Stability of peer victimization: A meta-analysis of longitudinal research

Behavioural Science Institute, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 25 October 2014
Revised 26 January 2016
Available online 11 March 2016

Keywords:
Peer victimization
Stability
Meta-analysis
Longitudinal

ABSTRACT

A meta-analysis was conducted of 77 longitudinal studies that contained at least one over-time correlation (range 1 to 36) between scores for peer victimization measured at different time points. The overall stability of self-reported peer victimization was determined at centered values (age 10, one-year interval). The effects of interval length, age, and type of informant (self, peer, teacher, other/combined) on the stability of victimization were also examined. Moderate overall stability of self-reported victimization at age 10 across a 1-year interval was found. Stability decreased across larger longitudinal intervals. Peer- and other/combined-reports of peer victimization yielded higher stability estimates than self-reports. Teacher-reports yielded stability estimates that were equal to those for self-reports. An interaction was found between age and informant type (peer vs. self), indicating a larger increase in the stability of victimization with age for peer-reports than for self-reports. Implications for further research and practice were discussed.

© 2016 Elsevier Inc. All rights reserved.

Introduction

Peer victimization is a common problem in schools. Children and adolescents are victimized when they are exposed to harmful behavior, repeatedly and over time, and when they are unable to defend themselves (Olweus, 1994; Solberg & Olweus, 2003). It has been estimated that 9 to 32% of youths in Western and non-Western countries are victims of bullying (Stassen Berger, 2007). These rates are...
alarming, because victimization is associated with serious negative outcomes, including internalizing and externalizing problems and poor school adjustment (Hawker & Boulton, 2000; Kochenderfer & Ladd, 1996; Reijntjes et al., 2011; Reijntjes, Kamphuis, Prinzie, & Telch, 2010).

Stable victimization is even more serious than incidental victimization. The chronic stress model states that children who are persistently exposed to stressful events are at greater risk than those who are temporarily exposed (Dohrenwend & Dohrenwend, 1981). In line with this model, stable victims show higher levels of internalizing problems (Burk et al., 2011; Juvonen, Nishina, & Graham, 2000; L. H. Rosen et al., 2009), reactive aggression (Camodeca, Goossens, Terwogt, & Schuengel, 2002), and social dissatisfaction (Kochenderfer-Ladd & Wardrop, 2001) than youths who are less persistently victimized.

An increasing number of longitudinal studies have examined the stability of peer victimization. The results have varied from low to high stability. For example, across studies the percentages of stable victims have ranged widely from 8 to 43% (e.g., Burk et al., 2011; Schäfer, 2005; Scholte, Burk, & Overbeek, 2013; Sourander, Helstelä, Helenius, & Piha, 2000); a variability that may very well be due to differences in study characteristics. Given the serious consequences of peer victimization in general and stable victimization in particular, a clear estimate of the stability of victimization is needed, as well as an understanding of the study factors that influence the estimate. To address these issues, we conducted a meta-analysis of a large number of previous studies to estimate the stability of self-reported victimization centered across 1 year at age 10. We also determined to what degree the following study characteristics influence this estimate: the length of the interval between data collection waves, the source of information on victimization (peer-, teacher-, other/combined- vs. self-reports), and the age of the participants.

Mechanisms of stable peer victimization

There may be different mechanisms of stable peer victimization. For some youths, victimization is stable due to the continuity of their dysfunctional social interaction patterns (Caspi, Elder, & Bern, 1987). When children enter a new peer group, bullies initially direct their aggression to different victims (Kochenderfer & Ladd, 1996; Perry, Perry, & Boldizar, 1990) and observe their responses. Bullies then restrict their aggression to a smaller group of victims who react in ways that are rewarding for the bullies, such as with reactive aggression, crying, or withdrawal. These peers are then likely to continue to be victimized (Kochenderfer & Ladd, 1996; Schwartz et al., 1998).

Continuity in the social environment also may contribute to the stability of victimization. The cumulative continuity model (Caspi et al., 1987) states that youths with maladaptive behavior tend to select environments that further reinforce their maladaptive responding. Victims are likely to become friends with other victims who also lack social skills (Browning, Cohen, & Warman, 2003; Hodges, Malone, & Perry, 1997; Salmivalli, Huttunen, & Lagerspetz, 1997). Thus victims limit their positive experiences with peers and opportunities to develop prosocial skills. Instead, they may acquire social skill deficits, dysfunctional interaction styles, and adjustment problems which contribute to continued victimization (Kochenderfer-Ladd & Wardrop, 2001; Scholte, Engels, Overbeek, de Kemp, & Haselager, 2007).

Further, victims often are rejected in the peer group and lack friend support (Hodges & Perry, 1999). The risks associated with bullying a low-status peer are low. Therefore, bullies can easily show their dominance by continuing to bully the same low-status peers – thus also contributing to the stability of victimization.

The effect of interval length on the stability of peer victimization

One study factor associated with the stability of victimization may be interval length. Some studies address children’s development within a school year (e.g., across 6 months); others over longer periods (e.g., across 12 to 132 months). In general, the stability of a developmental dimension is lower over longer time intervals (Wohlwill, 1973). There are different reasons why the stability of victimization also may decrease as time intervals increase. As time passes, there is more room for biological and environmental changes that contribute to variability in children’s behavior. Victimization also may change over time due to fluctuations in the composition of the peer group. Children, parents, and teachers may use strategies to reduce victimization, which also impact stability (Kochenderfer-Ladd & Skinner,
Finally, it is easier for participants to remember their answers to the previous assessment when the time interval is small (Benedict & Zgaljardic, 1998). Indeed, the stability of victimization decreased with increasing interval length in several studies. For example, for intervals of 6, 18, and 30 months, researchers have reported respective stability coefficients (Pearson's r) of .46, .33, and .14 (Giesbrecht, Leadbeater, & Macdonald, 2011), and .73, .36, and .28 (Kochenderfer-Ladd & Wardrop, 2001). As the decrease of stability with increasing interval length clearly differed between studies, we examined the effect of interval length on the stability of victimization meta-analytically, while also controlling for age and type of informant. Knowledge of this effect will help researchers to plan new data collections and school practitioners to plan student screenings (Rueger, Malecki, & Demaray, 2011).

The effect of informant type on the stability of peer victimization

By origin, self-reports are the most common measure of victimization (Stassen Berger, 2007). In addition, peers, teachers, and occasionally parents or independent observers or combinations of those, have been used. Self, peer, and teacher-reports tap into different aspects of victimization and are weakly to moderately related (Graham & Juvonen, 1998; Lohre, Lydersen, Paulsen, Maehle, & Vatten, 2011; Scholte et al., 2013). Self-reports tap into children's or adolescents' own subjective experiences, whereas peer and teacher-reports are based on daily interactions and observations (Hymel, Wagner, & Butler, 1990). When a child has established a certain position in the classroom, peers are likely to perceive this child’s behavior in ways that are consistent with it. Their reports may rely on the child’s previously established status as a victim rather than on recently witnessed interactions (Hymel et al., 1990; Scholte et al., 2013). In contrast, self-reports have been found to change within a period of two weeks based on the number of victimization incidents a student experienced (Nishina & Juvonen, 2005). Thus, peer-reports are less influenced by daily fluctuations in victimization than self-reports, leading to higher stability estimates. Indeed, previous work has shown higher stability coefficients for peer and teacher-reports ($r's = .63$ and $.76$) than for self-reports ($r = .44$) (e.g., Iyer, Kochenderfer-Ladd, Eisenberg, & Thompson, 2010). The current study also contributes to the literature by addressing the differences in the stability of peer victimization between types of informants.

The effect of age on the stability of peer victimization

A general developmental phenomenon is that (social) constructs become more stable with increasing age (e.g., Jiang & Cillessen, 2005; Roberts & DelVecchio, 2000). Plasticity decreases when children grow up, and it has been assumed that this also applies to peer victimization (Hymel & Swearer, 2015; Kochenderfer-Ladd & Wardrop, 2001; Perry, Kusel, & Perry, 1988; Wolke, Woods, & Samara, 2009). Victimization may become more stable with age because the structure of the peer group becomes more hierarchical in adolescence (Schäfer, Korn, Brodbeck, Wolke, & Schulz, 2005) and peer status carries a higher weight than in childhood (LaFontana & Cillessen, 2010). It is easier to escape victimization in peer groups with a low hierarchical structure than in groups with a high hierarchical structure (Schäfer et al., 2005). Victims also receive less friend support in adolescence than in childhood (Rigby & Slee, 1991) as peers risk to lose status when they support a victim. For these reasons it is more likely that victimized youths will remain victimized over time in adolescence than in childhood.

Previous studies have shown mixed evidence for a positive association between age and the stability of victimization. Some studies indeed found higher stability with age (Bellmore & Cillessen, 2006; Vaillancourt et al., 2011), but others did not (e.g., Salmivalli & Helteenvuori, 2007; Sweeting, Young, West, & Der, 2006). Given these mixed findings, the aggregation of results across studies in our meta-analysis clarifies the degree to which the stability of peer victimization changes with age. This also informs us of when it might be most optimal to target victimization. It is important to intervene at an age before victimization becomes highly resistant to change (Hanish & Guerra, 2004; Rueger et al., 2011).

The effect of age or developmental stage on the stability of victimization may depend on the type of informant. Understanding this will be important for researchers planning a new study – it may help them to decide which informants to use given the age of the planned participants. At younger ages,
social-cognitive development may impact how children perceive victimization. As bullying in early childhood is directed at many different victims, children may have difficulty distinguishing peers who are victimized incidentally versus repeatedly (Monks, Smith, & Swettenham, 2003). Perspective taking is needed to perceive victimization among peers, and it is not yet fully developed in early childhood (Marvin, Greenberg, & Mossler, 1976). Indeed, research using self- and teacher-reports has shown that some children may already be stable victims in kindergarten and first grade (Hanish et al., 2004; Kochenderfer & Ladd, 1996; Monks et al., 2003), while peer-reports of victimization are still unstable at this time (Monks et al., 2003; Pouwels & Cillessen, 2013).

As social cognitive skills develop, children may become more accurate in detecting peer victimization and the stability of their reports may increase as a result. Adolescents may rely on peers’ reputation and status which becomes more salient at this age when the peer group is more hierarchical (Schäfer, 2005). Thus, reputational bias may play a larger role in adolescence, contributing to a higher stability of peer-reports than self-reports (Hymel et al., 1990).

The present study

This study examined the stability of peer victimization in longitudinal research by means of a quantitative meta-analysis. Stability of peer victimization can be measured in different ways. In this study, it was operationalized as the correlation between victimization scores over time. This is in line with most studies reporting on the stability of peer victimization and meta-analytical studies on personality traits and social status (Jiang & Cillessen, 2005; Roberts & DelVecchio, 2000). The stability correlation indicates the extent to which the relative ordering of participants’ victimization within the reference group is maintained over time (Caspí, Roberts, & Shiner, 2005). Stability operationalized as a correlation does not discard the possibility of individual or mean-level change (Roberts & DelVecchio, 2000).

This study had three goals. The first goal was to estimate the stability of self-reported peer victimization centered at age 10 across a one-year interval. We were interested in self-reports as youths themselves are the most common informant of peer victimization (Stassen Berger, 2007). We focused on age 10, which marks the ending of elementary school and start of early adolescence, the period when victimization is expected to crystallize (Hymel & Swearer, 2015).

Our second goal was to test whether the stability of victimization systematically differs depending on interval length, type of informant (self, peer, teacher, other/combined), and age of the participants. Our hypotheses were that stability of victimization would be negatively associated with interval length and positively with participants’ age. We also expected that peer-, teacher-, and other/combined-reports of victimization would be more stable than self-reports.

Our third goal was to examine whether the association between the stability of victimization and age differs by informant, with self as the reference category (peer vs. self, teacher vs. self, other/combined vs. self). We expected a stronger positive association between age and stability for peer-reports than for self-reports. We expected peer-reports to be less stable than self-reports in early childhood but more stable than self-reports in adolescence. We did not expect that the effect of age for self-reports would differ from the effect of age for teacher- or other/combined-reports.

Method

Literature search

Fig. 1 provides a summary of the literature search. In January 2012, longitudinal studies on peer victimization were collected by means of a computerized literature search of the PsycINFO (from 1978 onward) and Web of Science (from 1991 onward) databases. All possible combinations of the keywords “victimization”, “harassment”, and “longitudinal” with “stability” were used. Further, the reference list of each identified study was reviewed to look for older studies that may have been relevant. Together, this first step resulted in 1212 studies.

In Step 2, the studies retrieved in Step 1 were reviewed based on the following inclusion criteria: (a) the construct of interest was peer victimization, including different types of peer victimization (e.g., verbal, relational, physical). Studies concerning other forms of victimization such as being a victim
of crime, violence, or partner victimization were excluded; (b) longitudinal studies measuring victimization at both the beginning and the end of at least one time period were included; (c) stability data had to be either Pearson’s $r$ for a continuous victimization measure or frequencies displayed in a $2 \times 2$ contingency table for a categorical measure for the study to be included; (d) victimization had to take place in a peer context, such as schools; studies of other contexts such as the workplace were excluded; (e) studies of clinical samples, intervention studies, or studies using a twin design were excluded. We also checked whether there was an overlap of samples between studies. In case of overlap, we only included the study that covered the largest time span. The abstracts of all 1212 studies first were scanned for these criteria, yielding 157 potential studies. These 157 were then evaluated in detail with the same criteria, yielding 53 studies that met the criteria for the meta-analysis.

In addition to these 53 studies, we contacted the authors from articles or dissertations that were not available online, but may have included a stability coefficient. We also contacted the authors of additional studies that did not meet the criteria initially because they did not report a stability coefficient but which would have been appropriate otherwise. We contacted them with a request to send
the additional results. When we did not receive a response, we sent reminders and searched for updated contact information or contacted co-authors. We sent an e-mail to the authors of 54 studies and authors of 30 studies responded. Based on their responses, 18 of the 54 studies were included. Some authors sent additional papers that they had published (n = 6), conference presentations (n = 1), and dissertations or manuscripts in preparation which covered a larger time span of the longitudinal study about which we contacted them (n = 5). These papers were evaluated against the inclusion and exclusion criteria and, subsequently, 6 studies were included. The final sample of the meta-analysis included 77 studies. Table 1 provides an overview of these studies.

Coding and entering of study characteristics and effect sizes

Stability of victimization

Stability of victimization was operationalized as Pearson’s r of continuous victimization scores at two time points (Caspi et al., 2005). Thirteen studies (17%) only included categorical data in 2 × 2 contingency tables (yes/no a victim at each time point). We computed tetrachoric correlations from these tables (Enzman, 2007; Hunter & Schmidt, 1990).

In 21 studies (27%), participants were assessed at more than two time points (waves). For these studies, the correlation coefficients for all possible combinations of test–retest intervals were entered. For example, when victimization was measured at three waves (say 1, 2, 3), \( r_{12} \), \( r_{13} \), and \( r_{23} \) were entered when they were available. In such cases, we controlled statistically for the dependency between such correlations as explained in Appendix A.

In four studies, correlation coefficients were reported for boys and girls separately. Because sex differences were reported only in a small number of studies we were not able to estimate the stability of victimization by sex. If correlations were reported by sex only, we computed the weighted stability coefficients for the entire sample and subsequently used those. Similarly, a composite stability coefficient was computed when stability was reported separately for different types of victimization or different type of informants within one study.

Study characteristics

For each stability coefficient, the following study characteristics were entered in the data file: (a) sample size, (b) interval length in months, and (c) mean age of the participants in years at the first time point of each interval. If only an age range was provided, the midpoint of the range was entered. If only grade was reported, age was estimated as the typical age of students in that grade level in their school system. Some study characteristics were entered that were fixed over different intervals within one study: (a) publication year, (b) percentage boys at the first wave of the study, (c) type of informant of victimization (self, peer, teacher, other/combined). Type of informant was coded as self-reported when participants reported on their own victimization, including different self-report scales such as the revised Olweus Bully–Victim Questionnaire (Solberg & Olweus, 2003). Type of informant was coded as peer-reported when peer nominations or peer ratings were used. Type of informant was coded as teacher-reported when teachers nominated students in their classroom for victimization or filled out student ratings. Finally, there was an other/combined-reported category which was chosen when the authors reported the stability of victimization for more than one informant in their article (combined). Most studies aggregated the reports from multiple informants into one aggregated measure. This made it impossible to know the stability coefficient for each informant type separately. In order to be consistent, we also combined the stability coefficients of the few studies that did present the estimates separately by type of informant. This category was also chosen when other informants, such as parents and independent observers (other) were used. The other informants always were part of the combined category.

All variables were entered by the first author. In addition, 20% of the included studies were randomly selected and two independent researchers identified the study characteristics and coded and entered them. In order to determine coder reliability, Cohen’s kappa was calculated for categorical data, and intra-class correlations for interval data. Agreement for the dependent variable was perfect (1.00). Agreement for the predictors was good to perfect (.81–1.00).
Table 1
Study characteristics.

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Sample Size (N)</th>
<th>Sample Age (T1)</th>
<th>% Male</th>
<th>Type of Informant</th>
<th>Nr. stability coefficients</th>
<th>Range interval length</th>
<th>Level</th>
<th>Collapsed Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biebl, DiLalla, Davis, Lynch, and Shinn (2011)</td>
<td>65</td>
<td>5</td>
<td>45</td>
<td>O/C: Self, Observation</td>
<td>1</td>
<td>132</td>
<td>Cat.</td>
<td></td>
</tr>
<tr>
<td>Boulton (1999)</td>
<td>89</td>
<td>9</td>
<td>49</td>
<td>Peer</td>
<td>1</td>
<td>5</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Browning et al. (2003)</td>
<td>162</td>
<td>Grade 3–5</td>
<td>46</td>
<td>Peer</td>
<td>1</td>
<td>12</td>
<td>Cat.</td>
<td></td>
</tr>
<tr>
<td>Camodeca et al. (2002)</td>
<td>215</td>
<td>7.6</td>
<td>47</td>
<td>Peer</td>
<td>1</td>
<td>12</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Davidson (2009)</td>
<td>98</td>
<td>Grade 6–8</td>
<td>52</td>
<td>Self</td>
<td>1</td>
<td>24</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Esbensen and Carson (2009)</td>
<td>1117</td>
<td>Grade 6–9</td>
<td>46</td>
<td>Self</td>
<td>1</td>
<td>6</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Fekkes (2005)</td>
<td>691–1118</td>
<td>10.5</td>
<td>50</td>
<td>Self</td>
<td>2</td>
<td>18</td>
<td>Cat.</td>
<td></td>
</tr>
<tr>
<td>Grouewald (2011)</td>
<td>346</td>
<td>5.5</td>
<td>51</td>
<td>O/C: Self, Peer</td>
<td>1</td>
<td>12</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Hanish et al. (2004)</td>
<td>126</td>
<td>4.3</td>
<td>54</td>
<td>Teacher</td>
<td>1</td>
<td>6</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Hodges, Boivin, Vitaro, and Bukowski (1999)</td>
<td>393</td>
<td>10.6</td>
<td>48</td>
<td>Peer</td>
<td>1</td>
<td>12</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Juvonen et al. (2000)</td>
<td>106</td>
<td>Grade 7–8</td>
<td>42</td>
<td>Self</td>
<td>1</td>
<td>12</td>
<td>Cont.</td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Sample Size (N)</th>
<th>Sample Age (T1)</th>
<th>% Male</th>
<th>Type of Informant</th>
<th>Nr. stability coefficients</th>
<th>Range interval length</th>
<th>Level</th>
<th>Collapsed Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kremer (2011)</td>
<td>153</td>
<td>Grade 3</td>
<td>53</td>
<td>Teacher</td>
<td>1</td>
<td>1</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>McLaughlin, Hatzenbuehler, and Hilt (2009)</td>
<td>848</td>
<td>Grade 6–8</td>
<td>51</td>
<td>Self</td>
<td>1</td>
<td>7</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Menzer, Oh, McDonald, Rubin, and Dashiell-Aje (2010)</td>
<td>518</td>
<td>Grade 6</td>
<td>52</td>
<td>Peer</td>
<td>1</td>
<td>7</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Miller et al. (2005)</td>
<td>113–201</td>
<td>KG–Grade 1</td>
<td>42</td>
<td>Self</td>
<td>3</td>
<td>6–18</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Monks et al. (2003)</td>
<td>102</td>
<td>5</td>
<td>67</td>
<td>O/C: Self, Peer</td>
<td>1</td>
<td>4</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Nishino et al. (2011)</td>
<td>524</td>
<td>12.2</td>
<td>44</td>
<td>Self</td>
<td>1</td>
<td>3</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Ostrov (2008)</td>
<td>120</td>
<td>3.7</td>
<td>43</td>
<td>Teacher</td>
<td>1</td>
<td>5</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Perren and Alsaker (2009)</td>
<td>378</td>
<td>KG</td>
<td>43</td>
<td>Teacher</td>
<td>1</td>
<td>18</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Rueger et al. (2011)</td>
<td>694</td>
<td>12.6</td>
<td>52</td>
<td>Self</td>
<td>1</td>
<td>6</td>
<td>Cat.</td>
<td></td>
</tr>
<tr>
<td>Salmivalli and Helteenvuori (2007)</td>
<td>238</td>
<td>11.5</td>
<td>48</td>
<td>Peer</td>
<td>3</td>
<td>4–12</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Salmivalli et al. (1998)</td>
<td>189</td>
<td>14.5</td>
<td>51</td>
<td>Peer</td>
<td>1</td>
<td>24</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Sapouna et al. (2012)</td>
<td>454</td>
<td>8.9</td>
<td>55</td>
<td>Self</td>
<td>1</td>
<td>2</td>
<td>Cat.</td>
<td></td>
</tr>
<tr>
<td>Authors (Year)</td>
<td>Sample Size (N)</td>
<td>Sample Age (T1)</td>
<td>% Male</td>
<td>Type of Informant</td>
<td>Nr. stability coefficients</td>
<td>Range interval length</td>
<td>Level</td>
<td>Collapsed Items</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-------------------</td>
<td>---------------------------</td>
<td>-----------------------</td>
<td>-------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Schäfer et al. (2005)</td>
<td>283</td>
<td>8.1</td>
<td>52</td>
<td>Self</td>
<td>1</td>
<td>72</td>
<td>Cat.</td>
<td></td>
</tr>
<tr>
<td>Schreier et al. (2009)</td>
<td>2100–5109</td>
<td>3.9</td>
<td>49</td>
<td>O/C: Self, Parent, Teacher</td>
<td>8</td>
<td>24–85</td>
<td>Cat.</td>
<td></td>
</tr>
<tr>
<td>Schwartz et al. (2009)</td>
<td>240</td>
<td>9</td>
<td>52</td>
<td>O/C: Peer, Teacher</td>
<td>1</td>
<td>12</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Sourander et al. (2000)</td>
<td>515</td>
<td>8.4</td>
<td>51</td>
<td>O/C: Parent, Child</td>
<td>1</td>
<td>96</td>
<td>Cat.</td>
<td></td>
</tr>
<tr>
<td>Sweeting et al. (2006)</td>
<td>100</td>
<td>16.4</td>
<td>41</td>
<td>Self</td>
<td>1</td>
<td>5</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Toper, Castellanos-Ryan, Mackie, and Conrod (2011)</td>
<td>82</td>
<td>12.5</td>
<td>40</td>
<td>Self</td>
<td>1</td>
<td>24</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Vail lancourt et al. (2011)</td>
<td>311</td>
<td>10.3</td>
<td>47</td>
<td>Peer</td>
<td>1</td>
<td>6</td>
<td>Cont.</td>
<td>Rel. &amp; Overt</td>
</tr>
<tr>
<td>Williford, Brisson, Bender, Jenson, and Forrest-Bank (2011)</td>
<td>139–156</td>
<td>12.3</td>
<td>54</td>
<td>Self</td>
<td>3</td>
<td>6–12</td>
<td>Cont.</td>
<td></td>
</tr>
<tr>
<td>Wolke et al. (2009)</td>
<td>484–501</td>
<td>10.2</td>
<td>47</td>
<td>Self</td>
<td>2</td>
<td>12–24</td>
<td>Cat.</td>
<td>V &amp; BV</td>
</tr>
<tr>
<td>Yeung and Leadbeater (2007)</td>
<td>432</td>
<td>6.5</td>
<td>50</td>
<td>Self</td>
<td>1</td>
<td>36</td>
<td>Cat.</td>
<td></td>
</tr>
</tbody>
</table>

O/C = Other/Combined-reported victimization. KG = Kindergarten. In the column “level”, we indicated whether the authors presented the stability by means of correlations (Cont. = Continuous) or 2x2 contingency tables (Cat. = Categorical). In the column “collapsed items” we indicated whether we calculated composite scores based on correlations that were reported separately by gender, type of victimization, ethnicity, or for victims and bully-victims. Rel. = Relational victimization, Phys. = Physical victimization, Verb. = Verbal victimization, Soc. = Social victimization, Rep. = Reputational victimization. V & BV: The authors made a distinction between the stability of victims and bully-victims.
Results

Demographics

Table 2 summarizes the characteristics of the 77 included studies. The number of stability coefficients (Pearson’s $r$) per study ranged from 1 to 36. The total number of available stability coefficients was 166 (two were excluded due to missing interval lengths).

Multi-level meta-analysis predicting stability of peer victimization

Given the hierarchical structure of the data (correlations nested within studies) (Raudenbush, 2002), a multilevel analysis was conducted in MIWin 2.23 (Rasbash, Charlton, Browne, Healy, & Cameron, 2011). This approach also made it possible to separate within and between study variance by means of a random effects model (Borenstein, Hedges, Higgins, & Rothstein, 2010; Hox, 2002).

Conceptual model

The conceptual model had two levels. Level 1 consisted of the correlations that were nested in Level 2, the studies. Stability of victimization (Level 1) was predicted by the variables interval length and age (Level 1) and type of informant (Level 2). Type of informant was a nominal variable consisting of self-, peer-, teacher-, and other/combined-reported victimization. It was coded by three dummy variables using the category ‘self’ as the reference category (Cohen & Cohen, 1983).

Statistical model

The conceptual model was implemented by a statistical model with some special features. We controlled for dependencies between correlations for multiple waves within one study, conducted a Fisher’s $r$-to-$Z$ transformation (Hayes, 1978), and modeled the known error variance (Maas, Hox, & Lensvelt-Mulders, 2004). We also checked for publication bias. Together, this resulted in a statistical model with three levels and several additional (dummy) variables. See Appendices A and B for the details and outcomes of the statistical model.

Final model

The results of the meta-analysis are presented in Table 3 (see Model 4). The first goal was to estimate the stability of self-reported peer victimization centered at age 10 across a one-year interval. The intercept showed that this stability coefficient was $r = .45$, indicating that self-reported victimization was moderately stable at age 10 across 1 year. The random effect at the correlation level was .14, indicating that correlations of about .59 were relatively high, and correlations of about .31 were relatively low.

<table>
<thead>
<tr>
<th>Number</th>
<th>Range</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Number of victimization stability coefficients within one study</td>
<td>1–36</td>
<td>1</td>
</tr>
<tr>
<td>Interval length ($m$)</td>
<td>1–132</td>
<td>12</td>
</tr>
<tr>
<td>Age at the first of the two waves ($y$)</td>
<td>3.7–19.5</td>
<td>11</td>
</tr>
<tr>
<td>Pearson’s $r$</td>
<td>.09–.93</td>
<td>.49</td>
</tr>
<tr>
<td>Sample size</td>
<td>65–5109</td>
<td>489</td>
</tr>
<tr>
<td>Number of studies with informant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>Peer</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Teacher</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Other/Combined</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Publication year</td>
<td>1995–2012</td>
<td>2009</td>
</tr>
<tr>
<td>% Male at first wave</td>
<td>21–67</td>
<td>49</td>
</tr>
</tbody>
</table>
Our second goal was to test whether the stability of victimization systematically differed by interval length, participants' age, and informant (Self, Peer, Teacher, Other/Combined). There was a negative effect of interval length on the stability of victimization. At age 10, across intervals of 6, 12, and 36 months, we found stability coefficients of self-reported victimization of .48, .45, and .33, respectively. Peer-reports ($r = .62$) and other/combined-reports ($r = .58$) differed significantly from self-reports ($r = .45$), but teacher-reports ($r = .57$) did not. Thus, peer- and other/combined-reports were more stable than self-reports at age 10 across a one-year interval. There was no significant effect for age on the stability of self-reported victimization.

Our third goal was to examine whether the association between the stability of victimization and age differed by informant, with self as the contrast (peer vs. self, teacher vs. self, other/combined vs. self). There was an interaction between type of informant (peer vs. self) and age, indicating that the increase in stability of peer-reported victimization with age was significantly stronger than the increase in stability of self-reported victimization with age. The change in teacher- and other/combined-reported victimization did not differ significantly from the change in self-reported victimization (see Fig. 2).

Simple effects analyses were run to determine the simple effects of age on the stability of victimization for informant self and peer. There was a significant positive age effect for peer-reports, $t(68) = 6.72$, $p < .001$, and no significant age effect for self-reports. The region of significance for the difference in stability between peer- and self-reports by age shows the ages at which both differed significantly from each other (see Appendix C for the analytic approach). The model estimated that self-reports were significantly more stable than peer-reports from ages 3.7 to 4.2 years, whereas peer-reports were significantly more stable than self-reports from ages 8.4 to 19.5 years (see Fig. 3).

**Discussion**

This study was a meta-analysis of the stability of peer victimization. The first goal was to estimate the stability of self-reported peer victimization centered at age 10 across a 1-year interval. Across 77 studies including 166 stability coefficients, this stability was estimated to be .45. Cohen (1977) placed
Fig. 2. Expected values of Pearson’s $r$ stability coefficient of peer victimization by age and type of informant across a one-year interval. Although we tested linear effects, curved lines are presented because of the Fisher’s $Z$ to $r$ transformation.

Fig. 3. Region of significance of the difference in stability between Peer- and Self-Reported Victimization (Peer minus Self).
the cut-off for a medium correlation at .30 and a high correlation at .50. While the estimate of .45 is within the moderate-to-high range, it approaches the latter value. Thus, we conclude that the 1-year stability of self-reported peer victimization is moderate to high at the end of elementary school (i.e., the beginning of early adolescence).

It should be noted that the stability of peer victimization, while substantial, is not quite as high as the estimated stability of other developmental constructs. The overall stability of victimization was lower than the one-year stability of peer acceptance and rejection ($r = .45–.60$) (Jiang & Cillessen, 2005), aggression ($r = .76$) (Olweus, 1979), internalizing behavior ($r = .50-.60$) (Reitz, Dekovic, & Meijer, 2005), and personality ($r = about .50$) (Roberts & DelVecchio, 2000). This might be because the stability of victimization depends not only on the stability of the victim’s behavior and personality, but also on the stability of the bully’s behavior.

The effect of interval length on the stability of peer victimization

Our second goal was to test whether the stability of victimization systematically differed depending on interval length, participants’ age, and type of informant. Our meta-analysis confirmed our first hypothesis that stability would be negatively associated with interval length. Stability decreased as interval length increased, similar to the stability of peer status, aggression, and personality (Jiang & Cillessen, 2005; Olweus, 1979; Roberts & DelVecchio, 2000). At age 10, across intervals of 6, 12, and 36 months, the model estimated the stability coefficients for self-reported victimization of .48, .45, and .33. This suggests that victimization is moderately stable within one school year and from one year to the other, but less stable over longer time intervals.

Stability of victimization may decrease because of biological and environmental changes that contribute to variability in developmental trajectories. This is a typical dynamic of social behavior as time passes. There also is a typical dynamic in victimization: levels of victimization have been shown to increase and decrease over time (Kochenderfer-Ladd & Wardrop, 2001) and aggression can be directed towards different victims. These changes contribute to a decrease in the stability over time. There also may be changes in victimization due to interventions. Teachers, parents, peers, and victims themselves may use explicit strategies to reduce victimization which may lead to mean-level changes and changes in the rank ordering of victimization in a classroom or school (e.g., Kochenderfer-Ladd & Skinner, 2002; Troop-Gordon & Ladd, 2015). Children also can change schools or classrooms, bringing them into new peer groups in which their position has to be redefined, possibly changing their victim status (Pellegrini & Bartini, 2000; Pellegrini & Long, 2002). When children form a new social group, different peers than before might become victimized, and it may take some time before victimization is directed towards a specific group of victims (Pellegrini & Long, 2002). This suggests that there may be a decrease in stability over school transitions. However, a few studies addressing this issue have shown that there is not a large decrease in stability over school or classrooms transitions (Paul & Cillessen, 2003; Salmivalli, Lappalainen, & Lagerspetz, 1998; Sweeting et al., 2006). Unfortunately, most studies in the current meta-analysis did not report whether children attended classes with different peers. And only a small part of the studies examined victimization across school transitions. Hence, we were not able to examine to what extent the stability of victimization decreases when peer groups change. Further research could compare the stability of victimization over school or classroom transitions with the stability of victimization within the same school or classroom.

Stability of victimization also may decrease with increasing intervals for methodological reasons, such as cumulative effects of measurement unreliability (Roberts & DelVecchio, 2000). It is also easier for participants to remember their answers on a previous assessment when the time interval is smaller (Benedict & Zgaljardic, 1998). Although these methodological reasons could not be disentangled from true change, the effect of interval length on stability coefficients of victimization may guide researchers in planning future data collection waves.

The effect of informant type on the stability of peer victimization

Our hypothesis was that peer-, teacher-, and other/combined-reported victimization would be more stable than self-reported victimization in early adolescence. In line with previous studies, this expectation
was confirmed for peer-reports and other/combined-reports (e.g., Hymel & Swearer, 2015; Iyer et al., 2010). Self-reports tap into children’s subjective experiences of victimization by examining the frequency of victimization. Peer-reports tap into youths’ reputation of being a victim in the classroom by examining their status relative to their peers. It has been argued that youths’ reputation among peers is more resistant to change than their own subjective experiences, leading to a higher stability of peer- than self-reports (Salmivalli, Kaukiainen, & Voeten, 2005). Indeed, when students have established a certain position in the classroom, peers are likely to perceive their behavior in ways that are congruent with the students’ position (Hymel et al., 1990; Scholte et al., 2013). In contrast, self-reports vary over short time-periods based on the number of victimization incidents youths experience (Nishina & Juvonen, 2005).

Although we expected that teacher-reports would be more stable than self-reports in early adolescence, this was not confirmed. It is possible that we were not able to find a significant effect due to limited power; the number of studies using teachers as informants was small. Although the effect of teacher-reports was comparable to the effect of peer- and other/combined-reports (.16 vs. .18; .23), it was not significant which is due to the large standard error. Future studies should compare self- and teacher-reports to determine whether their stabilities differ from each other.

The other/combined category consisted of studies with different type of informants within one study, sometimes including observations or parent reports. When the stability coefficient was based on combined reports, victimization was more stable than when based on self-reports only. One reason for the higher stability of other/combined- than self-reports could be that it may take some time before all informants have changed their perceptions when changes in victimization occur. It is important to note that this finding must be interpreted with caution because the type of informants that were combined in this category differed between studies.

**The effect of age by informant type on the stability of peer victimization**

Our next hypothesis was that stability of victimization would be positively associated with the age of the children or adolescents. This expectation was not confirmed for self-reports. The .45 estimate for the 1-year stability of self-reported peer victimization applied to an equal extent to younger and older participants. There was, however, a significant interaction between age and type of informant (Self vs. Peer) on stability that will be discussed below.

Our third goal was to examine whether the association between the stability of victimization and age differs between self- and peer-reports. As expected, the differential stability of peer- and self-reports was further qualified by age: the increase in stability of peer-reports was significantly stronger than the increase in the stability of self-reports. Peer-reports significantly increased in stability with age, but stability of self-reports remained the same.

In early childhood, self-reports were more stable than peer-reports of victimization. This is in line with previous studies (Hanish et al., 2004; Kochenderfer & Ladd, 1996; Monks et al., 2003). Children’s social and cognitive skills are still developing at this age. As bullying in early childhood is directed at many different victims, peers may have difficulty distinguishing children who are victimized incidentally from those who are victimized repeatedly (Monks et al., 2003). A certain degree of perspective taking is needed to perceive victimization among peers, and this skill is not yet fully developed in early childhood (Marvin et al., 1976).

Another explanation may be that victimization is not a stable phenomenon in early childhood. Instead, stability in self-reports at an early age may reflect a more general self-concept or internal working model. Children may base their self-reports of victimization on a general internal working model of social relationships. This model taps into more general self-views and is not necessarily based on actual victimization experiences. As general internal working models stabilize early, this might explain why self-reports are more stable than peer-reports in early childhood and remain a constant level of stability over time (Bretherton, Ridgeway, & Cassidy, 1990).

In middle childhood, the stability of peer-reports did not differ from the stability of self-reports. As children’s social-cognitive skills develop, they may become more accurate in detecting victimization among peers. As a consequence, the stability of peer-reports may increase.
In late childhood and adolescence, peer-reported victimization was more stable than self-reported victimization. Adolescents may rely more on a peer’s status and reputation because status becomes more salient at this age when the peer group is more hierarchical than in childhood (Schäfer, 2005). Therefore, reputational bias may play a larger role in adolescence, contributing to higher stability of peer- than self-reported victimization (Hymel et al., 1990). Another reason for the prominent role of reputation bias may be that it is difficult to detect victimization in adolescence. The frequency of victimization decreases by the end of high school (Hymel & Swearer, 2015), suggesting that victimization does not occur in every classroom. However, peer nominations often invite adolescents to nominate at least one student in their classroom. Further, the frequency of indirect victimization increases, whereas the frequency of direct victimization decreases (Crick et al., 2001; Schäfer, Werner, & Crick, 2002). As indirect victimization is difficult to detect and victimization may not occur in every classroom, adolescents may more strongly rely on previously established reputations than actual recent victimization incidents.

The interaction between type of informant (peer vs. self) and age may be important for research distinguishing groups of victims based on the overlap between self- and peer-reports. One of these groups is usually labeled “deniers,” consisting of youths who are identified as victims by peers but who do not see themselves as victims (Graham & Juvonen, 1998). Our meta-analysis suggests that it may be difficult to escape from victim status, even when youths themselves experience that they are no longer bullied. This would imply that victims correctly perceive that they are no longer bullied. Peers, however, may deny that the victim status of these children has been changed. The label ‘deniers’ may therefore be a misleading term for this group. It is interesting to further understand this discrepancy in future research. For example, future studies may examine the actual victimization experiences of this group by means of diary assessments. When peers are asked to report on a daily basis, it is likely that they still remember the victimization episodes that occurred during the day. Therefore, they may rely less on a child’s previous status as victim which reduces the impact of reputation biases (Bolger, Davis, & Rafaeli, 2003; Nishina & Juvonen, 2005). Thus, although peers may have the perception that a child is victimized according to his previous reputation, diary assessments may make them more aware that the child is not victimized anymore. If this diary assessment indeed leads to a reduction in the prevalence of daily diary peer-reported victimization, then this will support our explanation that global peer-reports are biased by a child’s reputation.

Strengths, limitations, and suggestions for further research

The current meta-analysis had several strengths, including the large sample size and the highly reliable entering and coding of studies and variables. The multilevel approach made it possible to include multiple correlations from each study and model the error variance. No association between sample size and the dependent variable (i.e. stability coefficient of victimization) was found, indicating no sign of publication bias.

Despite these strengths, there were also some limitations. Stability was operationalized as the Pearson correlation over time; therefore, certain changes could not be identified. For example, we could not detect overall increases or decreases in the level of victimization in the entire sample, or individual differences in absolute levels of victimization. To address such issues, further research should examine intra-individual changes and mean-level changes in victimization, or use an intra-class correlation as the measure of stability.

Informant self was the reference category because self-reports are the most commonly used measure of peer victimization. A limitation is that we could not directly compare whether peer-, teacher- and other/combined reports significantly differed from each other.

A substantial part of the variance of the stability of victimization was explained by the predictors age, interval length, and type of informant. Part of the variance remained unexplained, to some extent because some information was reported by only a small number of studies or was inconsistently reported across studies. For example, stability of victimization may differ between victims and bully-victims (Schäfer et al., 2005) and between different types of peer victimization (e.g., overt vs. relational) (Schwartz, Gorman, Nakamoto, & Toblin, 2005). In addition, there may be gender differences in the stability of victimization (Camodeca et al., 2002). We were not able to examine them because there was not enough
variation in the gender ratio across studies and only a few studies reported stability coefficients for boys and girls separately.

Another reason for the unexplained variance may be that there was variation in the methods used between studies that was not consistently reported. For example, self-reported victimization was examined with various questionnaires that differed in internal consistency (e.g., Crick & Grootpetre, 1996; Solberg & Olweus, 2003). Studies that displayed victimization in a $2 \times 2$ contingency table differed in their cut-off points for victimization. In addition, either peer nominations or ratings were used to examine peer-reported victimization. This also may affect stability, as peer ratings tend to be more stable than nominations (Jiang & Cillessen, 2005). Some studies using peer nominations did not indicate whether they were limited or unlimited, and whether a minimum number of nominations were required. Finally, in some studies, youths were given a definition of victimization but in other studies they were not. It has been shown that youths report significantly less victimization and bullying when they are provided with definitions than when they are not (Kert, Codding, Tryon, & Shiyko, 2010). To conclude, there was heterogeneity in the measures of victimization in our sample. Further research needs to examine whether some measures and methods yield more stable results than others.

**Practical implications**

The findings of this study have implications for the content and evaluation of anti-bullying programs. This meta-analysis demonstrated that peer victimization is moderately stable across one year when there is no intervention. Intercept variance showed that there was even negative stability of self-reports in some cases. Fortunately, these cases suggest that some victims decrease in their level of victimization over time. However, the average stability of peer victimization without intervention is sufficiently large to support the need for prevention and intervention efforts. In line with the chronic stress model (Dohrenwend & Dohrenwend, 1981), stable victims have poorer developmental outcomes than incidental victims (Burk et al., 2011; Camodeca et al., 2002; Juvonen et al., 2000; L. H. Rosen et al., 2009). More research is needed on the predictors of stable versus unstable victims in order to early identify students at risk.

A first step will be to screen for victimized students. As the effect of interval length shows that students’ victimization may change during one school year, it seems important to conduct such screening at least twice each year (Rueger et al., 2011). There was no main effect of age on the stability of self-reported victimization. Thus, some children already consistently indicate that they are a victim of bullying in early childhood. The stability of peer-reported victimization increased with age. This highlights the need for early intervention. This is in line with the idea that intervention should take place early before victimization crystallizes and becomes more difficult to change (Hanish & Guerra, 2004; Rueger et al., 2011; Smith, Ananiadou, & Cowie, 2003).

The effect of type of informant also has implications for practice. Although adolescents may no longer identify themselves as victims, they may maintain victim status in the peer group. This stable reputation may influence how peers respond to them. Peer-reported victimization is associated with rejection, low peer acceptance, and low friendship qualities (Scholte et al., 2013) – problematic correlations that persist over time (Hodges & Perry, 1999; Scholte et al., 2013). Another study further demonstrated the power of the victim reputation among peers. This study showed that girls were more aggressive in a noise-blast task to a peer who was seen as a victim in the classroom than to a peer who was not seen as a victim in the classroom, while the peer’s actual behavior could not have provoked the participants’ aggression due to the preprogrammed feedback the participants received during the noise-blast task. (Lansu, Cillessen, & Sandstrom, 2014). Further research on how adolescents’ perceptions of peer victimization may be colored by their knowledge of a peer’s previous victim status is needed.

The findings of this meta-analysis also have implications for the content of interventions. If there is indeed a ‘once a victim, always a victim’ bias among adolescents, this bias could potentially be changed. One reason for such a bias may be that adolescents are afraid to lose status when they affiliate with victims of bullying. Instead, they may assist or reinforce the bully to maintain or establish a higher position for themselves in the peer group (Juvonen & Galván, 2008). As they do not interact with the victim, they may not notice that the victim is no longer bullied. The interventions should challenge adolescents to rely less on their views of a peer’s previous victimization status and more on observations...
of actual bullying behavior. This may reduce a ‘once a victim, always a victim’ bias, and subsequently, (former) victims may have more opportunities to gain positive social experiences with peers.

In addition to assisting youths to perceive their peers more objectively, interventions also may focus on identity formation of victimized youths. An important developmental task in adolescence is the formation of a coherent sense of self based on one’s strengths and values (Erikson, 1986). Adolescents shift from pursuing relationships based on popularity and status to pursuing relationships that are supportive and intimate (Steinberg & Morris, 2001). As operational and abstract thinking improves, identity matures in adolescence (Meeus, Van De Schoot, Keijzers, Schwartz, & Branje, 2010). Adolescents also may have the cognitive abilities to reframe victimization in a way that is not destructive to the self. Therefore, stability of self-reported victimization may decrease in adolescence, as an increasing number of students may not identify themselves as victim anymore. However, we found that the stability of self-reported victimization did not vary by age. Perhaps not all adolescents benefit from their cognitive development. Victimized adolescents are especially at risk for developing a non-coherent sense of self, because they have difficulties integrating their identity across contexts (van Hoof, Raaijmakers, van Beek, Hale III, & Aleva, 2008). This may explain why self-reported victimization remains constant despite increased social cognitive skills. This is alarming because identity problems place victimized adolescents at risk for depressive symptoms (2008). Therefore, intervention programs should assist victimized adolescents with their identity development and reframing of victimization experiences.

The stability of victimization also has to be taken into account in the evaluation of anti-bullying programs. It is important to examine whether programs lead to more changes in which children are being victimized than that would have occurred without intervention, rather than only focusing on the absolute amount or frequency of victimization. Program evaluations also need to take into account the higher stability of peer-reports than self-reports of victimization in adolescence. Studies have shown that interventions are less successful according to peer-reports than self-reports (e.g., Salmivalli et al., 2005). One reason for this difference may be a sensitization effect (Salmivalli et al., 2005; Smith et al., 2003). Youths may become more sensitive to victimization when their awareness of bullying is increased by the intervention. They may also begin to recognize indirect aggression as bullying. This meta-analysis showed that peer-reports are more stable than self-reports in adolescence also when there is no intervention. Therefore, the reduced effectiveness of interventions according to peer-reports may be due not only to a sensitization bias, but also to stable peer reputations.

Conclusion

This meta-analysis on the stability of peer victimization showed moderate stability of self-reported victimization over a one-year interval. Stability decreased as interval length increased. The increase in the stability of victimization with age was larger for peer-reported victimization than for self-reported victimization. In new studies, complete reporting of measurements and sample characteristics is needed to facilitate further meta-analyses.

Acknowledgments

The first author was supported by a Research Assistant Grant from the Netherlands Royal Academy of Sciences. The authors are grateful to all researchers who have sent information about the stability of victimization in their studies. The authors are also grateful to dr. Thao Ha for her contribution to the coding and entering of studies.

Appendix A

Description Statistical Model

The conceptual model was implemented by means of a statistical model which had some special features. First, when stability correlations were available for three time points, say 1, 2, and 3, we expected that the correlation between 1 and 3, \( r_{13} \), would depend on \( r_{12} \) and \( r_{23} \). This is analogous to the
effect calculation in a path model, such as a mediation model (Hayes, 2013). To correct for this dependency we included \( r_{12} \) and \( r_{23} \) as predictors for \( r_{13} \) in the model by means of dummy coding. Simulations showed that using \( r_{12} \) and \( r_{23} \) or \( r_{12} \times r_{23} \) did not lead to very different outcomes. The latter would intuitively seem more appropriate but is technically more complex. Therefore, for the sake of simplicity we chose to include the separate correlations.

Second, in order to have a more suitably distributed variable for the stability of peer victimization than Pearson’s \( r \) itself, we performed Fisher’s \( r \)-to-\( Z \) transformations (Hayes, 1978). The \( Z \) transformation provided us with a measure for the standard error of the transformed stability of peer victimization. Specifically, this means that the Level 1 error variance of the dependent variable is known. We applied the method demonstrated by Maas et al. (2004) to model this known error variance. This has the advantage that heteroscedasticity could be modeled and that it was possible to distinguish within and between study variance, according to the random effects model (Hox, 2002). An additional level was added to the statistical model in order to model the error variance.

The Fisher \( r \)-to-\( Z \) transformation has another advantage: tetrachoric correlations do have substantially larger standard errors than Pearson correlations and a correction should be applied. However, Ogasawara (2010) indicated that a Fisher \( r \)-to-\( Z \) transformation is also valid for a tetrachoric correlation.

Appendix B

Results of the Statistical Model

Model 1 was a null model that contained only an intercept without any predictors. Results showed significant error variance at Level 2, indicating the necessity of a multilevel approach (Snijders & Bosker, 1999).

In Model 2, the null model was expanded with the \( r_{1,2} \) and \( r_{2,3} \) correlations to check for functional dependencies. Results indicated that \( r_{1,2} \) and \( r_{2,3} \) had a substantial effect size, although the effects were non-significant. In addition, sample size of the study was added to examine the possibility of publication bias (Hox, 2002, pp. 150–1). According to Card (2012), a significant effect of sample size might indicate a publication bias. Sample size had a very large range compared to the ranges of the other variables. Therefore, sample size was standardized so that very small estimates would not lead to estimation problems (Hox, 2002). Findings revealed a small and not significant effect of sample size, indicating that effect sizes were independent of study sample size, not suggesting publication bias (Hox, 2002).

In Model 3, the previous model was expanded with the predictors of interest: interval length (centered at 1 year), age (centered at 10 years), type of informant, and the interaction between age and type of informant. Results showed that the effects of \( r_{1,2} \) and \( r_{2,3} \) have become negligible, suggesting that the effects were due to the substantive predictors and not to the functional predictors.

Model 4 was the final model, in which \( r_{1,2} \) and \( r_{2,3} \) and sample size were removed because they were non-significant and negligible in effect size. The other effects barely changed after removal of these variables. Model 4 was accompanied by a test of whether the residuals met the assumptions. Normal plots (Rasbash, Steele, Browne, & Goldstein, 2012) showed that the residuals were normally distributed. Model 3 and 4 show negative deviances, because of the small scale of the dependent variable (Scientific Software International, 2014).

Appendix C

Determining the Region of Significance of the Difference in Peer- Minus Self-Reported Victimization by Age

The aim was to know for which centered age (Cage) the stability of victimization (\( Z_r \)) differs significantly between the Informants Peer and Self. Therefore, we calculated the confidence intervals for the difference in stability in Fisher’s \( Z \) transformed stability coefficient of Peer minus Self-reported victimization analogous to Preacher, Curran, and Bauer (2006).

The confidence interval for the difference, depending on age was:
\[
Cl_{age} = \left( \hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}} \right) \pm t_{\text{crit}} SE \hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}}
\]

where \( \hat{Z}_{r,s} \big|_{\text{peer}} \) and \( \hat{Z}_{r,s} \big|_{\text{self}} \) were the expected Fisher’s Z transformed stability coefficients for Peer and Self.

To find the \( SE \hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}} \) we started with the model for the expected value of the Fisher’s Z transformed stability coefficient based on fixed effects:

\[
\hat{Z}_{r,s} = \beta_1 \text{Cinterval} + \beta_2 \text{Self} + \beta_3 \text{Peer} + \beta_4 \text{Teacher} + \beta_5 \text{Other} + \beta_6 \text{Self Cage} + \beta_7 \text{Peer Cage} + \beta_8 \text{Teacher Cage} + \beta_9 \text{Other Cage}
\]

The equations for peer- and self-reports then were:

\[
\hat{Z}_{r,s} \big|_{\text{peer}} = \beta_1 \text{Cinterval} + \beta_2 \text{Peer} + \beta_3 \text{Peer Cage} = \beta_1 \text{Cinterval} + \beta_2 + \beta_3 \text{Cage}
\]

and

\[
\hat{Z}_{r,s} \big|_{\text{self}} = \beta_1 \text{Cinterval} + \beta_2 \text{Self} + \beta_3 \text{Self Cage} = \beta_1 \text{Cinterval} + \beta_2 + \beta_3 \text{Cage}
\]

The equation for the difference in stability between peer and self was:

\[
\hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}} = (\beta_1 \text{Cinterval} + \beta_3 + \beta_7 \text{Cage}) - (\beta_1 \text{Cinterval} + \beta_2 + \beta_6 \text{Cage})
\]

\[
= (\beta_3 - \beta_2) + (\beta_7 - \beta_6) \text{Cage}
\]

The \( SE \) for the confidence intervals, dependent on Cage was:

\[
SE \hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}} = \sqrt{\text{Var}(\hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}})}
\]

\[
= \sqrt{\text{Var}(\beta_3 - \beta_2 + \beta_7 \text{Cage} - \beta_6 \text{Cage})}
\]

where

\[
\text{Var}(\hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}}) = \text{Var}(\beta_3 - \beta_2 + \beta_7 \text{Cage} - \beta_6 \text{Cage})
\]

\[
= 2\text{Cage}[\text{Cov}(\beta_2, \beta_6) - \text{Cov}(\beta_2, \beta_7) - \text{Cov}(\beta_3, \beta_6) + \text{Cov}(\beta_3, \beta_7)]
\]

\[
+ \text{Cage}^2[\text{Var}(\beta_6) + \text{Var}(\beta_7) - 2\text{Cov}(\beta_6, \beta_7)]
\]

\[
+ \text{Var}(\beta_3) - 2\text{Cov}(\beta_2, \beta_3)
\]

All the variances and co-variances in the above formula were squared SE’s and covariances of the fixed effects (which could be found in column 1099 in Mlwin).

About \( t_{\text{crit}} \):

The number of degrees of freedom were the number of level 2 units (number of publications, 77) minus the total number of predictors (9) = 68. The critical t value for a (one sided) 95% confidence interval than was: \( t_{\text{crit}} (68) = 1.6676 \).

We used a one-sided test because the hypotheses were one sided.

This resulted in the following formula of the confidence interval for the of Fisher’s Z transformed stability coefficient for peer minus self-reports dependent on Cage:

\[
\{(\beta_3 - \beta_2) + (\beta_7 - \beta_6) \text{Cage}\} \pm t_{\text{crit}} SE \hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}} = \{(\beta_3 - \beta_2) + (\beta_7 - \beta_6) \text{Cage}\} \pm 1.6676SE \hat{Z}_{r,s} \big|_{\text{peer}} - \hat{Z}_{r,s} \big|_{\text{self}}
\]

The results were transformed into stability coefficients using an inverse Fisher Z transformation.
References


