

La gestion des informations dans le cadre de l'analyse criminelle de la délinquance sexuelle sérielle

Vers une réduction empirique de ViCLAS

Identifying links among serial sexual crimes

Toward reducing the empirical information required for ViCLAS

La gestión de las informaciones en el marco del análisis criminal de la delincuencia sexual serial

Hacia una reducción empírica del ViCLAS

Julien Chopin et Marcelo F. Aebi

La criminologie de l'information : état des lieux et perspectives
Volume 52, numéro 2, automne 2019

URI : <https://id.erudit.org/iderudit/1065860ar>

DOI : <https://doi.org/10.7202/1065860ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0316-0041 (imprimé)

1492-1367 (numérique)

[Découvrir la revue](#)

Citer cet article

Chopin, J. & Aebi, M. F. (2019). La gestion des informations dans le cadre de l'analyse criminelle de la délinquance sexuelle sérielle : vers une réduction empirique de ViCLAS. *Criminologie*, 52 (2), 164–192. <https://doi.org/10.7202/1065860ar>

Résumé de l'article

Cette recherche s'intéresse à la quantité d'information nécessaire aux systèmes de gestion des liens permettant d'identifier les cas sériels de criminalité violente. L'objectif de cette recherche est de tester empiriquement et de discuter de la pertinence de réduire le contenu de ces outils. Les données utilisées proviennent d'une extraction de la base de données du ViCLAS en France. Les analyses portent sur un échantillon de variables les plus utilisées dans le cadre des agressions sexuelles. Des analyses descriptives et factorielles sont utilisées pour déterminer quelles informations sont collectées par les policiers et comment elles pourraient être réduites empiriquement en agrégeant les données corrélées. Les résultats montrent qu'environ 65 % des modalités testées sont utilisées dans moins de 5 % des cas d'agressions sexuelles. Les résultats des analyses factorielles montrent qu'il est possible de réduire drastiquement l'outil. Les implications pratiques de ces résultats sont discutées dans une perspective d'investigation policière.

La gestion des informations dans le cadre de l'analyse criminelle de la délinquance sexuelle sérielle

Vers une réduction empirique de ViCLAS

Julien Chopin, Ph. D.^{1,2}

*Chercheur postdoctoral
École de criminologie, Université Simon Fraser
julien@c@sfu.ca*

Marcelo F. Aebi, Ph. D.

*Professeur
École des sciences criminelles, Université de Lausanne
Suisse
marcelo.aebi@unil.ch*

RÉSUMÉ • Cette recherche s'intéresse à la quantité d'information nécessaire aux systèmes de gestion des liens permettant d'identifier les cas sériels de criminalité violente. L'objectif de cette recherche est de tester empiriquement et de discuter de la pertinence de réduire le contenu de ces outils. Les données utilisées proviennent d'une extraction de la base de données du ViCLAS en France. Les analyses portent sur un échantillon de variables les plus utilisées dans le cadre des agressions sexuelles. Des analyses descriptives et factorielles sont utilisées pour déterminer quelles informations sont collectées par les policiers et comment elles pourraient être réduites empiriquement en agrégeant les données corrélées. Les résultats montrent qu'environ 65% des modalités testées sont utilisées dans moins de 5% des cas d'agressions sexuelles. Les résultats des analyses factorielles montrent qu'il est possible de réduire drastiquement l'outil. Les implications pratiques de ces résultats sont discutées dans une perspective d'investigation policière.

MOTS CLÉS • Système de gestion des liens, réduction de données, agression sexuelle, efficacité des systèmes.

1. School of Criminology, Simon Fraser University, 8888, University Drive, Burnaby (C.-B.), Canada, V5A 1S6.

2. Les auteurs souhaitent remercier le chef de l'Office central pour la répression des violences aux personnes (OCRVP), ses collaborateurs ainsi que Monsieur le directeur central de la Police judiciaire française. Les auteurs remercient également le Fonds national suisse de la recherche scientifique qui a soutenu la préparation et la rédaction de cet article (Fonds n° P2LAPI_178193).

Introduction

L'avènement de l'outil informatique au cours de la seconde moitié du xx^e siècle a constitué une évolution sociétale et évidemment une révolution pour la police dont l'impact est encore difficilement mesurable (voir Ribaux, 2014). La possibilité pour les forces de police de recenser une quantité quasi infinie de données et de pouvoir les analyser constitue une ère nouvelle dans l'investigation criminelle et rend l'utilisation des techniques computationnelles très séduisante pour les polices judiciaires (Grossrieder, Albertetti, Stoffel et Ribaux, 2013). Toutefois, l'augmentation de la quantité d'information stockée exige le développement de stratégies efficaces pour accéder à ces données et pour les mettre en relation. C'est dans ce contexte que, durant le dernier quart du xx^e siècle, s'est développée une nouvelle forme d'analyse criminelle basée sur l'utilisation d'outils informatiques et d'approches explicatives novatrices (Collins, Johnson, Choy, Davidson et MacKay, 1998; Martineau et Corey, 2008; Ribaux, 2014). L'un des domaines dans lesquels ce phénomène peut être le mieux observé est celui de l'analyse de la délinquance sexuelle sérielle, qui peut être opérationnalisé par la présence de plusieurs événements criminels, généralement au moins deux (voir par exemple Beauregard, 2005; Beauregard, Proulx, Rossmo, Leclerc et Allaire, 2007; Beauregard, Rossmo et Proulx, 2007; Chopin et Aebi, 2018; Deslauriers-Varin et Beauregard, 2013), séparés par un intervalle de temps et commis par le même auteur (Fox et Levin, 1994; Geberth, 1986). Dans ce domaine, des outils toujours plus puissants et enregistrant un nombre d'informations toujours plus important ont été mis sur pied afin d'aider les policiers dans les enquêtes les plus complexes. Actuellement, la police fait appel à ces outils principalement dans le but de créer des liens entre les affaires supposément sérielles de criminalité violente et tout particulièrement dans le cadre de la délinquance sexuelle. Les outils les plus connus ont été développés en Amérique du Nord, en premier lieu par le Federal Bureau of Investigation (FBI) avec la création du *Violent Criminal Apprehension Program* (ViCAP) en 1985, puis par la Gendarmerie royale du Canada (GRC) avec la création du *Violent Crime Linkage Analysis System* (ViCLAS) en 1992 (Collins *et al.*, 1998; Martineau et Corey, 2008). L'utilisation de ces outils a été perçue comme une avancée significative dans les méthodes d'investigation et, petit à petit, ils ont été incorporés dans l'éventail des méthodes de travail d'un certain nombre de polices judiciaires à travers le monde, suscitant néanmoins un certain nombre de questionnements

(Bourque, LeBlanc, Utzschneider et Wright, 2009; Margot, 2009; Ribaux, 2014; Snook, Cullen, Bennell, Taylor et Gendreau, 2008; Snook, Eastwood, Gendreau, Goggin et Cullen, 2007; Snook, Luther, House, Bennell et Taylor, 2012). En se basant sur la littérature empirique ayant évalué le système ViCLAS, la présente recherche essaie d'établir si l'accumulation de données dans ce système est profitable d'un point de vue opérationnel et si une réduction empirique du nombre d'informations requises actuellement par le système serait envisageable.

Revue de littérature

L'évaluation scientifique du ViCLAS

Il existe plusieurs manières d'évaluer un outil et, parmi les différentes possibilités qui s'offrent aux chercheurs, les plus classiques consistent à tester la fiabilité et la validité de l'outil en question (Aebi, 2006). De tels tests peuvent être conduits pour l'ensemble des instruments de mesure tels que les sondages de délinquance autoreportée (Aebi, 2006; Aebi et Jaquier, 2008; Killias, Aebi et Kuhn, 2012) ainsi que les outils d'analyse criminelle (Katz, Webb et Schaeffer, 2000). Néanmoins, comme le rappellent Deslauriers-Varin, Bennell et Bergeron (2018a), relativement peu de travaux empiriques ont évalué de manière indépendante le système ViCLAS puisqu'il est très difficile pour des chercheurs en provenance du milieu de la recherche d'accéder à ces données (Margot, 2009; Pakkanen, Santtila et Bosco, 2015).

L'analyse de la fiabilité du ViCLAS

La fiabilité peut être définie comme la capacité d'un outil à fournir des mesures reproductibles et intersubjectives (voir Aebi, 2006, p. 15). Les méthodologues des outils de mesure expliquent qu'un trop grand nombre de possibilités de réponses augmente inexorablement l'intervariabilité et diminue ainsi la fiabilité du questionnaire (Bradburn, Sudman et Wansink, 2004). Par ailleurs, la formulation des questions ainsi que le choix des réponses proposées auront aussi un impact non négligeable sur la fiabilité de l'outil (voir par exemple Villettaz, 1993). L'intervariabilité est donc une conséquence d'une diminution de la fiabilité et correspond au fait que le même phénomène étudié par plusieurs observateurs n'aboutit pas au même résultat. Ainsi, plus les éléments subjectifs de l'outil sont importants, plus le risque d'augmenter l'intervariabilité et donc de réduire la fiabilité est important.

En ce qui concerne l'outil ViCLAS, il a fallu attendre près de dix ans après sa première utilisation pour que les premières analyses concernant sa fiabilité soient publiées (Martineau et Corey, 2008; Snook *et al.*, 2012). Martineau et Corey (2008) ont testé la fiabilité de ViCLAS en remettant à 237 enquêteurs deux scénarios de crimes pour lesquels ils devaient introduire les informations sur le livret ViCLAS. Le but de cette expérimentation était de mesurer le niveau d'accord interjuge des praticiens du ViCLAS sur un même cas. Les chercheurs se basent sur trois mesures de la fiabilité, à savoir: le pourcentage d'accord entre les réponses (PA)³, le pourcentage d'occurrences concordantes (OPA)⁴ et le pourcentage d'absence d'occurrences concordantes (NOPA)⁵. Les résultats des PA obtenus indiquent en moyenne une excellente fiabilité de l'outil tandis que les OPA et les NOPA indiquent une fiabilité relativement moyenne de ViCLAS. Si Martineau et Corey (2008) s'attachent particulièrement à analyser les résultats des PA, il apparaît que ce sont les résultats des OPA et des NOPA qui reflètent le mieux la fiabilité de l'outil (Snook *et al.*, 2012). En effet, la plupart des questions sont construites de manière à obtenir plusieurs réponses et c'est ainsi en analysant l'occurrence de la combinaison des réponses plutôt que chaque réponse dans son individualité que la fiabilité sera la plus significative. Cette analyse concorde d'ailleurs avec l'approche de Snook et de ses collègues (2012) qui reproduisent l'expérience avec un processus de recherche similaire, mais avec un échantillon moins important. Leurs résultats concernant la fiabilité se rapprochent d'ailleurs plus de ceux obtenus avec l'OPA de Martineau et Corey (2008). En étudiant les hypothèses qui sont avancées par les auteurs pour expliquer cette fiabilité relativement faible, on trouve des explications méthodologiques et des explications de fond qui concernent la structure de l'outil. Les auteurs mettent directement en cause la lourdeur de l'outil à travers le nombre de ses variables et de ses modalités foisonnantes (Bennell, Snook, MacDonald, House et Taylor, 2012; Marin, Poirret, Quemener et Gallois, 2003; Snook *et al.*, 2012). Le trop grand nombre de variables (dérivées des 156 questions) et surtout le trop grand nombre de modalités (1140) sont, pour ViCLAS, largement responsables du manque de fiabilité des données, se traduisant par une grande inter-

3. *Percent agreement.*

4. *Occurrence percent agreement.*

5. *Non occurrence percent agreement.*

variabilité. Ils soulignent également le manque de connaissances des enquêteurs pour remplir le questionnaire et en connaître toutes les subtilités ainsi que parfois leur manque d'intérêt lié à la mécompréhension de son usage. Il faut cependant rappeler que ces différentes expériences ont été faites dans des conditions de « laboratoire » et que les résultats sont conditionnés par les limitations et les biais inhérents à une telle approche.

L'analyse de la validité du ViCLAS

À notre connaissance, deux études ont analysé empiriquement la validité du ViCLAS (Chopin et Aebi, 2018, 2019). La validité d'un outil peut être définie comme son aptitude à mesurer efficacement le phénomène étudié (voir Aebi, 2006, p. 15). L'étude de Chopin et Aebi (2017) s'est penchée sur la validité de construction de l'outil ViCLAS. La validité de construction présente un rapport au degré d'adéquation entre les résultats empiriques de la recherche et les prévisions théoriques (Aebi, 2006, p. 250). En se basant sur les données françaises, le but était de savoir si les modèles théoriques opérationnalisés à travers le questionnaire de l'outil étaient utilisés adéquatement par les policiers. Les résultats de Chopin et Aebi (2017) laissent supposer que les informations sociodémographiques et situationnelles sont beaucoup plus souvent utilisées et complétées par les policiers que les informations comportementales et psychologiques. Ces résultats semblent confirmer le fait que les données situationnelles jouent un rôle prépondérant dans l'analyse criminelle opérationnelle (Birrer, 2010; Boba, 2009; Grossrieder, 2017; Ribaux, 2014). Dans une deuxième étude, Chopin et Aebi (2018) ont analysé la validité concourante de ViCLAS, c'est-à-dire le « degré d'adéquation entre les résultats empiriques de la recherche et les prévisions théoriques » (Aebi, 2006, p. 250). En d'autres termes, il s'agissait de savoir si les modèles théoriques qui ont été opérationnalisés à travers les questions de l'outil étaient corroborés par les résultats obtenus avec cet outil. L'objectif était de comparer les résultats obtenus en matière d'identification de la sérialité par l'outil avec ceux obtenus précédemment avec des outils dont la validité a déjà été testée et approuvée, comme les sondages de délinquance autorapportée (Chopin et Aebi, 2018). Leurs résultats laissent entendre que le degré de sérialité détecté par l'outil ViCLAS est inférieur à celui obtenu avec d'autres outils (Chopin et Aebi, 2018).

Les critiques du ViCLAS

Les résultats des évaluations empiriques effectuées sur l'outil ViCLAS suggèrent un certain nombre de pistes pour améliorer le fonctionnement de l'outil et ainsi l'aide qu'il pourra apporter aux policiers dans le cadre des enquêtes criminelles en facilitant l'identification de séries (Chopin, 2017 ; Chopin et Aebi, 2018, 2019 ; Martineau et Corey, 2008 ; Snook *et al.*, 2012). Le volume de données requis par le ViCLAS est symptomatique d'une difficulté à dégager un objectif clair à cet outil (Margot, 2009). En effet, le ViCLAS vise un niveau d'analyse extrêmement précis afin de ne laisser au hasard aucun élément d'analyse du crime. Ce niveau de micro-analyse apparaît même encore réduit lorsque le questionnaire aborde des notions extrêmement spécifiques qui destinent l'outil à un niveau d'analyse clinique. Il est souligné d'ailleurs dans la littérature qu'il faut se méfier du phénomène dit « N = 1 » (Snook *et al.*, 2008). Snook *et al.* (2008) soulignent à travers ce terme que les événements exceptionnels sont séduisants, car ils marquent les esprits, mais qu'ils n'ont pas leur place dans une démarche scientifique. L'analyse doit venir de la généralisation et non d'anecdotes qui empêchent de lier les différents cas entre eux. En effet, les anecdotes et les éléments exceptionnels viennent plutôt singulariser chaque affaire, ce qui réduit considérablement les possibilités de les lier avec d'autres cas en les rendant quasiment uniques. Ce phénomène est à mettre en relation avec ce que Sheptycki (2004) appelle le « bruit » et qu'il identifie comme une pathologie des systèmes d'analyse criminelle. Le bruit correspond à une surabondance d'information de faible qualité qui noie l'information importante et empêche les systèmes de fonctionner convenablement (Sheptycki, 2004).

Le niveau d'analyse, la qualité ainsi que la quantité d'informations demandée et l'absence de données dont la fiabilité est bonne engendrent un phénomène qui est appelé le *linkage blindness* (Egger, 1984 ; Egger et Doney, 1990 ; Ribaux, 2014). Ce phénomène correspond à l'incapacité pour les forces de police de lier les cas non résolus entre eux afin de mettre à jour des crimes sériels et de les résoudre (Egger, 1984). Cette incapacité à mettre à jour des liens entre des cas qui sont pourtant liés est à mettre en relation avec la grande quantité d'information requise par le système qui engendre le phénomène « N = 1 » décrit par Snook *et al.* (2008). Witzig (2003) décrit le pendant américain du ViCLAS, le ViCAP, comme un monstre administratif engloutissant de grandes quantités de données sans pour autant parvenir à lier ensemble les cas

qui devraient l'être. D'ailleurs, Witzig (2003) signale que l'outil américain a fait l'objet d'une réduction de son contenu sans que son efficacité soit touchée.

Objectif de la recherche

La revue de littérature a mis en évidence les failles du système ViCLAS en montrant du doigt la taille de l'outil qui rend complexe la création de liens entre les affaires qui devraient l'être. En se basant sur l'exemple américain du ViCAP, qui a été réduit sans pour autant que son efficacité en soit affectée, la présente recherche s'intéresse à la réduction empirique du ViCLAS. Chopin et Aebi (2019) ont déjà montré qu'il était possible de réduire théoriquement le nombre de variables en se basant sur les informations situationnelles qui sont les plus exploitées par les forces de police. Cette recherche poursuit cette démarche en s'intéressant à la réduction empirique des modalités contenues dans les variables. Elle fonde son analyse sur deux questions de recherche : 1) Est-ce que l'ensemble du contenu des variables du ViCLAS est exploité par les policiers ? et 2) Est-il possible de réduire le contenu des variables du ViCLAS ?

Méthodologie

Échantillon

Cette recherche se fonde sur une extraction des données françaises du ViCLAS. Les affaires enregistrées par les analystes dans cette base de données sont des crimes sexuels et/ou violents extrafamiliaux ainsi que certains crimes intrafamiliaux qui sont intervenus dans un contexte spécifique⁶ et dont le mobile est inconnu. Il s'agit d'affaires qui se sont déroulées entre 1979 et 2013 sur le territoire français (métropole et territoires d'outre-mer). Cette étude se base sur un échantillon de 3965 affaires à caractère sexuel qui représente 71,64% du total des affaires résolues enregistrées dans la base de données au moment de l'extraction en 2013. Les homicides sexuels ont été exclus de l'analyse,

6. Il s'agit de crimes comprenant la commission d'actes de torture et de barbaries qui caractérisent le comportement de certains prédateurs sexuels. Le choix d'inclure ces crimes dans la base de données est fait par les analystes et non par les auteurs de cette recherche qui ont décidé de les conserver afin de proposer des résultats au plus près des impératifs de l'outil et des réalités policières.

car ces crimes présentent des caractéristiques propres en regard des agressions sexuelles (voir par exemple Beauregard et DeLisi, 2018a, 2018b; Beauregard, DeLisi et Hewitt, 2017; Chopin et Beauregard, 2019a).

Toutes les affaires incluses dans l'analyse sont des affaires résolues. Les données sont enregistrées dans le système par un groupe d'analystes qui tirent les informations des dossiers d'enquêtes et qui complètent eux-mêmes la base de données. À la différence d'autres pays utilisant ViCLAS et où les enquêteurs remplissent eux-mêmes les formulaires, cette procédure a été mise en place pour accroître la fiabilité des données. La base de données a été divisée en deux en fonction de l'âge des victimes avec d'une part celles de moins de 15 ans et d'autre part celles de 15 ans et plus. En effet, de nombreux travaux ont montré des différences importantes dans le processus criminel entre les affaires à caractère sexuel impliquant des enfants et celles impliquant des victimes adultes (voir par exemple Beauregard, Leclerc et Lussier, 2012; Chopin, 2017; Chopin et Beauregard, 2019b; Chopin et Caneppele, 2019a, 2019b; Ciavaldini, 1999; Gravier, Mezzo, Abbiati, Spagnoli et Waeny, 2010; Leclerc, Proulx et Beauregard, 2009). L'échantillon pour cette recherche comprend ainsi 2783 affaires concernant des victimes de 15 ans et plus ainsi que 1182 affaires concernant des victimes de moins de 15 ans au moment des faits.

Variables utilisées

Basées sur les résultats de Chopin et Aebi (2019), seules les variables les plus utilisées par les policiers ont été incluses dans l'analyse. L'accent a été mis sur les variables à choix multiples qui concentrent le plus grand nombre de modalités de réponse et qui concernent des informations descriptives ou situationnelles. Les analyses se sont donc portées sur 22 variables à choix multiples présentant un total de 287 modalités de réponse. Toutes les modalités de réponse des variables sont binaires, c'est-à-dire qu'il est possible d'y répondre par oui (codé avec 1) ou non (codé avec 0). Les variables utilisées dans l'analyse fournissent des informations sur les victimes, les agresseurs, la phase précrime du mode opératoire ainsi que la phase de commission du crime.

Les variables relatives aux victimes sont au nombre de deux : 1) les facteurs de risque entourant le crime (13 modalités différentes); et 2) les activités routinières de la victime avant la commission du crime (24 modalités différentes).

Les variables relatives aux agresseurs sont également au nombre de deux : 1) les facteurs de risque entourant le crime (13 modalités différentes); et 2) les préférences sexuelles de l'agresseur (17 modalités différentes).

Les variables relatives à la phase précrime sont au nombre de 15 : 1) relation étrangère entre l'auteur et la victime (3 modalités de réponse), 2) relation de connaissance entre l'auteur et la victime (18 modalités de réponse), 3) relation familiale entre l'auteur et la victime (6 modalités de réponse), 4) approche par la ruse opérée par l'agresseur (21 modalités de réponse), 5) approche par la surprise opérée par l'agresseur (6 modalités de réponse), 6) approche éclair⁷ opérée par l'agresseur (5 modalités de réponse), 7) type de lieu, extérieur (24 modalités de réponse), 8) type de lieu, résidence (13 modalités de réponse), 9) type de lieu, véhicule (9 modalités de réponse), 10) type de lieu, lieu public (24 modalités de réponse), 11) type de lieu, travail (6 modalités de réponse), 12) type de lieu, *parking* (6 modalités de réponse), 13) type de lieu, divertissement (4 modalités de réponse), 14) type de lieu, plan d'eau (7 modalités de réponse), 15) accès au lieu (6 modalités de réponse).

Les variables relatives à la phase de crime sont au nombre de trois : 1) précautions prises par l'auteur au moment du crime (23 modalités de réponse), 2) actes commis (type 1) (26 modalités de réponse), actes commis (type 2) (13 modalités de réponse).

Stratégie d'analyse

La stratégie analytique utilisée s'articule en deux étapes. Dans un premier temps, le but est d'évaluer le degré d'utilisation des modalités des 22 variables. Pour ce faire, des analyses descriptives détaillent la fréquence des modalités utilisées dans moins de 5 %, 2,5 % et 1 % des cas.

Dans un second temps, des analyses factorielles en composantes principales (ACP) (Hair, Black, Babin et Anderson, 2010; Tabachnick et Fidell, 2001; Tacq, 1997) sont utilisées afin d'opérer une réduction sur des variables multidimensionnelles. Cette technique permet de prendre en considération les corrélations entre les différentes modalités d'une même variable à choix multiple. Afin de déterminer si l'analyse factorielle est satisfaisante, deux tests sont appliqués de manière préli-

7. Correspond à une attaque coercitive, violente.

minaire : le test de Kaiser-Meyer-Olkin (KMO), qui permet de mesurer la qualité de l'échantillonnage, ainsi que le test de sphéricité de Bartlett, qui permet de connaître l'intérêt de procéder à une réduction des dimensions en analysant la dépendance ou l'indépendance de ces composantes initiales. Afin de pouvoir procéder à une analyse factorielle satisfaisante, le test de KMO doit atteindre le seuil de 0,5, tandis que le test de sphéricité doit être significatif (0,000). Pour définir la limite de l'information importante et du bruit statistique, le critère Kaiser⁸ est appliqué. En ce qui concerne les nouveaux modèles de réduction, seules les corrélations ayant un alpha supérieur ou égal à 0,3⁹ ont été retenues comme cela est suggéré par la littérature (Peres-Neto *et al.*, 2003 ; Richman, 1988 ; Sanders, Gugiu et Enciso, 2015 ; Tabachnick et Fidell, 2001). La somme des pourcentages de variance expliquée par les axes de la représentation indique la part de l'information restituée par l'ACP. Plus celle-ci est proche de 100%, meilleure est l'analyse. En sciences sociales, on considère comme correct un modèle qui restitue au moins 50% de la variance (voir Hair *et al.*, 2010 ; Tabachnick et Fidell, 2001 ; Tacq, 1997).

Résultats

Analyse descriptive du contenu des variables

Le Tableau 1 présente les résultats de l'analyse des réponses aux 22 variables à choix multiple sur lesquelles porte cette analyse. En ce qui concerne les agressions sexuelles sur les victimes de 15 ans et plus, l'analyse globale révèle que, sur les 287 modalités utilisées, 65% concernent moins de 5% des cas, 48% moins de 2,5% des cas et 30% moins de 1% des cas. Pour les variables concernant les victimes, 68% des 37 modalités utilisées représentent moins de 5% des réponses et 30% moins de 1% des réponses. Les variables concernant les auteurs contiennent 30 modalités. Parmi celles-ci, 73% représentent moins de 5% des réponses et 43% moins de 1% des réponses. Pour les variables incluses dans la phase de précrime et présentant 158 modalités de réponses, 60% représentent moins de 5% des réponses et 26% moins de 1% des réponses. Finalement, dans la phase de commission du crime,

8. Ce critère retient les axes dont l'inertie est supérieure à l'inertie moyenne ($I/p = 1$).

9. L'utilisation d'un alfa supérieur ou égal à 0,5 ne présente pas de différence importante dans la détermination des modèles finaux par rapport au seuil de 0,3 (voir par exemple Jackson, 1993 ; Peres-Neto, Jackson et Somers, 2003).

TAB LEAU 1
Analyse des modalités de réponse pour les variables à choix multiples (N = 22)

Variables	Modalités totales	Victimes de 15 ans et +						Victimes de moins de 15 ans					
		< 5 % des réponses		< 2,5 % des réponses		< 1 % des réponses		< 5 % des réponses		< 2,5 % des réponses		< 1 % des réponses	
		n	%	n	%	n	%	n	%	n	%	n	%
Style de vie/activité routinière de la victime													
Facteurs de risque (victime)	13	6	46 %	3	23 %	3	23 %	8	62 %	4	31 %	2	15 %
Activités de la victime avant l'agression	24	19	79 %	13	54 %	8	33 %	18	75 %	15	63 %	9	38 %
Sous-total	37	25	68 %	16	43 %	11	30 %	26	70 %	19	51 %	11	30 %
Style de vie/activité routinière de l'auteur													
Facteurs de risque (auteur)	13	7	54 %	6	46 %	6	46 %	7	54 %	6	46 %	5	38 %
Préférences sexuelles de l'agresseur	17	15	88 %	10	59 %	7	41 %	12	71 %	11	65 %	8	47 %
Sous-total	30	22	73 %	16	53 %	13	43 %	19	63 %	17	57 %	13	43 %
Phase précrime													
Relation entre l'auteur et la victime (étranger)	3	2	67 %	1	33 %	0	0 %	2	67 %	2	67 %	0	0 %
Relation entre l'auteur et la victime (connaissance)	18	15	83 %	12	67 %	7	39 %	15	83 %	15	83 %	14	78 %
Relation entre l'auteur et la victime (familiale)*	6	0	0 %	0	0 %	0	0 %	2	33 %	2	33 %	2	33 %
Approche opérée par l'auteur (ruse)	21	14	67 %	12	57 %	5	24 %	15	71 %	15	71 %	12	57 %
Approche opérée par l'auteur (surprise)	6	2	33 %	1	17 %	0	0 %	2	33 %	2	33 %	1	17 %
Approche opérée par l'auteur (éclair)	5	3	60 %	2	40 %	1	20 %	4	80 %	3	60 %	1	20 %
Type de lieu (extérieur)	24	20	83 %	13	54 %	9	38 %	17	71 %	13	54 %	6	25 %

TABLEAU 1 (suite)

Variables	Modalités totales	Victimes de 15 ans et +						Victimes de moins de 15 ans					
		< 5 % des réponses		< 2,5 % des réponses		< 1 % des réponses		< 5 % des réponses		< 2,5 % des réponses		< 1 % des réponses	
		n	%	n	%	n	%	n	%	n	%	n	%
Type de lieu (résidence)	13	8	62 %	8	62 %	6	46 %	8	62 %	7	54 %	6	46 %
Type de lieu (véhicule)	9	6	67 %	5	56 %	4	44 %	8	89 %	6	67 %	5	56 %
Type de lieu (lieu public)	24	18	75 %	14	58 %	7	29 %	16	67 %	16	67 %	10	42 %
Type de lieu (travail)	6	2	33 %	2	33 %	0	0 %	3	50 %	2	33 %	1	17 %
Type de lieu (<i>parking</i>)*	6	1	17 %	0	0 %	0	0 %	1	17 %	0	0 %	0	0 %
Type de lieu (divertissement)*	4	2	50 %	1	25 %	0	0 %	1	25 %	1	25 %	1	25 %
Type de lieu (plan d'eau)*	7	1	14 %	1	14 %	1	14 %	1	14 %	1	14 %	0	0 %
Accès au lieu	6	1	17 %	1	17 %	1	17 %	0	0 %	0	0 %	0	0 %
Sous-total	158	95	60 %	73	46 %	41	26 %	95	60 %	85	54 %	59	37 %
Phase de commission du crime													
Précautions prises par l'auteur	23	16	70 %	11	48 %	9	39 %	21	91 %	15	65 %	10	43 %
Actes commis (type 1)	26	19	73 %	17	65 %	11	42 %	19	73 %	17	65 %	13	50 %
Actes commis (type 2)*	13	10	77 %	4	31 %	1	8 %	8	62 %	5	38 %	2	15 %
Sous-total	62	45	73 %	32	52 %	21	34 %	48	77 %	37	60 %	25	40 %
Total	287	187	65 %	137	48 %	86	30 %	188	66 %	158	55 %	108	38 %

*Les variables représentent moins de 5 % des cas.

TABLEAU 2

Réduction des dimensions du contenu des variables à choix multiples pour les agressions sexuelles sur les victimes de 15 ans et plus

Variables	KMO	Bartlett Sig.	Composantes initiales	Composantes principales	% Taux d'évolution	% V(X)
Style de vie/activité routinière de la victime						
Facteurs de risque (victime)	0,5	0,000	13	4	-69,23%	70,41%
Activités de la victime avant l'agression	0,5	0,000	24	8	-66,67%	76,25%
Sous-total			37	12	-67,57%	
Style de vie/activité routinière de l'auteur						
Facteurs de risque (auteur)	0,6	0,000	14	6	-57,14%	87,94%
Préférences sexuelles de l'agresseur	0,7	0,000	16	5	-68,75%	82,92%
Sous-total			30	11	-63,33%	
Phase précrime						
Relation entre l'auteur et la victime (étranger)	0,5	0,000	3	1	-66,67%	82,97%
Relation entre l'auteur et la victime (connaissance)	0,5	0,000	18	2	-88,89%	66,40%
Relation entre l'auteur et la victime (familiale)	0,5	0,000	6	2	-66,67%	56,56%
Approche opérée par l'auteur (ruse)	0,5	0,000	21	7	-66,67%	77,33%
Approche opérée par l'auteur (surprise)	0,5	0,000	6	3	-50,00%	87,03%
Approche opérée par l'auteur (éclair)	0,5	0,000	5	2	-60,00%	95,16%
Type de lieu (extérieur)	0,5	0,000	23	5	-78,26%	67,77%
Type de lieu (résidence)	0,5	0,000	13	4	-69,23%	87,74%

TABLEAU 2 (suite)

Variables	<i>KMO</i>	<i>Bartlett Sig.</i>	<i>Composantes initiales</i>	<i>Composantes principales</i>	<i>% Taux d'évolution</i>	<i>% V(X)</i>
Type de lieu (véhicule)	0,5	0,000	9	2	-77,78%	86,31%
Type de lieu (lieu public)	0,5	0,000	22	6	-72,73%	71,47%
Type de lieu (travail)	0,5	0,000	6	2	-66,67%	84,05%
Type de lieu (<i>parking</i>)	0,5	0,000	6	3	-50,00%	79,45%
Type de lieu (divertissement)	0,5	N.S.	***	***	***	***
Type de lieu (plan d'eau)	0,5	0,000	7	3	-57,14%	75,84%
Accès au lieu	0,5	0,000	6	2	-66,67%	67,11%
Précautions prises par l'auteur	0,6	0,000	23	7	-69,57%	75,92%
Sous-total			174	51	-70,69%	
Phase de commission du crime						
Actes commis (type 1)	0,7	0,000	26	8	-69,23%	81,76%
Actes commis (type 2)	0,7	0,000	16	5	-68,75%	75,03%
Sous-total			42	13	-69,05%	
Total			283	87	-69,26%	

qui est la dernière à inclure des variables à choix multiples, sur 62 modalités, 73 % représentent moins de 5 % des réponses et 34 % moins de 1 % des réponses.

Pour les agressions sexuelles sur les victimes de moins de 15 ans, la distribution générale des modalités activées pour qualifier les cas révèle que 66 % de ces modalités représentent moins de 5 % des réponses, 55 % moins de 2,5 % des réponses et 38 % moins de 1 % des réponses. L'analyse des variables concernant les victimes indique que 70 % de ses modalités représentent moins de 5 % des réponses et 30 % moins de 1 % des réponses. Les variables décrivant les auteurs ont 63 % de leurs modalités qui représentent moins de 5 % des réponses et 43 % d'entre elles qui représentent moins de 1 % des réponses. Pour les variables qualifiant la phase précrime, 60 % de leurs modalités représentent moins de 5 % des réponses et 37 % moins de 1 % des réponses. Finalement, les variables à choix multiples incluses dans la phase de commission du crime voient 77 % de leurs modalités représenter moins de 5 % des réponses et 40 % moins de 1 %.

Cette analyse permet de mettre en évidence le fait que pour l'ensemble des variables à choix multiples, environ deux tiers de leurs modalités représentent une quantité de réponses statistiquement négligeable. La plupart des cas qui sont traités dans la base sont donc qualifiables avec seulement une petite proportion de modalités. Il est donc tout à fait opportun d'envisager une réduction de la taille de ces variables dont les nombreuses modalités constituent plus de bruit statistique qu'elles ne permettent de qualifier le phénomène qu'elles sont censées représenter.

Réduction de la dimensionnalité des informations

Le Tableau 2 et le Tableau 3 ainsi que l'Annexe 1 synthétisent les résultats sur la réduction des dimensions du contenu des variables à choix multiples. En ce qui concerne les agressions sexuelles sur les victimes de 15 ans et plus, les résultats ont permis de révéler que 21 variables sur 22 pouvaient être réduites de manière statistique en mettant en évidence les *patterns* criminels importants et cohérents. Dans toutes les phases, le taux de réduction est supérieur à 60 % avec une moyenne pour le modèle global de 69,3 %. Le taux de réduction le moins élevé concerne les variables en lien avec l'auteur avec un taux de

réduction de 63,3 % et le plus élevé concerne la phase précrime avec un taux de réduction moyen de 70,67 %. Finalement, sur les 287 modalités utilisées pour qualifier les crimes sexuels pour ce groupe d'étude spécifique, les résultats indiquent qu'à partir d'une réduction et d'une « re-dimensionnalisation », un modèle à 87 nouvelles composantes était satisfaisant. Les résultats mettent en évidence que 196 modalités sont sous-exploitées et n'apportent pas d'informations suffisamment importantes pour être intégrées. Ainsi, 68,3 % des 287 modalités produisent du bruit statistique plutôt que de l'information exploitable.

Pour les agressions sexuelles sur les mineurs de 15 ans, les résultats ont mis en évidence que 18 variables sur 22 pouvaient faire l'objet d'une réduction. Ces 18 variables présentent, après réduction, des *patterns* criminels clairs et exploitables. Dans toutes les phases, le taux de réduction est ici également supérieur à 60 % avec un taux moyen de réduction de 70 %. La catégorie où la réduction est la moins prononcée est également celle des auteurs avec une réduction de 64,3 %, tandis que la phase où elle est la plus prononcée est également la phase précrime avec une réduction de 72,3 %. Sur les 240 modalités effectivement utilisées pour qualifier ces crimes, les ACP permettent d'identifier que 72 nouvelles composantes sont importantes. Ces nouvelles composantes sont le fruit de la suppression de certaines modalités initiales et du regroupement d'autres qui sont corrélés. Ainsi, le modèle proposé révèle que 168 modalités initiales du modèle, soit 70 %, ne remplissent pas correctement leur rôle en n'étant pas exploitées, en ne produisant pas suffisamment d'information et en venant parasiter l'information véritablement importante.

Discussion

Cette recherche a pour but d'évaluer la possibilité de réduire empiriquement le contenu de certaines variables de l'outil ViCLAS déterminées par les recherches précédentes comme particulièrement utiles d'un point de vue théorique et pratique (voir Chopin, 2017; Chopin et Aebi, 2019). Des analyses descriptives ainsi que des analyses factorielles en composantes principales ont été réalisées afin de répondre aux objectifs de cette étude.

TABLEAU 3

Réduction des dimensions du contenu des variables à choix multiples pour les agressions sexuelles sur les victimes de moins de 15 ans

Variables	KMO	Bartlett Sig.	Composantes initiales	Composantes principales	% Taux d'évolution	% V(X)
Style de vie/activité routinière de la victime						
Facteurs de risque (victime)	0,6	0,000	11	4	-63,64 %	74,36 %
Activités de la victime avant l'agression	0,5	0,000	23	7	-69,57 %	79,35 %
Sous-total			34	11	-67,65 %	
Style de vie/activité routinière de l'auteur						
Facteurs de risque (auteur)	0,5	0,000	12	5	-58,33 %	81,85 %
Préférences sexuelles de l'agresseur	0,5	0,000	16	5	-68,75 %	87,48 %
Sous-total			28	10	-64,29 %	
Phase précrime						
Relation entre l'auteur et la victime (étranger)	0,5	N.S.	***	***	***	***
Relation entre l'auteur et la victime (connaissance)	0,5	0,000	17	2	-88,24 %	62,61 %
Relation entre l'auteur et la victime (familiale)	0,5	0,000	4	1	-75,00 %	65,76 %
Approche opérée par l'auteur (ruse)	0,5	0,000	21	6	-71,43 %	87,91 %
Approche opérée par l'auteur (surprise)	0,5	0,000	6	2	-66,67 %	80,15 %
Approche opérée par l'auteur (éclair)	0,5	0,000	4	1	-75,00 %	92,49 %
Type de lieu (extérieur)	0,6	0,000	24	7	-70,83 %	73,10 %
Type de lieu (résidence)	0,5	0,000	11	3	-72,73 %	78,36 %

TABLEAU 3 (suite)

Variables	KMO	Bartlett Sig.	Composantes initiales	Composantes principales	% Taux d'évolution	% V(X)
Type de lieu (véhicule)	0,5	0,000	7	1	-85,71%	87,80%
Type de lieu (lieu public)	0,6	0,000	19	8	-57,89%	85,11%
Type de lieu (travail)	0,5	0,000	5	1	-80,00%	71,91%
Type de lieu (<i>parking</i>)	0,5	0,000	6	2	-66,67%	65,68%
Type de lieu (divertissement)	0,5	N.S.	***	***	***	***
Type de lieu (plan d'eau)	0,5	0,000	7	4	-42,86%	85,08%
Accès au lieu	0,5	0,000	6	2	-66,67%	69,54%
Précautions prises par l'auteur	0,6	0,000	18	3	-83,33%	68,71%
Sous-total			155	43	-72,26%	
Phase de commission du crime						
Actes commis (type 1)	0,7	0,000	23	8	-65,22%	82,62%
Actes commis (type 2)	N.S.	N.S.	***	***	***	***
Sous-total			23	8	-65,22%	
Total			240	72	-70,00%	

Le phénomène N = 1

Les résultats concernant la répartition de l'utilisation des modalités de chaque variable testée dans cette étude ont montré qu'environ deux tiers d'entre elles représentent moins de 5% des réponses, que la moitié d'entre elles représentent moins de 2,5% des réponses et qu'environ un tiers d'entre elles représentent moins de 1% des réponses. Ces premiers résultats illustrent bien le concept de Snook *et al.* (2008) du phénomène de N = 1. Le fait qu'une minorité de réponses représente une majorité des cas est symptomatique des difficultés rencontrées par le système à trouver des liens entre les affaires. La volonté des créateurs de l'outil de vouloir en faire un outil couvrant toutes les situations dans les moindres détails empêche l'émergence de *patterns* suffisamment larges pour permettre le rapprochement d'affaires présentant des similitudes. Un mélange d'informations spécifiques et de données plus générales crée un certain trouble dans l'identification des liens entre les affaires.

Vers une réduction empirique du contenu de ViCLAS et la création de patterns généraux

Les résultats des analyses factorielles permettant d'obtenir des modèles épurés sont relativement éloquentes et confirment que toutes les données introduites dans le système ne revêtent pas la même importance et le même rôle. Comme mentionné par Witzig (2003) pour le ViCAP, une réduction du ViCLAS s'avère possible puisqu'une partie de l'information requise par le système n'est pas fondamentale. Sauf rares exceptions, liées à leur quasi-absence d'utilisation, toutes les variables utilisées dans l'analyse peuvent être réduites de façon significative puisque les résultats les moins ambitieux laissent voir la possibilité d'une réduction de moitié des modalités. La réduction qui est proposée ici au moyen des ACP permet de faire le tri entre les informations importantes et les informations secondaires relevant du détail et faisant office de bruit statistique (Sheptycki, 2004). L'analyse des combinaisons des différentes modalités a permis soit de regrouper les informations corrélées qui représentent un *pattern* criminel particulier, soit d'exclure ces informations dont la part de variance expliquée entraîne plus de confusion qu'elle n'apporte d'information. La réduction du système suggérée par les études précédentes (Chopin, 2017; Chopin et Aebi, 2018, 2019; Ribaux, 2014; Snook *et al.*, 2012; Witzig, 2003) s'avère donc être une possibilité réelle, vraisemblable.

Les résultats de cette étude ont montré qu'il était possible de relever des *patterns* plus généraux pour chacune des variables en regroupant les informations importantes et corrélées. L'étape suivante de l'amélioration du système pourrait passer par le regroupement de ces informations importantes (c.-à-d. l'information sur la victime, le modus operandi ainsi que les données situationnelles) en *patterns* généraux afin de vérifier la sérialité des affaires présentant des *patterns* similaires. Ce processus a déjà été mis en place en Suisse avec la Plateforme d'Information du CICOP pour l'Analyse et le Renseignement (PICAR) dont l'un des objectifs est la détection des crimes sériels et particulièrement ceux liés à la propriété (ex.: cambriolage d'habitation) (voir Birrer, 2010; Grossrieder, 2017; Grossrieder *et al.*, 2013; Grossrieder et Ribaux, 2017). Dans ce système, les crimes sont classifiés en *patterns* établis sur la base des théories situationnelles en criminologie (Birrer, 2010). Plus spécifiquement, ces *patterns* sont définis en fonction du mode opératoire, de la voie d'entrée, du moment de la journée ainsi que du type de cible (Grossrieder, 2017). Cet outil présente un modèle de simplification qu'il serait intéressant de transposer à ViCLAS pour les crimes sexuels et/ou violents sans mobile apparent. À partir des variables utilisées dans cette étude et de la réduction proposée, il serait possible de créer des *patterns* englobant différents types d'information contribuant à identifier la signature d'un auteur à travers plusieurs crimes.

L'art délicat du juste milieu : la nécessité d'une approche collaborative

Toute la difficulté des outils de gestion de liens est de trouver le juste milieu entre la quantité d'informations nécessaire à la création des liens et la qualité de ces informations. L'objectif est d'améliorer l'identification de « vrais positifs » (c.-à-d. l'identification de cas réellement sériels) et de « vrais négatifs » (c.-à-d. l'identification des cas qui sont réellement non sériels) dans les liens entre les affaires. Concrètement, cela revient à passer de la création de « faux négatifs » (c.-à-d. l'identification erronée d'absence de sérialité) liée à l'excès d'information à la création de « faux » et « vrais positifs » liée à la réduction des informations avec l'utilisation des *patterns* criminels identifiés. Les faux négatifs empêchent nettement la création de tout lien entre les affaires tandis que les faux positifs favorisent la création de liens hypothétiques entre les affaires.

S'il est tout à fait préférable pour un outil de type ViCLAS de s'engager dans une voie de création de faux positifs, il est toutefois nécessaire de rester attentif au niveau d'analyse choisi afin de ne pas créer trop de liens inexistantes à cause de *patterns* trop généraux.

Le calibrage d'un tel équilibre dans un système de gestion de données apparaît comme une tâche difficile et probablement utopique. Il faut avoir conscience qu'aucun outil ne parviendra à produire des liens parfaits entre les crimes et qu'à partir des cas hypothétiquement liés qui seront mis en avant par le système, la réflexion humaine des analystes et des enquêteurs devra dissocier les liens véritablement existants des liens qui sont artificiellement créés. Dans le cas d'un outil d'aide à la détection de séries comme l'est le ViCLAS, il semble préférable d'accroître la détection de liens hypothétiques entre les affaires qui seront ensuite analysés méticuleusement par les analystes. Ceux-ci pourront alors juger en regard de leur expérience et d'informations spécifiques complémentaires de la pertinence d'accroître ou non les investigations sur les crimes potentiellement sériels. Il serait intéressant à ce stade de l'analyse que des techniques alternatives de priorisation d'hypothèses et de suspects soient utilisées afin de valider ou invalider les séries identifiées par le ViCLAS. Dans ce contexte, le profilage géographique ainsi que le profilage comportemental pourraient s'avérer être des renforts très importants et complémentaires aux techniques d'investigation traditionnelles (voir Chopin et Beauregard, sous presse; Deslauriers-Varin, Bennell et Bergeron, 2018b).

Conclusion

Cette recherche n'est pas sans limites. La première qui peut être évoquée concerne la validité externe de cette étude qui ne peut être testée qu'en comparant les résultats de cette étude avec ceux obtenus par d'autres travaux sur le sujet (voir Aebi, 2006 avec références). En effet, à notre connaissance, aucune autre analyse de ce type n'a été réalisée ou, à tout le moins, rendue publique. Dans cette perspective, les résultats obtenus ne sont pas généralisables au système ViCLAS en général et demeurent limités au système exploité en France, avec les données des crimes qui se sont déroulés dans un contexte français. Il serait tout à fait envisageable que la méthode d'enregistrement, qui varie entre les différentes polices, ait une incidence sur les résultats obtenus. Il est aussi tout à fait imaginable que les *patterns* obtenus à l'issue de la réduction varient d'un

pays à l'autre dans la mesure ou le contexte des agressions pourrait aussi être différent. Cependant, plusieurs travaux portant sur les comparaisons internationales des homicides sexuels montrent que les différences de contexte entraînent très peu de différences en ce qui a trait au mode opératoire et au processus criminel (voir par exemple Chopin et Beauregard, 2019c; James, Proulx et Lussier, 2018; Skott, Beauregard et Darjee, soumis).

Dans un second temps, cette étude est limitée par son échantillon et les résultats présentés ne sont valables que pour les cas d'agressions sexuelles qui ont été résolus par la police en France. Il est tout à fait possible que d'autres *patterns* se dégagent pour des crimes différents tels que les homicides sexuels ou les homicides non sexuels (voir Beauregard *et al.*, 2017; Beauregard et DeLisi, 2018a, 2018b; Chopin et Beauregard, 2018) ainsi que pour les affaires non résolues (voir Chopin, Beauregard, Bitzer et Reale, 2019).

Les résultats de cette étude permettent d'ouvrir une réflexion sur le rôle d'un outil informatique d'analyse criminelle. Dans le cas de ViCLAS, l'interprétation qui peut être faite est que l'objectif pour lequel il a été conçu, à savoir la création de liens entre les affaires, a été combiné dans la pratique avec celui de développer une banque de données d'archive dans laquelle on pourrait stocker un maximum d'informations sur ces affaires. Cela mène l'outil à jouer deux rôles relativement antinomiques, tant ces objectifs sont différents et devraient par conséquent être atteints en utilisant des méthodes différenciées.

Dans le cadre de l'analyse criminelle sur la délinquance sérieuse, l'objectif de l'outil sera la création de liens entre les affaires à partir de modèles de *patterns* criminels. Dans le cas d'un outil ayant pour but d'archiver des informations concernant les affaires criminelles, l'objectif sera de recenser un maximum d'informations différentes. Le facteur discriminant entre ces deux outils sera la méthodologie employée pour atteindre l'objectif ainsi que le niveau d'analyse envisagé. Dans le cas de l'outil opérationnel qui va servir à effectuer l'analyse criminelle, le niveau d'information doit être suffisamment large pour faire ressortir des *patterns* criminels permettant d'établir des liens entre les crimes qui ont des aspects communs. Dans le cas d'un outil d'archive, les informations devront être suffisamment spécifiques et précises pour qualifier chaque cas en restituant la réalité de ce qui s'est passé de la façon la plus précise possible. Les résultats de cette recherche suggèrent qu'en simplifiant une partie des informations présentement collectées par

ViCLAS, il serait possible de l'éloigner du rôle d'une banque de données d'archive et de le rapprocher du rôle d'un outil efficace pour l'analyse criminelle.

Les prochaines études pourraient ainsi reproduire cette recherche en suivant une méthodologie similaire sur des jeux de données issues du ViCLAS d'autres pays afin d'évaluer la généralisabilité, c'est-à-dire la validité externe, des conclusions de cette recherche. Les recherches à venir pourraient s'intéresser à mesurer empiriquement l'apport en termes d'efficacité de procéder à la réduction du ViCLAS. Il serait également pertinent de s'intéresser à la création empirique de *patterns* transversaux (c.-à-d. incluant différents aspects prépondérants dans la qualification des crimes) en utilisant des techniques de classification telles que l'analyse de classes latentes. Finalement, il serait intéressant de tester cette approche sur d'autres crimes tels que les homicides sexuels, les homicides non sexuels et les affaires non résolues afin de déterminer si, une fois réduites, les variables présentent des *patterns* similaires aux agressions sexuelles.

Références

- Aebi, M. F. (2006). *Comment mesurer la délinquance?* Paris, France: Armand Colin.
- Aebi, M. F. et Jaquier, V. (2008). Les sondages de délinquance autoreportée: origines, fiabilités et validité. *Déviante et Société*, 32(2), 205-227.
- Beauregard, E. (2005). *Processus de prédation des agresseurs sexuels sériels: une approche du choix rationnel* (Thèse de doctorat, Université de Montréal). Repéré à <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/16598>
- Beauregard, E. et DeLisi, M. (2018a). Stepping stones to sexual murder: The role of developmental factors in the etiology of sexual homicide. *Journal of Criminal Psychology*, 8(3), 199-214.
- Beauregard, E. et DeLisi, M. (2018b). Unraveling the personality profile of the sexual murderer. *Journal of Interpersonal Violence*. doi: 10.1177/0886260518777012
- Beauregard, E., DeLisi, M. et Hewitt, A. N. (2017). Sexual murderers: Sex offender, murderer, or both? *Sexual Abuse*, 30(8), 932-950.
- Beauregard, E., Leclerc, B. et Lussier, P. (2012). Decision making in the crime commission process: Comparing rapists, child molesters, and victim-crossover sex offenders. *Criminal Justice and Behavior*, 39(10), 1275-1295.
- Beauregard, E., Proulx, J., Rossmo, K., Leclerc, B. et Allaire, J. F. (2007). Script analysis of the hunting process of serial sex offenders. *Criminal Justice and Behavior*, 34(8), 1069-1084. doi: 10.1177/0093854807300851
- Beauregard, E., Rossmo, K. et Proulx, J. (2007). A descriptive model of the hunting process of serial sex offender: A rational choice perspective. *Journal of Family Violence*, 22, 444-463.

- Bennell, C., Snook, B., MacDonald, S., House, J. C. et Taylor, P. J. (2012). Computerized crime linkage systems: A critical review and research Agenda. *Criminal Justice and Behavior*, 39(5), 620-634.
- Birrer, S. (2010). *Analyse systématique et permanente de la délinquance sérielle: place des statistiques criminelles; apport des approches situationnelles pour un système de classification; perspectives en matière de coopération* (Thèse de doctorat, Université de Lausanne). Repéré à https://serval.unil.ch/resource/serval:BIB_183D2A096F2C.P001/REF.pdf
- Boba, R. (2009). *Crime analysis with crime mapping* (2 éd.). Sage Publications.
- Bourque, J., LeBlanc, S., Utzschneider, A. et Wright, C. (2009). *Efficacité du profilage dans un contexte de sécurité nationale*. Commission canadienne des droits de la personne. Repéré à <https://chrc-ccdp.gc.ca/fra/content/efficacite-du-profilage-dans-le-contexte-de-la-securite-nationale>
- Bradburn, N., Sudman, S. et Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design* (1^{re} éd.). San Francisco, États-Unis: John Wiley & Sons.
- Chopin, J. (2017). *La gestion des liens entre les crimes sexuels de prédateurs: repenser ViCLAS sous la perspective du paradigme situationnel* (Thèse de doctorat inédite). Université de Lausanne.
- Chopin, J. et Aebi, M. F. (2018). Les données de police permettent-elles la détection des agresseurs sexuels sériels? Une analyse de la validité concurrente de ViCLAS. *Revue Internationale de Criminologie et de Police Technique et Scientifique*, 71(1), 21-37.
- Chopin, J. et Aebi, M. F. (2019). The tree that hides the forest? Testing the construct validity of ViCLAS through an empirical study of missing data. *Policing: A Journal of Policy and Practice*, 13(1), 55-65. doi: 10.1093/police/pax062
- Chopin, J. et Beauregard, E. (2019a). The sexual murderer is a distinct type of offender. *International Journal of Offender Therapy and Comparative Criminology*, 63(9), 1597-1620. doi: 10.1177/0306624X18817445
- Chopin, J. et Beauregard, E. (sous presse). Contributions of psychological science on police response to sexual assault. Dans P. Marques et M. Paulino (dir.), *Police psychology*. Elsevier.
- Chopin, J. et Beauregard, E. (2019c). Sexual homicide in France and Canada: An international comparison. *Journal of Interpersonal Violence*. Prépublication. doi: 10.1177/0886260519875547
- Chopin, J. et Beauregard, E. (2019b). Sexual homicide of children: A new classification. *International Journal of Offender Therapy and Comparative Criminology*, 63(9), 1681-1704. doi: 10.1177/0306624X19834419
- Chopin, J., Beauregard, E., Bitzer, S. et Reale, K. (2019). Rapists' behaviors to avoid police detection. *Journal of Criminal Justice*, 61, 81-89. doi: 10.1016/j.jcrimjus.2019.04.001
- Chopin, J. et Caneppele, S. (2019a). The mobility crime triangle for sexual offenders and the role of individual and environmental factors. *Sexual Abuse: A Journal of Research and Treatment*. 31(7), 812-836. doi: 10.1177/1079063218784558

- Chopin, J. et Caneppele, S. (2019b). Geocoding child sexual abuse: An explorative analysis on journey to crime and to victimization from French police data. *Child Abuse & Neglect*, 91, 116-130. doi: 10.1016/j.chiabu.2019.03.001
- Ciavaldini, A. (1999). *Psychopathologie des agresseurs sexuels*. Paris, France: Masson.
- Collins, P. I., Johnson, G. F., Choy, A., Davidson, K. T. et MacKay, R. E. (1998). Advances in violent crime analysis and law enforcement: The Canadian violent crime linkage analysis system. *Journal of Government Information*, 25(3), 277-284. doi: Doi 10.1016/S1352-0237(98)00008-2
- Deslauriers-Varin, N. et Beauregard, E. (2013). Investigating offending consistency of geographic and environmental factors among serial sex offenders: A comparison of multiple analytical strategies. *Criminal Justice and Behavior*, 40(2), 156-179.
- Deslauriers-Varin, N., Bennell, C. et Bergeron, A. (2018a). Crime linkage et profilage criminel. *Cahiers de la Sécurité et de la Justice*, 43, 81-90.
- Deslauriers-Varin, N., Bennell, C. et Bergeron, A. (2018b). Criminal investigation of sexual offenses. Dans P. Lussier et E. Beauregard (dir.), *Sexual offending: A criminological perspective*. New York, NY: Routledge.
- Egger, S. A. (1984). A working definition of serial murder and the reduction of linkage blindness. *Journal of Police Science and Administration*, 12(2), 348-355.
- Egger, S. A. et Doney, R. H. (1990). *Serial murder: An elusive phenomenon*. Westport, CT: Praeger.
- Fox, J. A. et Levin, J. (1994). *Overkill: Mass murder and serial killing*. New York, NY: Plenum Press.
- Geberth, V. J. (1986). Mass, serial and sensational homicides: The investigative perspective. *Bulletin of the New York Academy of Medicine*, 62, 492-496.
- Gravier, B., Mezzo, B., Abbiati, M., Spagnoli, J. et Waeny, J. (2010). *Prise en charge thérapeutique des délinquants sexuels dans le système pénal vaudois, étude critique*.
- Grossrieder, L. (2017). *De la trace aux connaissances: une approche interdisciplinaire dans l'intégration des méthodes computationnelles en analyse et renseignement criminel* (Thèse de doctorat inédite). Université de Lausanne.
- Grossrieder, L., Albertetti, F., Stoffel, K. et Ribaux, R. (2013). Des données aux connaissances, un chemin difficile: réflexion sur la place du data mining en analyse criminelle. *Revue Internationale de Criminologie et de Police Technique et Scientifique*, 66, 99-116.
- Grossrieder, L. et Ribaux, O. (2017). Towards forensic whistleblowing? From traces to intelligence. *Policing: A Journal of Policy and Practice*, 13(1), 80-93.
- Hair, J. F., Black, W. C., Babin, B. J. et Anderson, R. E. (2010). *Multivariate data analysis: A global perspective*. Saddle River, NJ: Pearson.
- Jackson, D. A. (1993). Stopping rules in principal components analysis: A comparison of heuristical and statistical approaches. *Ecology*, 74(8), 2204-2214.
- James, J., Proulx, J. et Lussier, P. (2018). A cross-national study of sexual murderers in France and Canada. Dans J. Proulx, E. Beauregard, A. J. Carter, A. Mokros, R. Darjee et J. James (dir.), *Routledge International Handbook of Sexual Homicide Studies* (p. 171-195). New York, NY: Routledge.

- Katz, C., Webb, V. J. et Schaeffer, D. R. (2000). The validity of police gang intelligence lists: Examining differences in delinquency between documented gang members and nondocumented delinquent youth. *Police Quarterly*, 3(4), 413-437.
- Killias, M., Aebi, M. F. et Kuhn, A. (2012). *Précis de criminologie*. Berne, Suisse: Stämpfli.
- Leclerc, B., Proulx, J. et Beaugard, E. (2009). Examining the modus operandi of sexual offenders against children and its practical implications. *Aggression and Violent Behavior*, 14(1), 5-12.
- Margot, P. (2009). *VICLAS – SALCV – SALVAC: Violent Crime Linkage System/ Système d'analyse de liens dans les crimes violents/ de la violence associées aux crimes. Rapport d'évaluation sur les fondements scientifiques du système VICLAS destiné aux autorités policières du Canton de Vaud*. Lausanne, Suisse: Université de Lausanne
- Marin, J. C., Poirret, P., Quemener, M. et Gallois, A. (2003). *Analyse criminelle et analyse comportementale: rapport du groupe interministériel*. Ministère de la Justice.
- Martineau, M. M. et Corey, S. (2008). Investigating the reliability of the Violent Crime Linkage Analysis System (VICLAS) crime report. *Journal of Police and Criminal Psychology*, 60, 51-60.
- Pakkanen, T., Santtila, P. et Bosco, D. (2015). Crime linkage as expert evidence in serial offense trials. Dans J. Woodhams et C. Bennell (dir.), *Crime linkage: Theory, research, and practice* (p. 225-250). Boca Raton, FL: CRC Press.
- Peres-Neto, P. R., Jackson, D. A. et Somers, K. M. (2003). Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. *Ecology*, 84(9), 2347-2363.
- Ribaux, O. (2014). *Police scientifique: le renseignement par la trace*. Lausanne, Suisse: Presses polytechniques et universitaires romandes.
- Richman, M. B. (1988). A cautionary note concerning a commonly applied eigenanalysis procedure. *Tellus B*, 40(1), 50-58.
- Sanders, M., Gugiu, P. C. et Enciso, P. (2015). How good are our measures? Investigating the appropriate use of factor analysis for survey instruments. *Journal of Multidisciplinary Evaluation*, 11(25), 22-33.
- Sheptycki, J. (2004). Organizational pathologies in police intelligence systems: Some contributions to the lexicon of intelligence-led policing. *European Journal of Criminology*, 1(3), 307-332.
- Skott, S., Beaugard, E. et Darjee, R. (soumis). The consistency of sexual homicide characteristics and typologies across countries: A comparison of Canadian and Scottish sexual homicides.
- Snook, B., Cullen, R. M., Bennell, C., Taylor, P. J. et Gendreau, P. (2008). The criminal profiling illusion: What's behind the smoke and mirrors? *Criminal Justice and Behavior*, 35(10), 1257-1276. doi: 10.1177/0093854808321528
- Snook, B., Eastwood, J., Gendreau, P., Goggin, C. et Cullen, R. M. (2007). Taking stock of criminal profiling: A narrative review and meta-analysis. *Criminal Justice and Behavior*, 34(4), 437-453.

- Snook, B., Luther, K., House, J. C., Bennell, C. et Taylor, P. J. (2012). The violent crime linkage analysis system: A test of interrater reliability. *Criminal Justice and Behavior*, 39(5), 607-619. doi: 10.1177/0093854811435208
- Tabachnick, B. G. et Fidell, L. S. (2001). *Using multivariate statistics* (4^e éd.). Needham heights, MA: Allyn & Bacon.
- Tacq, J. (1997). *Multivariate analysis technique in social science research*. Londres, Royaume-Uni: Sage Publications.
- Villetta, P. (1993). Le libellé des items de délinquance: son effet sur les réponses. *Bulletin de criminologie*, 19(1), 100-132.
- Witzig, W. E. (2003). The new VICAP. *FBI Law enforcement Bulletin*, 72(6).

Identifying links among serial sexual crimes: Toward reducing the empirical information required for ViCLAS

ABSTRACT • *This study focuses on obtaining a better understanding of the amount of data required by a crime linkage system to identify serial violent crimes. The key objective of the research was to empirically test and discuss the relevance of reducing the content necessary to use these tools. Data for the current study was extracted from a French violent crime linkage database (ViCLAS). Analysis focused on the most used variables in the context of a sexual assault. Descriptive and factorial analysis were used to identify what kind of information is collected by the police and how it could be reduced by aggregating correlated data. Findings show that approximately 65% of the types of information collected are used in less than 5% of sexual assault cases. Results of factorial analysis show that it is possible to drastically reduce the information needed to use the tool effectively. Practical implications of these findings for sexual assault investigations are discussed.*

KEYWORDS • *Crime linkage system, data reduction, sexual crime, efficiency of systems.*

La gestión de las informaciones en el marco del análisis criminal de la delincuencia sexual serial: Hacia una reducción empírica del ViCLAS

RESUMEN • *Esta investigación se interesa por la cantidad de información necesitada por los sistemas de gestión de los vínculos que permiten identificar los casos seriales de criminalidad violenta. El objetivo de esta investigación es testear empíricamente y examinar la pertinencia de reducir el contenido de estos sistemas. Los datos utilizados provienen de una extracción de la base de datos ViCLAS en Francia. Los análisis fueron realizados sobre una muestra de las variables más utilizadas en el marco de las agresiones sexuales. Análisis descriptivos y factoriales son utilizados para identificar qué informaciones son recogidas por los policías y cómo estas podrían ser reducidas empíricamente, agregando los datos correlacionados. Los resultados muestran que alrededor del 65% de las modalidades testeadas son utilizadas en menos del 5% de los casos de agresiones sexuales. Los resultados de los análisis factoriales muestran*

que es posible reducir drásticamente el mecanismo. Las implicaciones prácticas de estos resultados son discutidos dentro de una perspectiva de investigación policial.

PALABRAS CLAVE • *Sistema de gestión de los vínculos, reducción de datos, agresión sexual, eficacia del sistema.*

ANNEXE 1
Nuage de points des ACP

