

Parti Katalin:<sup>1</sup>

## What works in anti-bullying programs? Evaluative assessment best practices and recommendations

### I. Objective

The study gives an overview of the five most popular school violence programs containing an anti-bullying component running in the U.S. today. The description of the major content elements and evaluation of the programs have a critical approach: we use the experience gained during the impact study of the programs to set up a standard for evaluation criteria as far as bullying assessment and implementation evaluative assessment are concerned. This document is a methodological analysis of the programs, and explains on the one hand (1) based on what principles and methods the *phenomenon* of bullying was examined, and (2) what evaluative methods were used to assess the *effectiveness*, applicability and adaptability to the given population *of the programs* on the other.

We formulate our proposals with the *ideal* methods for assessing the effect of the phenomenon and the programs in mind. In actual reality, conditions are only in the rarest cases truly ideal for the introduction of programs – e.g. the decisions of authorities or the school management, the involvement of educators and parents, the motivation of the children, or financial resources may be lacking. Despite that, we considered it important to paint an ideal picture, as we have to know the goals that we should strive for to be able to get closer to achieving them.

### II. The selected programs and their commonalities

It is important to overview "what works" strategies, i.e. programs that have been proven to work in practice, because after the theoretical foundations are laid, it is practice that may show us whether we can achieve the desired effect with the individual program elements. At the same time, choosing an evaluation method for assessing impacts that is not suitable may bring misleading results. For this reason, not only strategies, but also evaluation techniques set up to assess results need to be elaborated with the greatest care and thoroughness.

The study builds up the ideal evaluation criteria based on the review of the characteristics and results of the following five programs:

- Olweus Bullying Prevention Program (OBPP);
- Positive Action (PA);

---

<sup>1</sup> Katalin Parti is a lawyer, sociologist, Ph.D. in legal and political sciences, senior research fellow with the Hungarian National Institute of Criminology. This study was conducted as a preparatory work for the Institute of Digital Media and Child Development, a project supported by the National Academy of Sciences to define the nation's research agenda regarding children and digital technology in the U.S. Author contact: parti@okri.hu

- Second Step (SS) and Second Step: Student Success Through Prevention middle school program (SS-SSTP) (Committee for Children);
- Steps To Respect (STR) (Committee for Children);
- KiVa, an anti-bullying program developed in Finland.

These five programs were selected for the analysis because all of them are programs in use in the U.S. (among other countries), that:

- Look back on many years or even a decade of *experience*.
- Several *impact studies* linked to the programs provide evidence of the operation of the programs in practice, their applicability, and also the less successful program elements.
- They are school-based intervention programs with a *reliable* (experimental or quasi-experimental<sup>2</sup>) *group design*. Each program reports data in a statistical format necessary to calculate effect size (e.g. means, standard deviation, group sizes, percentages).
- They are also *implemented internationally*, and they have been successfully adapted in several countries and on different continents (e.g. Europe, America, Australia).
- They contain *components developed for bullying situations*: a curriculum, training and other sensitization tools for educators and students. The research reports of the chosen programs indicate that intervention for bullying behavior is either the primary focus or a main component of the program.
- They contain *components developed for cyberbullying situations* – as part of the curriculum (case solution, lessons focused on cyberbullying) or in the form of a "cyberbullying toolkit".
- They are complex program packages aimed at the whole community – teachers, other school staff, students, parents – based on the *whole-school approach*. This is essential, because only those programs managed to achieve a positive impact, that were aimed at the whole school and the community outside the school beyond the classroom curricula and social skills training.<sup>3</sup>
- They are complex module systems, which approach the targeted groups of population with *different methods*, and contain e.g. teacher training, parental meetings, student curricula, and sensitizing conferences addressed to the community coming into regular contact with the children.
- They use a *terminology* widely accepted in international literature on the topic. The definition of bullying used by the reviewed programs can be described as

---

<sup>2</sup> An experimental study is a type of evaluation that seeks to determine whether a program or intervention had the intended causal effect on program participants. A quasi-experimental study is a type of evaluation, which aims to determine whether a program or intervention has the intended effect on a study's participants. Quasi-experimental studies take on many forms, but may best be defined as lacking key components of a true experiment. While a true experiment includes (1) pre-post test design, (2) a treatment group and a control group, and (3) random assignment of study participants, quasi-experimental studies lack one or more of these design elements.

<sup>3</sup> J. David Smith – Barry H. Schneider – Peter K. Smith – Katerina Ananiadou: The effectiveness of whole-school antibullying programs: A synthesis of evaluation research. *School Psychology Review*, 2004/4., 547-560. pp.

follows. "The definition of school bullying includes several key elements: (1) physical, verbal, or psychological attack or intimidation that is intended to cause fear, distress, or harm to the victim; (2) an imbalance of power (psychological or physical) with a more powerful child (or children) oppressing less powerful ones; (3) and repeated incidents between the same children over a prolonged period of time."<sup>4</sup> Consequently, it is not considered bullying when two persons of equal power (physical, psychological, or verbal) or in equal positions of power come into a one-off conflict. A further attribute of school bullying is that it occurs on the premises of the school, or outside that on the way to school, or from school home<sup>5</sup> or between the students of a school, whose effect is felt at school – in the form of ruining school climate, disrupting the peaceful learning atmosphere, or even weakening the cohesion of a smaller community (school learning groups, classes, etc.).

Characteristically these programs include prevention and intervention elements or a combination of the two: they exert their influence through general sensitization and individual skills-building. Within that, they put more emphasis on proactive, preventive techniques, than on reactive, subsequent response. Consequently, they contain education of not only the bully and the victim, but also the bystanders, children standing up for each other and in some cases peer mentor training are integral parts of the programs as well.

The programs consist of several components on several levels. A typical element is the classroom curriculum, which may range from a few sessions to a systematic offer of classes for a whole semester. The written curriculum of classes is sometimes supported by video spots, and in some cases (KiVa) an online video game helps understand the incidents. In addition to the informative-sensitizing classes held for students, direct training for educators also appears as a compulsory program element, which may be complemented by indirect training for a trainer diploma (train the trainer). Educators with such a diploma are entitled to hold trainings for other educators at their own or other schools. The programs also include parent trainings/meetings on specific issues, and school conferences. The programs also support the teaching staff with complete protocols for the resolution of cases, and responses. All written materials of the programs (curriculum, black lines, procedural protocols, reporting forms, information materials for parents, etc.) are available and can usually be downloaded for a fee from the program websites. Programs also include improved playground supervision and a whole-school anti-bullying policy, disciplinary (but non-punitive) methods, classroom rules, and cooperative group-work.

---

<sup>4</sup> Referred by David P. Farrington: Understanding and preventing bullying. In: Michael Tonry (Ed.): *Crime and Justice*, 17. University of Chicago Press, Chicago, 1993, 381-458. pp.; Dan Olweus: *Bullying at School. What We Know and We Can Do*. Blackwell, Oxford, 1993; Maria Ttofi – David P. Farrington: Effectiveness of school-based programs to reduce bullying: A systematic and meta-analytic review. *Journal of Experimental Criminology*, 2011/7., 27-56. pp.

<sup>5</sup> Ttofi – Farrington: Id.

Of the meta-analyses overviewing anti-bullying programs<sup>6</sup> the most rigorous so far is the Ttofi & Farrington study<sup>7</sup>, aggregating 44 different program evaluations based on the procedure of the Campbell Collaboration Systematic Review<sup>8</sup>, which compared the findings of randomized clinical trials. It established that the analyzed programs based on the procedure of appropriately documented, randomized controlled trials reduced the incidence of bullying by an average 20 to 23%, and victimization by 17 to 20%.

It has also become clear that the sole use of an anti-bullying curriculum shows no significant correlation with either perpetration or victimization.<sup>9</sup> The different program elements can change school climate and improve bullying indices in an interdependent and synergistic way.

Programs focusing on older children (above the age of 11) achieved better results than programs aimed at younger children.<sup>10</sup> Admittedly, according to Smith and his colleagues this is possible, because Ttofi and Farrington applied between-program comparisons instead of within-program comparisons in their meta-analysis. Any one program, however, is aimed at several age groups, and contains age-specific components. For this reason, the effects of age-specific program components need to be assessed for each age group within the same program. Smith and colleagues have proven that if the comparison is made between the age groups, it is exactly the age group six to 11 years where the most positive effect can be achieved.<sup>11</sup>

Anti-bullying programs introduced in the U.S. rely on the elements and methods of the Olweus Bullying Prevention Program (OBPP) for the assessment, prevention and management of the phenomenon. In contrast to the original OBPP implemented in Norway on a nationwide sample, these research projects produced statistically not significant, in certain cases negative, but at best mixed results. Evans and his colleagues, who compared the findings of research published between 2011 and 2014 following the meta-analysis of Ttofi and Farrington<sup>12</sup> establish, that the programs implemented in

---

<sup>6</sup> See Smith – Schneider – Smith – Ananiadou: Id.; Friedrich Lösel – Andreas Beelmann: Effects of child skills training in preventing antisocial behavior: A systematic review of randomized evaluations. *Annals of the American Academy of Political and Social Science*, 587, (2003), 84-109. pp.; Kenneth W. Merrell – Duane Isava – Barbara A. Gueldner – Scott W. Ross: How effective are school bullying intervention programs? A meta-analysis of intervention research. *School Psychology Quarterly*, 2008/1., 26-42. pp.; Ttofi – Farrington: Id.; Caroline B.R. Evans – Mark W. Frazer – Kattie L. Cotter: The effectiveness of school-based bullying prevention programs: A systematic review. *Aggression and Violent Behaviour*, 19 (2014) 532-544. pp.; David Scott Yeager – Carlton J. Fong – Hae Yeon Lee – Dorothy L. Espelage: Declines in efficacy of anti-bullying programs among older adolescents: A developmental theory and a three-level meta-analysis. *Journal of Applied Developmental Psychology*, 37 (2015) 36-51. pp.

<sup>7</sup> Ttofi – Farrington: Id.

<sup>8</sup> *Campbell collaboration systematic reviews: Policies and guidelines (V 1.0)*. [www.campbellcollaboration.org](http://www.campbellcollaboration.org); (last accessed: 30.04.2015)

<sup>9</sup> Ttofi – Farrington: Id.

<sup>10</sup> Id.

<sup>11</sup> Peter K. Smith – Christina Salmivalli – Helen Cowie: Effectiveness of school-based programs to reduce bullying: A commentary. *Journal of Experimental Criminology*, 2012/8., 433-441. pp.

<sup>12</sup> Caroline B.R. Evans – Mark W. Frazer – Kattie L. Cotter: The effectiveness of school-based bullying prevention programs: A systematic review. *Aggression and Violent Behaviour*, 19 (2014) 532-544. pp.

Europe may have produced better results, because the assessed population and the sample taken were more homogenous.

### **III. Evaluative assessment best practices**

Evaluation may focus on reviewing the content requirements of the program. This is content assessment (i.e. bullying assessment), which measures the degree of implementation of program components and the program's intensity. If the result of evaluation is negative, the program was not implemented correctly, and the phenomenon was not measured using the criteria defined in the program description. In such a case, there may be problems with the degree of intensity, the frequency of monitoring, and the degree of implementation. In contrast to that, program assessment (i.e. implementation assessment) examines whether the desired results were achieved with the program, i.e. whether the program is suitable for triggering the changes that it is expected to achieve. If so, the program can be continued, but it is recommended to adjust the components, measuring tools and methods to arising new needs identified by way of continuous needs assessment. If as a result of the program the desired results did not materialize (e.g. bullying and victimization indices did not fall, the school staff and the children did not become more sensitive towards the issue, and they did not learn any new tools to help them break out of their bystander role, etc.), the program's faults need to be resolved after an investigation of causes (method, intensity, personal or financial conditions, etc.), and the program can be continued accordingly, in a developed form.

#### **1. Recommendations for bullying assessment. What to measure?**

##### **Targeted techniques by subgroups**

It has become clear from evaluative assessments that it is easier to teach victims protection mechanisms, than convincing perpetrators to leave the path that they have chosen. In this context, program developers have the responsibility in the future to examine how reliably the strategy they have developed is suitable for the re-education of the perpetrators. On the other hand, desistance-enhancing techniques need to be elaborated specifically for perpetrators and they need to be incorporated into programs. The curriculum needs to include strategies for popular bullies, who may be able to replace their violent behavior with peaceful, socially accepted methods, and still keep their high status in the community. Research is needed into how the coaching of children (by teachers) who bully differs from coaching of those who are victims or bully-victims. Is coaching differentially effective among these groups? Are such interventions effective if they do not occur in the context of a comprehensive program?<sup>13</sup> The strategies need to be tested, and methods that truly work need to be identified.

Yeager and colleagues recommend that the programs should introduce different language use suited to the different age groups while conveying the same set of rules. In younger

---

<sup>13</sup> Miriam K. Hirschstein – Leihua Van Schoiack Edstrom – Karin S. Frey – Jennie L. Snell – Elizabeth P. MacKenzie: Walking the Talk in Bullying Prevention: Teacher Implementation Variables Related to Initial Impact of the Steps to Respect® Program. *School Psychology Review*, 2007/1, 3–21. pp.

age groups it is controlling language that is more effective, while later it is autonomy-supportive language.<sup>14</sup> The effectiveness of programs can be increased with this simple technique when working with adolescents in their late teens.

It is also Yeager and colleagues who recommend that it should be examined in each age group separately what components have an effect and why. So for example among three- to 11-year-olds popularity is a strong and significant predictor of bullying, but among older adolescents aged 12 to 18, popularity is a weaker, non-significant predictor of becoming a bully.<sup>15</sup> They also establish, however, that among moderate-to-high popular students the programs may be effective, but they show no effect on the most popular students (of the programs reviewed, only KiVa examined this phenomenon, and came to a similar conclusion.<sup>16</sup> The suitability of social-emotional skill instructions for children who demonstrate a high level of social intelligence and are capable of manipulating others is questionable. Skills need to be developed therefore for emotion regulation and assertiveness, which are specifically targeted at popular children.<sup>17</sup>

A striking gap in the literature on anti-bullying programs is that while students with disabilities have been identified as one subgroup that is potentially at a higher risk of experiencing and engaging in bullying, bullying prevention programs have not been developed for use with this specific population. None of the programs assessed have been researched with students with disabilities and none of the program guidelines include information on how to adapt or modify the program to make it more accessible for these populations.

Intensive bullying activity is only characteristic of a small group.<sup>18</sup> Most of those involved in bullying are bystanders, or are on rare occasions victims or perpetrators. More research is needed into how the "hard core" – those who regularly involved in bullying – can be approached, what skills they need to be taught to achieve desistance. According to research, African American, Hispanic and other ethnic minorities show different characteristics in school violence.<sup>19</sup> It needs further investigation what characteristics school bullying has along ethnic and socio-economic variables and

---

<sup>14</sup> Yeager – Fong – Lee – Espelage: Id.

<sup>15</sup> Clayton R. Cook – Kirk R. Williams – Nancy G. Guerra – Tia E. Kim – Shelly Sadek: Predictors of bullying and victimization in childhood and adolescence: A meta-analytic investigation. *School Psychology Quarterly*, 2010/25, 65–83. pp.

<sup>16</sup> Claire F. Garandea – Ihno A. Lee – Christina Salmivalli: Differential effects of the KiVa anti-bullying program on popular and unpopular bullies. *Journal of Applied Developmental Psychology*, 35, (2014), 44–50. pp.

<http://www.sciencedirect.com/science/article/pii/S0193397313000828>; (last accessed: 08.05.2015)

<sup>17</sup> Karin S. Frey – Susan Bobbitt Nolen – Leihua Van Schoiack Edstrom – Miriam K. Hirschstein: Effect of a school-based social-emotional competence program: Linking children's goals, attributions, and behavior. *Applied Developmental Psychology*, 26 (2005a), 171–200. pp.

<sup>18</sup> See for example Karin S. Frey – Miriam K. Hirschstein – Jennie L. Snell – Leihua Van Schoiack Edstrom – Elizabeth P. MacKenzie – Carole J. Broderick: Reducing playground bullying and supporting beliefs: An experimental trial to the Steps To Respect program. *Developmental Psychology*, 41 (2005b), 479–491. pp.

<sup>19</sup> Sandra Graham – Jaana Juvonen: Ethnicity, peer harassment, and adjustment in middle school: An exploratory study. *Journal of Early Adolescence*, 22, (2002), 173–199. pp.

matching program elements need to be developed. Although researchers in recent years have begun investigating racially motivated bullying and harassment<sup>20</sup>, little is known about intra-racial/ethnic bullying.

Sexual maturation may further influence the reaction of age groups to the program. It may be a catalyst especially in relation to bullying manifesting as relational aggression.<sup>21</sup> The evaluative assessments of several programs were not performed for single age groups. For example the relatively recent KiVa is currently being tested in the lowest age group, and PA has never been assessed in the age group 15 to 18.<sup>22</sup> In order to target high school populations, we need to identify first what factors make programs more successful in older age groups.<sup>23</sup> The curriculum is required to be adjusted to the needs of the older age groups. KiVa, for example achieves this by offering separate online games for younger children (KiVa Game with anima-style characters) and adolescents (KiVa Street, featuring an edgy teen with skateboarder clothes and stocking cap).<sup>24</sup>

Even though programs act against bullying on many different levels – classroom, individual, community level –, there is currently very little research into the correlation of classroom or individual level implementation of bullying prevention programs *and* program effects. E.g. STR examined measures against bullying at classroom level (“talk the talk”: frontal presentations) and at individual level (“walk the talk”: coaching, one-on-one interactions), but the assessment was not suitable for separating the effects of the two levels.<sup>25</sup> These questions also need attention in the future.

### **Finding a definition**

According to research, if the questionnaire determines the definition, we can expect a lower prevalence as a result of the survey, which suggests that it could be a way to ensure validity and avoid term overuse.<sup>26</sup> If we want to measure with a single question (one-item measure), without a definition, the object of the survey will be compromised, and it is therefore recommended to ask questions about specified behaviors. In such a case, most

---

<sup>20</sup> Stephen L. Wessler – Leila L. De Andrade: Slurs, Stereotypes, and Student Interventions: Examining the Dynamics, Impact, and Prevention of Harassment in Middle and High School. *Journal of Social Issues*, 2006/3, 511–532. pp.

<sup>21</sup> Tracy Vaillancourt – Jessie L. Miller – Aanchal Sharma: “Tripping the prom queen”: Female intrasexual competition and indirect aggression. In: Karin Österman (Ed.): *Indirect and direct aggression*. Peter Lang Publishing, Frankfurt, Germany, 2010.

<sup>22</sup> Kin-Kit Li – Isaac Washburn – David L. DuBois – Samuel Vuchinich – Peter Ji – Vanessa Brechling – Brian Flay: Effects of the Positive Action programme on problem behaviors in elementary school students: A matched-pair, randomized control trial in Chicago. *Psychology & Health*, 2011/2., 187-204. pp.

<sup>23</sup> Yeager – Fong – Lee – Espelage: Id.

<sup>24</sup> Antti Kärnä – Rinus Voeten – Todd D. Little – Elisa Poskiparta – Erkki Alanen – Christina Salmivalli: Going to scale: A nonrandomized nationwide trial of the KiVa antibullying program for grades 1-9. *Journal of Consulting and Clinical Psychology*, 2011/6 (2011a), 796-805. pp.; Antti Kärnä – Rinus Voeten – Todd D. Little – Elisa Poskiparta – Anne Kaljonen – Erkki Alanen – Christina Salmivalli: A large-scale evaluation of the KiVa antibullying program. *Child Development*, 82 (2011b), 311-330. pp.

<sup>25</sup> Hirschstein – Van Schoiack Edstrom – Frey – Snell – MacKenzie: Id.

<sup>26</sup> Susan M. Swearer – Dorothy L. Espelage – Tracy Vaillancourt – Shelley Hymel: What can be done about school bullying? Linking research to educational practice. *Educational Researcher*, 39, (2010), 38–47. pp.

<http://digitalcommons.unl.edu/edpsychpapers/141>; (Last accessed: 15.05.2015)

questionnaires do not provide a definition,<sup>27</sup> but it would still be desirable for more precision. All of the programs assessed use some modified version of the Olweus Bullying/Victimization Questionnaire,<sup>28</sup> but there are several anti-bullying programs running in the U.S., which attempt to measure bullying by listing the elements of violence, which however does not fully correspond to the phenomenon they intend to measure. Also, anti-bullying research should include as an essential element the measurement of the dynamics of bullying, i.e. the role of the group. Despite that, Second Step used the University of Illinois Bullying/victimization Scale,<sup>29</sup> Steps to Respect the School Experiences Survey,<sup>30</sup> which contain items for measuring violence, even though both programs refer to action against bullying. As an example not discussed here in detail, the School Experiences Survey concentrates so much on violence, that it lacks any items on power imbalance, repetition and intent. There are also other standards beside Olweus' basic definition, which may serve as a reference point for research, such as e.g. the manual developed by Centers for Disease Control and Prevention.<sup>31</sup>

According to social-ecological and evolutionary psychology research, bullying is a behavior rooted in evolution, it is more intentional than accidental, and bears the characteristics of proactive aggression to maintain a position in a community rather than reactive aggression.<sup>32</sup> We should consider integrating these newly developed elements (goal-directedness, proactive aggression) into the definition of bullying. Prevention and intervention programs need to be amended with these elements in mind, i.e. program elements adjusted to the aim of bullying need to be added.<sup>33</sup>

The use of the terms 'bully' and 'bullying' in questionnaires and other measurement tools is also another disputed issue. Some research uses the term bullying, while others list all the behaviors that bullying includes. Research proves that use of the term bullying may elicit socially desirable responses.<sup>34</sup> Others suggest that sensitization works as a backlash of prevention programs: overusing the term 'bullying' may compromise validity of

---

<sup>27</sup> Evans – Frazer – Cotter: Id.

<sup>28</sup> *Olweus Bullying Questionnaire Script*, Hazelden Foundation, 2009, <http://www.sisd.net/cms/lib/TX01001452/Centricity/Domain/56/Teacher%20Script%20for%20Olweus%20Questionnaire.pdf>; (last accessed: 30.05.2015); Mona E. Solberg – Dan Olweus: Prevalence estimation of school bullying with the Olweus bully/victim questionnaire. *Aggressive Behavior*, 29, (2003), 239–268. pp.

<sup>29</sup> Dorothy Espelage – Melissa K. Holt: Bullying and victimization during early adolescence: Peer influences and psychosocial correlates. *Journal of Emotional Abuse*, 2001/2, 123-142. pp.

<sup>30</sup> Eric C. Brown – Sabina Low – Brian H. Smith – Kevin P. Haggerty: Outcomes from a school-randomized controlled trial of Steps to respect: A bullying prevention program. *School Psychology Review*, 40, (2011), 423-443. pp.

<sup>31</sup> Matthew R. Gladden – Alana M. Vivolo-Kantor – Merle E. Hamburger – Corey D. Lumpkin: *Bullying surveillance among youths: Uniform definitions for public health and recommended data elements, Version 1.0*, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention and U.S. Department of Education, 2014.

<sup>32</sup> Anthony A. Volk – Andrew V. Dane – Zopito A. Marini: What is bullying? A theoretical redefinition. *Developmental Review*, 2014/4., 327-343. pp.

<sup>33</sup> Ttofi - Farrington: Id.

<sup>34</sup> Dorothy L. Espelage – Susan M. Swearer: Research on school bullying and victimization: What have we learned and where do we go from here? *School Psychology Review*, 32, (2003), 365–383. pp.

data.<sup>35</sup> Further research is needed to examine the strength and limitations of each approach.

### **Culture and adaptability**

It needs to be examined whether the positive program effects can be implemented in other countries. It is possible that different countries will show significant differences in culture, school types, school homogeneity, and the school system. For this reason, the possibilities of adaptation, and the barriers to implementation between countries need to be investigated before adapting any program in a foreign country. Forecasts need to be compiled about the different needs (needs assessment), and about how attitudes related to bullying influence the effect of the new program. Before adaptation, the changes potentially needed in the program are to be mapped. To accommodate that, programs need to be designed in a way that makes it possible for them to be flexibly adjusted in other cultures and school systems. The rules need to be concrete, but suitably flexible. For example in Eastern Europe, where today's top-down controlled, authoritarian school system bears the marks of the historical Prussian educational system, and norm crisis has become wide-spread after the fall of communism, it is necessary to examine how these effects can influence the effects of the program to be adapted.<sup>36</sup>

Care must be given to the assessment of such culture-dependent environmental factors as poverty incidence, the everyday survival issues of disadvantaged families, problems from the social welfare system (e.g. the work-load of parents prevents them from properly supervising the leisure activities of their children), traditions of discipline in the family and at school, the structure of the school system, basic education of educators, equipment and infrastructures of schools, ethnic setup of students. It is also necessary to measure what problems students are facing at school and in their living environment.

The Olweus Bullying/Victimization Questionnaire needs to be amended and enable it to measure everyday stress factors in low-income neighborhoods and also peer aggression/bullying. The stress factors of educators also need to be measured. In the knowledge of these, school- of school district-specific bullying prevention programs can be amended with the suitable strategic elements.

The community outreach component of programs needs to be strengthened, which is particularly needed in regions grappling with poverty. Opportunities need to be created for community leaders to get involved in school-based anti-bullying programs using their own resources.

---

<sup>35</sup> Elizabeth K. Englander: *Bullying and Cyberbullying: What every educator needs to know*. Harvard Education Press, Cambridge, MA, 2013.

<sup>36</sup> Péter Fóti: *Poroszos-e a mai magyar iskolarendszer? [Is the current Hungarian school system Prussian?]* 2009.

<http://www.foti-peter.hu/porosz.html>; (last accessed: 16.05.2015)

William S.F. Pickering – Geoffrey Walford: *Durkheim's Suicide: A century of research and debate*. British Centre for Durkheimian Studies. Routledge, London and New York, 2000, 25. p.

### **Causal research and the research of latent and complex bullying phenomena**

None of the reviewed programs managed to achieve a drop in the non-physical bullying component. It is possible that the programs are only able to manage the simple, visible forms of bullying, and are not suitable for the prevention and reduction of more complex and latent phenomena. Evaluations lag behind in terms of causal research, what is more, they do not examine the development of cyberbullying, as a form of non-physical bullying. In the future, more energy needs to be invested in the examination of the phenomenon of cyberbullying. Positive changes in indirect – verbal, cyber- or relational – bullying demand a longer time. Changes are necessary in the sensitivity and attitude of not only students, but also educators.<sup>37</sup> Care must be given to the development of program elements, which are capable of changing attitudes, and the testing of these.

Research has not assessed the correlation of lack of empathy and bullying. Descriptive work suggests a possible link between lack of empathy for victims and perpetrating or watching bullying.<sup>38</sup> The link between lack of empathy and bullying needs to be researched in the future.

Malicious gossip may be a starting point of bullying, or one of its components. Only a few studies have investigated this phenomenon, its spread and role in changing group status<sup>39</sup> this needs to be corrected in the future.

### **Practical solutions for educators**

Beyond the universal approach (targeting whole schools and classrooms) – which every anti-bullying program contains – we also need to offer options for solving current incidents. Indicated actions (which was elaborated the most thoroughly by KiVa of all the reviewed programs) are needed to intervene in ongoing bullying (for example systematic discussion techniques for addressing bullying cases; simulation exercises during teacher training). Educators usually say they miss techniques the most in the programs that they can use in every day practice.<sup>40</sup> Children are looking for role models, while teachers need to be taught field intervention techniques, which they can use to actually intervene in incidents (e.g. coaching, reinforcement in the moment, which STR uses successfully) and may show children an example for learning supportive attitudes. We need to assess which practice can be used in which situation the most effectively.

---

<sup>37</sup> Dorothy L. Espelage – Sabina Low – Joshua R. Polanin – Eric Brown: The impact of middle school program to reduce aggression, victimization, and sexual violence. *Journal of Adolescent Health*, 53, (2013), 180-186. pp.

<sup>38</sup> Inger M. Endresen – Dan Olweus: Self-reported empathy in Norwegian adolescents: Sex differences, age trends, and relationship to bullying. In Arthur C. Bohart – Deborah J. Stipek (Eds.): *Constructive & destructive behavior: Implications for family, school, and society*. American Psychological Association, Washington, D.C., 2001, 147–165. pp.; Anthony D. Pellegrini – Jeffrey D. Long: A longitudinal study of bullying, dominance, and victimization during the transition from primary school through secondary school. *British Journal of Developmental Psychology*, 20, (2002), 259–280. pp.

<sup>39</sup> Frey – Hirschstein – Snell – Van Schoiack Edstrom – MacKenzie – Broderick 2005a; Zsolt Boda – Bálint Néray: Inter-ethnic friendship and negative ties in secondary school. *Social Networks*, 43, (2015), 57-72. pp.

<sup>40</sup> Elizabeth K. Englander – Katalin Parti – Meghan McCoy: Evaluation of A University-Based Bullying and Cyberbullying Prevention Program. *Journal of Modern Education Review*, 2015/10., 937-950. pp.

## The rationale for research on cyberbullying

There are several studies available on the rationale for the research of the cyberbullying phenomenon.<sup>41</sup> Even though studies invariably state that cyberbullying is a low frequency phenomenon compared to school bullying, they also agree that (1) the term needs to be used restrictively – Peter K. Smith uses the definition of Olweus for traditional bullying and extends it to electronic perpetration; (2) the term needs unification, otherwise the findings of research will not be comparable and the research will not be reproducible; (3) it needs to be labeled as a form of bullying, and not as a completely different phenomenon. They however also acknowledge that there are phenomena – e.g. online bystander roles and dynamics, and the process of online bullying on social networking sites –, which are more complex and require separate assessment.<sup>42</sup> Olweus argues that traditional bullying and cyberbullying overlap – since the majority of cyberbullied children are also exposed to traditional bullying –, and consequently, if we treat traditional bullying properly, we can also combat its forms in cyberspace.<sup>43</sup> Others<sup>44</sup> would like to see separate program packages for the treatment of cyberbullying. In fact, there are no reliable surveys into these questions. Research has failed to identify overlaps between school and online bullying, and also typically co-occurring forms of appearance. According to Gradinger and colleagues, the perpetration and victimization forms of both traditional and cyberbullying have to be identified, for us to be able to establish the characteristics of risk groups, which were not receptive to programs mainly treating traditional bullying.<sup>45</sup> Olweus also agrees however that there are cardinal questions in relation to cyberbullying that demand an answer. These are:<sup>46</sup>

- To what extent can the traditional criteria of bullying, intentionality, repetition, and power imbalance, be applied to cyberbullying?
- Do cyber-victimization / bullying items differ in important ways from traditional victimization / bullying items measuring physical, verbal and indirect / relational forms of bullying?
- Do cyber-victimization / bullying items and other victimization / bullying items go together in one factor and, in case, have roughly similar loadings?
- Do cyber-victimization / bullying variables relate in the same way as traditional victimization / bullying variables to psychosocial adjustment dimensions they can be expected to be associated with?

---

<sup>41</sup> Dan Olweus: Cyberbullying: An overrated phenomenon? *European Journal of Developmental Psychology*, 2012/5, (2012a), 520–538. pp.; Dan Olweus: Comments on cyberbullying article: A rejoinder. *European Journal of Developmental Psychology*, 2012/5, (2012b), 559-568. pp.; Sameer Hinduja – Justin W. Patchin: Cyberbullying: Neither an epidemic nor a rarity. *European Journal of Developmental Psychology*, 2012/5., 539–543. pp.; Peter K. Smith: Cyberbullying: Challenges and opportunities for a research program. A response to Olweus (2012). *European Journal of Developmental Psychology*, 2012/5., 553–558. pp.; Ersilia Menesini: Cyberbullying: The right value of the phenomenon. Comment on the paper: Cyberbullying: An overrated phenomenon? *European Journal of Developmental Psychology*, 2012/5., 544-552. pp.

<sup>42</sup> Smith: Id.

<sup>43</sup> Olweus: Id.

<sup>44</sup> Hinduja - Patchin: Id.; Smith: Id.

<sup>45</sup> Petra Gradinger – Dagmar Strohmeier – Christiane Spiel: Traditional bullying and cyberbullying. *Zeitschrift für Psychologie/Journal of Psychology*, 217, (2009), 205–213. pp., 212. p.

<sup>46</sup> Olweus: Id.

Cyberbullying is still less frequent than school (offline) bullying. Cyberbullying is a low-frequency and high-intensity incident, which may be more damaging than school bullying.<sup>47</sup> Compared to the low prevalence, it has the potential to have such a profound effect on the wellbeing and social status of individuals, that this in itself justifies the existence of research into the forms of appearance and effect of cyberbullying.<sup>48</sup>

## **2. Proposals for program evaluative assessment. How to be more effective?**

### **Improve the qualitative indices of the degree of implementation**

In order to improve the qualitative indices of the degree of implementation, it is recommended to organize meetings and conferences with stakeholders (educators, guidance counselors), and to continuously monitor participating schools. The close connection is necessary *before* and *during* the implementation of the program as well, as it makes it possible to control the suitability of the degree of implementation. KiVa can be regarded as a model in this respect, as the program provides detailed guides to schools, and prepares them at conferences and interactive meetings.

### **Fidelity monitoring**

Several studies mention the effect of systematic monitoring on implementation fidelity.<sup>49</sup> The STR program, which expressly emphasizes the importance of fidelity monitoring in the interest of maintaining the program effect, also mentions among the financial conditions that program fidelity monitoring is free of charge. This program guarantees continuous supervising making sure that the schools implement the program according to requirements. Other anti-bullying programs do not concentrate so much on program fidelity, what is more, some programs actually allow schools to choose those program elements freely that best suit local needs and teachers to only hold the classes of the curriculum that they like. In order to ensure validity, it would be more fortunate to find a middle ground and offer the option of program monitoring to these programs, even if for a minimal fee.

### **The importance of consistency and continuity**

Programs started in preschool need to be continued into the late teens (if conditions are given and the program contains components for older age groups). It is however necessary to check the components for age-appropriateness, i.e. whether they meet the needs of the older adolescent age groups, and whether rules can be made to fit the adolescents' drive for independence, and their psychological and psychosexual development processes.

---

<sup>47</sup> Volk – Dane – Marini: Id.

<sup>48</sup> Sonia Livingstone – Lucyna Kirwil – Cristina Ponte – Elisabeth Staksrud: *In their own words: What bothers children online?* 2013  
<http://www.lse.ac.uk/media@lse/research/EUKidsOnline/EU%20Kids%20III/Reports/Intheirownwords020213.pdf>; (last accessed: 03.05.2015)

<sup>49</sup> Smith – Schneider – Smith – Ananiadou: Id.

### **Combined use of quantitative and qualitative methods may promote the expression of attitudes towards the program**

Mostly quantitative methods are used to measure program effectiveness – e.g. pre- and post self-report test with students, and self-report perception tests with teachers and parents (school administrator reports, parent reports). The PA program also recorded qualitative reports: parents and children had the opportunity to express their opinion about the program and to what an extent its effects met with their expectations in comments and essays.<sup>50</sup> This method complements multiple choice tests very well, and with its help we may learn about impressions on issues, which cannot be measured with multiple choice tests. The research EU Kids Online for example recorded one-on-one interviews with parents, who explained to what an extent they think their children are threatened online, or to what degree their child is involved in school bullying.<sup>51</sup> The researchers of EU Kids Online also asked open questions to measure such sensitive issues as to what an extent the children felt disturbed by the incidents that happened with them online, and how much time they needed to process their psychological injuries.<sup>52</sup> When performing the evaluative data collection of the anti-bullying program of the Massachusetts Aggression Reduction Center, I also opted for the technique of interviews surveying the educators of the participating schools. During the one-on-one interviews, I gained information on program implementation, which would not have been possible with the limited number of responses in multiple choice tests.<sup>53</sup>

### **We need to repeat the surveys every three years to check program results**

The program evaluations prove that some program effects may change – strengthen or weaken – over the years. For example the positive effect of the reviewed programs seen in the first year was reduced or disappeared in the second year of intervention, and also with older children, who have been participating in the program for longer.<sup>54</sup> Ryan and Smith<sup>55</sup> recommend a three-year follow-up period to properly investigate the effects of the intervention program. The developers of KiVa emphasize that this is necessary because the participating teachers should have enough time to solve the problems arising during implementation, and for program effects to manifest to the fullest extent possible.

---

<sup>50</sup> Positive Action Evaluation Report, Proposed Mental Health Services Act – Positive Action Evaluation Report Data from August 2013 - June 2014.

<http://www.co.shasta.ca.us/docs/HHSA/mental-wellness/positive-action-evaluation-report-final-oct-14.pdf>; (last accessed: 03.05.2015)

<sup>51</sup> EU Kids Online: findings, methods, recommendations 2014. EU Kids Online, LSE, London, UK. <http://eprints.lse.ac.uk/60512/>; (last accessed: 03.05.2015)

<sup>52</sup> Livingstone – Kirwil – Ponte – Staksrud: Id.

<sup>53</sup> Englander – Parti – McCoy: Id.

<sup>54</sup> Dan Olweus – Françoise Delange Alsaker: Assessing change in cohort-longitudinal study with hierarchical data. In: David Magnusson – Lars Bergman – Georg Rudinger – Bertil Törestad (Eds.): *Problems and Methods in Longitudinal Research: Stability and Change*, Cambridge University Press: New York, NY, 1991, 107-132. pp.

<sup>55</sup> Wendy Ryan – J. David Smith: Antbullying programs in schools: How effective are evaluation practices? *Prevention Science*, 2009/10., 248-259. pp.

### **Program effect can be assessed between-program and within-programs**

Meta-analyses are good example for the former, which are needed to filter out well working and effective program components. Nonetheless, if we want to examine the effect of the single components of a program on an age group, we need to do a within-program analysis. This may be necessary to establish in which age group the strategies offered in the program work the best, and in which grade they achieve the greatest possible effect. When we make statements about program effectiveness, we should not forget about the possibility of some program elements being effective only at a certain age.<sup>56</sup>

### **Modern info-communication tools' effect to be measured**

We have to examine the effect of modern info-communication tools – e.g. the use of virtual reality computer games. Despite the fact that these components have been introduced by several programs so far, there is no data on how and to what an extent they can help convey the prevention messages.<sup>57</sup>

### **Subsequent surveys to be conducted**

Subsequent surveys need to be conducted to check the validity of implementation evaluation results. A focus group discussion may be suitable for that, or with teachers the informal interview technique, which should be ideally conducted by an independent researcher.

## **IV. Recommendations for data collection and evaluative assessment studies**

The basic rules of measuring bullying can be summarized as follows:

- a) Bullying assessments need to be systematic and well-planned as a result of the cooperation of the professional team outside and inside schools. Programs need to be elaborated by an independent professional team after the assessment of local characteristics, needs and problems, but the implementation of the program also requires work from the teams at the institutions (administrators, counselors, teachers, guidance counselors and psychologists). The program should be piloted before implementation, if possible.
- b) Before implementation, racial, cultural, ethnic and socio-economic have to be mapped within the given school community, so that stratification can be ensured. This is the only way any program can be suitable for assessing the problems of the heterogeneous populations of schools, and for giving an adequate response.
- c) Before deciding on the use of any particular measure method, it needs to be examined whether the given method is suitable for measuring the desired phenomenon in the given environment (technical adequacy).

---

<sup>56</sup> Smith – Salmivalli – Cowie: Id.

<sup>57</sup> Ttofi – Farrington: Id.

- d) In the interest of reliability and validity, data need to be collected routinely and systematically over a long period, within a set timeframe.<sup>58</sup>
- e) Data have to be collected using more than one method. Qualitative techniques, like observation, interviews or teacher rating scale may complement or control the quantitative techniques, such as surveys, self-report and peer-report questionnaires.
- f) Parents are to be involved if possible both in the assessment and the implementation phase. Even though reaching out to parents is the most difficult aspect, their involvement in the program is necessary both to protect children's right to privacy and in the interest of program fidelity.
- g) The highest possible number of school staff must be involved in data collection. Recurrent trainings and quality checks should be used to ensure the competence of the personnel and the consistency of research.
- h) The school staff has to be informed about the findings of assessment and implementation, obviously without any prejudice to the privacy or personal data of the children. Schools can only be expected to understand problems and cooperate during implementation in this way.

The golden rules formulated in methodology literature and confirmed by anti-bullying programs need to be incorporated into the impact study of the programs:

- a) Care must be taken to conduct the programs using a unified and controllable, strictly documented methodology, so that the results can be compared between school districts or even states.
- b) Control schools need to be selected that match the program schools in the sample: intervention and control schools need to be matched based on grades, gender, settlement type, and other factors important for measuring; it is crucial that intervention and control schools should not show great differences along these variables.
- c) Schools must be randomly selected, by equal possibility selection into the intervention or the control group.
- d) More than one method has to be enlisted for the assessment of results (multi-method outcome assessment).
- e) The perseverance and continued participation of schools that started using the program is a must for longitudinal research, as is maintaining the initial enthusiasm and activity levels. This ensures the homogeneity of the sample on the one hand, and on the other lasting results can only be expected from consistent implementation ongoing for several years. Anti-bullying research has proven that a drop in self-reported and peer-reported bullying shows a statistically significant correlation with the degree of implementation.
- f) To achieve and maintain a high level of implementation, both the program provider and the school management need to assure educators of their support. This requires the preparation of suitably detailed manuals appropriate for practical use, the regular monitoring of teachers' activities (systemic monitoring analysis)

---

<sup>58</sup> Anthony D. Pellegrini – Maria Bartini: An empirical comparison of methods of sampling aggression and victimization in school settings. *Journal of Educational Psychology*, 92, (2000), 360–366. pp.

from the part of the program provider, and creating the conditions for a whole-school approach from the school management. If we expect feedback from educators, it may increase motivation just as much as material support (incentives).

- g) Last but not least, programs do not only need to be assessed, but the findings also need to be incorporated into the programs in a strictly documented manner. Program development and communication with the target group (school policy makers, school leadership, school staff) need to be continuous.

These methodological assumptions all have to be observed for a program's suitability for the reduction of bullying and victimization to be measured, and also for assessing whether the changes that occurred were actually triggered by the program.