with particular focus on differences between the UK and Finland that are relevant to burglary crime.

Cross-National Differences between the UK and Finland

The UK and Finland differ in terms of population density, with Finland averaging approximately 16 persons per square kilometre compared to 254 persons per square kilometre in the UK¹. Housing is, therefore, much more dispersed in Finland than the

¹All statistics included in this section of the manuscript were obtained from:
http://en.wikipedia.org/wiki/Main_Page
<http://www.ymparisto.fi/default.asp?node=10078&lan=en>
<http://www.communities.gov.uk/documents/corporate/pdf/971061.pdf>

UK, which would be expected to impact considerably on offender spatial behaviour such as journey-to-crime and the intercrime distance.

There are also differences between the UK and Finland in terms of housing. The predominant type of residential accommodation in the UK is a house or bungalow (82% of households in 2006), whereas in Finland the majority of housing is split across two types (44% are flats and 40% detached housing in 2003). The slightly wider variation in housing that is evident in Finland might impact on offender consistency and distinctiveness because there would be more scope for offenders to target different types of house and they may need to employ a wider range of entry behaviours. Thus, it might be hypothesised that case linkage performance for target and entry behaviours would be enhanced in a Finnish compared to a UK sample.

Also, it is not unreasonable to suggest that police forces in Finland will differ from those in the UK in terms of how they record information about burglary crime. There may be additional behaviours recorded in Finland that are not recorded in the UK and/or the same behaviours may be recorded in different ways. Differences such as these have the potential to impact on case linkage performance.

Methodology in Behavioural Case Linkage Research

Much of the research currently conducted into behavioural case linkage has followed a methodology originally proposed by Bennell (2002). This involves creating pairs of linked and unlinked crimes and coding each crime for the presence or absence of a range of offender behaviours. The ability of these behaviours to distinguish between linked and unlinked offences is then tested by logistic regression and Receiver Operating

Characteristic (ROC) analysis. The hypothesis is that linked crimes will display greater behavioural similarity than unlinked crimes, which leads to statistically significant discrimination accuracy as measured by these two types of analysis (e.g. Bennell & Canter, 2002).

Researchers adopting this approach have recently begun to explore how this methodology might be improved, including whether more appropriate statistical measures of consistency and distinctiveness exist (e.g. Bennell et al., 2010) and whether alternative approaches to forming the linked and unlinked pairs might be utilised (Woodhams, 2008). In terms of the latter, Woodhams (2008) has argued that the current practice of forming the unlinked pairs from the same sample of crimes that are used to form the linked pairs may be problematic for several reasons. We will focus on the statistical criticisms she presents. Woodhams argues that such an approach leads to the assumption of statistical independence becoming violated during logistic regression. The impact of violating this assumption is that the confidence interval is spuriously inflated and the subsequent *p*-value of any statistical test diminished (Hopkins, 2001). Consequently, the statistical significance of certain offender behaviours may have been underestimated in previous research, thus leading to them being inappropriately rejected as valuable case linkage features.

Another issue that arises from forming the unlinked pairs from the same sample as the linked pairs is that it leads to a sample consisting solely of serial offences (Bennell & Canter, 2002; Woodhams, 2008). This is problematic because police crime analysts who are charged with conducting behavioural case linkage in real life must distinguish linked crimes from a backdrop of not just serial crimes but non-serial crimes as well. The

current methodology for testing case linkage is not, therefore, a true reflection of the reality in which case linkage is expected to perform. This is problematic for an area of applied research such as this that aims to be relevant to police practice.

The Current Study

The current study, therefore, aimed to explore whether existing case linkage findings for residential burglary in the UK would replicate cross-nationally with a sample of Finnish burglaries. Based on previous research, it was hypothesised that the intercrime distance and temporal proximity would achieve the highest levels of discrimination accuracy. However, it was also predicted that target and entry behaviours would perform more successfully than they have in previous UK-based research. The analyses were initially conducted using Bennell's (2002) original methodology to ensure comparability with previous research, but they were also conducted using a methodology whereby the unlinked pairs were formed from a statistically independent sample of serial and non-serial crimes. The outcomes from these two methodological approaches were then compared to determine whether the findings altered when using an approach that is potentially more ecologically valid and statistically sound. Given the paucity of research in this area it was not possible to make specific predictions regarding these comparisons.

Method

The Data

To facilitate the replication aspect of this study, 234 solved residential burglary crimes committed by 117 serial burglars in the Greater Helsinki region of Finland¹ (between 1990 and 2001) were extracted from a dataset that had been established during a previous project (Laukkanen, Santtila, Jern, & Sandnabba, 2008; Santtila, Ritvanen, & Mokros, 2004). The 234 crimes represent a random selection of solved residential burglary crimes committed during this period. These data were originally collected to facilitate an investigation of offender and geographical profiling in Finland. Two offences were randomly selected from the series of each offender, which was necessary to prevent highly prolific offenders with unusually consistent or inconsistent offence behaviour having an undue influence on the findings (Bennell, 2002). This dataset is referred to as dataset one.

To facilitate the analysis of methodology, 508 serial and non-serial burglaries were extracted from the same dataset described above (Laukkanen et al., 2008; Santtila, Ritvanen et al., 2004). None of these 508 crimes were included in the 234 burglaries that were extracted for the purposes of replication, so these two datasets can be considered statistically independent. Due to the possibility that serial burglaries are more common in any jurisdiction than non-serial burglaries (Bennell & Canter, 2002; Goodwill & Alison, 2006), the 508 crimes contained a disproportionate number of serial to non-serial burglaries. In the absence of any published literature to suggest exactly how disproportionate serial and non-serial burglaries are a ratio of approximately 3:1 was used. It was hoped that this approach would contribute to creating a more realistic

¹ The greater Helsinki region of Finland covers an area of approximately 815KM² that contains the capital of Finland, Helsinki, and the neighbouring cities of Espoo and Vantaa.

research environment in which to test behavioural case linkage. This dataset is referred to as dataset two.

For each of the crimes in these two datasets a range of behavioural data existed, including the location of the crime (x, y coordinates indicating the offence location to the nearest metre), the date the crime was committed (in many cases this was the mid-point between an “earliest crime date” and a “latest crime date” because the exact offence time was unknown, which is not unusual for burglary crime; Ratcliffe, 2002), the type of property burgled, the method of entry, the search behaviour, and the type and cost of property stolen (see the Appendix for a full list of behavioural data included in this study).

Apart from the location and temporal information, the data were stored in a binary format (1 = present in the crime; 0 = absent). The use of binary data is consistent with previous literature on behavioural case linkage and is justified by findings suggesting that more complex coding schemes are unreliable with police data (Canter & Heritage, 1990). Satisfactory inter-rater reliability has been reported for the larger dataset from which the current data were selected (*Mdn* case-by-case $\kappa = 0.78$ and *Mdn* variable-by-variable $\kappa = 0.88$; Santtila, Ritvanen et al., 2004).

Procedure

The offence behaviours were first grouped into behavioural domains that contained behaviours that either served a similar function during the offence (e.g. they facilitated entry into the property), or that occurred at a similar stage of the offence (e.g. they occurred at the start of the offence when a burglar was selecting the target), or that

represented one ‘type’ of offender behaviour (e.g. spatial behaviour) (see the Appendix for a full listing of which behaviours comprised each domain). Seven behavioural domains were created: 1) Target Characteristics (e.g. the type of property burgled); 2) Entry Behaviours (e.g. the point and method of entry); 3) Internal Behaviours (e.g. search behaviour); 4) Property Stolen (e.g. cash, keys etc.); 5) The Intercrime Distance; 6) Temporal Proximity; 7) A combined behavioural domain, which included all behaviours in the target, entry, internal, and property domains. These domains were derived from previous case linkage studies on burglary and the behaviours were placed into domains according to their placement in previous research (Bennell, 2002; Bennell & Canter, 2002; Markson et al., 2010).

Pairs of crimes were then created from the two burglary datasets. Initially, 117 linked pairs of crimes were created (one for each offender) from dataset one. Each pair contained two crimes committed by the same offender that were taken randomly from each offender’s series. One-hundred-and-seventeen pairs of unlinked crimes were then created from dataset one, which contained two crimes committed by different offenders. Finally, a further set of 117 unlinked pairs were created from dataset two.

Having created these crime pairs, each group of pairs (linked dataset one; unlinked dataset one; and unlinked dataset two) was split into two halves to form a development sample and a test sample. So, 58 of the 117 linked pairs and 58 of the 117 unlinked pairs from dataset one were used to create a development sample and the remaining 59 linked and 59 unlinked pairs from dataset one were used to create a test sample. These development and test samples were used to examine discrimination accuracy using Bennell’s (2002) original methodology. Also, 58 of the 117 unlinked pairs

from dataset two were used to form a further development sample and the remaining 59 pairs formed a further test sample. These development and test samples were used in combination with the samples formed from the linked pairs in dataset one to test discrimination accuracy using a new methodology that addresses statistical and practical issues with Bennell's (2002) existing approach. It should be noted that the larger number of pairs in the test samples is due to there being an uneven number of offenders.

The procedure of splitting data into development and test samples is known as split-half validation (Efron, 1982; Gong, 1986), and is discussed by Bennell and colleagues as a way of reducing the potential bias that might arise from developing and testing linkage models on the same sample (e.g. Bennell & Jones, 2005).

Data Analysis

The first stage of analysis was to calculate the degree of behavioural similarity between the linked and unlinked crime pairs. To achieve this, a measure called Jaccard's coefficient was used, which ranges from 0 (indicating no behavioural similarity) to 1 (indicating complete behavioural similarity). Jaccard's coefficient has been used extensively in previous studies of behavioural case linkage (e.g. Bennell & Canter, 2002; Markson et al., 2010). Jaccard's coefficients were calculated for each crime pair in terms of target, entry, internal, and property behaviours separately, as well as for the combination of these behaviours. In addition to this, the kilometre-distance and number of days between the two crimes in each pair were calculated.

The potential value of these seven measures of offender behaviour for distinguishing between linked and unlinked crimes was then assessed using logistic

regression and ROC analysis (e.g. Bennell, 2002). In order to test the cross-national replicability of case linkage research, seven direct logistic regression analyses were conducted on the development sample from dataset one (one regression for each of the seven linkage features) with linkage status (linked versus unlinked) as the dependent variable and the linkage features as the independent variables. These analyses allowed the case linkage performance of each linkage feature to be judged independently from the others (Woodhams & Toye, 2007). A forward stepwise logistic regression was then conducted, where all linkage features were entered into the model simultaneously, thus allowing the optimal combination of linkage features to be identified (Bennell & Canter, 2002). However, the combined domain was not included in these analyses because this domain was comprised of a combination of behaviours from the target, entry, internal and property domains. Consequently, the inclusion of this variable in the same regression model as the other domains would risk violating the assumption of multicollinearity, which can lead to reduced p -values, incorrect regression coefficients and, ultimately, to incorrect conclusions (Field, 2005). Furthermore, the decision to exclude the combined domain was consistent with previous research (e.g. Bennell, 2002), which is important given that one of the primary aims of the current study was to replicate previous work.

Having developed regression models on the development sample, these same models were used to produce predicted probabilities (ranging from 0 to 1) for each crime pair in the test samples from dataset one (Bennell & Canter, 2002). These predicted probabilities were then used as the test variables and linkage status (linked, unlinked) as the state variable to produce ROC curves for each of the seven single-feature behavioural

domains and for the optimal combination of domains. The Statistical Package for the Social Sciences, version 17.0, was used to produce these ROC curves.

The associated ROC statistic is the Area Under the Curve (AUC), which gives a measure of discrimination accuracy for the linkage feature/s used to construct the curve (i.e. how well the offender behaviour discriminated between linked and unlinked crimes). The AUC can range from 0 (indicating perfect inaccuracy) to 1 (indicating perfect accuracy), with a value of 0.5 indicating a chance level of accuracy (Bennell & Jones, 2005). AUC values of 0.50-0.70 are described as a low level of discriminative accuracy, values of 0.70-0.90 are moderate, and values of 0.90-1.00 are high (Swets, 1988). The regression statistics and AUC values were compared visually with those obtained in previous UK-based burglary studies to allow the cross-national replicability of case linkage findings to be examined.

To explore the impact of using a statistically independent sample of serial and non-serial burglaries to form the unlinked pairs, the same regression and ROC analyses were run using the linked pairs from dataset one and the unlinked pairs formed from dataset two. The regression statistics obtained from the first set of analyses were then compared with these analyses visually and the AUC statistics compared statistically using ROCKIT 1.1B2 (© University of Chicago). SPSS version 17.0 was used to produce the ROC curves, but it is currently not possible to compare ROC curves statistically using SPSS, so ROCKIT was used for this purpose. Any differences would suggest that the choice of methodology impacts on case linkage findings.

Results

A Cross-National Replication of UK Burglary Case Linkage Research

The results from seven direct logistic regression analyses using linked and unlinked pairs from dataset one are summarised in Table 1. A degree of success was evident for all seven models of offence behaviour, although some models clearly outperformed others. The most successful single-feature models were intercrime distance, followed by the combined domain, then temporal proximity. These models all had highly significant model χ^2 values and Wald statistics ($p < .001$), with between 24% and 57% of the variability in linkage status explained individually by each of these three behavioural domains (Brace, Kemp, & Snelgar, 2003; Kinnear & Gray, 2000). Furthermore, all three models offered an improvement in predictive accuracy above the level one would expect through chance, with each model offering an approximate 23% to 30% improvement (see Table 2). In contrast to these models, however, the target, entry, and internal models performed less favourably. These models explained between 11% and 25% of the variability in linkage status and offered an approximate 15% improvement in predictive accuracy above chance. The poorest performance was for the property domain, with just 5% to 6% of the variability accounted for by this model and a 6% improvement in predictive accuracy.

The signs of the logit coefficients indicated that linked crimes were characterised by greater behavioural similarity in terms of combined, target, entry, internal, and property behaviours and shorter intercrime distance and temporal proximity values than unlinked crimes.

To determine whether these individual domains could be combined to produce superior discriminative performance, a forward stepwise logistic regression was conducted. The stepwise regression proceeded through three steps before it converged on a final model. The final model (referred to as optimal 1 in Tables 1 and 2 below) contained three domains (intercrime distance, temporal proximity, and target characteristics), which accounted for between 56% and 75% of the variance in linkage status and facilitated an improvement in predictive accuracy of almost 30% above chance (see Tables 1 and 2). These results indicate that optimal model 1 was superior to any of the single-feature regression models in terms of discriminative performance.

It should be noted that a model combining the intercrime distance and temporal proximity was able to perform at a similar level to optimal 1 in terms of predictive accuracy and the percentage of variance explained (see Optimal 2 in Tables 1 and 2 below).

--- INSERT TABLES 1 AND 2 HERE ---

To facilitate further comparisons, nine empirical ROC curves were produced (one for each of the seven single-feature regression models and two for the optimal regression models). The results are summarised in Table 3.

The ROC results are largely consistent with those obtained from the logistic regression analyses, with the combined, intercrime distance and temporal proximity domains performing well in comparison to the entry, internal and property domains. Also, the optimal models performed marginally better than any of the single-feature

models. The fact that similar findings emerged from both the development and test samples suggests that the current findings are robust and potentially have wider applicability to other burglary crimes committed in this region of Finland.

However, there were several areas of slight inconsistency between the regression and ROC findings that are worthy of comment. First, the target domain performed slightly better in the ROC analyses than would have been expected from the regression analyses. This suggests that it might be premature to discount the potential independent value of target characteristics in the linkage task. The second area of slight inconsistency was the equivalent performance of the two optimal models in terms of the AUC. Based on the stepwise regression results one would predict that performance would be superior in optimal model 1. But, it is important to remember that the difference between these two optimal models was small in the regression analyses, so it is perhaps unsurprising that they achieved a similar level of discrimination accuracy in the ROC analyses.

--- INSERT TABLE 3 HERE ---

The Impact of Methodological Variation in Case Linkage Research

The findings from seven direct logistic regression analyses using linked pairs from dataset one and unlinked pairs from dataset two are summarised in Tables 4 and 5. When these findings are compared with those obtained previously, we see that there is a trend towards reduced discrimination accuracy in the current set of analyses, with less substantial model χ^2 , Wald and R^2 statistics for all domains except the target domain. But, the magnitude of these differences is small. Indeed, when the predictive accuracies from

these two analyses are compared (Tables 2 and 5) none of the domains differ by more than 5.20%.

A forward stepwise logistic regression was conducted to facilitate further comparisons. The stepwise regression proceeded through the same previous three steps before converging on a final model, which contained the same domains (target, intercrime distance, and temporal proximity). The only slight difference was in terms of the performance of the optimal models, whereby a reduced performance was observed in the current set of analyses (as indicated by the model χ^2 , Wald and R^2 statistics). However, it should be noted that the predictive accuracies in Tables 2 and 5 indicate an improved rather than a reduced performance. The reason for this contradiction is probably due to the way in which these measures of model performance are calculated (Field, 2005).

--- INSERT TABLES 4 AND 5 HERE ---

The analyses thus far indicate that minor differences exist as a function of how the unlinked pairs are formed. To further examine this issue, nine empirical ROC curves were created as before. The results are summarised in Table 6 and compared statistically with the previous ROC results in Table 7. There were no statistically significant differences in terms of the AUC statistics produced in the two sets of analysis.

--- INSERT TABLES 6 AND 7 HERE ---

Discussion

In the current study the cross-national replicability of case linkage findings relating to residential burglary was examined using a sample of burglaries committed in Finland. There was evidence to suggest that a range of offender behaviours can be used to distinguish between linked and unlinked crimes. The most successful behaviours were the intercrime distance, temporal proximity, and the combined domain, which is somewhat consistent with previous research that has shown the value of the intercrime distance and temporal proximity for linking burglary crimes committed in the UK (e.g. Bennell, 2002; Bennell & Canter, 2002; Ewart et al., 2005; Goodwill & Alison, 2006; Markson et al., 2010).

However, the magnitude of discrimination accuracy in the current study was larger for the combined, target, entry, and internal domains than in previous UK-based work. Most notably, the combined and target domains both achieved AUC values in excess of 0.70, which indicates a moderate degree of discrimination accuracy (Swets, 1988). In previous work the AUC values obtained in over ten UK police jurisdictions have never exceeded 0.69 for these domains (Mean combined AUC = 0.65; Mean target AUC = 0.60; Bennell, 2002; Bennell & Canter, 2002; Bennell & Jones, 2005; Markson et al., 2010). Likewise, the entry and internal domains (AUCs = 0.66) compare favourably to previous research on UK data (Mean Entry AUC = 0.58; Mean Internal AUC = 0.51). These findings suggest that a wider range of offender behaviours in Finland than the UK demonstrate the relative consistency and distinctiveness required to facilitate successful behavioural case linkage.

There are several potential explanations for these differences. First, it is possible that Finnish burglars are more consistent and distinctive in their offence behaviour than burglars from the UK. This might be due to individual differences (such as the presence of particularly rigid and unique behavioural scripts for offending) and/or due to environmental differences (such as the availability and diversity of potential targets with which to offend against). In terms of the former, there is no theoretical basis to suggest that Finnish and English burglars possess different characteristics that would be expected to impact on consistency and distinctiveness. In terms of the latter, one potential environmental factor was discussed earlier that might partially account for the observed differences. There is a wider variety of housing in Finland than in the UK, which could allow for between-offender differences in burglary behaviour to emerge more readily among Finnish than English offenders. A comparison between the current data and those from Markson et al. (2010) support this suggestion. In the current sample, 65% of the crimes in dataset one targeted detached housing or second floor apartments (two separate categories) and the remaining 35% were split across the other four categories of housing. This compares with 84% of crimes in Markson et al.'s sample that fell under one category of housing. The wider variation in types of housing targeted by Finnish burglars might partially account for the superior discrimination accuracy observed in the current study for target characteristics.

Another explanation is that the UK and Finnish police may differ in terms of their data recording and storage practices. A comparison between the data available for the current and previous studies suggests that there may be some value in this explanation. The internal domain in this study, for example, included the number of offenders

responsible for the crime and how the offender/s exited the crime scene. These variables were not included in previous research on UK data (Bennell, 2002). It is plausible that additional behaviours such as these led to the improved case linkage performance observed in the current study. It is also plausible that the behaviours were operationalised in a more appropriate way in the current study compared with past research. For example, the target domain in this study included several variables related to the owner's occupancy, whereas in previous research occupancy has been defined simply in terms of one variable. Differences such as these may also have contributed to the improved discrimination accuracy in the current study. If these explanations are valid, then it suggests that the UK police may be able to enhance case linkage performance by altering the types and nature of information that is recorded on police databases.

However, it is difficult to tease out and test these potential explanations using the current set of data, which means that it is not possible to draw any definitive conclusions.

Despite the difficulty in definitively explaining these findings, the potential implications are clear. From a practical perspective, they suggest that the Finnish police can use a range of offender behaviours to identify linked residential burglary crime series. But, it seems that they should prioritise the use of the intercrime distance and temporal proximity in this process because these two features offer the greatest level of predictive accuracy with a minimum of cost in terms of the time necessary to use these features in practice. However, in the absence of data relating to these features, it seems that there is scope for the Finnish police to rely on target, entry and internal behaviours to link burglaries.

It is worth highlighting at this point, though, that the AUC values obtained in this study were not perfect (1.00), so a degree of error can be expected when linking burglary crimes in practice using these behaviours. Indeed, given the low base rate of crimes that are linked in real life, high AUC values may not necessarily translate into error-free prediction (Bennell & Jones, 2005; Szmukler, 2001). This is an important point for practitioners who are involved in linking crimes to consider.

From a theoretical point of view, the consistency and distinctiveness of intercrime distance in this study provides support for several seminal theories of offender spatial behaviour, which suggest that offenders seek to minimise the efforts and risks involved in offending (e.g. by returning to the same places that are familiar to them and by not travelling great distances to offend) (Brantingham & Brantingham, 1981, 1984; Clarke & Felson, 1993; Cohen & Felson, 1979; Cornish & Clarke, 1986; Felson, 1986, 1994).

Furthermore, these findings support previous case linkage research in the UK that has shown burglary, robbery and car theft offenders tend to offend in somewhat distinct, non-overlapping geographical areas (e.g. Bennell & Canter, 2002; Bennell & Jones, 2005; Tonkin et al., 2008; Woodhams & Toye, 2007). That is, the geographical locations that one offender chooses to offend in are somewhat different from the areas that a different offender may choose. A potential explanation for this finding comes from previous research using this dataset (Laukkanen et al., 2008), which has shown that Finnish burglars do not travel far from home to offend (a median of 3.88KM). Thus, it may be that the current sample chose to offend close to home and, by virtue of the fact that the offenders live in different areas, somewhat distinct, non-overlapping patterns of offender spatial behaviour emerged.

However, as noted above, the AUC values observed in the current study were not perfect (1.00), so the offending “territories” of the burglars in this sample were not completely non-overlapping. Furthermore, the geographical area studied here was relatively large (815KM²), so it is unclear whether these findings would be replicated at a smaller geographical scale. But, we might be cautiously optimistic given that similar findings have been observed with burglars in the UK using study areas that are much smaller in size (112-230KM²; Bennell & Jones, 2005).

Another main aim of the current study was to explore the impact of using different methodological approaches to forming the unlinked crime pairs. The regression and ROC analyses revealed minor differences in discriminative accuracy (none of which reached statistical significance) when the ‘traditional’ methodology (Bennell, 2002) was compared with a new methodology that offers a potentially more realistic and statistically sound approach to testing case linkage. This is reassuring because it suggests that previous research is valid and that the current set of findings is robust across different methodological approaches.

Based on these findings and the significant amount of additional work that is required to utilise an independent sample of crimes, it might be argued that researchers can continue with Bennell’s (2002) original methodology at present. However, it is important to conduct further comparisons between the two methodologies using different datasets.

In terms of limitations, the current study shares many of the limitations associated with previous case linkage research, including the fact that these findings are only applicable to the geographical and temporal period studied. Although the use of split-half

validation suggests that these findings may be applicable to other similar areas in Finland. Another limitation is that the sample was restricted to solved crime. Researchers have discussed the fact that this may artificially inflate discrimination accuracy and that this approach moves research away from the reality in which case linkage is expected to perform (i.e. with unsolved crimes) (Bennell, 2002). Future research should address this issue by testing case linkage with samples of unsolved crime. Nonetheless, the current study has contributed to the growing body of case linkage literature by extending the evidence to a new country and by exploring new methodological issues in a systematic way.

The challenge now, however, is to test findings such as these in a prospective way by applying them in practice to ongoing police investigations to determine whether they can truly yield improved behavioural case linkage performance. Indeed, this would seem particularly pertinent given recent evidence that statistically significant AUC statistics do not necessarily translate into substantial practical accuracy in discriminative tasks (Szmukler, 2001).

References

- Bennell, C. (2002). *Behavioural consistency and discrimination in serial burglary* (Unpublished doctoral dissertation). University of Liverpool, Liverpool, UK.
- 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., Gauthier, D., Gauthier, D., Melnyk, T., & Musolino, E. (2010). The impact of data degradation and sample size on the performance of two similarity coefficients used in behavioural linkage analysis. *Forensic Science International, 199*, 85-92. doi: 10.1016/j.forsciint.2010.03.017
- 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

- Brace, N., Kemp, R., & Snelgar, R. (2003). *SPSS for psychologists: A guide to data analysis using SPSS for Windows* (2nd ed.). Basingstoke: Macmillan.
- Brantingham, P. J., & Brantingham, P. L. (1981). *Environmental criminology*. Beverly Hills: Sage.
- Brantingham, P. J., & Brantingham, P. L. (1984). *Patterns in crime*. New York: Macmillan.
- Canter, D., & Heritage, R. (1990). A multivariate model of sexual offences behaviour: Developments in 'offender profiling'. *Journal of Forensic Psychiatry, 1*, 185-212. doi: 10.1080/09585189008408469
- Clarke, R. V., & Felson, M. (1993). *Routine activity and rational choice*. New Brunswick, NJ: Transaction.
- Cohen, L. E., & Felson, M. (1979). Social change and crime-rate trends: A routine activity approach. *American Sociological Review, 44*, 588-608. url: <http://www.jstor.org/stable/2094589>
- Cornish, D. B., & Clarke, R. V. (1986). *The reasoning criminal: Rational choice perspectives on offending*. New York: Springer-Verlag.

- Efron, B. (1982). *The jackknife, the bootstrap and other re-sampling plans*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Ewart, B. W., Oatley, G. C., & Burn, K. (2005). Matching crimes using burglars' modus operandi: A test of three models. *International Journal of Police Science and Management*, 7, 160-174. Retrieved from:
<http://www.vathek.com/ijpsm/contents.php?vi=7.3>
- Felson, M. (1986). Linking criminals' choices, routine activities, informal control, and criminal outcomes. In D. Cornish & R. V. Clarke (Eds.), *The reasoning criminal: Rational choice perspectives on offending* (pp. 119-128). New York: Springer-Verlag.
- Felson, M. (1994). *Crime and everyday life: Insights and implications for society*. Thousand Oaks, CA: Pine Forge Press.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
- Gong, G. (1986). Cross-validation, the jackknife, and the bootstrap: Excess error estimation in forward logistic regression. *Journal of the American Statistical Association*, 81, 108-113.
url: <http://www.jstor.org/stable/2287975>

- Goodwill, A. M., & Alison, L. J. (2006). The development of a filter model for prioritizing suspects in burglary offences. *Psychology, Crime & Law, 12*, 395-416. doi: 10.1080/10683160500056945
- Grubin, D., Kelly, P., & Brunson, C. (2001). *Linking serious sexual assaults through behaviour* (Home Office Research Study 215). London: Home Office Research, Development and Statistics Directorate.
- Hopkins, W. G. (2001). *A new view of statistics*. Retrieved from: <http://www.sportsci.org/resource/stats/modeldetail.html>
- Kinnear, P. R., & Gray, C. D. (2000). *SPSS for Windows made simple*. Hove: Psychology Press.
- Laukkanen, M., Santtila, P., Jern, P., & Sandnabba, K. (2008). Predicting offender home location in urban burglary series. *Forensic Science International, 176*, 224-235. doi: 10.1016/j.forsciint.2007.09.011
- Markson, L., Woodhams, J., & Bond, J. W. (2010). Linking serial residential burglary: Comparing the utility of *modus operandi* behaviours, geographical proximity, and temporal proximity. *Journal of Investigative Psychology and Offender Profiling, 7*, 91-107. doi: 10.1002/jip.120

- Ratcliffe, J. H. (2002). Aoristic signatures and the spatio-temporal analysis of high volume crime patterns. *Journal of Quantitative Criminology*, 18, 23-43. doi: 10.1023/A:1013240828824
- Salfati, C. G., & Bateman, A. L. (2005). Serial homicide: An investigation of behavioural consistency. *Journal of Investigative Psychology and Offender Profiling*, 2, 121-144. doi: 10.1002/jip.27
- 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., Ritvanen, A., & Mokros, A. (2004). Predicting burglar characteristics from crime scene behaviour. *International Journal of Police Science & Management*, 6, 136-154. Retrieved from: <http://www.vathek.com/ijpsm/contents.php?vi=6.3>
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 1285-1293. Retrieved from: <http://www.ncbi.nlm.nih.gov/pubmed/3287615>
- Szmukler, G. (2001). Violence risk prediction in practice. *The British Journal of Psychiatry*, 178, 84-85. url: <http://bjp.rcpsych.org/cgi/reprint/178/1/84>

- 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
- Woodhams, J. (2008). *Understanding juvenile sexual offending: An investigative perspective* (Unpublished doctoral dissertation). University of Leicester, Leicester, UK.
- Woodhams, J., Grant, T. D., & Price, A. R. G. (2007). From marine ecology to crime analysis: Improving the detection of serial sexual offences using a taxonomic similarity measure. *Journal of Investigative Psychology and Offender Profiling*, *4*, 17-27. doi: 10.1002/jip.55
- 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
- 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

Appendix- Content Dictionary of Offence Behaviours and Behavioural Domains**Target Characteristics**

A detached house

1st floor of a multi-storey building

2nd floor or above of a multi-storey building

A studio flat

A terraced or semi-detached house

Other target

Owners present

Owners temporarily away (i.e. less than 24 hours)

Owners away 1-3 days

Owners away 3+ days

Target has safeguards present (alarm, security light, dog etc.)

Target in an urban city area

Entry Behaviours

Door

Back door

Balcony door

Window

Mailbox

Open or unlocked door

Manual force

Breaking glass

Climbing (above street level)

Lock

Key

Tool

Crowbar

Hook

Sharp weapon

Garden tool

Screwdriver or spike

Brick or stone

Tool brought to the scene

Tool used from the scene or the immediate vicinity

Internal Behaviours

Interrupted

Interrupted by a guard or the owner

Fingerprints, footprints, or DNA left at the scene

Tools used in the burglary left at the scene

1 offender

2 offenders

3+ offenders

Tidy search

Untidy search

Only first room entered was searched

Whole target searched

Drawers/cabinets opened and searched

Drawers pulled out and contents possibly thrown on floor

Inner doors opened using force

Property piled up to be carried away

Stolen items hidden close by

Stolen items abandoned

Used facilities (consumed food/drink, used toilet/shower, defecated/urinated)

Exit by car

Exit on foot

Property Stolen

Cash

Credit or bank cards, cheques, bank book, shares

Firearms, ammunition, explosives

Sharp weapons (not cutlery)

Watches, wristwatches

Small-size consumer electrical items

Large electrical equipment, musical instruments that need to be carried with both hands

Tapes, CD's LP's, videotapes

Jewellery

Fake jewellery (costume)

Prescription medication

Tobacco products/smoking tools

Cosmetic, hygiene products

Alcohol

Plates, cups, cutlery and other utensils

Food

Clothes

Purses, hand bags, suitcases, backpacks

Wallet

Keys (home, car)

Identity documents (e.g. passport, driving licence, library card etc.)

Spectacles, sunglasses or other optical items

Antique or art objects

Construction tools or materials

Porcelain, crystal glass, silverware

Games or sports equipment

Vehicle

Items stolen could be carried by one person

Stolen items worth less than 170 Euros

Stolen items worth up to 1700 Euros

Stolen items worth more than 8400 Euros

Combined

Contains all behaviours listed above under target characteristics, entry behaviours, internal behaviours, and property stolen.

Intercrime Distance (in kilometres)

The distance in kilometres between two crime sites.

Temporal Proximity (in days)

The number of days between two crime dates.

Tables

Table 1

*Nine Logistic Regression Models for a Sample of Finnish Burglars: Bennell's (2002)**Methodology*

Model	Constant (SE)	Logit (SE)	χ^2 (df)	Wald (df)	R ² (Cox & Snell- Nagelkerke)
Combined	-2.98 (0.65)	9.57 (2.03)	35.68 (1)***	22.13 (1)***	0.27-0.35
Target	-0.94 (0.29)	2.96 (0.71)	23.99 (1)***	17.61 (1)***	0.19-0.25
Entry	-1.17 (0.33)	3.14 (0.74)	22.37 (1)***	17.80 (1)***	0.18-0.23
Internal	-1.14 (0.39)	2.86 (0.87)	12.97 (1)***	10.93 (1)**	0.11-0.14
Property	-0.76 (0.39)	2.89 (1.27)	5.48 (1)*	5.14 (1)*	0.05-0.06
Intercrime Distance	2.19 (0.51)	-0.32 (0.06)	54.77 (1)***	25.95 (1)***	0.43-0.57
Temporal Proximity	1.01 (0.28)	-0.00 (0.00)	31.96 (1)***	20.59 (1)***	0.24-0.32
Optimal 1	1.80 (0.69)		80.41 (3)***		0.56-0.75
Intercrime		-0.24 (0.07)		12.68 (1)***	
Temporal		-0.00 (0.00)		9.31 (1)**	
Target		4.65 (1.60)		8.44 (1)**	
Optimal 2	2.77 (0.60)		67.62 (2)***		0.50-0.67
Intercrime		-0.26 (0.06)		16.51 (1)***	
Temporal		-0.00 (0.00)		8.39 (1)**	

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 2

Predictive Accuracy of the Models (%): Bennell's (2002) Methodology

	Combined	Target	Entry	Internal	Property	Intercrime Distance	Temporal Proximity	Optimal 1	Optimal 2
Random	50.00	50.00	50.00	50.00	50.00	58.60	50.00	58.60	58.60
Model	75.90	65.50	66.40	65.50	56.00	79.80	73.30	86.90	85.90

Table 3

Summary of the Receiver Operating Characteristic (ROC) Analyses: Bennell's (2002)

Methodology

Model	AUC (SE)	95% Confidence Interval	Classification Category
Combined	0.72 (0.05)***	0.63-0.81	Moderate
Target	0.73 (0.05)***	0.64-0.82	Moderate
Entry	0.66 (0.05)**	0.56-0.76	Low
Internal	0.66 (0.05)**	0.56-0.76	Low
Property	0.58 (0.05)	0.48-0.69	Low
Intercrime Distance	0.84 (0.04)***	0.75-0.93	Moderate
Temporal Proximity	0.82 (0.04)***	0.74-0.90	Moderate
Optimal 1	0.86 (0.04)***	0.78-0.93	Moderate
Optimal 2	0.86 (0.04)***	0.79-0.94	Moderate

Note. AUC = Area Under the Curve

Classification categories are according to Swets (1988), where an AUC value of 0.50 is non-informative, a value of 0.50-0.70 is low, 0.70-0.90 is moderate, and 0.90-1.00 is high

** $p < .01$; *** $p < .001$

Table 4

Nine Logistic Regression Models for a Sample of Finnish Burglars: New Methodology

Model	Constant (SE)	Logit (SE)	χ^2 (df)	Wald (df)	R^2 (Cox & Snell- Nagelkerke)
Combined	-2.52 (0.60)	7.80 (1.77)	27.88 (1)***	19.43 (1)***	0.21-0.29
Target	-0.98 (0.29)	3.18 (0.73)	26.79 (1)***	19.02 (1)***	0.21-0.28
Entry	-1.02 (0.33)	2.60 (0.69)	16.83 (1)***	14.33 (1)***	0.14-0.18
Internal	-1.09 (0.39)	2.69 (0.85)	11.55 (1)**	9.91 (1)**	0.10-0.13
Property	-0.55 (0.38)	1.99 (1.22)	2.73 (1)	2.64 (1)	0.02-0.03
Intercrime Distance	1.73 (0.45)	-0.25 (0.06)	36.23 (1)***	19.16 (1)***	0.34-0.46
Temporal Proximity	0.93 (0.27)	-0.00 (0.00)	28.97 (1)***	17.71 (1)***	0.22-0.30
Optimal 1	2.13		68.02 (3)***		0.55-0.73
Intercrime	(0.75)	-0.24 (0.07)		14.00 (1)***	
Temporal		-0.00 (0.00)		9.08 (1)**	
Target		3.74 (1.36)		7.57 (1)**	

Optimal 2	2.93		58.16 (2)***		0.49-0.66
Intercrime	(0.65)	-0.23 (0.06)		14.56 (1)***	
Temporal		-0.00 (0.00)		11.16 (1)**	

** $p < .01$; *** $p < .001$

Table 5

Predictive Accuracy of the Models (%): New Methodology

	Combined	Target	Entry	Internal	Property	Intercrime Distance	Temporal Proximity	Optimal 1	Optimal 2
Random	50.00	50.00	50.00	50.00	50.00	52.30	50.00	52.30	52.30
Model	70.70	67.20	66.40	61.20	56.90	76.70	74.10	87.20	87.20

Table 6

Summary of the Receiver Operating Characteristic (ROC) Analyses: New Methodology

Model	AUC ^a (SE)	95% Confidence Interval	Classification Category ^b
Combined	0.73 (0.05)***	0.64-0.82	Moderate
Target	0.71 (0.05)***	0.61-0.80	Moderate
Entry	0.66 (0.05)**	0.56-0.76	Low
Internal	0.72 (0.05)***	0.63-0.81	Moderate
Property	0.55 (0.05)	0.44-0.66	Low
Intercrime Distance	0.85 (0.04)***	0.76-0.93	Moderate
Temporal Proximity	0.82 (0.04)***	0.74-0.90	Moderate
Optimal 1	0.88 (0.04)***	0.81-0.95	Moderate
Optimal 2	0.89 (0.03)***	0.82-0.96	Moderate

Note. AUC = Area Under the Curve

Classification categories are according to Swets (1988), where an AUC value of 0.50 is non-informative, a value of 0.50-0.70 is low, 0.70-0.90 is moderate, and 0.90-1.00 is high

** $p < .01$; *** $p < .001$

Table 7

A Comparison of the Area Under the Curve (AUC) Statistics Produced Using Bennell's (2002) Methodology and a New Methodology

Model	Bennell Approach AUC	New Approach AUC
Combined	0.72	0.73
Target	0.73	0.71
Entry	0.66	0.66
Internal	0.66	0.72
Property	0.58	0.55
Intercrime Distance	0.84	0.85
Temporal Proximity	0.82	0.82
Optimal 1	0.86	0.88
Optimal 2	0.86	0.89

Note. All statistical comparisons were non-significant ($p > .05$).