Sibling violence: validating a 2-factor model of severity in non-offender populations.

Roxanne Khan

School of Psychology
University of Central Lancashire

Correspondence concerning this article should be addressed to Dr. Roxanne Khan,
School of Psychology, Darwin Building, University of Central Lancashire, Preston, PR1 2HE, UK. E-mail: RKhan2@uclan.ac.uk
Abstract

Objective: Despite a recent surge of academic and clinical interest in sibling violence (SV), valid measures of severity have not been psychometrically established using non-offender populations. This study examined the factor-structure of intentional SV severity in a non-forensic sample considered to be not at ‘high-risk’ for violence, using the only existing empirically-driven model of severe SV committed with intent (Khan & Cooke, 2013). The prior model was established in a high-risk for violence, young offender sample (N=111; mean age=14.53) and revealed two underlying factors: ‘SV with weapon use’ and ‘SV without weapon use’. Method: This study examined data from an older, mixed community and student sample (N=899; mean=22.53) to test the factor structure and reliability of the existing severity model. Results: Participants reported a wide range of violent acts against their sibling(s) with aim of injuring them, including weapon use. Using exploratory factor analyses and confirmatory factor analyses, the prior 2-factor model was empirically supported using this non-correctional population. The new model comprised Factor 1 (potentially lethal SV) and Factor 2 (non-life threatening SV). Conclusion: The generalizability of the original 2-factor model, established using an offender sample, was demonstrated in this non-offender sample designated not at ‘high risk’ for violence.

Keywords: Abuse; Assessment; Bullying, Family Violence; Psychopathology; Sibling Aggression; Victimization; Weapon Use
Sibling violence: validating a 2-factor model of severity

The measurement inconsistencies underlying sibling violence (SV) research are a cause for concern, not least due to reports of elevated victimization rates during early siblinghood. Despite an increased awareness of SV being pervasive and injurious, research efforts are typically hindered by methodological constraints that result from lax definitions and weak or inconsistent measurement (Eriksen & Jensen, 2009). Due to these assessment limitations, researchers are forced to struggle with SV measurement at a conceptual level; this undermines the confidence held in empirical research to produce generalizable results that are applicable across populations, and reduces their ability to reliably inform policy and practice that aims to protect victims from familial abuse. The present study takes steps to address these concerns by testing the generalizability of an existing 2-factor model of SV severity, developed using a young offender sample, and validating its factor-structure in non-correctional populations who are not at ‘high risk’ for violence.

Prevalence of Sibling Violence

Although siblicide represents an extreme consequence of SV, it is perhaps surprising that national data reveal a frequency range of between only 1% and 8% in the United States (Gebo, 2002; Peck & Heide, 2012; Underwood & Patch, 1999). In 2002, the US Bureau of Justice recorded only 119 siblicide cases from a total of 9,102 family homicides (Harlow, Langan, Motivans, Rantala, & Smith, 2005). More recent data examinations revealed 1,002 siblicides in the Federal Bureau Investigation’s (FBI) Supplement Homicide Report between 2000 and 2007 (Walsh & Krienert, 2014). These figures can be misleading however, as they are clearly at odds with non-lethal SV rates; for example, numerous studies report a high frequency of SV victimization during childhood (Duncan, 1999), adolescence (Goodwin & Roscoe, 1990), and early adulthood (Button & Gealt, 2010; Khan & Rogers, 2015; Reese-Weber, 2008), with rates that range as high as 70% to 96% for those who claimed to have...
physically assaulted their sibling(s) (e.g., Eriksen & Jensen, 2009; Kettrey & Emery, 2006; Roscoe, Goodwin, & Kennedy, 1987). Although these are some of the highest prevalence rates reported in family violence research, it is not uncommon to find general estimates for SV that range between 30 to 60 percent (e.g., Hardy, 2001; Khan & Cooke, 2004; Tucker, Finkelhor, Shattuck & Turner, 2013; Tippett & Wolke, 2014; Rothman et al. 2010). These SV rates are not exclusive to European-American or British populations either, and are also reported in ethnic minority groups in the United States (Rapoza, Cook, Zaveri, & Malley-Morrison, 2010) and the United Kingdom (Irfan & Cowburn, 2004), and in comparable populations in other parts of the world, including Portugal (Relva, Fernandes, & Mota, 2013), Finland, Puerto Rico, Israel, and Canada (Steinmetz, 1981).

These studies form a volume of research that demarcate SV as an important area of family violence to investigate, with reports of both minor wounds (e.g., cuts and bruises; Straus & Gelles, 1990) and serious injuries (e.g., burns, puncture wounds, and broken limbs; Khan & Cooke, 2008) as well as long-term psychological effects of SV, including anxiety symptoms (Graham-Bermann, Cutler, Litzenberger, & Schwartz, 1994), depression (Hoffman & Edwards, 2004; Stocker, Burwell, & Briggs, 2002), substance abuse (Button & Gealt, 2010), eating disorders, and attempts at suicide (Wiehe, 1997). These findings have encouraged a theoretical shift from individual or psychoanalytical explanations of why SV might occur (for a review, see Whiteman, McHale, & Soli, 2011) to more testable evolutionary perspectives (e.g., Archer, 2013; Khan et al. 2016) and prevailing conflict, feminist, and social learning theories (Hoffman & Edwards, 2004).

**Conceptualizing Violence: Problems with Measurement and Assessment**

Despite the increased awareness of these detrimental victimization experiences, a persistent concern underlying extant SV research is a lack of uniformity in assessment that results from the divergent approaches used to investigate its occurrence across different
populations. The challenges of measurement are not isolated to SV incidents, and are well
documented in the literature pertaining to both spousal assault in normative or community
(non-offender) populations who are not at ‘high risk’ for violence (Barling, O’Leary, Jouriles,
Vivian, & MacEwan, 1987; Hornung, McCollough, & Sugimoto, 1981; Pan, Neidig, &
O’Leary, 1994), and instrumental aggression and violence in higher risk adult male prisoners
(Cooke, Michie, De Brito, Hodgins, & Sparkes, 2010; Michie & Cooke, 2006). Typically,
assessment concerns arise from the ‘one size fits all’ approach to measuring violent acts. That
is, a measure that ascertains one broad behavioral outcome (i.e., conflict or aggression)
experienced within a particular victim-offender relationship (that of a spouse, for example)
might be used to measure another, more specific, type of aggression (e.g., physical assault
with weapons) in an incomparable relational-dyad (e.g., strangers). Even within a same
victim-offender relationship (e.g., siblings), the measurable outcome may not be explicitly
defined (e.g., intentional aggression vs. play fighting vs. accidental harm) and thus, imprecise
measures will not adequately distinguish behaviors that might overlap due to ambiguity.
Tools used to assess SV may therefore lack utility, as they were designed to measure other,
unspecific behavioral constructs and are thus, too broad in scope.

Pertinently, the precision of measurement, despite being acknowledged as
fundamental element of valid violence assessment, is often overlooked in the selection of
instruments used to measure SV in psychological research. In a review of sibling aggression
studies, Archer (2013) usefully summarized the key methodological differences across
twenty studies published from 1960 to 2010. These inconsistences invariably resulted from
the use of different aggression measures. For instance, measures of sibling aggression ranged
from ‘fought: moderately to constantly’ (i.e., Koch, 1960; Kratcoski, 1985) and ‘high levels
of conflict’ (i.e., Graham-Bermann et al. 1994) to ‘pushing around or hitting: pretty often to
very often’ adapted from the bullying Peer Relations Questionnaire (Rigby & Slee, 1993)
Validating a 2-factor model of sibling violence

(i.e., Duncan, 1999). It is notable that the Conflict Tactics Scale [CTS] (Straus, 1979), the CTS-2 (Straus, Hamby, Boney-McCoy, & Sugarman, 1996), or a measure based on either of these two scales were used in twelve of the 20 studies in Archer’s review (e.g. Hardy, Beers, Burgess, & Taylor, 2010) as well as numerous other studies not included in the review (e.g., Eriksen & Jensen, 2009; Hendy, Burns, Can, & Scherer, 2011; Mangold & Koski, 1990; Noland, Liller, McDermott, Coulter, & Seraphine, 2004; Reese-Weber, 2008; Relva et al. 2013; Rothman et al. 2010; Simonelli, Mullis, Elliott, & Pierce, 2002). This is not surprising as the CTS/CTS-2 are used extensively in the aggression literature to measure dyadic familial conflict, and physical assault measures based on these two scales (albeit not explicitly the same) are also used in other SV studies (e.g., Felson, 1983; Felson & Russo, 1988; Finkelhor, Turner, & Ormrod, 2006; Khan & Cooke, 2004). Yet despite the use of CTS/CTS-2 based measures in these studies, other methodological anomalies were noted, such as sample characteristics (e.g., community vs. college/university vs antisocial youth populations), the data source (e.g., self-report vs. parental-report) and the time frame of recorded incidents (e.g., lifetime or when residing with sibling vs. previous 12 months) (cf. Archer, 2013).

Furthermore, studies that considered more serious acts of physical SV, and thus used the CTS Severe Violence sub-scale (e.g., Ketteray & Emery, 2006; Mackey, Fromuth, & Kelly, 2010) or a variant of this (e.g., Khan & Cooke, 2004), do not always distinguish intentional SV acts intended to cause harm from those committed accidently, or resulting from play-fighting.

These measurement inconsistencies will explain, to some degree, the widely disparate SV prevalence rates noted in Archer’s review, which at its most prolific, ranged from 18.8% (c.f. Hardy et al. 2010) to 82% (c.f. Mackey et al. 2010) for two seemingly comparable self-report questionnaire studies, both using large samples from American university-student populations, with an approximate mean age of 20 years. Upon closer scrutiny, it is also perplexing that the higher rate was recorded for SV incidents occurring during a restricted 12-
month period, at age 13 years, compared with the former study, which gathered data on incidents perpetrated during participants’ lifetime or while living with their siblings. While there will be explanations for the discrepancy in estimates, it is reasonable to contend that the diversity of the measures used in SV research demonstrates an imprecision, heterogeneity, and seemingly arbitrary approach to assessing SV, despite the portentous consequences reported by victims in cases of severe violence.

Rationale for Establishing a 2-Factor Model of SV Severity

The importance of distinguishing SV on the basis of severity was raised by Caffaro and Conn-Caffaro (1998) who, from their clinical experience, recognized differences between less serious forms of sibling aggression that can be more readily defined as ‘rivalry’, from more severe and intentionally harmful acts that could be recognized as ‘assault’. Eriksen and Jensen (2009) also maintained that the lack of distinction of SV severity in assessment has had a detrimental effect on the generalizability of research findings from one study to another, and across populations. Using the CTS in a sample of 994 families, they differentiated between SV that was “less severe” (79.1%) and more “severe” (14%). The lesser SV acts were conceptually distinguishable as less harmful and widely experienced as a developmentally-related consequence of siblinghood, while more serious acts included acute victimization experiences that were potentially injurious (i.e., beating up and the use, or threat of, weapons). Different risk factors were also found for “severe” SV perpetration (explained by parent-to-child violence and unpredictability) than for “less severe” SV (e.g., contextual factors, such as family environment) – a variance that is reported in other SV studies using samples from normative (Khan & Cooke, 2004), clinically-referred (Tompsett, Mahoney, & Lackey, 2016), and forensic (Khan & Cooke, 2008) populations. These findings reinforce a need for empirical validation of this conceptual distinction; that is, to find support for the differentiation between intentional SV that may be rooted in dysfunctional-
Validating a 2-factor model of sibling violence

psychopathology, and less harmful acts of aggression that might occur as a result of other circumstantial influences (Eriksen & Jensen, 2009; Rosenthal & Doherty, 1984). This lack of distinction not only brings into question the validity and significance of reported prevalence rates, but is also likely to impinge on clinicians’ attempts to target intervention for SV cases that may be driven by a pathological etiology. This greatly reduces the ability of research efforts that explore SV risk or protective factors, to reliably inform policy and practice and thus, ultimately, any legal or clinical attempts made to protect children and young people at risk from physical harm or serious injury when living with abusive siblings.

Drawing on these concerns, Khan and Cooke (2013) examined this conceptualized distinction and established an empirically-driven model of SV severity. This resulted in a clear 2-factor model reflecting (1) more serious SV (i.e., SV with weapon use) and (2) less serious SV (i.e., SV without weapon use) using a higher risk for violence sample of young people being dealt with by the criminal justice system for their criminal and/or antisocial behavior, and whom, therefore might be expected to engage in SV to a greater extent (both frequency and severity). Perkins (2014) notes, however, that this model limited its focus on severe violence in an ‘at risk’ sample of young offenders. This provided a strong rationale to conduct the current study, which aimed to test the existing 2-factor model of SV severity using a sample not at ‘high risk’ for violence, to validate its generalizability across populations.

**Research Aims**

This study aims to cross-validate Khan and Cooke’s (c.f. 2013) 2-factor model of intentional SV severity (1-SV with weapon use and 2-SV without weapon use) which, based on a thorough review of this literature, is the only empirically established model of SV severity currently available. The 2-factor model was originally validated in a ‘high-risk’ young offender sample (n=111) as a result of being generalized from an adult sample of
validating a 2-factor model of sibling violence (n=250; mean age: 26.8 years, SD=5.9). Therefore, the present study used a non-offender sample (n=899) recruited from community-dwelling and student populations, and were thus designated to represent a population not at ‘high-risk’ for violence. Two analytical procedures were used to determine the best fitting model. In Analyses 1, exploratory factor analyses were conducted on the same 10 SV items used in Khan and Cooke’s (c.f. 2013) model, then with an additional 3 items (n=13) to reflect less severe SV in a non-offender sample. In Analyses 2, confirmatory factor analyses were run on both the 10 SV items then 13 SV items, with a series of errors covariances being added to these models based on the modification indices obtained in the initial analyses conducted. Internal consistency was also tested to validate the model structure’s reliability.

**Methods**

**Participants**

A total of 899 males (n=373) and females (n=526) volunteered to participate in the current study. Participants’ ages ranged from 16 to 55 years (mean=22.53; SD=5.99; median=21.00; mode=20). With exception of number of siblings, the current sample’s characteristics corresponded approximately with Khan and Cooke’s sample (denoted by an *). A majority of the participants were Caucasian (85%:100%*), and part of a sibling–dyad (43.3%:18.9%*), while the remainder formed a sibling-triad (30.2%:33.3%*) or had three or more siblings (26.5%:47.7%*). Although a large proportion of participants were raised by either both or one of their birth parents (82.7%:90%*), a small number were raised by their birth father (11.7%:2.7%*) or extended family (e.g., grandparent(s)) (5.6%:2.7%*).

In line with Khan and Cooke’s selection criteria, volunteers in the present study were required to have at least one sibling who was raised in the same household as them; participants were asked to report on SV experiences during that time period of living with their sibling(s). It is recognized that this timeframe would vary per participant according to
their age and period of time in the same residence, yet, it is consistent with measures used in Khan and Cooke’s study. Also, this timeframe was used to record SV instances in the majority of studies in Archer’s review (cf. 2013).

Measures

Each participant completed a questionnaire booklet, placed within an information briefing sheet and debriefing sheet, to provide information on their use of SV against any one of their siblings. Instructions asked participants to report on any incidents that had occurred in which they had intentionally and purposefully (but not accidentally or playfully) committed an act of violence against their sibling(s) whilst living with them when growing up.

In selecting this study’s measures, inconsistencies across more commonly used scales were noted as such: the CTS uses 9 physical assault items (3 minor; 6 severe), the CTS-2 uses 12 physical assault items (5 minor; 7 severe) and the Severe Violence Index uses 5 items of SV severity. Additionally, these three scales combine several SV acts and present them as one item (i.e., “kicked, bit, or hit with a fist” and “threatened with a gun or knife”). Thus, to separate SV acts listed as one item in these three scales, the present study used a total of 13 severe physical assault items, from which 7 reflected severe SV acts without weapon use (denoted by an *), and 6 items characteristic of severe SV with weapon use. The final 13 severe items aimed to measure intentional SV, and comprised the following 10 items used in Khan and Cooke’s (c.f. 2013) study: (1*) Have you kicked or bitten your sibling(s)? (2*) Have you punched your sibling(s)? (3*) Have you thrown something heavy or sharp at your sibling(s)? (4*) Have you battered or beaten your sibling(s)? (5) Have you attempted to strangle your sibling(s)? (6) Have you threatened your sibling(s) with a knife? (7) Have you wounded your sibling(s) with a knife? (8) Have you threatened your sibling(s) with a gun? (9) Have you fired a gun at your sibling(s)? (10) Have you used another serious form of aggression against your sibling(s)? (e.g., hanging, burning them).
An additional 3 items were also included to reflect the less severe violence expected in non-offender sibling conflict experiences: (11*) Have you pushed your sibling(s)? (12*) Have you grabbed your sibling(s)? (13*) Have you slapped your sibling(s)? Consistency was maintained with previous studies (c.f. Khan & Cooke, 2013) by recording responses on a 6-point scale (0=“never”; 1= “once”; 2= “rarely”; 3=“sometimes”; 4=“often”; 5=“very often”).

**Procedures**

Subsequent to institutional ethics committee approval, which adheres to the British Psychological Society (BPS) guidelines, participants were recruited via opportunity sampling. Four research assistants distributed questionnaires to potential participants. Students were recruited via undergraduate and postgraduate classes plus common areas within a large University in the north of England. Potential participants were also approached in the local community, inside shopping centers and outside colleges. As part of a standardized recruitment procedure, when approached, potential volunteers were informed that in order to take part in this study, they must have at least one sibling with whom they were raised in the same household with. Those who fulfilled this criteria and agreed to take part were given the questionnaire booklet placed inside a self-addressed envelope; this was a precautionary measure taken to ensure the confidentiality and safety of respondents recruited within the community so that they could return questionnaires if they wished, anonymously and without pressure from the researchers. In order to obtain informed consent, potential participants were verbally informed of the research topic, plus the anonymous and voluntary nature of the questionnaire; this was also clearly stated on the briefing sheet, which participants were asked to closely read in their own time, to decide if they were willing and able to take part in the study. This also informed participants to complete the questionnaire honestly and without conferring, and that returning their questionnaire indicated consent for their information to be used. On campus, participants were told they could return completed
questionnaire in their own time to a secured drop-in box located in a Student Resources room or via the university’s internal post system. No tokens or rewards were offered for participation in this study, yet a vast majority of participants returned their completed questionnaires with a high response rate of 75 percent recorded.

**Statistical Analyses**

The analyses conducted consists of exploratory factor analyses (EFAs), confirmatory factor analyses (CFAs), and reliability analyses. The EFAs and reliability analyses were conducted in SPSS v. 22, while the CFAs were conducted in Amos v. 21. Additional descriptive statistics were also conducted using SPSS 22. In order to mix the university sample (n=584) and community sample (n=315), the full dataset was split approximately into half through the use of a random number generator in SPSS to create two 50:50 samples.

In Analyses 1, EFAs were conducted on the first sample (n=473), and in Analyses 2, CFAs were conducted on the second sample (n=426). In Analyses 2, two separate CFAs were conducted that contained slight differences with respect to the second latent variable included within these models. The first CFA incorporated the following six measures into the first latent variable: hang/burn, fired a gun, threatened with a gun, used a knife, threatened with a knife, and attempted to strangle. The second latent variable incorporated the following four measures: kicked/bitten, punched, hit with a heavy/sharp object, and battered/beaten. For the second CFA conducted, this same set of ten indicators were incorporated, with the following additional three indicators: slapped, pushed, and grabbed. Both models incorporated a covariance between the two factors included within each model; after these models were initially run, modification indices were calculated and implemented in order to improve model fit. All modifications made consisted of covariances specified between the errors associated with the latent variable indicators. No covariances were specified between indicators from two separate latent variables.
Results

The distributions of all 13 SV measures analyzed in this study are presented in Table 1. These show that while the distributions differ, more severe forms of physical violence were found to be least common, with less severe forms of SV being much more frequent. It was noted that although this data was from a non-offender population, just under one-fifth (16.7%: n=150) to over two-thirds (70.2%: n=631) of this not ‘at risk’ for violence sample reported committing each of the 10 severe acts on one occasion or more. Roughly between one-quarter (27%; n=243) to one-third of these participants (37.5%; n=337) reported never having committed any of the three additional acts (i.e. slapped, pushed, and grabbed) considered to be more likely to occur in lower risk populations.

Analyses 1: Exploratory Factor Analyses

The initial EFA focused upon the primary set of 10 SV items, to initially determine an appropriate factor structure with these new data before running the CFA. First, the KMO measure of sampling adequacy was found to be .840, with Bartlett's Test of Sphericity found to achieve statistical significance, $\chi^2(45) = 2375.156$, $p < .001$. These results indicate the appropriateness of factor analysis with respect to these data. The communalities for this analysis showed the extracted factors explain between approximately 53% and 78% of the variance in these items. Using the Kaiser criterion, a total of two factors would be selected, that explain approximately 65% of the variance in these items. These two initial factors were retained in this analysis and the component matrix following varimax rotation is presented in Table 2.

As shown, Factor 1 consisted of six items with factor loadings ranging from .61 to .88 that explained 37.56% of the total variance. Factor 1 was found to be associated with the
following items: attempted to strangle, threatened with a knife, used a knife, threatened with a gun, used a gun, and hang/burn. Factor 2 contained four items, with loadings ranging from .67 to .88 that explained 27.4% of the total variance. Factor 2 was associated with having: kicked/bitten, punched, thrown something heavy or sharp, and battered/beaten. Factor 1 focused upon potentially lethal physical violence, whilst those on Factor 2 focused upon acts of non-life-threatening physical violence. Overall, this EFA with 10 SV items produced a strong factor solution with clear separation between these factors.

Following this, a second EFA was conducted focusing upon all 13 items; these comprised the original 10 severe items used in Khan and Cooke’s study, plus the three additional items selected to reflect the SV acts more likely to occur in normative samples, not high risk for violence. In this analysis, the KMO measure of sampling adequacy was found to be .876, with Bartlett's Test of Sphericity found to achieve statistical significance, $\chi^2(78) = 3344.312, p < .001$. These results indicate the appropriateness of factor analysis with respect to these data. The communalities for this analysis showed the extracted factors explained between approximately 52% and 77% of the variance in these items. Using the Kaiser criterion, a total of two factors were retained in this analysis which explained close to 62% of the variance in these items. Table 3 presents the varimax rotated component matrix which shows that Factor 1 consisted of six items with factor loadings ranging from .61 to .87 that explained 32.3% of the total variance.

Factor 1 was associated with the following items: attempted to strangle, threatened with a knife, used a knife, threatened with a gun, fired a gun, and hang/burn. Factor 2 contained seven items with loadings that ranged from .63 to .83 that explained 29.4% of the total variance. Factor 2 was associated with: kicked/bitten, punched, thrown something heavy or sharp, and battered/beaten, along with the new measures of slapped, pushed, and grabbed.
Factor 1 items reflected acts of potentially lethal physical violence, whilst Factor 2 focused upon acts of non-life-threatening physical violence. This EFA with 13 SV items produced a strong factor solution with clear separation between these factors.

**Analyses 2: Confirmatory Factor Analyses**

Figure 1 presents the initial CFA run on these data, focusing upon the initial set of 10 SV items. As shown, the first factor was specified to load upon the following items: hang/burn, used a knife, fired a gun, threatened with a gun, threatened with a knife, and attempted to strangle. The second factor was specified to load on battered/beaten, kicked/bitten, thrown something heavy or sharp object, and punched.

The Bentler Comparative Fit Index (CFI; Bentler, 1980), the Bentler-Bonnett Non-Normed Fit Index (NNFI; Bentler, 1980), the root mean square of approximation (RMSEA; Steiger & Lind, 1990), and normed chi-square were utilized to evaluate the fit of the model to the data. The normed chi-square statistic for the proposed model was $\chi^2/df = 5.171$. Standardized estimates were all found to be high and all factor loadings that were not constrained to be equal to one, were found to achieve statistical significance at the .001 alpha level. While the measures of model fit generally indicated good model fit (CFI = .981, TLI = .963, RMSEA = .068), overall, this factor structure was found to be appropriate (Oishi, 2007).

Figure 2 presents the second CFA conducted on all 13 SV items. In this model, the additional three items of slapped, pushed, and grabbed were all specified to be associated with Factor 2. Other than this addition, the factor structure remained the same.

The normed chi-square for the proposed model was $\chi^2/df = 5.677$, and again, standardized estimates were found to be high in all cases with statistical significance found in all cases in which the path was not constrained to be equal to one. Additionally, the measures
of model fit again indicated good model fit (CFI = .970, TLI = .948, RMSEA = .072), while overall these results indicate an appropriate factor structure being present (Oishi, 2007).

While non-normality produces maximum likelihood parameter estimates that are less likely to be biased, this tends to increase standard errors of the maximum likelihood parameters. This means that there will be a lower likelihood of finding significant results, with model fit indices also likely to be underestimated, and model chi-square statistics being inflated (Wang & Wang, 2012). However, in Analyses 2, statistical significance was found in all cases, along with good model fit. For these reasons, the normality of the variables was not considered a concern with respect to these two CFAs.

**Reliability Analyses**

A series of reliability analyses were conducted focusing upon the factors identified in the previous EFAs and CFAs. Cronbach’s alpha coefficients were used as a measure of the internal consistency reliability of these scale measures. Cronbach’s alpha coefficients of these scales were as follows: 10-item scale, non-life-threatening SV (Cronbach’s α = .845), 10-item scale, potentially lethal SV (Cronbach’s α = .883), 13-item scale, non-life threatening (Cronbach’s α = .892), 13-item scale, potentially lethal SV (Cronbach’s α = .883).

**Discussion**

This study examined if a 2-factor model of sibling violence severity, previously established in a young offender population, could be generalized to an older, non-offender sample not at ‘high risk’ for violence. Several interesting findings emerged as a result of this investigation. Most saliently, self-reports from this mixed community-student sample illustrated that committing severe violence against a sibling was not necessarily associated with having a criminological history, or possibly any engagement in delinquency at all. For example, over one-fifth of this non-forensic sample designated to be not at ‘high risk’ for violence had sometimes, often, or very often kicked/bitten, punched, and/or thrown a sharp or
heavy object at their sibling(s) with the intention of causing them harm whilst living with them. Although more commonly reported on just one occasion, a smaller proportion still had threatened to, or had used a weapon (i.e., knife or gun), attempted to hang/burn, or strangle their siblings(s) with aim of injuring them. These prevalence rates add to the growing research that highlights the injurious nature of this often minimized and normalized form of family abuse.

These findings also suggest that despite the pervasiveness of severe SV in siblinghood, there may be an important distinction between the types of acts committed when there is an intent to cause harm. In Analyses 1, EFAs were conducted with the same 10 items used in Khan and Cooke’s original model, two dimensions underlying severe SV were revealed: Factor 1 (potentially lethal SV) and Factor 2 (non-life threatening SV). The first factor captured facets of violent behavior against siblings that included attempted to strangle, threatened with a knife, used a knife, threatened with a gun, fired a gun, and hang/burn. The second factor encapsulated the following items that represented non-life threatening physical SV: having kicked/bitten, punched, thrown something heavy or sharp, and battered/beaten a sibling(s). When three additional items were included to reflect violent acts reportedly more typical in non-correctional populations, the two factors contained the same items and all three additional factors loaded onto Factor 2 (non-life threatening SV). For Analyses 2, CFAs validated the acceptability of the 2-factor model. Overall, these findings cross-validated the previously established 2-factor model of intentional severe SV originally found with an ‘at risk’ sample of young offenders to an older, non-offender population not at ‘high risk’ for violence. This demonstrated generalizability of the duel-structure underlying the 2-factor model in terms of SV severity. Inferentially, these findings provide empirical support for investigations that had conceptually distinguished between potentially pathologically-rooted SV, from physical aggression against siblings that may be a common developmental-

**Limitations:** Although the increased research interest in SV has improved awareness of this abuse, research efforts are often hindered by a myriad of methodological limitations that reduce the generalizability of existing findings, from one study to another. The current study is not without such problems. While this study made efforts to reliably test the 2-factor structure of the existing SV model using a non-offender sample, adhering to Khan and Cooke’s (c.f. 2013) research protocol may have hindered, rather than strengthened, the current study’s design. For example, the timeframe in which SV incidents occurred was not recorded, so the influence of developmental-phases could not be examined. As the deleterious effect of time on accurate memory recall has been established, it would be advantageous for future studies’ designs to allow for model invariance to be tested across sample characteristics, such as current age and age (or developmental-stage) at time of victimization. While it is desirable for further replication studies to better match samples to boost the model’s generalizability, it is reasonable to expect some reliability issues associated with retrospective, self-report data validity (i.e., social desirability and memory bias) to impact on all participants, regardless of their age or personal experience. Although efforts were made to recruit participants from the community so as not to conduct the study solely on a relatively young, well-educated sample of university students, future studies would benefit from examining the validity of this 2-factor model in wider, mixed or demographically-focused (e.g. gender-specific) populations. It would also be advantageous to measure offense history. The current study was unable to obtain ethical clearance to ask participants directly about their involvement in criminal activity, and there may indeed be individuals with criminal justice histories in this sample designated not at ‘high risk’ for violence. However, as the
current model provided a good fit despite these limitations, it is a good indicator for the robustness of this 2-factor model.

**Research Implications:** A positive aspect of these findings is the attempt to provide clarity and uniformity, at an empirical level, to conceptually-based SV assessment. This can improve measurement precision in future research studies that attempt to investigate SV prevalence rates across different populations, correlates in terms of severity, and explorations into protective and risk factors that serve to guide valid intervention attempts. These findings also add weight to the importance of classifying violence in terms of severity, which has been noted in the criminological literature on young offenders (Kenny & Press, 2006) and research that has examined the efficacy of dating violence measures, in which parallels have been made between the legal classifications of assault (“simple” and “aggravated”) and the “minor” and “severe” violence classifications in the CTS (Straus & Mickey, 2012).

**Clinical and Policy Implications:** The present findings therefore add to the growing volume of research that emphasizes a need for severe SV to be addressed, within a criminological context and led by the authorities who develop evidence-based policies to combat family violence. Aggression researchers have argued that if the victim-perpetrator relationship were any other than that of siblings, many of the violent acts reported in psychology studies using non-offenders samples would readily be classified within a legal-context as criminal assault (Eriksen & Jensen, 2009; Khan & Rogers, 2015; Krienert & Walsh, 2011a). It is pertinent to note that while legal safeguards do not exist to protect against physical SV (Stock, 1993), criminal legislature is in place to protect victims from sibling sexual abuse, thus research in this area is more visible in the formulation of social policy (Perkins & O’Connor, 2016).

Partly explained by minimization and normalization, the high physical SV prevalence rates are to some extent not wholly unsurprising; more so perhaps, when considered in the
context of estimates that globally, the majority of people are raised with siblings during childhood (Whiteman et al. 2011). Motivations for SV are often explained as reflecting the stages of cognitive, affect-motor, moral and ethical development (Rosenthal & Doherty, 1984). Therefore in populations not ‘at high’ risk for violence, SV has been reported to result from a range of issues in which siblings are forced to negotiate, practice, and develop their conflict resolution skills, including caretaking responsibilities, division of labour, shared resources, property disputes, perceived favouritism, and power issues (Caffaro & Conn-Caffaro, 1998; Felson, 1983; Raffaelli, 1992; Wiehe, 1997). Evidence from the evolutionary literature also suggests that the inherent power imbalance and non-elective nature of sibling relationships plus demand for resources during this period, may prime siblings for violent conflict (Black, Mock & Parker, 1997). To this end, if SV is deemed to be characteristic of most normative sibling relationships, the current findings have potential to guide the development of education programs that assist families in their use of successful conflict resolution (Tucker & Finkelhor, 2015).

Maybe future research and practice interventions would be better aligned if focused on the predictors and correlates associated with the different degrees of SV severity (e.g., potentially lethal vs non-life threatening) in different populations. Previous research in this area, although currently scarce, has found different risk factors in relation to SV severity in youths from lower risk community samples (cf. Eriksen & Jensen, 2009; Khan & Cooke, 2004) and higher risk samples from clinically-referred (cf. Tompsett et al. 2016) and offender populations (cf. Khan & Cooke, 2008). Distinctly, severe SV in the criminal group was symptomatic of a pathological etiology, as predicted by personality disorder (i.e., primary psychopathy), arson, animal abuse, plus physical assaults and verbal abuse against school staff. The association between severe SV and psychopathology is supported in a recent study that makes a valuable comparison between higher risk forensic and clinical samples (Kuay et
al., 2016). In an examination of case notes, the former group were found to have a history of fire-setting and disruptive behavior disorders, were more aggressive and likely to use weapons. These findings suggest that there may be value in examining the association between SV severity and its co-occurrence with other types of sibling abuse, within this context of psychopathology. For example, it is possible that different degrees of physical SV might reflect concurrent familial sexual abuse perpetration, in terms of severity, including weapon use. Although sibling sexual abuse is more commonly associated with coercion and manipulation, a large-scale examination of 13,013 incidents using National Incident-Based Reporting System data (2000–2007) revealed that only a minority of incidents (6%) occurred in the presence of weapons (excluding hands and feet), and that injurious outcomes were reported in only a proportion of these cases (11%) (Krienert & Walsh, 2011b). These findings indicate a potentially pathological element to more extreme types of sibling abuse (both physical and sexual), and thus, the present 2-factor distinction of SV severity should be of interest to clinical practitioners involved in risk assessment and designing or delivering interventions, in child abuse cases and as part of family therapy.

To conclude, the study of sibling violence is a relatively recent advancement in the aggression literature, and this to some extent explains some of the assessment difficulties currently associated with investigations into this underestimated form of family abuse. The current study makes an important step in providing some clarity on the assessment to an area of violence research that has been, to date, fraught with difficulties.

References


representative sample of high school students in Boston, Massachusetts. *Archives of Pediatrics & Adolescent Medicine, 164*(12), 1118-1124. doi.10.1001/archpediatrics.2010


Validating a 2-factor model of sibling violence


Table 1: The Percentage and Frequency of 13 SV Items Reported by Participants (n=899)

<table>
<thead>
<tr>
<th>Sibling Violence Items</th>
<th>Never</th>
<th>Once</th>
<th>Rarely</th>
<th>Sometimes</th>
<th>Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kicked or bitten</td>
<td>29.8%</td>
<td>17.9%</td>
<td>23.2%</td>
<td>19.5%</td>
<td>6.5%</td>
<td>3.1%</td>
</tr>
<tr>
<td></td>
<td>268</td>
<td>161</td>
<td>209</td>
<td>175</td>
<td>58</td>
<td>28</td>
</tr>
<tr>
<td>Punched</td>
<td>31.1%</td>
<td>19.8%</td>
<td>20.2%</td>
<td>17.7%</td>
<td>7.8%</td>
<td>3.3%</td>
</tr>
<tr>
<td></td>
<td>280</td>
<td>178</td>
<td>182</td>
<td>159</td>
<td>70</td>
<td>30</td>
</tr>
<tr>
<td>Thrown heavy/sharp</td>
<td>44.0%</td>
<td>19.4%</td>
<td>16.5%</td>
<td>12.5%</td>
<td>4.9%</td>
<td>2.8%</td>
</tr>
<tr>
<td></td>
<td>396</td>
<td>174</td>
<td>148</td>
<td>112</td>
<td>44</td>
<td>25</td>
</tr>
<tr>
<td>Battered/Beaten</td>
<td>64.2%</td>
<td>16.6%</td>
<td>8.8%</td>
<td>5.7%</td>
<td>3.6%</td>
<td>1.2%</td>
</tr>
<tr>
<td></td>
<td>577</td>
<td>149</td>
<td>79</td>
<td>51</td>
<td>32</td>
<td>11</td>
</tr>
<tr>
<td>Attempted to strangle</td>
<td>74.9%</td>
<td>17.6%</td>
<td>3.1%</td>
<td>1.9%</td>
<td>1.7%</td>
<td>.8%</td>
</tr>
<tr>
<td></td>
<td>673</td>
<td>158</td>
<td>28</td>
<td>17</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>Threatened with knife</td>
<td>79.4%</td>
<td>17.5%</td>
<td>1.3%</td>
<td>1.0%</td>
<td>.6%</td>
<td>.2%</td>
</tr>
<tr>
<td></td>
<td>714</td>
<td>157</td>
<td>12</td>
<td>9</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Used a knife</td>
<td>82.6%</td>
<td>14.5%</td>
<td>1.0%</td>
<td>1.0%</td>
<td>.3%</td>
<td>.4%</td>
</tr>
<tr>
<td></td>
<td>743</td>
<td>130</td>
<td>9</td>
<td>9</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Threatened with gun</td>
<td>80.8%</td>
<td>14.8%</td>
<td>.7%</td>
<td>1.9%</td>
<td>1.0%</td>
<td>.8%</td>
</tr>
<tr>
<td></td>
<td>726</td>
<td>133</td>
<td>6</td>
<td>17</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>Used a gun</td>
<td>83.3%</td>
<td>14.3%</td>
<td>1.0%</td>
<td>.7%</td>
<td>.2%</td>
<td>.4%</td>
</tr>
<tr>
<td></td>
<td>749</td>
<td>129</td>
<td>9</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Hang/Burn</td>
<td>81.2%</td>
<td>15.6%</td>
<td>1.2%</td>
<td>.9%</td>
<td>.6%</td>
<td>.6%</td>
</tr>
<tr>
<td></td>
<td>730</td>
<td>140</td>
<td>11</td>
<td>8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Slapped</td>
<td>37.0%</td>
<td>19.2%</td>
<td>18.6%</td>
<td>13.6%</td>
<td>7.8%</td>
<td>3.7%</td>
</tr>
<tr>
<td></td>
<td>333</td>
<td>173</td>
<td>167</td>
<td>122</td>
<td>70</td>
<td>33</td>
</tr>
<tr>
<td>Pushed</td>
<td>27.0%</td>
<td>12.1%</td>
<td>21.8%</td>
<td>23.5%</td>
<td>9.6%</td>
<td>6.0%</td>
</tr>
<tr>
<td></td>
<td>243</td>
<td>109</td>
<td>196</td>
<td>211</td>
<td>86</td>
<td>54</td>
</tr>
<tr>
<td>Grabbed</td>
<td>37.5%</td>
<td>15.5%</td>
<td>18.7%</td>
<td>17.4%</td>
<td>6.5%</td>
<td>4.6%</td>
</tr>
<tr>
<td></td>
<td>337</td>
<td>139</td>
<td>168</td>
<td>156</td>
<td>58</td>
<td>41</td>
</tr>
</tbody>
</table>
Table 2: Analyses 1 - Exploratory Factor Analysis on 10 items: Factor Loadings, Eigenvalues, and Percentage of Variance Explained

*Rotated Component Matrix*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Component</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
<td></td>
</tr>
<tr>
<td>Kicked or bitten</td>
<td>.016</td>
<td>.871</td>
<td></td>
</tr>
<tr>
<td>Punched</td>
<td>.011</td>
<td>.876</td>
<td></td>
</tr>
<tr>
<td>Thrown something heavy or sharp</td>
<td>.199</td>
<td>.754</td>
<td></td>
</tr>
<tr>
<td>Battered/Beaten</td>
<td>.304</td>
<td>.665</td>
<td></td>
</tr>
<tr>
<td>Attempted to strangle</td>
<td>.613</td>
<td>.328</td>
<td></td>
</tr>
<tr>
<td>Threatened with knife</td>
<td>.743</td>
<td>.243</td>
<td></td>
</tr>
<tr>
<td>Used a knife</td>
<td>.880</td>
<td>-.025</td>
<td></td>
</tr>
<tr>
<td>Threatened with gun</td>
<td>.731</td>
<td>.155</td>
<td></td>
</tr>
<tr>
<td>Used a gun</td>
<td>.865</td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td>Hang/Burn</td>
<td>.799</td>
<td>.126</td>
<td></td>
</tr>
<tr>
<td>Eigenvalues</td>
<td>4.331</td>
<td>2.168</td>
<td></td>
</tr>
<tr>
<td>Percentage of variance accounted for</td>
<td>37.564</td>
<td>27.432</td>
<td></td>
</tr>
</tbody>
</table>

*Note: Factor loadings on the corresponding factor are in boldface*
Table 3: Analyses 1 - Exploratory Factor Analysis on 13 items: Factor Loadings, Eigenvalues, and Percentage of Variance Explained.

Rotated Component Matrix

<table>
<thead>
<tr>
<th>Measure</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kicked or bitten</td>
<td>.011</td>
<td>.833</td>
</tr>
<tr>
<td>Punched</td>
<td>.007</td>
<td>.828</td>
</tr>
<tr>
<td>Thrown something heavy or sharp</td>
<td>.200</td>
<td>.695</td>
</tr>
<tr>
<td>Battered/Beaten</td>
<td>.297</td>
<td>.634</td>
</tr>
<tr>
<td>Attempted to strangle</td>
<td>.613</td>
<td>.310</td>
</tr>
<tr>
<td>Threatened with knife</td>
<td>.739</td>
<td>.260</td>
</tr>
<tr>
<td>Used a knife</td>
<td>.876</td>
<td>.014</td>
</tr>
<tr>
<td>Threatened with gun</td>
<td>.738</td>
<td>.097</td>
</tr>
<tr>
<td>Used a gun</td>
<td>.863</td>
<td>.031</td>
</tr>
<tr>
<td>Hang/Burn</td>
<td>.796</td>
<td>.149</td>
</tr>
<tr>
<td>Slapped</td>
<td>.057</td>
<td>.802</td>
</tr>
<tr>
<td>Pushed</td>
<td>.198</td>
<td>.748</td>
</tr>
<tr>
<td>Grabbed</td>
<td>.201</td>
<td>.735</td>
</tr>
</tbody>
</table>

Eigenvalues  
5.323  2.704

Percentage of variance accounted for:  
32.341  29.411

Note: Factor loadings on the corresponding factor are in boldface
Figure 1: Analyses 2 - Initial Confirmatory Factor Analysis: Structural Model of 10 SV Items.

Note. $\chi^2$ (23) = 118.929, p < .001; Normed $\chi^2 = 5.171$, TLI = .963, CFI = .981, RMSEA = .068, 90% CI [.056, .081]
Validating a 2-factor model of sibling violence

Figure 2: Analyses 2 - Second Confirmatory Factor Analysis: Structural Model of 13 SV Items.

Note. $\chi^2 (45) = 255.443, p < .001$; Normed $\chi^2 = 5.677$, TLI = .948, CFI = .970, RMSEA = .072, 90% CI [.064, .081]