Epidemiological pattern of bullying using a multi-assessment approach: results from the Bullying and Youth Mental Health Naples Study (BYMHNS)

Authors: Gennaro Catone¹, Simona Signoriello², Simone Pisano³, Margherita Siciliano², Katia Russo¹, Roberta Marotta⁴, Marco Carotenuto², Matthew R Broome⁵, Antonella Gritti¹, Vincenzo Paolo Senese⁶*, Antonio Pascotto²*.

1 Faculty of Educational Science, Suor Orsola Benicasa University, Naples, Italy
2 Department of Mental and Physical Health and Preventive Medicine, University of Campania, Naples, Italy
3 Department of Medicine and Surgery, Clinic of Child and Adolescent Neuropsychiatry, University of Salerno, Baronissi, Italy
4 IRCCS Stella Maris, Pisa
5 Institute for Mental Health, University of Birmingham, Birmingham, United Kingdom
6 Department of Psychology, University of Campania "Luigi Vanvitelli"

* joint last authors

Corresponding Author: Dr Gennaro Catone, Faculty of Educational Science, Suor Orsola Benicasa University; Via Suor Orsola 10, Naples, Italy; catoge@libero.it
Abstract

Bullying is a widespread phenomenon that has captured attention from mental health researchers. Several studies assessed bullying prevalence with some methodological concerns.

In this paper we estimated bullying prevalence in the context of the metropolitan city of Naples. Preliminarily we evaluated the psychometric properties of two bullying scales for victimization (the multidimensional peer victimization scale - MPVS) and for perpetration (the bully subscale of the Illinois bully scale - IBS). Finally we evaluated the effect of gender and classroom for the phenomenon.

The study was cross sectional. 2959 students constituted the sample. Data collection was obtained adopting a multi-assessment approach that included both single-item questions and intensity scales in order to compare the two methods and to found eventually correspondence or differences.

The two scales showed good reliability and validity properties. The MPVS displayed a 1-factor second order model. The IBS-P had a mono-fatorial structure. Both showed full invariance for gender and classroom. Prevalence of victimization was about 37% whereas that for perpetration about 21%. As expected we obtained several bullying prevalence results depending on the specificity of questions and in particular repetitiveness of episodes. There was a good correspondence between results of single-item questions and multi-item scales. Finally results demonstrated several differences for gender and classroom attended.

In this epidemiological study the multi-assessment approach identified different but complementary features of bullying phenomena. The use of the two measurement approaches allowed us to obtain more precise and exhaustive information on bullying prevalence and compare it with previous findings.

Keywords: bullying, prevalence, Naples, assessment, methods
Introduction

Since Dan Olweus in the far 1983 in Norway completed the first national survey on the bullying phenomenon in Norway, we have learned that the study of bullying events may represent a valid subject of inquiry in mental health and related disciplines because in a social psychiatry perspective it contributes to the understanding of how environmental factors interact with the biological ones in the pathogenesis of the mental stress-related disorders. Moreover it allows implementing the possibility of preventive strategies in youth (Arseneault, 2018).

In his definition of bullying, Olweus emphasized three characteristics: 1) it is an “intentional” aggressive behaviour, this excludes friendly and playful action without the intention of offend or harm another person; 2) negative actions should occur “repeatedly”, isolated circumstances may not refer to bullying; 3) bullying encompasses an imbalance of power between the bullies and victims, with the result that the victims can not easily defend themselves. This asymmetry in the relationship regards physically or mental weakness, difference in numbers or in retrieving the source of perpetration (Peter K. Smith, 1999).

Bullying can manifest in a direct and indirect way. Direct episodes are physical or verbal attack such as kicking, punching or pushing and threats, insults, name-calling. Indirect episodes indicate covert forms of aggression/victimization which can take the form of social exclusion, spreading rumours and damage other’s property (Marini, Dane, Bosacki, & Ylc, 2006). It was well-established than male were more likely to be bullies and involved in physical episodes than females, whereas indirect victimization occurred in late years of adolescence (Juvonen & Graham, 2014). Bullying episodes demonstrated temporal stability across years although there is a gradual decrease in them with age (Juvonen & Graham, 2014).

Another facet of bullying is the cyber-bullying, which represents a form of aggression carried out in the electronic context (e.g. web, blog, chat, phone). It has become an interesting topic of study since, despite sharing peculiar elements with traditional bullying, it differentiates in many features due to the peculiarity of the of virtual space functioning (Kowalski, Giumetti, Schroeder,
During the last years, research on bullying has undertaken two directions: epidemiological and association/causation studies. As regards the former, researchers investigated how widespread is the phenomenon (prevalence). On the other side, researchers were interested in demonstrating the relationship between bullying, as a stress factor, and the onset of mental symptoms/disorders. This last trend demonstrated a strong association between bullying (both victims and bullies) and mental health problems such as anxiety, depression (Zwierzynska, Wolke, & Lereya, 2013), psychosis (Catone et al., 2015; Catone, Marwaha, Lennox, & Broome, 2017; Moffa et al., 2017) and suicide attempts (Kim, Koh, & Leventhal, 2005). Consequently, institutions are striving to promote security aspects of schools and bullying can be included as a concern among children, parents and teachers (Arseneault, Bowes, & Shakoor, 2010).

About the prevalence, epidemiological studies have been conducted more or less around the world (Liang, Flisher, & Lombard, 2007; Nansel et al., 2001; Olweus, 1993; Rigby & Slee, 1991; Wang et al., 2012; Whitney I., 1993). An updated systematic review considered more than 80 studies reporting a prevalence of traditional bullying involvement around 35% among adolescents recruited in the studies (age range 12-18 years) (Modecki et al., 2014). In Italy there have been virtuous examples of epidemiological research on bullying (Bacchini, 2009; Fonzi, 1997; M. L. Genta, Menesini, Fonzi, Costabile, & Smith, 1996; M. L. Genta, Menesini, E., Fonzi, A., Costabile, A, 1996). Genta et al. (Fonzi, 1997) were the first that described bullying prevalence in two Italian cities (Firenze and Cosenza, centre and south Italy respectively); studies then continued in different regions and almost all agreed that prevalence estimates were higher in Italy than in other countries. In a study (M. L. Genta, Menesini, E., Fonzi, A., Costabile, A, 1996), considering 1379 students aged 8-14 years, victims of bullying were 42% and 28% respectively in the primary and middle school and perpetrators were 20% and 16% in the primary and middle schools. Another study (M. L. Genta et al., 1996) conducted in six Italian regions, confirmed that victims and bullies were
respectively around the 40% and 22% of the school sample. In the Campania region, Bacchini and colleagues (Bacchini, 2009) promoted an epidemiological study on bullying considering 4760 students of primary, middle and secondary school grade, showing that around 30% of participants were involved in bullying; among victims and perpetrators male were more involved in direct (physical) behaviours whereas female in indirect behaviours. Most of these studies in Italy found no gender difference on victimization experience, showed that males were more involved in perpetration behaviour than females, and confirmed that the bullying decreases with age. All together these percentages were higher than those in other countries and this has been attributed to cultural and social differences (e.g., language, habits) [19-22].

Despite a multitude of study have been conducted, nowadays some epidemiological aspects remained unclear, one possible explanation of the observed difference in the prevalence could be related to the different definitions of bullying used by the researchers and because of the multi-cultural comparison. Moreover, there are some methodological limits that can be highlighted in the bullying prevalence studies that are related to the assessment procedures considered. These aspects have prevented researchers interested to describe the prevalence of bullying to capture the real extent of the phenomenon (Modecki et al., 2014). In fact the studies were carried out from different geographical areas and from populations with different languages often considering small samples, whereas large representative samples were needed (in the order of thousands of participants) to get a more valid description of the phenomenon. Furthermore it is also difficult to compare the various studies carried out since the measurement tools vary across researches and the study designs may concern the prevalence or effectiveness of prevention programs (Zych, Ortega-Ruiz, & Marin-Lopez, 2016). For Example Vivolo-Kantor et al. (Vivolo-Kantor, Martell, Holland, & Westby, 2014) in a comprehensive systematic review on bullying measurement found several methodological differences among studies and indicated the need for an exhaustive bullying definition and shared terminology, defined timeframe for the study design, and on the importance of providing bullying examples to study participants. As showed by Vivolo-Kantor et al. [24], a
quarter of the studies evaluated self-reported bullying behaviours by means of a dichotomous single-item, whereas three-quarter of studies used multi-item scales, with a consequent difficulty in comparing results. An issue raised from several researchers that contributed to the research on bullying was how to detect the phenomenon and whether it would be preferable to use a single-item question or multiple item scale; while the former are easier to use the second seem to be more accurate and reliable (Menesini & Nocentini, 2009). Menesini and Nocentini criticized the single-item approach preferring the multi item assessments for their higher validity and reliability (Menesini & Nocentini, 2009). Zych et al. (Zych et al., 2016) affirmed that each methodology had its advantages and disadvantages, for example, single-item question included a concept that is more clearly expressed and was shorter, easier and more efficient instrument however with a lower reliability and a higher margin of error if compared with the multi-item approach. Some authors pointed out the need of more detailed comparison between different assessment measures (Card & Hodges, 2008; Chan, Myron, & Crawshaw, 2005), but until now this gap in the literature remains poor investigated, and more important, the psychometric properties of the traditionally used measures were seldom assessed. Felix et al. pointed out that accurate assessment in bullying behaviours is essential and in the attempt of develop and validate an assessment instrument (the California Bullying Victimisation scale – CBVS, a multi-item scale) compared their “multi item” instrument with "single item" detection method in the same sample. Authors used the expressions “definition based strategy” (youth had to respond if they experienced victimization or perpetration based on a working bullying definition, this corresponded to the depicted single-item assessment), and “behaviour based strategy” (a list of victimization/perpetration related behaviors was presented and participants had to respond how often they have experienced them, this corresponded to the multi-item questions assessment), showing a 77.4% agreement rate between the two measures (Felix, Sharkey, Green, Furlong, & Tanigawa, 2011).

Given the extent of the phenomenon and the methodological concerns it is important to estimate prevalence data in a valid way and to use and develop measure tools with appropriate
psychometric properties adapted for the different populations. Starting from the abovementioned
considerations, the aim of this study was to estimate bullying phenomena prevalence in the context
of the metropolitan city of Naples (Campania Region, Italy). Because in the literature different
assessment approaches have been used, to better compare our results with previous epidemiological
studies, in this study a multi-assessment approach has been considered. Consequently, the
preliminary aim of this study was to adapt for the Italian population and evaluate their psychometric
properties two intensity scales of bullying: one for the victimization, the Multidimensional Peer
Victimization Scale (MPVS), and one for the perpetration, the perpetration subscale of the Illinois
bullying scale (IBS-P). In detail, we intended to verify the factorial structure of the MPVS and the
IBS-P; verify the invariance of both scales as a function to gender and age; and verify the reliability
and validity of the two scales, by computing alpha and split-half coefficients and the correlation
between MPVS and the IBS-P scores with criterion measures.

Once verified the psychometric properties of the two multi-item scales, the prevalence of
bullying phenomenon was investigated. We have paid particular attention to the assessment
procedures used to detect phenomena and therefore we included both single-item questions and
multi-item questions (intensity scales) and compared them. At the best of our knowledge the
assessment of bullying prevalence with direct comparison of single-item questions and multi-item
approach was original and innovative and carried out for the first time in the Italian population.
Finally, we wanted to outline several bullying features such as the role of sex, age and type both for
victims and perpetrators phenomenon.

In line with the previous literature, we hypothesized that:

1. MPVS has 1-factor second order structure, the scale is invariant across gender and
classroom and has good reliability and validity;

2. IBS-P has a unidimensional structure, the scale is invariant across gender and
classroom and has good reliability and validity;
3. single-item and multi-item assessment to detect bullying prevalence tend to be associated even if they evaluate different and complementary aspects of the phenomenon;

4. victimization experiences decrease as a function of the age, and that the perpetration behaviours are more frequent in males than females.

Methods

Study design

This study is a part of larger project that involves a network of multiple Universities with multiple aims; in fact, along with estimates of bullying and cyber-bullying phenomena, it also covers other aspects of mental health (not reported here, see supplementary material). The design is cross-sectional, a school population based-study settled in the metropolitan city of Naples. The study setting has been determinate in order to link the research with a possible intervention within the territory in which the university institutions, involved in the project, operate. According to the 2016 National Institute for Statistics (ISTAT) national survey, the metropolitan city of Naples had 3.107.006 people and its capital Naples had 970.185 people. 10-14 years population in the metropolitan city of Naples was 181.686 (5.8%) youth of which 52.163 (5.4%) in its capital Naples. Twelve middle school across the area of interest (age range of students were 10-14 years approximately, three level of instruction= 1st, 2nd and 3rd grade), six in the city of Naples and six in suburban area (SA), to ensure representativeness (geographical criterion), were recruited and accepted to participate. The ethical committee of the Campania University “Luigi Vanvitelli” approved the study protocol (N. 500, 29/04/2016). Data collection was obtained throughout the administration of self-report questionnaire to participants.

Measures

For data presented here we used a section, including bullying and cyber-bullying victimization and perpetration, of the questionnaire used for the whole data collection (for detail of the whole questionnaire battery see the online supplementary material).
**Bullying phenomena measured with single-item question**

Participants were asked if they have ever suffered or perpetrated bullying according to a definition (in line with the above definition of Olweus (Olweus, 1993)) explained from a researcher before starting the self-report assessment. Moreover, bullying definition information was presented at the beginning of the questionnaires battery in the form of explicative vignettes and examples in which it was clarified what is considered bullying and what not, and also presented hints on repetition and imbalance of power factors. The single-item question for victimization was “*Have you ever been bullied?*” (possible answers were: “no”, score 0, or “yes”, score 1), while the question for perpetration was “*Have you ever made bullying?*” (Possible answers were: “no”, score 0, or “yes”, score 1). From these two questions we derived the bullying lifetime variables (victims/bullies). Furthermore if respondents endorsed yes for both questions they had the possibility to answer to additional questions. They were “*if you have been bullied how many times happened in the last six months?*” and “*if you have bullied others how many times happened in the last six months*” respectively for victimization and perpetration current bullying. Students could have responded among four categories: “once”, “twice”, “3-4 times”, “5 or 6 times or more” (we pointed out on frequency such as an important factor to understand the phenomenon on the basis of previous literature (Olweus, 1993)). From these two last questions we derived two variables for current bullying both for victimization and for perpetration: *absence*, that included participants who responded “no” to the first question (score 0); *broad* victimization/perpetration, that included subjects who responded “once” or “twice” to the questions of current bullying presented above (score 1); *narrow* victimization/perpetration, that covered subjects who responded “3-4 times” or “5-6 times or more” to the questions of current bullying presented above (score 2). All these questions were based on the “Questionario sulle prepotenze tra ragazzi a scuola”, original versions by Dan Olweus and Whitney and Smith; Italian translation and adaptation by Fonzi, Genta e Menesini (M. L. Genta, Menesini, E., Fonzi, A., Costabile, A, 1996).

**Bullying phenomena measured with multi-item questions**
**Victimization.** The multidimensional peer victimization scale (MPVS) has been translated into Italian by an expert and the translation has been verified by an independent expert using a back-translation methodology. It consists of four subscales: 1) Physical victimization scale; 2) Verbal victimization scale; 3) Social manipulation scale and 4) Attacks on property scale (scores on each of the four subscales have a possible range of 0 to 8) and total scale which score is computed by summing item responses (range 0-32). The questionnaire has 16 item with answers on a 3 point Likert scale (range 0-2): “Not at all” (score 0), “Once” (score 1), “More than once” (score 2). The MPVS has been scored in order that higher scores indicated higher victimization. The MPVS showed a good reliability and validity in measuring victimization, in previous studies (Mynard & Joseph, 2000), and was recently used in a clinical study on bullying and psychotic like experiences even though the psychometric features were not yet analyzed (Catone, Marotta, et al., 2017).

**Perpetration.** The Illinois bully scale is a questionnaire composed of 18 items and subdivided into three different subscales: 1) bully subscale, 2) victim subscale and 3) fighting subscale. In this study we used only the bully subscale because we are interested exclusively in bullying perpetration. The IBS-P bully subscale has been translated into Italian by an expert and the translation has been verified by an independent expert using a back-translation methodology. This subscale consists of 9 items with 5 different response options with a maximum score of 0-36: "Never" (score 0), "1 or 2 times" (score 1), "3 or 4 times" (score 2), "5 or 6 times" (score 3), "7 or more times” (score 4). The scale showed previously good validity (Espelage & Holt, 2001). We have used this tool because it is short and easy to use and has a reliable theoretical construct and its utilization has been replicated by other studies (Espelage, Basile, & Hamburger, 2012; Rose, Espelage, Monda-Amaya, Shogren, & Aragon, 2015).

**Cyber-bullying phenomena measured with multi-item questions**

For cyber-bullying assessment we used, the Smith cyber-bullying self-report questionnaire, adapted by Bacchini et al. for their study in Campania region (Bacchini, 2009). Different cyber-bullying experiences, in terms of victimization or perpetration, have been collected asking
participants how many times eight types of cyber-bullying (victimization/perpetration) episodes happened since the beginning of the school year. The types examined were: 1) through text messaging; 2) pictures/photos or video clips; 3) phone calls; 4) email; 5) chat rooms; 6) instant messaging; 7) websites and 8) others. Frequencies were: “never” (score 0), “1 or 2 times” (score 1), “2-3 times at month” (score 2), “once a week” (score 3), “several times a week” (score 4) (P. K. Smith et al., 2008).

Data analysis

Preliminarily, the psychometric properties of the multi-item MPVS and the IBS-P were investigated. The factorial structure of the MPVS and the IBS scales was investigated using a confirmatory factor analyses. In line with the literature (Mynard & Joseph, 2000) and the hypotheses, for MPVS we compared three models: a one latent-factor model, a four correlated latent factors model and a 1-factor second order model; for IBS we tested a one latent factor model (for more information see the data analysis section in the supplementary material). Measurement invariance of MPVS and IBS measures was tested as a function of gender and classrooms. Reliability of the scales was examined using Cronbach’s alpha and split-half, whereas the concurrent validity of the scales was tested by computing the correlation coefficients between the MPVS, IBS and the Cyber-victimization and Cyber-bullying scores of the Smith cyber-bullying self-report questionnaire respectively. To control the familywise type I error, the False Discovery Rate (FDR) correction was applied to p-values (Benjamini, Drai, Elmer, Kafkafi, & Golani, 2001).

For bullying prevalence, we analysed results of both single-item questions and multi-items assessment; about the single questions, for lifetime and current bullying we obtained frequencies distribution and percentages; for multi items assessment prevalence data were presented as means and standard deviations.

For single-item question the effect of gender and age and their interaction were evaluated by means of two 2-step hierarchical logistic regressions, one for the victimization item and the other for the perpetration item. In both regressions, in the first step, gender (male = 1, females = 0) and
classroom (1st grade = 1; 2nd grade = 2; 3rd grade = 3) were introduced; whereas in the second step the interaction term (gender × classroom) was entered. The presence of victimization/perpetration was used as dependent variable (no = 0; yes = 1).

As regard the multi item assessment, to evaluate the effect of gender and age on MPVS subscales and total scale score we executed a 2×3 between subject MANOVA that treated gender as a 2-level between factor (Males vs. Females), and Classroom as a 3-level between factor (1st, 2nd and 3rd grade). To analyse main effect and interactive effects of Gender and Classroom on the IBS-P score, we executed a 2×3 between subject ANOVA that treated gender as a 2-level between factor (Males vs. Females), and Classroom as a 2-level between factor (1st, 2nd and 3rd grade). In all the analyses, the Bonferroni correction was applied to analyse post hoc effects, and the magnitude of significant effects was indicated by partial eta squared (\( \eta^2_p \)) (see the supplementary material).

Finally, to investigate the congruency between the single-item and the multi item assessment the non-parametric Spearman correlations were computed. To control the familywise type I error, the False Discovery Rate (FDR) correction was applied to p-values (Benjamini & Hochberg, 1995).

Statistical analyses were performed with R software version 3.0.1 (team, 2013). Confirmatory factor analyses (CFAs) and measurement invariance (MI) analyses were performed with Lisrel 8.71 software (Jöreskog, Sörbom, & SPSS Inc., 1988).

Results

Final sample

The study was conducted in the school year 2015/2016. Globally from the school records the 12 schools had 4444 students. 1116 (25.1%) did not gave consent to participate and 457 (10.3%) were absent the day of data collection. Total sample was constituted by 2959 students (66.6%). 1298 (44%) belonged to the schools of the city of Naples and 1661 (56.1%) to those of the suburban area. 1426 (48.2%) were females and 1533 (51.8%) were males. In the 1st grade classrooms 1048 (35.4%) students participated whereas in the 2nd grade 995 (33.6%) and in the 3rd 916 (31%).
Psychometric properties of the MPVS and IBS

All the analyses of factorial structures, invariance, reliability and validity of the MPVS and IBS were reported in the supplementary material. Results confirmed that for the MPVS the 1-factor second order model showed the best-fit indices (see supplementary materials: Table S1). The same factorial structure was observed when the analysis was repeated in the different subsamples extracted as a function of the Gender (Males and Females) and the Classroom (1st, 2nd and 3rd grade). Moreover, data showed that MPVS was characterized by full measurement invariance as a function of both factors (gender and classroom; see supplementary materials: Table S2 and S3).

The CFA analysis on IBS-P showed that the mono-factorial structure had the best fit to the data (see supplementary materials: Table S4). Also for this scale, the same latent structure was observed when the analysis was repeated in the different subsamples extracted as a function of the Gender (Males and Females) and the Classroom (1st, 2nd and 3rd grade). The IBS showed full measurement invariance as a function of both factors (gender and classroom; see supplementary materials: Table S4 and S5).

All subscales of the MPVS and the total score, and the IBS proved to have acceptable levels of internal consistency (Cronbach’s alphas > .69 and Split-Halfs > .68; see supplementary materials: Table S7). As regards the validity analysis, results showed that both the MPVS and the IBS had adequate concurrent and discriminant validity. Indeed, as expected the results showed that all subscales of the MPVS and the total score correlated more with Cyber-victimization score then Cyber-bullying score, and that the reverse pattern was observed for the IBS (see supplementary materials: Table S7).

Bullying categories

Lifetime bullying. According to lifetime bullying definition, 1204 (40.7%) students had reported having suffered at least once in their life an episode of victimization, while 651 (22%) students reported that they had perpetrated bullying at least once in their life.
Current bullying. According to current bullying definition, 1099 (37.1%) students reported they had suffered an episode of victimization in the last six months, of which 762 subjects only 1 or 2 time in the last six months (25.7%, broad category) and 337 subjects three or more times in the last six months (11.4%, narrow category). 618 (20.9%) subjects declared an event of perpetration in the last six months of which 466 students only 1 or 2 times (15.7%, broad category) and 152 three or more times (5.1%, narrow category). These results showed that there was a difference of 105 (3.5%) respondents (subjects who were involved in bullying victimization in lifetime but not in the last six months) between lifetime and current victims, while there was difference of 33 (1.1%) subjects between lifetime and current perpetrators, that is respondents who were involved in bullying perpetration in lifetime but not in the last six months. Figure1 highlighted characteristics of the sample.

Bullying dimensions and relation between single- and multi-item assessment

Table 1 showed MPVS and IBS-P results (means and standard deviations for broad and narrow categories) both for victimization and perpetration. The correlation analysis showed a substantial congruence between the two measurements approach. Though the effect size indicated a medium association between the two measurement approaches. In particular, data showed that single-item victimization index was strongly correlated with the multi-item victimization score \( r = .477, N = 2959, FDR \text{ adjusted } p\text{-value < .001} \) than multi-item perpetration score \( r = .290, N = 2959, FDR \text{ adjusted } p\text{-value < .001} \); and that the single-item perpetration index was strongly correlated with the multi-item perpetration score \( r = .443, N = 2959, FDR \text{ adjusted } p\text{-value < .001} \) than multi-item victimization score \( r = .180, N = 2959, FDR \text{ adjusted } p\text{-value < .001} \); see Table 2).
No gender effect was observed. The model parameters analysis (see Table 3) indicated that, over and above the gender, the victimization decreased as a function of the classroom, \( b = -0.222 \), odds ratio = 0.801, \( p < .001 \), though the effect size was very weak. The higher was the class the lower was the likelihood that students declared to be exposed to bullying episodes. Finally, the interaction term (step 2) did not improve the fit of the model, \( \chi^2(1, N = 2959) = 3.24, p = .072, R^2_{\text{diff}} = .001 \).

**Perpetration.** The results of the hierarchical logistic regression showed that, over and above the classroom, the gender factor was associated with perpetration, \( \chi^2 (2, N = 2959) = 50.62, p < .001, R^2 = .017 \). No classroom effect was observed. The model parameters analysis (see Table 3) indicated that, over and above the classroom, the perpetration was related to the gender, \( b = 0.654 \), odds ratio = 1.924, \( p < .001 \), though the effect size was very weak. Males declared themselves as expressing more Bullying behaviours than females. Finally, the interaction term (step 2) did not improve the fit of the model, \( \chi^2(1, N = 2959) = 0.01, p = .978, R^2_{\text{diff}} = .000 \).

**Effect of gender and age on bullying – multi-item**

**Victimization.** Results of the MANOVA executed on the MPVS subscale revealed significant main effects of Gender (Males versus Females), Wilks’ lambda = .909, \( F(4, 2950) = 73.58, p < .001, \eta^2_p = .091 \), and Classroom (1st, 2nd and 3rd grade), Wilks’ lambda = .983, \( F(8, 5900) = 6.32, p < .001, \eta^2_p = .008 \). The Gender × Classroom interaction was not significant, Wilks’ lambda = .997, \( F(8, 5900) = 1.05, p = .392, \eta^2_p = .001 \). The follow-up one-way ANOVAs showed that Gender significantly affected Physical Victimization, \( F(1, 2953) = 177.54, p < .001, \eta^2_p = .057 \), and Social Manipulation, \( F(1, 2953) = 37.24, p < .001, \eta^2_p = .012 \); whereas not significant were the effect on the Verbal Victimization dimension, \( F(1, 2953) = 1.79, p = .168, \eta^2_p = .001 \), and Attacks on Property, \( F(1, 2953) = 2.78, p = .096, \eta^2_p = .001 \). Males rated themselves as receiving more Physical Victimization (\( M = 1.35 \)) and Attacks on Property (\( M = 2.35 \)) than females (\( M = 0.56 \) and
whereas females rated themselves as receiving more Social Manipulation ($M = 2.62$) than males ($M = 2.11$). 

As regards the Classroom main effect, the follow-up one-way ANOVAs showed a significant but weak effect on Physical Victimization, $F(2, 2953) = 8.30, p < .001, \eta^2_p = .006$, and Verbal Victimization dimensions, $F(2, 2953) = 6.53, p < .001, \eta^2_p = .004$; whereas not significant were the effect on the Social Manipulation, $F(2, 2953) = 2.11, p = .413, \eta^2_p < .001$, and Attacks on Property, $F(2, 2953) = 0.47, p = .623, \eta^2_p < .001$. Adjusted mean comparisons indicated that student of the 1st grade rated themselves as receiving more Physical Victimization ($M = 1.12$) than student of the 2nd and 3rd grade ($M = 0.91$ and $M = 0.83$ respectively); whereas student of the 2nd and 3rd grade ($M = 3.43$ and $M = 3.46$ respectively) rated themselves as receiving more Verbal Victimization than student of the 1st grade ($M = 3.10$; see table 4).

Perpetration. Results of the ANOVA executed on the IBS-P revealed significant main effects of Gender (Males versus Females), $F(1, 2953) = 99.42, p < .001, \eta^2_p = .033$, Classroom (1st, 2nd and 3rd grade), $F(2, 2953) = 17.62, p < .001, \eta^2_p = .012$, and the Gender × Classroom interaction, $F(2, 2953) = 4.19, p = .015, \eta^2_p = .003$. Adjusted mean comparisons showed that males rated themselves as expressing more Bullying behaviours ($M = 3.79$) than females ($M = 2.31$); student of the 2nd and 3rd grade ($M = 3.20$ and $M = 3.50$ respectively) rated themselves as expressing more Bullying behaviours than student of the 1st grade ($M = 2.45$). Finally, the interaction effect showed that gender differences increased as a function of the classroom (see table 4).

Discussion

In this paper we presented the data on bullying phenomena in the Metropolitan city of Naples (Campania, Italy) by adopting for the first time a multi-assessment approach. To this aim we adapted for the Italian population two intensity scales of bullying, one for the victimization, the Multidimensional Peer Victimization Scale (MPVS), and one for the perpetration, the perpetration
subscales of the Illinois Bullying Scale (IBS-P), and their psychometric properties were investigated. As regards the MPVS scale, data confirmed a 4-Factor latent structure, with a general second order factor, which is fully invariant across Gender and Classroom factors. As regards the IBS-P, data confirmed a mono dimensional latent structure, which is fully invariant across Gender and Classroom factors. Moreover, data showed that both scales had an adequate reliability and concurrent and discriminating validity.

Results demonstrated the temporal stability of the bullying phenomena because differences between lifetime and current bullying were negligible (Espelage, Bosworth, & Simon, 2001).

Results on bullying prevalence provided interesting findings. The use of single-item question allowed us the expose results in a categorical manner (categories and percentages of presence of victimization and perpetration). Our study design pointed out the repetitiveness features such as an important object of study in the evaluation of epidemiological patterns of bullying phenomena together with the intentionality and the imbalance of strength (which represent the fundamental bullying characteristics described by Olweus). We achieved a progressive narrowing of the prevalence of victimization episodes (41% - 37% - 28% - 11%) and of perpetration episodes (22% - 21% - 16% - 5%) observing with the lens of repetitiveness. In our opinion this point could explain some heterogeneity derived from prevalence studies on bullying phenomena, beyond the methodological differences of detection. In Campania region, Bacchini found a prevalence of about 10% - 40% among various victimization forms and a prevalence of 5% - 35% among several perpetration practices. Looking at his data, the criterion of repetitiveness restricts the gap (Bacchini, 2009). Genta et al. found that among victims the 27.4% - 29.6% reported episodes sometimes or more (comparable to our broad category) and around 10% reported episodes once a week and more (comparable to our narrow category) in middle schools of two cities in Italy (Florence and Cosenza) (M. L. Genta et al., 1996). Authors affirmed that Italian bullying has higher percentage than other European countries and they elegantly examined cultural and methodological issues to explain these differences. Widening the field, in UK, recent results from the Bullying Intervention
Group (a survey that involved England, Wales, Scotland and Northern Ireland) found that 33%, among 11,000 participants, were “sometimes” involved in bullying phenomena and 11% were “a lot” (Zych, Farrington, Liorent, & Ttofi, 2017). Ortega et al. compared prevalence data of bullying in three different countries. 5,862 students constituted the sample (Italy, $N = 1,964$; Spain, $N = 1,671$; England, $N = 2,227$). Direct bullying ranged from 7.5% to 11.6% occasionally (“once or twice”) and from 3.2% to 7.1% frequently (“more frequently”): indirect bullying varied from 12.4% occasionally (“once or twice”) and from 3.4% to 6.2% frequently (“more frequently”) (lower estimates in Spain, higher estimates in the UK) (Ortega et al., 2012).

On the other hand, the use of multi-item assessment displayed intensity of the phenomena and distinction between different types (only for victimization). Verbal bullying (direct) was the most prevalent followed by social manipulation, attack on property (indirect) and physical bullying (direct); we found the same distribution observed in a previous study that used the same assessment methodology (Catone, Marotta, et al., 2017).

Our findings indicate a substantial congruency between the two forms of assessment. Indeed, victimization evaluated by the single-item approach was more correlated with victimization than perpetration evaluated by the multi-item assessment. The opposite was valid for the perpetration. Previously, Felix et al (Felix et al., 2011), used a similar strategy in a study that examined the psychometric properties of the California Bully Victim Scale (CBVS), a multi-item scale on bullying phenomena compared with a single-item question. In their study authors found evidence for concurrent validity between the two measures and concluded that prevalence rate was similar between the two methods despite they identified different set of student being/having bullied and this meant the possibility for the used method of influencing the status of bully/victim for the participants. For the authors, in the multi item model the lack of the term bullying can facilitate the affirmative response of a victim, on the contrary, the absence of the concept of power imbalance can limit the affirmative answer of who did not subjectively perceive this state. Thomas et al. argued that the choice should be definitively guided by the study aims, in fact single-item should be
used for a general prevalence purpose whereas multi-item for detecting the different types of victimization/perpetration, to capture any changes in the phenomenon over time and to get a measure congruent with the multidimensional model. Furthermore, the simultaneous use of both measurement approaches could allow to estimate the global prevalence of the phenomenon and the opportunity to examine also the information obtained from further follow-up questions (Thomas). We substantially agree with these conclusions and our results confirmed that the single question assessment easily handled the prevalence data in the sample, whereas the multi items assessment allowed the differentiation of the forms of victimization. Moreover, the simultaneous use of the two measurement approaches allowed us to get a clearer picture of the bullying prevalence and directly compare our results with the previous studies.

As regards the effect of gender and classroom on bullying, our data are in line with the previous literature showing that victimization decreased over time (with the increase of classroom) and that independently of the measurement approach, categorical or dimensional, perpetration behaviours were more frequent for males than females. Seals et al found that male were more involved in bullying phenomena than females in 454 students of the public school (Seals & Young, 2003). Another study that included German students showed that males reported more perpetration than females and that victimisation episodes were more frequently experienced by younger students (Scheithauer, Hayer, Petermann, & Jugert, 2006). Kumpulainen et al, in a 4 year follow up study showed a progressive decline of children involved in bullying and also a prevalence of males among bullies; despite this they conclude that for someone (especially the bullies/victims) the phenomenon can last for several years (Kumpulainen, Rasanen, & Henttonen, 1999). The use of the multi-item assessment allowed us to analyse more in detail the role of gender and age both for victimization and perpetration. Indeed results highlighted that male experienced more likely physical victimization and attack on property whereas female experienced higher social manipulation episodes. In the first classroom the physical victimisation was more prevalent but in the second and third classroom the verbal victimisation increased. Perpetration intensity seemed to
increase from the first to the third classroom. These results confirmed others present in literature that demonstrated difference regarding gender and age for direct and indirect victimisation (cit.).

In our study, we used a multi-assessment approach. The multi item question displayed data on dimensional construct. On the other hand the use of single item question produced several prevalence of bullying victimization and perpetration, which apparently show dichotomous categories (presence/absence of victimisation/perpetration). However, also for single item approach it emerged a progressive dimensionality that depended on the repetitiveness of the phenomenon (i.e. the 11% of the sample belonged to the narrow category but this does not mean that 89% belonged to the “no victims” category but rather that we have outside of this 11% broader victims or lifetime victims, and this also applied to perpetrators). Indeed, we cannot affirm that certain prevalence is the correct one and therefore corresponds strictly to the observed phenomenon while the others cannot be used. Certainly we are tempted to affirm that the narrow victimization and perpetration are of higher interest because of the links with the psychopathology and the implications about the forms of intervention and prevention; on the other hand, broad categories were seldom studied and we can only speculate, but cannot demonstrate right now, a dose response link to psychopathology. It emerges that a widespread dimension rather than dichotomous categories could better capture bullying phenomena (that is linked to, but does not reflect entirely to the dimensional score) both when a single-item or multi-item assessment were used. This speculative hypothesis would need empirical validation although some aspects suggest the dimensional nature of the phenomenon. A dimension in fact recognizes the participation of many influences (genetic, environmental) while a category usually recognizes a more net causal factor, a cumulative or threshold effect (Ruscio & Ruscio, 2008). Bullying recognize many factors that interact between them and increase the probability of its occurrence (Juvonen & Graham, 2014).

This paper had several strengths: 1) the sample is large and it is rather homogeneous and cut into a specific age group; 2) we used both single-item questions and multi-item scale for traditional bullying. Limitations regarded the cross-sectional design that prevents us to test for other measures,
such as test retest reliability of the scales. Furthermore, we need to investigate the relationship between traditional and cyber-bullying and our research group has already working on this issue.

Acknowledgments: The research project was partially funded by “Fondazione Banco di Napoli”. We would thanks the head of the schools that participate to the research projects and all teachers, school staff, parents and students involved.

GC, SP, VPS, MC, MS, KR, SS, AG, AP and MRB are all part of a research group whose aim is the study of the relationship between psychosocial phenomena and psychopathological symptoms (The Bullying and Youth Mental Health Naples Study, BYMHNS)
References


