Linking serial sexual offences

Woodhams, Jessica; Tonkin, Matthew; Imre, Hanne; Winter, Jan M.; Lam, Eva; ten Brinke, Gert-Jan; Webb, Mark; Labuschagne, Gerard; Bennell, Craig; Ashmore-Hills, Leah; van der Kemp, Jasper; Lipponen, Sami; Pakkanen, Tom; Rainbow, Lee; Salfati, C. Gabrielle; Santtila, Pekka; Burrell, Amy

DOI:
10.1111/lcrp.12144

Legal and Criminological Psychology (2018)

The British Psychological Society

www.wileyonlinelibrary.com

Linking serial sexual offences: Moving towards an ecologically valid test of the principles of crime linkage

Jessica Woodhams[1]*, Matthew Tonkin[2], Amy Burrell[3], Hanne Imre[4], Jan M. Winter[5,6], Eva K. M. Lam[5], Gert Jan ten Brinke[5], Mark Webb[7], Gerard Labuschagne[8], Craig Bennell[9], Leah Ashmore-Hills[10], Jasper van der Kemp[11], Sami Lipponen[12], Tom Pakkanen[13], Lee Rainbow[7], C. Gabrielle Salfati[14] and Pekka Santtila[15]

[1]University of Birmingham, UK
[2]University of Leicester, UK
[3]Coventry University, UK
[4]Belgian Federal Police, Brussels, Belgium
[5]Dutch National Police, The Hague, The Netherlands
[6]Department of Clinical and Lifespan Psychology (KLEP), Vrije Universiteit Brussel, Belgium
[7]National Crime Agency, London, UK
[8]L&S Threat Management, Pretoria, South Africa
[9]Carleton University, Ottawa, Ontario, Canada
[10]Birmingham City University, UK
[11]Vrije Universiteit Amsterdam, The Netherlands
[12]National Bureau of Investigation, Vantaa, Finland
[13]Åbo Akademi University, Turku, Finland
[14]John Jay College of Criminal Justice, New York City, New York, USA
[15]New York University, Shanghai, China

Purpose. To conduct a test of the principles underpinning crime linkage (behavioural consistency and distinctiveness) with a sample more closely reflecting the volume and nature of sexual crimes with which practitioners work, and to assess whether solved series are characterized by greater behavioural similarity than unsolved series.

Method. A sample of 3,364 sexual crimes (including 668 series) was collated from five countries. For the first time, the sample included solved *and* unsolved but linked-by-DNA sexual offence series, *as well as* solved one-off offences. All possible crime pairings in the

**Correspondence should be addressed to Jessica Woodhams, School of Psychology, University of Birmingham, Edgbaston B15 2TT, UK (email: j.woodhams@bham.ac.uk).*

DOI:10.1111/lcrp.12144

data set were created, and the degree of similarity in crime scene behaviour shared by the crimes in each pair was quantified using Jaccard's coefficient. The ability to distinguish same-offender and different-offender pairs using similarity in crime scene behaviour was assessed using Receiver Operating Characteristic analysis. The relative amount of behavioural similarity and distinctiveness seen in solved and unsolved crime pairs was assessed.

Results. An Area Under the Curve of .86 was found, which represents an excellent level of discrimination accuracy. This decreased to .85 when using a data set that contained one-off offences, and both one-off offences and unsolved crime series. Discrimination accuracy also decreased when using a sample composed solely of unsolved but linked-by-DNA series (AUC = .79).

Conclusions. Crime linkage is practised by police forces globally, and its use in legal proceedings requires demonstration that its underlying principles are reliable. Support was found for its two underpinning principles with a more ecologically valid sample.

Crime linkage[1] refers to a group of practices where the crime scene behaviour displayed in multiple crimes is analysed for similarity and distinctiveness to assess the likelihood of those crimes being committed by the same offender. Where similar yet distinctive behaviour is observed, greater confidence is attributed to the crimes being the work of the same perpetrator (Woodhams, Bull, & Hollin, 2007). The underlying principles of crime linkage are therefore that offenders will show a degree of consistency in their crime scene behaviour over time (the Consistency Hypothesis; Canter, 1995) and that offenders will show a degree of distinctiveness in their crime scene behaviour (Bennell & Canter, 2002), allowing the crimes of one offender to be distinguished from those of another offender committing a similar sort of crime.[2]

In many countries, police units exist that specialize in this behavioural analysis for the most serious forms of crime (e.g., sexual offences and homicides) (Bennell, Snook, MacDonald, House, & Taylor, 2012). This analysis informs police investigations and can have several benefits such as identifying crime series where physical trace evidence is lacking or is costly or time-consuming to process, pooling evidence from multiple crime scenes, and enhancing victim credibility (Davies, 1991; Grubin, Kelly, & Brunsdon, 2001; Labuschagne, 2015). However, errors in linkage prediction can misdirect investigative efforts and unnecessarily increase public fear of a serial offender being active in the area (Grubin *et al.*, 2001).

Crime linkage analysis can also inform legal decision-making and has been admitted as similar fact evidence for robbery, burglary, homicide, kidnapping, and rape prosecutions in *State v. Mogale* (2012), *State v. Nyauza* (2007), *State v. Steyn* (2012), *State v. Sukude* (2006), and *State v. van Rooyen* (2007) in South Africa, in *R v. R.B.* (2003) and *R. v. Burlingham* (1993) in Canada (Labuschagne, 2015), and in *Pennell v. State* (1991), *State v. Russell* (1994), *People v. Prince* (2007) and *State v. Yates* (2007) in the United States (Pakkanen, Santtila, & Bosco, 2015). However, such evidence has also been ruled inadmissible in some cases due to concerns about the reliability of its underlying principles and the methods used (*Her Majesty's Advocate v. Young*, 2013; *State of New*

[1] *Crime linkage is also referred to as linkage analysis (Hazelwood & Warren, 2004), case linkage (Woodhams & Grant, 2006), and comparative case analysis (Bennell & Canter, 2002).*

[2] *The assumption of consistency is operationalized in practice and research as an evaluation of the similarity in crime scene behaviour between two or more crimes. Consistency is used in this paper when referring to the behaviour displayed by the same individual over time/events, and similarity is used when referring to linked/unlinked crime pairs and predicting linkage status because, in practice, an analyst would not know for certain whether a set of crimes were committed by the same person or not.*

Jersey v. Fortin, 2000). Regarding the latter, when making their assessments, these courts have been guided by legal standards for the admissibility of scientific expert evidence including the Daubert criteria (*Daubert v. Merrell Dow Pharmaceuticals Inc.,* 1993) and the Federal Rules of Evidence (2011, 702). These standards require that the testimony is the product of reliable principles and methods and that there needs to be a known or potential error rate for the practice. In *HMA v. Young* (2013), for example, a *voir dire* admissibility hearing was held to consider the empirical support for crime linkage analysis and its principles; crime linkage analysis evidence was ultimately ruled inadmissible.

Crime linkage can, therefore, have a potentially significant impact on police investigations and legal outcomes (whether prosecutions or appeals). As such, it is important that research seek to test the viability of crime linkage and to test this in the most realistic way possible.

Paradigms for assessing the principles of crime linkage

The basic tenet of studies of the crime linkage principles is to assess the accuracy with which quantitative measures of similarity in crime scene behaviour (i.e., similarity coefficients) can be used to predict whether two or more crimes are linked.[3] The similarity coefficients are usually calculated from binary codings of offender crime scene behaviour (e.g., Did the offender kiss the victim? – Yes/No?). These codings can be pre-existing, having been completed by trained police staff as part of their routine practice (see Method), or the coding is completed by researchers based on police files documenting each offence. These data are then subject to statistical analysis.

There are two common analytical approaches used: The first rank orders crimes, offenders, or series in order of similarity in behaviour to the ‘query’ crime and assesses the accuracy of prediction though comparison to actual series membership and compares this level of accuracy to what would be expected by chance alone (e.g., Santtila, Junkkila, & Sandnabba, 2005). The second approach (e.g., Bennell & Canter, 2002) assesses the degree of behavioural similarity shared by a given pair in the data set and determines, based on whether this is high or low, whether the pair was likely committed by the same offender (linked), or whether the two crimes in the pair are by two different offenders (unlinked), respectively. Both of these approaches simultaneously assess the two principles of crime linkage – behavioural consistency and distinctiveness.

Receiver Operating Characteristic (ROC) analysis is the preferred measure of predictive accuracy in forensic psychology (Harris & Rice, 1995) and is commonly used to quantify the accuracy with which behavioural similarity can be used to predict series membership or linkage status (linked/unlinked) (see Bennell, Mugford, Ellingwood, & Woodhams, 2014; Winter *et al.*, 2013). It has four possible outcomes: a hit (where a pair is predicted to have been committed by the same offender and was), a false alarm (where a pair is predicted to have been committed by the same offender but was not), a correct rejection (where the two crimes in a pair are predicted to have been committed by two different offenders and they were), and a miss (where the two crimes in a pair are predicted to have been committed by two different offenders and were committed by the same offender). Predicting which series a given crime belongs to (same series/different series) can be conceptualized in the same way (e.g., a hit would be where a crime is correctly predicted to be a member of a series) (Winter *et al.*, 2013). A ROC analysis plots

[3] *Other quantitative metrics can also be used in linkage predictions (e.g., the inter-crime distance).*

the proportion of hits against the proportion of false alarms at every possible decision threshold (in this case at each predicted probability value) from the most stringent threshold to the most lenient. This produces a ROC curve, and the Area Under the Curve (AUC) represents the predictive accuracy of the decision task. The AUC ranges from 0 to 1, with 0.5 representing chance level accuracy and values closer to 1.0 representing high levels of predictive accuracy.

To assess the predictive or diagnostic accuracy of a method or tool, the outcome being predicted needs to be known (or become known) for the cases to which the method/tool is applied. In the context of crime linkage research, this equates to using a sample where the series membership of crimes is known; for example, offender 1 is known to be responsible for crimes 1, 2, and 3 in the data set. A robust test of the crime linkage principles necessitates confidence in such attributions, and studies have typically used offender conviction and/or scene-to-scene DNA hits as confirmation of series membership. It follows that the conditions under which the principles are tested will never represent the exact conditions under which police analysts conduct crime linkage: Police analysts search for crime series within data sets of series and one-off offences, where series membership is known in some cases but not in others, and where their predictions of series membership may not be confirmed due to a lack of feedback or to investigative efforts not yielding an outcome (Davies, Alrajeh, & Woodhams, 2018). However, the ecological validity of studies designed to test the crime linkage principles can be improved by designing studies that more closely resemble the data searched by analysts.

A critical reflection on studies of the crime linkage principles

More than a decade of research testing the crime linkage principles exists, and the general conclusion from this body of research is that the principles are empirically supported *to an extent* (Bennell *et al.*, 2014): Some serial offenders show sufficient behavioural consistency and distinctiveness for their crimes to be linked; however, some offenders and some series are characterized by inconsistent and/or indistinct behaviour (Slater, Woodhams, & Hamilton-Giachritsis, 2015). However, most of these research studies have sampled series confirmed by conviction. Only sampling series confirmed by conviction does not reflect the data searched by analysts and may artificially inflate the accuracy with which linked crime pairs can be distinguished from unlinked crime pairs, or with which crimes can be attributed to the correct series. This is because convicted series might have been solved and convicted, in part, due to the distinctive and consistent behaviour of the offender (Bennell & Canter, 2002). Improving ecological validity by extending the sampling frame to include unsolved crime series that are linked by DNA allows researchers to establish ground truth without biasing the sample in this way (Woodhams *et al.*, 2007). To date, a handful of studies have adopted this design, but these remain the minority (Pakkanen *et al.*, 2015). Only one study exists with sexual offences: Woodhams and Labuschagne (2012a) included in their sample of 599 linked crime pairs, 19 linked pairs that were unsolved but linked by DNA (representing 3% of the linked pairs). Linked crime pairs could be distinguished from unlinked crime pairs with an AUC of .88, thereby providing empirical support for the crime linkage principles. A larger AUC was found than had been reported in previous studies (e.g., 0.75, Bennell, Jones, & Melnyk, 2009).

Two further studies of the crime linkage principles with sexual offences have improved the ecological validity of their samples by extending their sampling frame to include one-off sexual offences alongside serial offences. Winter *et al.* (2013) sampled 90 serial sexual offences and 129 one-off offences and found AUCs ranging from .80 to .89. Slater *et al.* (2015) found an AUC of .86 with a sample of 144 convicted serial offences and 50 convicted one-off offences.

Despite these improvements in methodological design, the sample sizes of these studies remain small. Indeed, this criticism applies to most studies of the crime linkage principles with sexual offences. Sample sizes range from 43 to 244 offences (Bennell *et al.*, 2009; Santtila *et al.*, 2005; Slater *et al.*, 2015; Winter *et al.*, 2013; Woodhams & Labuschagne, 2012a).[4] This can be contrasted with the volume of sexual crimes searched by police analysts in countries that use the Violent Crime Linkage Analysis System (ViCLAS) (e.g., approximately 8,000 cases are on the ViCLAS database in Belgium and 30,000 in the United Kingdom; Davies *et al.*, 2018).

The current study was therefore designed to test the principles of crime linkage using a research design with improved ecological validity, by, for the first time, utilizing a much larger sample of crimes and sampling convicted and unsolved but linked-by-DNA series, as well as convicted one-off offences. Our research questions were as follows:

(1) Are crimes committed by the same offender ('linked' crime pairs) characterized by greater behavioural similarity than crimes committed by different offenders ('unlinked' crime pairs), which would imply both greater behavioural consistency and greater distinctiveness?
(2) At what level of accuracy could linked crime pairs be differentiated from unlinked crime pairs as assessed by ROC analysis?
(3) Would the inclusion of unsolved series and one-off crimes in the sample reduce the ability to distinguish linked from unlinked crime pairs?

Method

Data

The study utilized police crime data relating to 3,364 sexual offences committed by 3,018 offenders (mean number of crimes per series = 3.25, range = 2–32 crimes). These data were provided by police units from five countries that specialize in crime linkage with sexual offences: (1) the Serious Crime Analysis Section (SCAS, United Kingdom, n = 2,579 offences); (2) the Investigative Psychology Section of the South African Police Service (n = 245 offences); (3) the National Bureau of Investigation, Finnish National Police (n = 123 offences); (4) the Central Unit-Team ViCLAS, Dutch National Police (n = 173 offences); and (5) the Zeden-Analyse-Moeurs unit, Belgian Federal Police (n = 244 offences). Within these data, there were solved serial crimes (n = 2,081) and solved apparent one-off crimes (n = 1,191) that had resulted in a conviction, and unsolved serial crimes that were linked by DNA (n = 92).[5] A breakdown of the data from each country is included in Table 1.

[4] *Yokota, Fujita, Watanabe, Yoshimoto, and Wachi (2007) are the exception having sampled 1,252 offences by 868 offenders.*
[5] *In this study, unsolved crime series consisted of crimes that had been linked via DNA. Thus, while they remain unsolved, we can be confident that the same offender was responsible. Apparent one-off crimes consisted of crimes committed by an offender who only had one recorded conviction for sexual offending at the time of data collection. This does not preclude the possibility that the offenders have committed other sexual offences for which they have not been convicted, but this limitation is unavoidable. No cases in our analyses were offences 'taken into consideration' (TICs). In England and Wales, during sentencing procedures, an offender can admit to other offences to 'wipe the slate clean' and ask that the Court take these into consideration (Sentencing Council, 2012).*

Table 1. Characteristics of the offences contained within the sample per country

Country	Number of series/cases	Number of offenders	Time frame	Length of series	Number of victims, gender, and age (in years) where known	Gender and age (in years) of offenders where known
South Africa	35 series 245 serial cases	Serial = 36[c]	1998–2012	2–32	N = 356 M = 45, F = 285 Age range = 0–68	M = 85 Age range = 17–55 M = 16[b]
Finland	16 series 43 serial cases 80 one-offs	Serial = 17[a] One-off = 85[a]	1983–2001	2–8	N = 124[a] F = 43[b] Age range = 15–62	Age range = 16–49[b] M = 1,549, F = 6 Age range = 12–77
The United Kingdom	534 series 1,579 serial cases 1,000 one-offs	Serial = 1,612 One-off = 1,000	1966–2013	2–20	N = 2643 M = 149, F = 2,486 Age range = 1–94	M = 89, F = 0 Age range = 13–55 M = 124, F = 1
The Netherlands	38 series 123 serial cases 50 one-offs	Serial = 39[a] One-off = 52[a]	1989–2014	2–10	N = 178 M = 5, F = 172 Age range = 4–96	Age range = 15–69
Belgium	45 series 183 serial cases 61 one-offs	Serial = 47[a] One-off = 80	1985–2014	2–12	N = 259 M = 11, F = 247 Age range = 3–84	

Notes. M = male, F = female. Age range for offenders includes the offender's age at each offence, where known, and therefore is based on the number of crimes rather than the number of offenders.

[a]The data provided indicated where there were multiple offenders (or victims) per incident but not the actual number. These will therefore be the minimum number of offenders (or victims) in the subsample.

[b]The data for the one-off offences were not available; therefore, the figures here are solely for the serial sample.

[c]This is the figure for the offenders confirmed to be serial offenders (i.e., convicted or DNA-linked to two or more offences). There were additional, unverified suspects in some offences who had not been identified. This will therefore be the minimum number of offenders in the sample.

Three data sets (the United Kingdom, Belgium, and the Netherlands) were collated from data already stored on the ViCLAS (see Collins, Johnson, Choy, Davidson, & MacKay, 1998). ViCLAS stores records of serious crimes including the crime scene behaviour engaged in by the offender in a standardized manner. It is used to support the process of crime linkage in Belgium, the Czech Republic, France, Germany, Ireland, the Netherlands, New Zealand, Switzerland, and the United Kingdom (Wilson & Bruer, n.d.). In Belgium, the Netherlands, and the United Kingdom, police investigators submit the case papers for each offence to be included on the database to the analytical units. The types of cases submitted to the three analytical units include stranger sexual offences and sexual homicides. In the United Kingdom, the data were extracted directly from ViCLAS by an analyst from the SCAS. In Belgium and the Netherlands, crime analysts employed in the ViCLAS units manually extracted the data from ViCLAS and other relevant systems (e.g., crime records to identify solved and unsolved cases). In both countries, all data retrieved from ViCLAS were reviewed by the analysts against the original paper files to ensure the coding was in accordance with the coding dictionary and quality control was assessed using the current quality assurance manual. These data sets were encrypted and sent to the third author.

The data from Finland were already coded due to its use in previous research studies (Häkkänen, Lindlöf, & Santtila, 2004; Santtila *et al.*, 2005). The South African data were collected by the third author *in situ* at the Investigative Psychology Section of the South African Police Service (SAPS) over a three-month period. Information was extracted directly from hard copy case files.

The crime linkage practitioners from the United Kingdom, Belgium, and the Netherlands assessed the comparability of a large set of variables across the different countries resulting in a common coding dictionary of 166 variables that could be considered comparable. For each crime in the data set, information pertaining to these 166[6] binary behavioural variables was, therefore, collated. These variables represent the type and quality of information stored regarding crimes on ViCLAS. Our data sharing agreements preclude the disclosure of the exact variables; however, they encompassed behaviours designed to gain and maintain control over the victim (e.g., how the victim was approached, whether a weapon was used and how, the instrumental use of violence), behaviours associated with exiting the crime scene or evading capture (e.g., wearing gloves, a mask or a disguise, giving a false name, taking forensic precautions), sexual behaviours (e.g., whether the victim was penetrated and how, whether the offender ejaculated, if and how clothing was removed), target selection variables (e.g., the time and day of the offence, the age and gender of the victim, whether the victim was physically or mentally impaired), and behaviours thought to reflect the offence 'style' of the offender and that 'are not directly necessary for the success of the attack' (Grubin *et al.*, 2001, p. 26) (e.g., the offender complimenting the victim, showing concern or revealing personal information).

To assess the reliability with which these 166 variables could be coded, the first five series from South Africa (n = 20 cases) were dual-coded by the first and third author for inter-rater reliability analysis (representing 3,320 discrete codes). Both are experienced coders of crime scene behaviours. Kappa and/or percentage agreement was calculated for 161 of the 166 variables. The remaining five variables related to objective characteristics of a crime scene/crime (day of the week and time of the day split into four categories).

[6] *For Finland, information on 42 rather than 166 behavioural variables was present. The data for Finland were historic and, therefore, the case files could not be revisited to code additional variables. Instead, a 0 was entered for these additional variables for each Finnish case. This was not considered problematic due to the use of Jaccard's coefficient, which does not include joint non-occurrences in its calculation of the similarity between a pair.*

Seventy variables were coded as present by at least one of the coders; therefore, it was possible to calculate a Kappa statistic for these. Kappa values for these variables ranged from .74 to 1.00 with 52 of the 70 variables achieving a Kappa value of 1.00. The remaining 96 variables all achieved 100% percentage non-occurrence agreement. It is just as important to demonstrate the reliable coding of non-occurrence since joint non-occurrence is considered by analysts in the linking of crimes (Davies *et al.*, 2018) and is used in the calculation of some similarity coefficients (although not Jaccard's coefficient).

While the researchers coded the data in South Africa, the variables had already been coded for the other countries, preventing further tests of coding reliability; however, it still stands that the coding of these variables was demonstrated to be reliable on South African case files. For the United Kingdom, Belgium, and the Netherlands, a rigorous data coding and quality assurance process is used: Data are entered onto ViCLAS by trained analysts who work with such data on a daily basis. The training of analysts is a lengthy process, typically lasting several months (but it can last as long as a year, or longer if necessary), and involving close supervision by an experienced analyst. In each country, data entry onto ViCLAS is closely supervised by senior analysts and guided by a detailed quality control guide/coding manual, which explains the meaning of individual ViCLAS variables and gives examples of how these variables should and should not be coded. Consequently, all analysts entering data onto the ViCLAS system are following the same coding rules. Furthermore, before analysis begins on any case, the case is reviewed to ensure that the information entered on the ViCLAS system matches the original police files. Any inconsistencies are fed back to the analyst who entered the data onto the system and amended within the ViCLAS database itself.

Finally, inter-rater reliability (IRR) had already been assessed for the Finnish data. As is published in the respective papers, a mean K of .77 was found for Santtila *et al.* (2005). All variables also yielded a $K > 0.61$ for Häkkänen *et al.* (2004) with two exceptions and only one of these variables featured in our datasets – that of revealing personal information. While this did not reach an acceptable level of inter-rater agreement for Häkkänen *et al.* (2004), it was coded reliably in our assessment of IRR with the South African data ($K = .83$).

Once all five datasets had been received, they were reformatted into one row per offence[7] and manually joined together by the third and eleventh authors.

Analytic strategy

Our analysis followed a method designed by Professor Craig Bennell in 2002 (Bennell, 2002), which has been used in many empirical tests of the crime linkage principles since (see Bennell *et al.*, 2014; for a review). Using a specially designed piece of software, B-LINK (Bennell, 2002), four separate data sets of linked and unlinked crime pairs were created (see Table 2).[8] Using the binary coded behavioural data for each crime (the 166 variables), B-LINK calculates the Jaccard's coefficient for every pair in the data set, thereby providing a quantitative measure of how similar the two crimes are in terms of offender crime scene behaviour.

---

[7] *The binary coding was at the offence, rather than the offender, level (for offences committed by groups); therefore, no attempts were made to attribute specific behaviours to individual offenders.*

[8] *Only unlinked crimes were paired that occurred within the same country since initial analyses indicated that a significantly larger AUC was obtained when contrasting linked crime pairs with unlinked crime pairs that included two crimes from different countries (AUC = .91) than when contrasting them to unlinked pairs composed of crimes only from the same country (AUC = .86) (D = .005,* p *< .001).*

Table 2. The composition of the four datasets subject to analysis

Data set number	1	2	3	4
Types of crime included	Solved serial crimes, unsolved serial crimes, and solved apparent one-off crimes	Solved and unsolved serial crimes only (apparent one-off offences removed)	Solved serial crimes and solved apparent one-off crimes (unsolved serial crimes removed)	Solved serial crimes only (unsolved serial and apparent one-off crimes removed)
Number of crimes	3,364	2,173	3,272	2,081
Number of linked/ unlinked pairs	4,569 linked pairs and 3,401,679 unlinked pairs	4,569 linked pairs and 1,296,211 unlinked pairs	4,006 linked pairs and 3,363,884 unlinked pairs	4,006 linked pairs and 1,267,648 unlinked pairs

The approach of contrasting the behavioural similarity of linked and unlinked crime pairs, whether using tests of difference or ROC analysis, simultaneously tests both the assumption of behavioural consistency and the assumption of behavioural distinctiveness. If offenders are consistent in their crime scene behaviour, the level of behavioural similarity for linked pairs is relatively high. If offenders commit their crimes in a distinctive manner, the pairing of two crimes by two different, distinctive individuals means unlinked crimes pairs share few behaviours and thus the level of behavioural similarity is low. Therefore, to distinguish linked from unlinked crime pairs based on relative behavioural similarity with a high degree of accuracy requires both assumptions to be valid.

Three stages of analysis were conducted separately on the four data sets. This allowed us to examine whether behavioural similarity, distinctiveness, and discrimination accuracy varied as a function of whether apparent one-off crimes and/or unsolved serial crimes were included in the sample under analysis: (1) Mann–Whitney *U*-tests assessed whether the Jaccard's coefficients for the linked crime pairs were significantly larger than those for the unlinked crime pairs. Significance tests were accompanied by effect size calculations; (2) binary logistic regression using a leave-one-out classification method[9] (LOOCV; Woodhams & Labuschagne, 2012a) with linkage status (linked or unlinked) as the outcome variable and Jaccard's coefficient as the predictor variable was used to produce predicted probabilities that were entered into; a (3) ROC analysis. As outlined above, ROC curves give an indication of discrimination accuracy via the AUC. The AUC is an effect size (Harris & Rice, 1995) and is therefore independent of sample size (Sullivan & Feinn, 2012).

The findings produced using these four data sets were then compared. A key comparison was between data set 1 (which contained solved, unsolved, serial, and

[9] *A LOOCV logistic regression includes a cross-validation step and involves removing a given case from the data set and developing a logistic regression model on the remaining cases. The model is then applied to the extracted case to yield a predicted probability value. This process is then repeated for each case in the data set. Cross-validation such as this ensures that models constructed will generalize to new data.*

apparent one-off crimes) and data set 4 (which contained just solved, serial crimes). Data set 1 more closely represents the data that might be used in practice when analysts are linking crimes, whereas data set 4 is comparable to the data used in most previous studies of the crime linkage principles, which is characterized by the limitations outlined above.

While the proportion of linked crime pairs formed from series that were unsolved but linked-by-DNA was much higher in this study (12%) compared to that in Slater *et al.* (2015; 3%), it was possible that their removal in data sets 3 and 4 might have little impact due to the size of the samples or be obscured by the inclusion of the one-off crimes. Consequently, meaningful differences between solved and unsolved crime series might be obscured. An additional analytic approach was, therefore, developed whereby a subset of linked and unlinked crime pairs were generated from the unsolved but linked-by-DNA crime series and the three stages of analysis repeated. This allowed for comparison in findings between crime pairs generated from two solved serial offences (n = 4,006 linked pairs and n = 1,267,648 unlinked pairs) and from two unsolved serial offences (n = 563 linked pairs and n = 1,467 unlinked pairs). This was an alternative way of examining whether the principles of consistency and distinctiveness were supported when including unsolved crime data in samples.

It is also important to note that, although a large AUC value indicates support for the principles underpinning crime linkage, it can still be associated with a considerable number of decision-making errors, particularly when there is an imbalance in the ratio of ‘positive’ (linked) to ‘negative’ (unlinked) cases, which is certainly the case with these data (see Longadge, Dongre, & Malik, 2013; for a review of the so-called ‘class imbalance problem’).[10] This issue is not unique to crime linkage and applies in other classification domains (e.g., risk prediction, diagnosis of rare diseases). Therefore, a final step in the analysis was to illustrate the number and type of errors made when adopting a particular decision threshold (i.e., a specific level of similarity used to determine when two crimes are similar and distinctive enough to warrant being linked). Based on discussions with crime linkage practitioners, we selected the false alarm rate of 15% for these illustrations since, in practice, it is preferable to minimize the number of false alarms.[11] With the false alarm rate fixed at 15%, the proportions of hits, correct rejections, and misses were calculated using the full data set (i.e., solved series, unsolved series, and one-off offences).

The ROC analysis was also repeated for each country individually using the full data set for each country. The compositions of these samples can be seen in Table 1.

Results

Mann–Whitney U-tests for international sample

The behavioural similarity of the linked crime pairs was significantly larger than that of the unlinked crime pairs ($p < .001$) across all four data sets (see Table 3), thereby demonstrating comparable support for the principles of crime linkage across data sets.

[10] *It is important to note that the class imbalance problem arises from the methodology of creating all possible (linked and unlinked) pairs; therefore, it will impact on any statistical technique used alongside this method.*

[11] *The impact of choosing different decision thresholds is an entire research question in itself and certainly something that should be subject to empirical study and cost-benefit analysis; however, this is beyond the scope of the current article.*

Table 3. Statistical comparisons of linked and unlinked crime pairs in terms of behavioural similarity

Dataset	Linked Crime Pairs Median Jaccard (Min.–Max.)	Unlinked Crime Pairs Median Jaccard (Min.–Max.)	Test statistics
All Data Included	.44 (.00–1.00)	.24 (.00–1.00)	$Z = 82.36, p < .001, r = .04$
Apparent One-Off Crimes Removed	.44 (.00–1.00)	.23 (.00–1.00)	$Z = 85.14, p < .001, r = .07$
Unsolved Crimes Removed	.44 (.00–1.00)	.24 (.00–1.00)	$Z = 76.66, p < .001, r = .04$
Both Apparent One-Off and Unsolved Crimes Removed	.44 (.00–1.00)	.23 (.00–1.00)	$Z = 79.51, p < .001, r = .07$

The effect size r was approximated using the formula from Pallant (2007) resulting in effect sizes ranging from .04 to .07.

ROC analysis for international sample

For the sake of brevity, only the ROC analyses are presented here, but a summary of the binary logistic regressions using LOOCV can be obtained from the first author upon request. Table 4 displays the AUC values and Figure 1 the ROC curves. All AUCs represent an excellent level of predictive accuracy (Hosmer & Lemeshow, 2000). Furthermore, the AUCs were similar across all four data sets. The inclusion of one-off offences in the sample reduced discrimination accuracy (as measured by the AUC) significantly, $D = 1.99$; $p < .05$, although the change in the AUC was small (from .86 to .85). The change in discrimination accuracy (AUC of .86 to .85) when both unsolved and one-off offences were added to the sample (data set 1) compared to when they were absent (data set 4) approached significance, $D = 1.93, p = .05$.

Separate analyses for solved versus unsolved serial crime pairs for international sample

When sampling only solved series, the AUC was .86 ($p < .001$, $SE = .003$, 95% CI = .86–.87, as per Table 4), whereas when sampling only unsolved series, the AUC was .79 ($p < .001$, $SE = .011$, 95% CI = .77–.81) representing an adequate level of discrimination accuracy (Hosmer & Lemeshow, 2000). The difference between these two values was statistically significant ($D = 5.47, p < .000001$).

The number and types of correct/incorrect decisions at a 15% false alarm rate threshold for international sample

Table 5 summarizes the proportion of hits, misses, and correct rejections when the threshold of a 15% false alarm rate was applied.

ROC analysis for each country separately

A ROC analysis was also run for each country's data separately. The results can be seen in Table 6.

Table 4. Receiver Operating Characteristic analysis testing discrimination accuracy across the four data sets

Data set	Area Under the Curve (*SE*)	95% confidence interval
All Data Included	.85 (.003)*	.84–.86
Apparent One-Off Crimes Removed (Series only)	.86 (.003)*	.86–.87
Unsolved Crimes Removed (Solved only)	.85 (.003)*	.84–.85
Both Apparent One-Off and Unsolved Crimes Removed (Solved Series Only)	.86 (.003)*	.86–.87

Note. *p < .001.

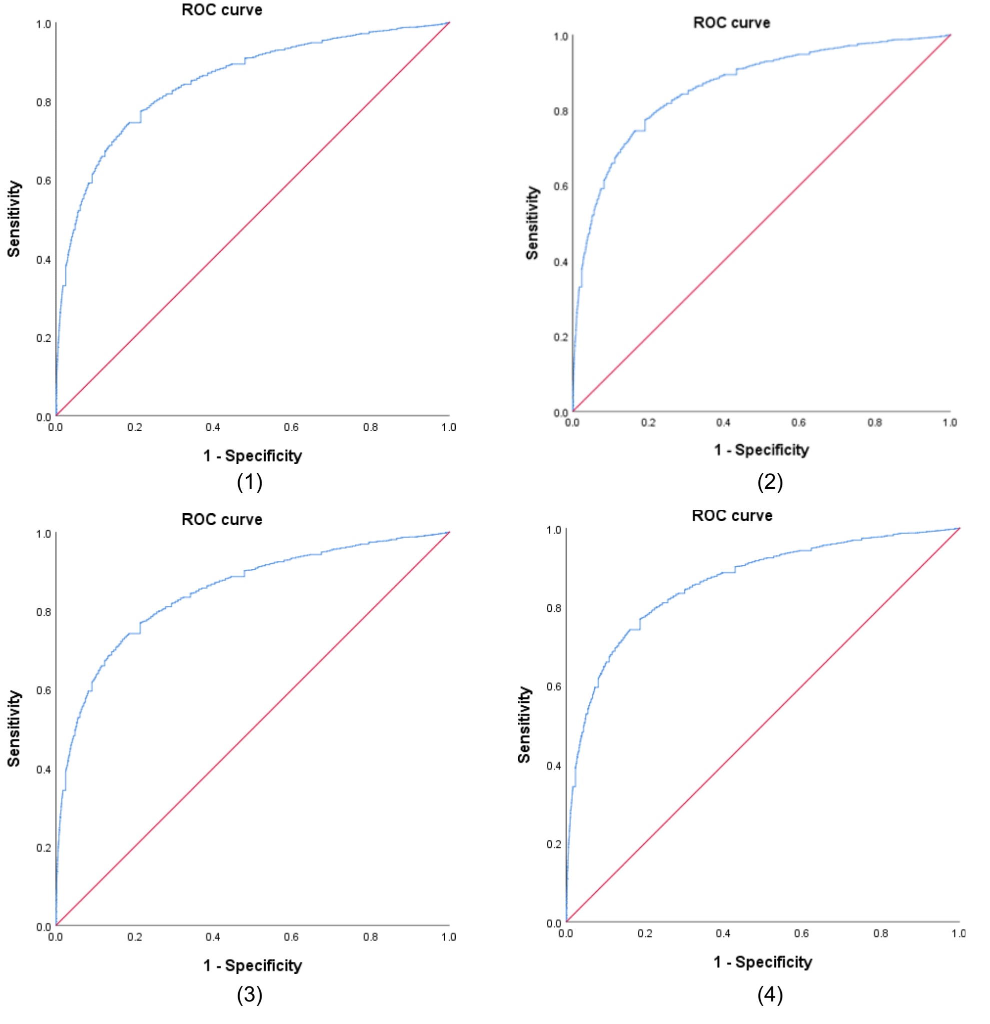

Figure 1. The ROC curves which correspond with the AUCs in Table 4 for (1) all data included; (2) apparent one-off crimes removed; (3) unsolved crimes removed; and (4) both apparent one-off and unsolved crimes removed.

Table 5. Number of hits, misses, correct rejections, and false alarms using a decision threshold of 15% false alarms

	Predicted linked	Predicted unlinked
Linked in Reality	71% Hit Rate (3,247 linked crime pairs were correctly identified)	29% Miss Rate (1,322 linked crime pairs were incorrectly classified as unlinked)
Unlinked in Reality	15% False Alarm Rate (532,170 unlinked crime pairs were incorrectly classified as linked)	85% Correct Rejection Rate (2,869,509 unlinked crime pairs were correctly identified)

Table 6. Receiver Operating Characteristic analysis testing discrimination accuracy across the five different countries

Country	Area Under the Curve (*SE*)	95% confidence interval
The United Kingdom[a]	.83 (.005)*	.82–.84
Belgium[b]	.85 (.012)*	.82–.87
Finland[c]	.56 (.039)	.49–.64
The Netherlands[d]	.76 (.019)*	.73–.80
South Africa[e]	.79 (.007)*	.78–.80

Notes. [a]Serial, one-off, solved, and unsolved crimes (linked pairs n = 2,537, unlinked pairs n = 3,321,794).
[b]Serial, one-off, solved, and unsolved crimes (linked pairs n = 400, unlinked pairs n = 29,246).
[c]Serial, one-off, solved, and unsolved crimes (linked pairs n = 55, unlinked pairs n = 7,448).
[d]Serial, one-off, solved, and unsolved crimes (linked pairs n = 189, unlinked pairs n = 14,689).
[e]Serial, solved, and unsolved crimes (linked pairs n = 1,388, unlinked pairs n = 28,502).
*$p < .001$.

Discussion

There is a growing trend of international courts viewing crime linkage analysis as a form of behavioural science and thus qualifying for assessment against legal standards governing the admission of scientific evidence (Pakkanen *et al.*, 2015). This, alongside its use to inform police decision-making, makes the reliability of its underlying principles an important subject for empirical research.

We tested the reliability of its underlying principles simultaneously using ROC analysis to assess the accuracy with which linked crime pairs could be distinguished from unlinked crime pairs based on quantitative measures of behavioural similarity. The AUCs obtained (.79–.86) are similar in size to those seen in past, smaller scale studies (e.g., Slater *et al.*, 2015; Winter *et al.*, 2013; Woodhams & Labuschagne, 2012a) and represent an adequate (.79) to excellent (.80 and above) level of discrimination accuracy. Even the AUC obtained when sampling only from unsolved but linked-by-DNA series (.79) was larger than AUCs reported in previous studies (e.g., Bennell *et al.*, 2009).

These previous studies demonstrated little impact of including either one-off crimes, or unsolved but linked-by-DNA series, on the AUC values obtained. However, their small samples sizes and the fact that none of these studies included confirmed series alongside one-off offences *and* unsolved crime series meant less confidence could be placed in their findings. Through the cooperation of police and academics from seven countries, a much larger sample was collated allowing for a more rigorous and ecologically valid test of the crime linkage principles. Our findings mirror those of previous studies; the inclusion of one-off crimes and unsolved crime series had little impact on the AUCs when using the full sample.

These findings are of global significance bearing in mind the use of crime linkage to inform police decision-making around the world regarding the most serious types of crimes (Bennell *et al.*, 2014; Wilson & Bruer, n.d.). Our results also provide the sorts of research findings regarding the principles of crime linkage which have been sought by the courts in the past, and which will likely be sought in the future, when deciding on the admissibility of crime linkage analysis as a form of expert evidence.

There are, however, important caveats to these generally positive findings. Our final phase of analysis considered the scale and type of decision errors that would be made if a decision threshold was utilized that capped the false alarm rate at 15%. This illustrated that, despite our logistic regression models achieving high AUCs, a considerable number of errors in linkage predictions can occur when using these statistical models. For example, due to the relative base rates of linked versus unlinked pairs in our data set, a 15% false alarm rate corresponds with more than 500,000 false alarm predictions being made. The number of misses is much smaller at just over 1,000. Such errors arise because within the data set there are linked crime pairs that are characterized by inconsistency and indistinctiveness, and unlinked crime pairs that are highly similar with respect to crime scene behaviour (see the Min and Max values in Table 3). Therefore, the principles of crime linkage do not hold for all cases.

Bearing in mind the police resources that might be put into further analytical and investigative work with this number of false alarms, it is likely that a more stringent false alarm rate would be needed in practice (this would, of course, result in a reduced hit rate).[12] While the paper does not provide a definitive answer as to the error rate associated with crime linkage in practice, it still aids the courts and researchers/practitioners by allowing them to appreciate the volume of errors that *can* occur even when specific linking strategies are associated with high AUCs. An important next step would be to establish the base rates of linked and unlinked pairs in databases such as ViCLAS to estimate the extent of the class imbalance problem in practice.[13] This, combined with a full cost-benefit analysis that considers the human and financial savings/costs associated with the four decision outcomes of the linkage task, would help inform future decisions regarding the most appropriate decision threshold to use.

In addition, we found a significant difference in AUC when contrasting linked and unlinked pairs using a sample generated from solved series versus unsolved but linked-by-DNA series. This finding is similar to that reported by Woodhams and Labuschagne (2012a) with a much smaller sample. They observed that linked crime pairs first identified as a series on the basis of DNA were characterized by less behavioural similarity (a smaller Jaccard's coefficient) than those first identified on the basis of similar *modus operandi*.

The actual composition of crime types in databases used for crime linkage, such as ViCLAS, is not currently known (e.g., ratios of serial, one-off, solved, and unsolved). However, our findings highlight the importance of such studies since the trends seen in

[12] *However, it should be noted that such large figures would only apply if you are comparing all crimes in a given database at the same time. In practice, certain filters to reduce the number of cases retrieved would be applied in addition (e.g., offender ethnicity, time, place, geography). For example, a case linked by DNA but where the specific DNA profile is not in the national database will lead to the decision to exclude all cases with a known offender as a first filter (Davies* et al., *2018).*

[13] *The volume of unsolved crimes in such databases would make it impossible to know the* real *base rates of linked and unlinked pairs.*

our data of decreasing discrimination accuracy with the addition of one-off offences and with unsolved but linked-by-DNA series could be more pronounced if databases contain many more offences of these types. One study has assessed how varying proportions might affect the discrimination accuracy yielded from statistical analyses; Haginoya (2016) found no effect of varying the ratio of one-off offences to series on the ability to link crimes; however, this analysis was limited to the linking features of geographical and temporal proximity. The optimum approach would be to conduct a study on the entire police database in each country. Where this is not possible, it is important in the future to (a) conduct a study where the proportion of serial to one-off offences is systematically varied to determine how this impacts on discrimination accuracy using offender crime scene behaviours and to replicate Haginoya; and (b) to determine what the ratio is on existing databases so that researchers can evaluate how much the proportions in their datasets reflect reality. This ratio would only be an estimate as it cannot be known for definite that a crime is truly a one-off offence or part of an undetected series. However, an estimate with these limitations in mind would still help inform the sampling frames of future studies where a full database cannot be used for analysis.

It is important to note that our study is a test of the *principles* of crime linkage and is not a test of the *practice* of crime linkage. This does not invalidate our findings because we set out to answer legal questions facing international courts surrounding the admissibility of crime linkage evidence and to inform an evidence-based policing approach to crime linkage (Rainbow, 2015). However, the accuracy of practitioner decision-making with and without the aid of statistical models to support their decision-making is a topic in need of study.

Finally, it is also important to recognize that the sample of crimes utilized in this study was dominated by UK crimes as the UK analytical unit, SCAS, contributed the largest number of cases. It was not possible to repeat all statistical manipulations conducted with the international data set with the data set from each country individually because the numbers of solved versus unsolved series, or series versus one-off offences, were insufficient. However, one overall ROC analysis was conducted on the full data set available per country using the steps described above. The AUCs per country (.76 to .85) were all within the range observed for previous studies of the crime linkage principles with serial sexual offences (i.e., .75 to .89; Bennell *et al.*, 2009; Slater *et al.*, 2014; Winter *et al.*, 2013; Woodhams & Labuschagne, 2012a) for all countries with the exception of Finland.[14] The variation in discrimination accuracy across the countries is interesting, but it is difficult to draw any firm conclusions from this. It is possible that they result from differences in the relative compositions of the samples (e.g., solved vs. unsolved or serial vs. one-off). There may be optimal sets of *modus operandi* behaviours per country, and identification of these may improve discrimination accuracy. Authors have previously commented that behaviours may vary in their relative distinctiveness from country to country (Woodhams & Labuschagne, 2012b). Alternatively, the differences observed may be due to the series sampled from each country and with a different set of series the findings might vary. This underscores the importance of future research studies aiming for a large, realistic sample of crimes when investigating crime linkage within a country. As noted above, ideally, where they exist and where permission is given, studies should

---

[14] *The AUC of .85 obtained with the multi-country sample with data set 1 was unchanged with the removal of the Finland subset of cases from the overall sample. The smaller AUC for Finland may reflect the reduction in behavioural information available for linkage predictions with 42 versus 166 behaviours.*

utilize the entire data set of crimes on databases that assist with crime linkage in that country (e.g., ViCLAS).

Conclusion

The paper reported a test of the reliability of the principles underlying crime linkage with the largest and most ecologically valid sample of sexual offences to date made possible by international police–academic cooperation. A sample of several thousand crimes, which included convicted series, unsolved but linked-by-DNA series, and convicted one-off sexual offences, was collated and subject to LOOCV logistic regression and ROC analysis. Support for the reliability of the underlying principles of crime linkage analysis was found. However, our calculations indicate that despite the large AUC values achieved by the regression models, there is still the potential for a large number of decision-making errors to be made due to the low base rate of same-offender crime pairs in the samples.

References

Bennell, C. (2002). *Behavioural consistency and discrimination in serial burglary* (Unpublished doctoral thesis). Department of Psychology, University of Liverpool, UK.

Bennell, C., & Canter, D. V. (2002). Linking commercial burglaries by modus operandi: Tests using regression and ROC analysis. *Science & Justice*, *42*, 153–164. https://doi.org/10.1016/S1355-0306(02)71820-0

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. https://doi.org/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. https://doi.org/10.1002/jip.1395

Bennell, C., Snook, B., MacDonald, S., House, J., & Taylor, P. (2012). Computerized crime linkage systems: A critical review and research agenda. *Criminal Justice and Behavior*, *39*, 620–637. https://doi.org/10.1177/0093854811435210

Canter, D. (1995). Psychology of offender profiling. In R. Bull & D. Carson (Eds.), *Handbook of psychology in legal contexts* (pp. 343–355). Chichester, UK: Wiley.

Collins, P. I., Johnson, G. F., Choy, A., Davidson, K. T., & MacKay, R. E. (1998). Advances in violent crime analysis and law enforcement: The Canadian violent crime linkage analysis system. *Journal of Government Information*, *25*, 277–284. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S1352023798000082

Daubert v. Merrell Dow Pharmaceuticals, Inc. (1993). 509 U.S. 579, 113 S. Ct. 2786.

Davies, A. (1991). The use of DNA profiling and behavioural science in the investigation of sexual offences. *Medicine, Science & Law*, *31*, 95–101. https://doi.org/10.1177/002580249103100202

Davies, K., Alrajeh, D., & Woodhams, J. (2018). *An investigation into the process of comparative case analysis by analysts working in the Serious Crime Analysis Section in the United Kingdom*. Official report submitted to the Serious Crime Analysis Section, UK.

Federal Rules of Evidence (2011). Retrieved from https://www.law.cornell.edu/rules/fre/rule_702

Grubin, D., Kelly, P., & Brunsdon, C. (2001). *Linking serious sexual assaults through behaviour*. London, UK. Retrieved from http://webarchive.nationalarchives.gov.uk/20110314171826/http:/rds.homeoffice.gov.uk/rds/pdfs/hors215.pdf

Haginoya, S. (2016). Examining linkage analysis in serial and non-serial crimes. *Japanese Journal of Applied Psychology*, *42*, 112–120.

Häkkänen, H., Lindlöf, P., & Santtila, P. (2004). Crime scene actions and offender characteristics in a sample of Finnish stranger rapes. *Journal of Investigative Psychology and Offender Profiling*, *1*, 17–32. https://doi.org/10.1002/jip.1

Harris, M. E., & Rice, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology*, *63*, 737–748.

Hazelwood, R. R., & Warren, J. L. (2004). Linkage analysis: Modus operandi, ritual, and signature in serial sexual crime. *Aggression and Violent Behavior*, *9*, 307–318. https://doi.org/10.1016/j.avb.2004.02.002

Her Majesty's Advocate v. Thomas Ross Young (2013). HCJAC 145.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York, NY: Wiley.

Labuschagne, G. N. (2015). The use of linkage analysis evidence in serial offense trials. In J. Woodhams & C. Bennell (Eds.), *Crime linkage: Theory, research, and practice* (pp. 197–224). Boca Raton, FL: CRC Press.

Longadge, R., Dongre, S., & Malik, L. (2013). Multi-cluster based approach for skewed data in data mining. *Journal of Computer Engineering*, *12*, 66–73. Retrieved from http://www.iosrjournals.org/iosr-jce/papers/Vol12-issue6/K01266673.pdf?id=7431

Pakkanen, T., Santtila, P., & Bosco, D. (2015). Crime linkage as expert evidence - making a case for the Daubert Standard. In J. Woodhams & C. Bennell (Eds.), *Crime linkage: Theory, research, and practice* (pp. 225–249). Boca Raton, FL: CRC Press.

Pallant, J. (2007). *SPSS survival manual* (3rd ed.). Maidenhead, UK: Open University Press.

People v Prince (2007). 40 C4th 1179, 1225.

R. v. Burlingham (1993). B.C.J. No. 1986 Vancouver Registry: CA006715.

R. v. R.B. (2003). No 4589. Canada.

Rainbow, L. (2015). A practitioner's perspective: Theory, practice and research. In J. Woodhams & C. Bennell (Eds.), *Crime linkage: Theory, research, and practice* (pp. 173–196). Boca Raton, FL: CRC Press.

Santtila, P., Junkkila, J., & Sandnabba, N. K. (2005). Behavioural linking of stranger rapes. *Journal of Investigative Psychology and Offender Profiling*, *2*, 87–103. https://doi.org/10.1002/jip.26

Sentencing Council (2012). *Offences taken into consideration and totality: Definitive guideline*. London, UK: The Crown Office.

Slater, C., Woodhams, J., & Hamilton-Giachritsis, C. (2015). Testing the assumptions of crime linkage with stranger sex offenses: A more ecologically-valid study. *Journal of Police and Criminal Psychology*, *30*, 261–273. https://doi.org/10.1007/s11896-014-9160-3

State of New Jersey v. Fortin (2000). 745 A.2d 509 N.J.

State v. Mogale (2012). Case No 36/2009. South Gauteng High Court, South Africa.

State v. Nyauza (2007). Case No CC97/2007. Pretoria High Court, South Africa.

State v. Pennell, Del. Supr. (1991). 602 A.2d 48

State v. Russell (1994). 125 Wn.2d, 882 P.2d 747

State v. Steyn (2012). Case No 47/2012. South Gauteng High Court, South Africa.

State v. Sukude (2006). Case No 34/2006. Pietermaritzburg High Court, South Africa.

State v. van Rooyen (2007). Case No SS 55/2007. Cape High Court, South Africa.

State v. Yates (2007). 161 Wash.2d 714, 728–33, 168 P.3d 359.

Sullivan, G. M., & Feinn, R. (2012). Using effect size – or why the *p*-value is not enough. *Journal of Graduate Medical Education*, *4*, 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Wilson, L., & Bruer, C. (n.d.). *Violent crime linkage system (ViCLAS)*. Retrieved from http://www.rcmp-grc.gc.ca/to-ot/cpcmec-ccpede/bs-sc/viclas-salvac-eng.htm#countries

Winter, J. M., Lemeire, J., Meganck, S., Geboers, J., Rossi, G., & Mokros, A. (2013). Comparing the predictive accuracy of case linkage methods in serious sexual assaults. *Journal of Investigative Psychology and Offender Profiling*, *56*, 28–56. https://doi.org/10.1002/jip.1372

Woodhams, J., Bull, R., & Hollin, C. R. (2007). Case linkage: Identifying crimes committed by the same offender. In R. Kocsis (Ed.). *Criminal profiling: International theory, research, and practice* (pp. 117–133). Totowa, NJ: Humana Press.

Woodhams, J., & Grant, T. D. (2006). Developing a categorization system for rapists' speech. *Psychology, Crime and Law*, *12*, 245–260. https://doi.org/10.1080/10683160500151134

Woodhams, J., & Labuschagne, G. N. (2012a). A test of case linkage principles with solved and unsolved serial rapes. *Journal of Police and Criminal Psychology*, *27*, 85–89. https://doi.org/10.1002/s11896-011-9091-1

Woodhams, J., & Labuschagne, G. N. (2012b). South African serial rapists: The offenders, their victims and their offenses. *Sexual Abuse: A Journal of Research and Treatment*, *24*, 544–574. https://doi.org/10.1177/1079063212438921

Yokota, K., Fujita, G., Watanabe, K., Yoshimoto, K., & Wachi, T. (2007). Application of the behavioral investigative support system for profiling perpetrators of serial sexual assaults. *Behavioral Sciences and the Law*, *25*, 841–856. https://doi.org/10.4992/jjpsy.86.13086

Received 27 November 2017; revised version received 18 October 2018