

International Journal of Conflict and Violence

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Open Section

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

International Journal of Conflict and Violence – IJCv

The International Journal of Conflict and Violence (IJCv) is a peer-reviewed periodical for scientific exchange and public dissemination of the latest academic research on conflict and violence. It was included in the Social Sciences Citation Index (SSCI) in March 2011. The subjects on which the IJCv concentrates have always been the subject of interest in many different areas of academic life. Consequently, the journal encompasses contributions from a wide range of disciplines including sociology, political science, education, social psychology, criminology, ethnology, history, political philosophy, urban studies, economics, and the study of religions. The IJCv is open-access: All text of the IJCv is subject to the terms of the Creative Commons Attribution-NoDerivatives License. The IJCv is published twice a year, in spring and in fall. Each issue will focus on one specific topic while also including articles on other issues.

Editors

Prof. Dr. Wilhelm Heitmeyer, Editor-in-Chief (University of Bielefeld)
 Prof. Douglas S. Massey, Ph.D. (Princeton University)
 Prof. Steven F. Messner, Ph.D. (University at Albany, NY)
 Prof. Dr. James Sidanius (Harvard University)
 Prof. Dr. Michel Wieviorka (École des Hautes Études en Sciences Sociales, Paris)

Advisory Board

Prof. Tore Bjørgo, Ph.D. (Norwegian Police University College, Oslo)
 Prof. Ronald Crelinsten, Ph.D. (University of Victoria)
 Prof. Robert D. Crutchfield, Ph.D. (University of Washington, Seattle)
 Prof. Donatella della Porta, Ph.D. (European University Institute, Florence)
 Prof. Dr. Julia Eckert (University of Bern, Switzerland)
 Prof. Dr. Manuel Eisner (University of Cambridge)
 Prof. Richard B. Felson, Ph.D. (Pennsylvania State University)
 Prof. Gideon Fishman, Ph.D. (University of Haifa)
 Prof. Ted Robert Gurr, Ph.D. (University of Maryland)
 Prof. Dr. Heinz-Gerhard Haupt (University of Bielefeld)
 Prof. Miles Hewstone, Ph.D. (University of Oxford)
 Prof. Rowell Huesmann, Ph.D. (University of Michigan)
 Prof. Dr. Barbara Krahé (University of Potsdam)
 Prof. Gary LaFree, Ph.D. (University of Maryland)
 Prof. Jianhong Liu, Ph.D. (University of Macau, China)
 Prof. Jens Ludwig, Ph.D. (University of Chicago)
 Prof. Dr. Jitka Malečková (Charles University Prague)
 Dr. Nonna Mayer (Centre de Recherches Politiques de Sciences Po, Paris)
 Prof. Dr. Friedhelm Neidhardt (Social Science Research Center Berlin)
 Prof. Thomas Pettigrew, Ph.D. (University of California Santa Cruz, CA)
 Prof. Dr. Ulrich Schneckener (University of Osnabrück)
 Prof. Dr. Rashmi Singh (University of St Andrews)
 Dr. Ekaterina Stepanova (Russian Academy of Sciences, Moscow)
 Prof. Dr. Helmut Thome (Martin-Luther-University of Halle-Wittenberg)
 Prof. Jorge Vala (Universidade de Lisboa)
 Prof. Dr. Ulrich Wagner (Philipps-University of Marburg)
 Prof. Dr. Andreas Zick (University of Bielefeld)

Editorial Staff (University of Bielefeld)

Dipl.-Soz. Julia Marth
 Dipl.-Psych. Friederike Sadowski
 Dr. Kurt Salentin
 Dr. Peter Sitzer
 Dipl.-Pol. Boris Wilke

Managing Editor

Julia Marth
 University of Bielefeld
 Institute for Interdisciplinary Research on Conflict and Violence
 P.O. Box 100131
 33501 Bielefeld
 Germany
 editorial.office@ijcv.org
 www.ijcv.org

Copy-Editing

Meredith Dale, Berlin
 medale@t-online.de

Typesetting

Florian Rudt, Bielefeld
 frudt@rudtpublishing.de

Design by

meier stracke gbr, büro für gestaltung
 Ernst-Rein-Str. 40
 33613 Bielefeld
 Germany
 post@meier-stracke.de
 www.meier-stracke.de

Editorial

Letter from the Editors

Dear Reader,

This issue of the journal focuses on the question of bullying prevention, with a collection of articles put together by Manuel Eisner and Tina Malti. We are very grateful to them for the hard work they put in as focus section editors – and in their contributions to the section. The open section this time takes us to North America for a study of identity and in-group superiority and Africa for a review of the question of youth and violence.

The next issue, to appear in spring 2013, will feature a double focus for the first time, presenting collections on transitional justice and on qualitative research on prejudices.

December 2012

Wilhelm Heitmeyer Douglas S. Massey Steven F. Messner James Sidanius Michel Wieviorka

The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section

Manuel Eisner, Institute of Criminology, University of Cambridge, United Kingdom
Tina Malti, Department of Psychology, University of Toronto, Canada

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

▶ **Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section** Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section

Manuel Eisner, Institute of Criminology, University of Cambridge, United Kingdom
Tina Malti, Department of Psychology, University of Toronto, Canada

Across Europe, there is an increasing demand for good evidence that can inform policies aimed at reducing violence against and among children and adolescents. However, there is still a paucity of high-quality research on effective prevention of bullying and violence, and researchers from different parts of Europe rarely discuss their findings. The focus section of this issue of the *International Journal of Conflict and Violence* brings together work by prominent prevention scholars from across Europe, who show that significant progress is being made. The introduction presents nine recommendations about how prevention research could be further strengthened in Europe.

Across Europe, there is an increasing demand for good evidence that can inform policies aimed at reducing violence against and among children and adolescents. However, there are wide differences between countries in the extent to which research supports prevention policy: In some countries evidence-based principles have become an important basis for policy implementation. In others, the underlying principles of evidence-based prevention are hardly known among policy-makers.

Overall, significant progress has been made: Across northern Europe, in particular, the past ten years have seen policy-makers increasingly interested in evidence-based prevention and intervention. In the United Kingdom, for example, the recent *Allen Report on Early Intervention* (Allen 2011) – which makes a strong case for evidence-based early prevention of child maladjustments – demonstrates broad support for research-based strategies to promote children's development. Also, centres such as the *Centre for Evidence-Based Intervention* (Oxford), the *Centre*

for Evidence-Based Early Intervention (Bangor), the *National Evaluation of Sure Start* (Birkbeck College), and the *Centre of Experimental Criminology* (Cambridge) are home to internationally recognized prevention research conducted in the United Kingdom. Major foundations such as the *Dartington Foundation* in the United Kingdom, *Atlantic Philanthropies* in Ireland, and the *Jacobs Foundation* in Switzerland have also committed significant resources to supporting research on evidence-based prevention. Scandinavian countries, as so often, lead the way. In Sweden, for example, the government has identified the dissemination of evidence-based research knowledge into mainstream services as a major challenge, and the Swedish government now considers evidence-based practice as an essential vehicle for improving the quality of care and services. Finally, there are encouraging signs of increased European co-operation: the *European Crime Prevention Network*, founded in 2001, is committed to identifying and disseminating good practice in crime prevention. Since 2006, the *Stockholm Symposium of Criminol-*

Acknowledgement: We would like to thank Michael Little, Maria Ttofi, and Simon Sommer for helpful comments on previous drafts of this paper. We would also like to thank all participants of the con-

ference on Evidence-Based Prevention of Bullying and Youth Violence: European Innovations and Experiences held in 2011 at the University of Cambridge for their contributions to the initial ideas be-

hind this paper. We particularly thank the Jacobs Foundation and the European Science Foundation for the generous support of the conference.

ogy has brought together policy-makers, practitioners, and researchers with the goal of finding better ways of reducing violence and crime. And in 2009, almost twenty years after its American sister organisation, the *European Society of Prevention Research* was founded.

Despite undeniable progress and increasing interest amongst governments in understanding how violence prevention can be made more effective, daunting challenges persist. To address some of these the *Institute of Criminology at the University of Cambridge* organized a conference on Evidence-Based Prevention of Bullying and Youth Violence: European Innovations and Experiences on 5 and 6 July 2011. Supported by the *European Science Foundation* and the *Jacobs Foundation*, its purpose was to bring together researchers, policy-makers, and practitioners to discuss innovative research. The conference also sought to identify areas where progress is essential to provide policy-makers with better knowledge about how to support positive child development and reduce the substantial harm resulting from violence and aggression.

1. What is the Issue?

The perpetration of bullying and aggression by young people is a widespread problem in Europe. According to the 2005/6 *Health Behaviour of School-Aged Children* survey, which covers almost all countries in Europe, an average of 42 percent of eleven-year olds and 35 percent of fifteen-year olds reported having been involved in a physical fight at least once during the previous twelve months (Currie et al. 2008). Aggressive behaviour can have serious and long-term negative effects on young people's health and emotional well-being. For example, children and adolescents actively involved in bullying and violence are at a significantly greater risk of later problem behaviours such as substance abuse, academic failure, unemployment, and criminal convictions (Fergusson, Horwood, and Ridder 2005; Loeber and Hay 1997).

Violence is also an important source of suffering amongst victims. According to the same *Health Behaviour of School-Aged Children* survey, 37 percent of eleven-year olds and 27 percent of fifteen-year olds reported having

been the victim of bullying at least once during the previous couple of months. Experiences of violent victimisation have been found to be associated with a range of negative effects including social withdrawal, academic difficulties, substance use, and future anxiety and depressive symptoms (Averdijk et al. 2009; Ttofi et al. 2011).

Over the past ten years, new forms of coercive and threatening behaviour have emerged while others may have declined. For example, cyber-bullying (threatening or hurtful behaviour towards the victim via electronic media) has become a serious problem in line with increasing use of social media and mobile telephones (Perren et al. 2012; Slonje and Smith 2008). Also, sexually coercive behaviours among adolescents are emerging as a pressing issue (Averdijk, Mueller-Johnson, and Eisner 2011).

2. General Principles of Effective Prevention

Due to the high numbers of children and adolescents involved in violence, the significant negative consequences for victims and perpetrators, and the emergence of new manifestations of bullying and violence, prevention of violence should be high on the agenda of public health policies. But what is needed to make the prevention of bullying and youth violence more effective?

Evidence-based prevention needs to be based on the correct identification of the causal risk factors and mechanisms that lead to violence and aggressive behaviour, as well as knowledge about the mechanisms that impede the manifestation of problem behaviours even where risk factors are present (i.e., protective factors). Prevention is likely to be effective if it reduces risk factors and/or builds up protective factors (Coie et al. 1993). Recent research, in particular, has shifted away from the more traditional concern with risk factors to paying more attention to protective factors, and how a better understanding of protective factors can help to build resilience and inform prevention policy [pic](Lösel and Farrington 2012; Pardini et al. 2012; Rutter 2012). Table 1 gives examples for risk and protective factors at the level of the individual, family, school, and neighbourhood/society at large.

Table 1: Examples of risk and protective factors underlying bullying and violence

	Risk factor	Protective factor
Individual	perinatal complications impulsivity restlessness and irritability low empathy social-cognitive biases low academic achievement antisocial beliefs alcohol and other drug use	positive mood low irritability emotion regulation skills self-efficacy high academic achievement social competencies
Parents and family	child abuse and neglect poor parental monitoring erratic parenting partner conflict and separation parental and sibling antisocial behaviour	parental support secure attachment and bonding intensive supervision parental disapproval of antisocial behaviour
School and peers	truancy poor teacher-child bond high school disorder association with delinquent peers negative school climate	positive teacher-child bonds academic motivation and success high school-level discipline and clear rules non-deviant best friends involvement in structured prosocial activities
Neighbourhood and society	social inequality and deprivation	high social cohesion and trust community involvement and access to social support

See Lösel and Farrington (2012) for a more extensive discussion.

There is now widespread agreement amongst prevention specialists about the general principles that underlie effective prevention of aggression, bullying, and violence across the life-course. These principles include (Allen 2011; Eisner, Ribeaud and Locher, 2009; Krug et al. 2002; World Health Organization 2010):

1. The need to start prevention during the first years of life by reducing risk factors and promoting protective factors during a time when humans have a high degree of plasticity (“start early in life”).
2. The need to have developmentally adequate prevention strategies in place across the whole life course from conception to adulthood (“developmentally adequate provision across the life course”).
3. The principle of embedding violence prevention into a general public health strategy that aims at reducing a range of negative outcomes including school dropout, teen pregnancy, substance abuse, delinquency and violence, unhealthy eating, and physical inactivity. These behaviours share many risk factors and should hence be considered as elements of a larger prevention strategy (“a public health perspective”).
4. The combining of universal, indicated, and selective prevention so that the largest resources reach the children and adolescents with the greatest needs (“adapt intervention intensity to risk exposure”).
5. The consideration of a socio-ecological model that recognizes the interplay of influences at the levels of the individual, the family, the school, peers and leisure-time activities, the neighbourhood, and the wider social, cultural and political context (“an ecological perspective of multi-layered prevention”).
6. An approach that integrates policy-making and research by using high-quality basic research to guide innovation in prevention programmes and strategies, by rigorously testing prevention strategies in methodologically sound outcome evaluations, and by working with governments and policy-makers to achieve real-world effects (“an evidence-based approach to policy change”).

We believe that governments could achieve noticeable population-wide reductions in bullying and aggressive behaviour by adopting an evidence-based prevention and intervention policy (Cartwright and Hardie 2012). This requires close co-operation between local and national governments and prevention researchers. Currently many European countries do not have the requisite research capacity or the evidence base to provide effective support in their societies. In the following postulates, we propose nine domains where research is needed to contribute to more effective violence prevention.

3. Nine Recommendations for Future Priorities

3.1. Expanding the Evidence Base

A move towards more effective prevention of aggression and violence requires efforts to expand the scientific evidence on what works (Sherman et al. 2002). The creation of a better evidence-base entails a staged process that includes small-scale efficacy trials of innovations or adaptations, effectiveness trials of the most promising approaches, and large-scale field trials of programmes that are planned to be taken to scale. Despite progress over the past twenty years the current knowledge base is generally still thin in Europe (Lösel and Beelmann 2003). Also, significant differences remain between European countries in the amount of research done.

More and better evaluation research is needed to create the knowledge base required for achieving a major population-level reduction in youth violence. This demands more coherent European financial and organisational support for high-quality evaluations and the encouragement of collaboration between academic institutions and practitioners. Also, systematic reviews for different types of preventive interventions suggest that more knowledge has been accumulated in respect of short-term effects and effects found in relatively small efficacy trials (Lösel and Beelmann 2003; Tofsi and Farrington 2011). In contrast, there are several areas where the lack of studies is particularly acute. These include field trials examining whether violence prevention programmes work under real-life conditions and studies examining long-term effects over months or even years. For this reason the present volume includes several studies that contribute to closing this

gap. In particular, the studies by Lösel und Stemmler (2012) on long-term outcomes of an early intervention, the overview by Hutchings on the implementation and evaluation of Incredible Years in Wales, the study by Goossens, Gooren, Orobio de Castro, Van Overveld, Buijs, Monshouwer, Onrust, and Paulussen (2012) on a routine implementation of PATHS in the Netherlands, the article by Little, Berry, Morpeth, Blower, Axford, Taylor, Bywater, Lehtonen, and Tobin (2012) on the large scale evaluation of PATHS, Triple-P, and Incredible Years in Birmingham, and the paper by Salmivalli and Poskiparta (2012) on the national evaluation of the KiVa bullying prevention programme in Finland represent remarkable progress in knowledge about what is required to make interventions work under real-world conditions.

3.2. Promoting Innovation in Programme Development

Progress in effective prevention depends on the development of interventions that reflect advances in research. Over the past two decades many impulses for evidence-based prevention strategies – such as parent training programmes, early support for at-risk mothers, and school-based social skills programmes – have come to Europe from elsewhere. As a result, many evaluations have examined whether existing products can be transferred into the European context (e.g. Hutchings 2012). In contrast, few innovations in research-based prevention have been initiated in Europe (but see Kärnä et al. 2011; Lösel and Stemmler 2012).

Testing the transportability of interventions will remain important in the future. The paper by Hutchings (2012) provides insight on the critical issues that need to be considered for the successful introduction of a programme in a new context. However, there is also potential for developing new approaches that have a better fit to the structure of social services, education systems, and cultural expectations in European societies. In the present volume, articles by Loesel and Stemmler (2012), Salmivalli and Poskiparta (2012), Ortega-Ruiz, Del Rey, and Casas (2012), and Menešini, Nocentini, and Palladino (2012) present evaluations of innovative programmes developed in Europe. Future funding should support the further development of innovative interventions for individuals, schools, families, and neigh-

bourhoods. These interventions should be tailored to meet the needs of different systems of services, specific target groups, and diverse groups of children with diverse manifestations of aggression and violence (Perren et al. 2012).

3.3. A Better Link Between Basic and Applied Research

Preventive interventions are more likely to be effective if they are based on empirically validated models of the causation of violence. There is therefore an important link between basic research on the causes of youth violence and the development of more effective interventions (see Stokes 1997). Too many preventive programmes in Europe are still implemented with little basis in developmental research. This increases the risk that significant resources will be invested in ineffective programmes.

We believe that improved collaboration between basic research and applied prevention research will produce a better knowledge base for effective youth violence prevention. Examples where this potential is particularly clear include the preventive implications of the link between developmental neuroscience and aggression (Bradshaw et al. 2012; Séguin et al. 2004), the implications of research on social networks for group-based prevention (Salmivalli, Huttunen, and Lagerspetz 1997), the lessons for violence prevention to be learned from research on moral development (Malti and Krettenauer 2012), or the ways in which research on judgement and decision-making can inform prevention strategies (Nagin 2007; Wikström et al. 2012). In the present volume, the contribution by Perren, Corcoran, Cowie, Dehue, Garcia, Mc Guckin, Sevcikova, Tsatsou, and Völlink (2012) shows how high-quality basic research on the responses of parents, teachers, and victims to cyberbullying can inform the development of better intervention and prevention strategies.

3.4. Evaluation of Embedded Practices and System Change

Much prevention research has examined the effects of standardized programmes that are added to an existing system. However, social services and education systems comprise many activities with a preventative purpose (Little 2010). For example, if a pupil shows disruptive behaviour in a classroom, teachers, head-teachers, and social workers may intervene in various ways. However, we lack knowledge

about the effectiveness of these interventions, and how they can be improved. Also, many evaluations test commercially distributed products. Yet local and national authorities often deliver services that are similar in purpose and structure (e.g. support for young mothers, parenting advice, anti-bullying programmes, social competencies in school curricula). Little is currently known about the effectiveness of practices embedded in mainstream services. But some findings suggest that interventions delivered as part of mainstream services may sometimes be as effective as new products (de Graaf et al. 2008). Finally, most policy changes in education, social welfare, family affairs, and policing and youth justice are implemented without any consideration of their effectiveness, and very few studies have attempted to assess whether new policies achieve their goals.

A better understanding of how whole systems can be made more effective could have considerable benefits for youth violence reduction (Little 2010). However, good research on this question requires that prevention science partly moves beyond classical randomized controlled experiments and broadens its methodological scope. Also, we believe that substantial progress could be made by building evaluation components into the process of policy change (Cartwright and Hardie 2012). For example, the paper by Spiel, Wagner, and Strohmeier (2012) in this volume presents a research-led violence prevention strategy for Austria that incorporated evaluation components during the roll-out phase.

3.5. Integrate Situational and Developmental Approaches to Violence Prevention

Researchers often distinguish between *developmental* approaches that try to influence the propensity to engage in violent acts over the life-course (i.e. change the person and his or her social, emotional, cognitive, and moral development; see Tremblay and Craig 1995) and *situational* approaches that try to influence the likelihood of a violent act happening. Situational approaches include CCTV cameras in public space, targeted police patrols in crime and violence hot-spots, firearm controls, school-surveillance in corridors, strengthening peer interventions against bullying, surveillance mechanisms on the internet, and alcohol sales policies (Clarke 1995). For historic reasons situational and

developmental approaches to violence prevention have been seen as opposites rather than as complementary strategies.

We believe that the most promising approach to violence prevention combines developmental and situational interventions. However, evaluation research that addresses both components has been rare, both in Europe and internationally. Strategic support for innovative research that combines situational and developmental components is likely to yield highly interesting findings with a direct impact on policy making across areas such as policing, urban planning, social and family policies and education.

3.6. Developing and Testing Tailored Prevention Strategies

Many risk and protective factors are similar for different types of aggression and violence. Also, most risk factors are relevant in different cultures and societies rather than being specific to any particular society. This suggests that an effective prevention strategy should be based on similar principles across all of Europe and that it should target a broad range of problem behaviours rather than being highly specific.

However, there is controversy about the extent to which delivery format, recruitment, and framing need cultural adaptation. For example, some evidence suggests that regular parent training programmes may be less effective for single parents than for two-parent families (Gardner et al. 2009). Also, children and adolescents differ in the extent to which they are exposed to specific risk factors, and different combinations of environmental and individual risks may require different approaches. For example, the approach required for socially isolated adolescents with concurrent attention deficits and academic difficulties may differ from the approach required for more dominant, sociable, and academically successful bullies. Future research should therefore examine how prevention programmes can be tailored to the specific needs of different risk groups or different types of aggression (Malti and Noam 2009). In the present volume, the article by Noam, Malti, and Guhn (2012) proposes a new measurement tool for assessing levels of resilience amongst children, which could facilitate the implementation of targeted intervention strategies.

3.7. Improving Quality Standards in Prevention Evaluation Research

Reviews suggest much variation in the methodological quality of outcome evaluations. While some studies meet high methodological standards, the methodological limitations of many make it difficult to draw firm conclusions about genuine treatment effects (Eisner 2009). Such limitations include poor overall study design, low validity of core outcome measures, limited or no measures of the implementation process, and insufficient reporting of study characteristics and analytic approaches.

There is significant scope for improving the quality standards of outcome evaluations conducted in Europe. Better-quality studies would provide more valid and generalizable information for policy-makers and practitioners on what works and what does not. For example, the study by Forster, Kling, and Sundell (2012, in this volume) shows the importance of developing uniform standards for assessing the clinical relevance of treatment effects when different studies are compared. Other measures for improving methodological standards include compulsory registration of all outcome evaluations, guidelines on the design and reporting of outcome studies, training in evaluation design, and greater transparency concerning potential conflicts of interest. Where there is likely to be a conflict of interest between the role of evaluator and of programme provider funding agencies should request an independent review of the study design and the data analyses.

Progress in evidence-based prevention is often hampered by obstacles to co-operation between researchers, intervention providers, and local stakeholders. Introducing evidence-led development and design into education, public health policy, social services, or family services requires that policy-makers and practitioners have a good understanding of the principles of evaluation research.

3.8. Improving Knowledge of Mechanisms and Active Components

Despite some success in identifying effective programmes, we still have a very limited understanding of the *causal mechanisms* that make them work. Also, we know little about the active components that render a preventive intervention effective. A better understanding of the active components of preventive interventions is essential for further

progress. Only if we understand the principles of *why* some interventions work can we make progress in designing the next generation of prevention approaches.

Progress on these issues has been difficult. The most frequent approach is to conduct analyses of mediators (mechanisms transporting the causal effect from the intervention to the outcome) and moderators (factors that are associated with variation in the achieved effect). For example, in the present volume Malti, Ribeaud, and Eisner (2012) examine whether a school-based intervention was more or less effective for children from different socio-economic backgrounds. At the level of meta-analyses, Hahn Fox, Ttofi, and Farrington (2012, in this volume) present important results on the factors that influence the effectiveness of anti-bullying programmes. It shows, for example, that bullying prevention programmes tend to be more effective if they are more intensive and if they include a parent training component (Hahn Fox, Ttofi, and Farrington 2012). However, we believe that further progress requires a new and innovative type of evaluation research. Rather than randomly allocating participants to whole packages of interventions (“programmes”) researchers will need to improve their capacity to isolate, on the basis of prior findings and theoretical considerations, promising elements of an intervention whose effects can then be examined. To the extent that innovative research could identify the active building blocks of prevention activities it could help to progressively tailor more effective interventions.

3.9. Upscaling and Mainstreaming

While a lot has been learned about how prevention approaches can be made to work in efficacy trials, much less is known about how programmes can be taken to scale without losing their effectiveness. Several studies in this

volume suggest that certain evidence-based programmes fail to produce desirable effects when examined in large field trials (Goossens et al. 2012; Little et al. 2012). We therefore believe that more well-designed, large-scale field trials that assess long term-effects are necessary (Farrington and Welsh 2007). Such trials can provide policy makers with realistic estimates of effects that are replicable at the level of whole populations. Often, such evaluations should be conducted as independent evaluations, in which the role of the evaluators and programme developers are institutionally separated. Large-scale dissemination trials are costly and it is essential that they are carefully planned and adequately resourced, and that their findings are effectively communicated amongst researchers and policy-makers in Europe. Also, more *translational research* on programmes and policies that can effectively be inserted into mainstream services is necessary (Woolf 2008).

4. Conclusion

In the past, the development and implementation of more effective violence prevention supported by research evidence has often been hampered by a lack of regular research collaboration across Europe.

The contributions in the present volume represent an attempt to bridge this gap and to encourage exchange amongst researchers from different academic backgrounds across Europe. Taken together, they show that violence prevention in Europe has become a dynamic field of research where knowledge is increasingly consolidated. In particular, there is growing evidence that high-quality prevention research may help to achieve substantial population-wide reductions in youth violence over the coming decade.

References

- Allen, Graham. 2011. *Early Intervention: The Next Steps (An Independent Report to Her Majesty's Government)*. London: HM Government.
- Averdijk, Margit, Katrin Mueller-Johnson, and Manuel Eisner. 2011. *Sexual Victimization of Children and Adolescents in Switzerland*. Zurich: Optimus Foundation.
- Averdijk, Margit, Barbara Müller, Manuel Eisner, and Denis Ribeaud. 2009. Bullying Victimization and Later Anxiety and Depression among Pre-Adolescents in Switzerland. *Journal of Aggression, Conflict and Peace Research* 3 (2): 103–109.
- Beelmann, Andreas. 2011. The Scientific Foundation of Prevention. The Status Quo and Future Challenges for Developmental Crime Prevention. In *Anti-social Behavior and Crime: Developmental and Evaluation Research to Prevention and Intervention*, ed. Thomas Bliesener, Andreas Beelmann, and Mark Stemmler, 137–64. Cambridge, MA: Hogrefe.
- Bradshaw, Catherine P., Asha Goldweber, Diana Fishbein, and Mark T. Greenberg. 2012. Infusing Developmental Neuroscience into School-Based Preventive Interventions: Implications and Future Directions. *Journal of Adolescent Health* 51(2, supplement): S41–S47. doi:10.1016/j.jadohealth.2012.04.020
- Cartwright, Nancy, and Jeremy Hardie. 2012. *Evidence-based Policy*. Oxford: Oxford University Press.
- Clarke, Ronald V. 1995. Situational Crime Prevention. *Crime and Justice* 19:91–150.
- Coie, John D., Norman F. Watt, Stephen G. West, J. David Hawkins, Joan R. Asarnow, Howard J. Markman, Sharon L. Ramey, Myrna B. Shure, and Beverly Long. 1993. The Science of Prevention. A Conceptual Framework and Some Directions for a National Research Program. *American Psychologist* 48 (10): 1013–22.
- Currie, Candace, Saoirse Nic Gabhainn, Emmanuelle Godeau, Chris Roberts, Rebecca Smith, Dorothy Currie, Will Pickett, Matthias Richter, Anthony Morgan, and Vivian Bernekow, eds. 2008. *Inequalities in Young People's Health: HBSC International Report from the 2005/06 Survey*. Copenhagen: WHO Regional Office for Europe.
- de Graaf, Ireen, Paula Speetjens, Filip Smit, Marianne de Wolff, and Louis Tavecchio. 2008. Effectiveness of the Triple P Positive Parenting Program on Behavioral Problems in Children: A Meta-Analysis. *Behavior Modification* 32 (5): 714–35.
- Eisner, Manuel, Denis Ribeaud, and Rachel Locher. 2009. *Prävention von Jugendgewalt* (report for the Swiss Ministry of Social Affairs). Berne: Bundesamt für Sozialversicherung.
- Eisner, Manuel. 2009. No Effects in Independent Prevention Trials: Can We Reject the Cynical View? *Journal of Experimental Criminology* 5 (2): 163–83.
- Farrington, David P., and Brandon C. Welsh. 2007. *Saving Children from a Life of Crime: Early Risk Factors and Effective Interventions*. Oxford: Oxford University Press.
- Fergusson, David M., L. John Horwood, and Elizabeth L. Ridder. 2005. Show Me the Child at Seven: The Consequences of Conduct Problems in Childhood for Psychosocial Functioning in Adulthood. *Journal of Child Psychology and Psychiatry* 46 (8): 837–49.
- Forster, Martin, Åsa Kling, and Knut Sundell. 2012. Clinical Significance of Parent Training for Children with Conduct Problems. *International Journal of Conflict and Violence* 6 (2): 187–200.
- Gardner, Frances, Arin Connell, Christopher J. Trentacosta, Daniel S. Shaw, Thomas J. Dishion, and Melvin N. Wilson. 2009. Moderators of Outcome in a Brief Family-Centered Intervention for Preventing Early Problem Behavior. *Journal of Consulting and Clinical Psychology* 77 (3): 543–53.
- Goossens, Ferry X., Evelien M. J. C. Gooren, Bram Orobio de Castro, Kees W. Van Overveld, Goof J. Buijs, Karin Monshouwer, Simone A. Onrust, and Theo G. W. M. Paulussen. 2012. Implementation of PATHS Through Dutch Municipal Health Services: A Quasi Experiment. *International Journal of Conflict and Violence* 6 (2): 234–248.
- Hahn Fox, Brianna, Maria Ttofi, and David P. Farrington. 2012. Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components. *International Journal of Conflict and Violence* 6 (2): 273–282.
- Hutchings, Judy. 2012. Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales. *International Journal of Conflict and Violence* 6 (2): 225–233.
- Kärnä, Anti, Marinus Voeten, Todd D. Little, Elisa Poskiparta, Anne Kaljonen, and Christina Salmivalli. 2011. A Large-Scale Evaluation of the KiVa Antibullying Program: Grades 4–6. *Child Development* 82 (1): 311–30. doi:10.1111/j.1467-8624.2010.01557.x
- Krug, Etienne G., Linda L. Dahlberg, James A. Mercy, Anthony B. Zwi, and Rafael Lozano. 2002. *World Report on Violence and Health*. Geneva: World Health Organisation.
- Little, Michael. 2010 *Improving Children's Outcomes Depends on Systemising Evidence-based Practice*. London: Demos.
- Little, Michael, Vashti Berry, Louise Morpeth, Sarah Blower, Nick Axford, Rod Taylor, Tracey Bywater, Minna Lehtonen, and Kate Tobin. 2012. The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK. *International Journal of Conflict and Violence* 6 (2): 260–272.
- Loeber, Rolf, and Dale Hay. 1997. Key Issues in the Development of Aggression and Violence from Childhood to Early Adulthood. *Annual Review of Psychology* 48:371–410.
- Lösel, Friedrich, and Andreas Beelmann. 2003. 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 (1):84–109.
- Lösel, Friedrich, and David P. Farrington. 2012. Direct Protective and Buffering Protective Factors in the Development of Youth Violence. *American Journal of Preventive Medicine* 43 (2): 8–23.
- Lösel, Friedrich, and Mark Stemmler. 2012. Preventing Child Behavior Problems in the Erlangen-Nuremberg Development and Prevention Study: Results from Preschool to Secondary School Age. *International Journal of Conflict and Violence* 6(2): 214–224
- Malti, Tina, and Tobias Krettenauer. 2012. The Relation of Moral Emotion Attributions to Prosocial and Antisocial Behavior: A Meta-Analysis. *Child Development*, early online publication, September 24, 2012. doi:10.1111/j.1467-8624.2012.01851.x
- Malti, Tina, and Gil G. Noam. 2009. A Developmental Approach to the Prevention of Adolescents' Aggressive Behavior and the Promotion of Resilience. *European Journal of Developmental Science* 3:235–46.
- Malti, Tina, Denis Ribeaud, and Manuel P. Eisner. 2012. Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors. *International Journal of Conflict and Violence* 6 (2): 249–259.
- Menesini, Ersilia., Annalaura Nocentini, and Benedetta Emanuela Palladino. 2012. Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model. *International Journal of Conflict and Violence* 6 (2): 313–320.
- Nagin, Daniel S. 2007. Moving Choice to Center Stage in Criminological Research and Theory: The American Society of Criminology 2006 Sutherland Address. *Criminology* 45 (2): 259–72.
- Noam, Gil, Tina Malti, and Martin Guhn. 2012. From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool. *International Journal of Conflict and Violence* 6 (2): 201–213.
- Ortega-Ruiz, Rosario, Rosario Del Rey, and José A. Casas. 2012. Knowing, Building, and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program. *International Journal of Conflict and Violence* 6 (2): 302–312.
- Pardini, Dustin A., Rolf Loeber, David P. Farrington, and Magda Stouthamer-Loeber. 2012. Identifying Direct Protective Factors for Nonviolence. *American Journal of Preventive Medicine* 43 (2, supplement 1): S28–S40.
- Perren, Sonja, Lucie Corcoran, Helen Cowie, Francine Dehue, D'Jamila Garcia, Conor Mc Guckin, Anna Sevcikova, Panayioti Tsatsou, and Trijntje Völlink. 2012. Tackling Cyberbullying: Review of Empirical Evidence Regarding Suc-

- successful Responses by Students, Parents, and Schools. *International Journal of Conflict and Violence* 6 (2): 283–292.
- Rutter, Michael. 2012. Annual Research Review: Resilience: Clinical Implications. *Journal of Child Psychology and Psychiatry* (online first).doi:10.1111/j.1469-7610.2012.02615.x
- Salmivalli, Christina, Arja Huttunen, and Kirsti M. J. Lagerspetz. 1997. Peer Networks and Bullying in Schools. *Scandinavian Journal of Psychology* 38 (4): 305–312.
- Salmivalli, Christina, and Elisa Poskipart. 2012. KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland. *International Journal of Conflict and Violence* 6 (2): 293–301.
- Séguin, Jean R., Daniel Nagin, Jean-Marc Assaad, and Richard E. Tremblay. 2004. Cognitive-Neuropsychological Function in Chronic Physical Aggression and Hyperactivity. *Journal of Abnormal Psychology* 113 (4): 603–13.
- Sherman, Lawrence W., David P. Farrington, Brandon C. Welsh, and Doris L. MacKenzie, eds. 2002. *Evidence-Based Crime Prevention*. London: Routledge.
- Slonje, Robert, and Peter K. Smith. 2008. Cyberbullying: Another Main Type of Bullying? *Scandinavian Journal of Psychology* 49 (2): 147–54.
- Spiel, Christiane, Petra Wagner, and Dagmar Strohmeier. 2012. Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy. *International Journal of Conflict and Violence* 6 (2): 176–186.
- Stokes, Donald E. 1997. Pasteur's Quadrant – Basic Science and Technological Innovation. Washington: The Brookings Institution.
- Tremblay, Richard E., and Wendy M. Craig. 1995. Developmental Crime Prevention. In *Building a Safer Society: Strategic Approaches to Crime Prevention, Crime and Justice: An Annual Review of Research* 19, ed. Michael Tonry and David P. Farrington. Chicago: University of Chicago Press.
- Ttofi, Maria, and David P. Farrington. 2011. Effectiveness of School-Based Programs to Reduce Bullying: A Systematic and Meta-Analytic Review. *Journal of Experimental Criminology* 7 (1): 27–56.
- Ttofi, Maria, David P. Farrington, Friedrich Lösel, and Rolf Loeber. 2011. Do the Victims of School Bullies Tend to Become Depressed Later in Life? A Systematic Review and Meta-Analysis of Longitudinal Studies. *Journal of Aggression, Conflict and Peace Research* 3 (2): 63–73.
- Wikström, Per-Olof, Dietrich Oberwittler, Kyle Treiber, and Beth Hardie. 2012. *Breaking Rules: The Social and Situational Dynamics of Young People's Urban Crime*. Oxford: Oxford University Press.
- Wolf, Steven H. 2008. The Meaning of Translational Research and Why It Matters. *JAMA: The Journal of the American Medical Association* 299 (2): 211–13. doi:10.1001/jama.2007.26
- World Health Organization. 2010. *Violence Prevention: The Evidence*. Geneva: World Health Organisation.

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy

Christiane Spiel, Faculty of Psychology, University of Vienna, Austria

Petra Wagner, Faculty of Applied Health and Social Sciences, University of Applied Sciences of Upper Austria

Dagmar Strohmeier, Faculty of Applied Health and Social Sciences, University of Applied Sciences of Upper Austria

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy

Christiane Spiel, Faculty of Psychology, University of Vienna, Austria

Petra Wagner, Faculty of Applied Health and Social Sciences, University of Applied Sciences of Upper Austria

Dagmar Strohmeier, Faculty of Applied Health and Social Sciences, University of Applied Sciences of Upper Austria

A qualitative study of Austria's national strategy against violence in the public school system introduced in 2008. The national strategy developed by researchers consists of six activity domains with specific goals and projects defined for each. The evaluation (1) analyzes how the realized projects contributed to the six activity domains, (2) evaluates the national strategy at a general level, and (3) provides future recommendations. Eight members of the steering committee were interviewed at two points in the implementation process. The systematic interviews were coded according to the goals of the activity domains. According to the interviewees most of the projects have been satisfactorily implemented. Networking and cooperation with the different actors in the field of violence prevention and cooperation among steering committee members have been improved. However, the national strategy has not achieved the intended public recognition. The lessons learned from the evaluation and its results are discussed.

As a consequence of the public recognition that violence is a severe problem in schools all over the world (Currie et al. 2012) many prevention and intervention programs have been developed and evaluated in numerous efficacy and effectiveness trials (e.g., Ferguson et al. 2007; Ttofi and Farrington 2009). However, the development of national or regional strategies supported by governments is rare (examples of exceptions are Cross et al. 2011; Roland 2011; Salmivalli, Kärnä, and Poskiparta 2011), although research indicates that such strategies might be a key factor for successful and sustainable violence prevention in schools (Ogden, Kärki, and Teigen 2010; Olweus 2004; Roland 2000; see also Spiel, Salmivalli, and Smith 2011). Austria is one case where a national strategy has been systematically developed and implemented. This paper describes its implementation in Austria and an evaluation of the implementation efforts at a general level.

1. The Austrian National Strategy "Together Against Violence"

The Austrian national strategy for violence prevention in the public school system differs in several aspects from strategies in other countries (for example PREVNet in Canada: Pepler and Craig 2011; the KiVa program in Finland: Salmivalli et al. 2011; the safe schools framework in Australia: Cross et al.): (1) it was introduced subsequent to national or regional strategies in other countries and was therefore able to benefit from experiences in other countries; (2) it seeks to integrate pre-existing activities and to bring the relevant stakeholders together; (3) it activates a variety of projects designed to ensure sustainability (e.g., violence prevention and social competence promotion are defined as obligatory components of basic teacher education).

At the beginning of 2007, in the wake of a quick succession of significant events in Austrian schools Spiel and Strohmeier were commissioned by the Federal Ministry of Education, the

Acknowledgement: The development of the national strategy for violence prevention and the evaluation of its implementation were financially supported by the Austrian Federal Ministry of Education, the Arts, and Culture.

Arts, and Culture to develop a national violence prevention strategy. In the process of developing the national strategy “Together Against Violence” (“Gemeinsam gegen Gewalt”; for details about the development of the national strategy and its aims see <http://www.gemeinsam-gegen-gewalt.at>; Spiel and Strohmeier 2007; see also Spiel and Strohmeier 2011, 2012), there was an intensive exchange with international colleagues who have been involved in similar national strategies in their own countries (Canada: Pepler and Craig 2011; Norway: Roland 2011; Australia: Cross et al. 2011). Furthermore, as suggested in the prevention literature (Datnow 2002, 2005; Shokoff and Bales 2011; Spoth and Greenberg 2011), the perspectives of different stakeholder groups already involved in violence prevention in Austria (school psychologists, social workers, teaching unions) were systematically integrated in the strategy development (Spiel and Strohmeier 2007).

Spiel and Strohmeier (2007) defined three goals in the national strategy for students, teachers, and parents, as well as for society as a whole (inspired by Christina Salmivalli’s KiVa game): (1) Increased awareness and knowledge about violence: *I know, we know*; (2) Increased social competence skills and strategies to cope with violence: *I can, we can*; (3) Increased responsibility and civil engagement: *I do, we do*. The strategy consisted of six activity domains: (1) policy and advocacy, (2) prevention and intervention, (3) knowledge transfer and education, (4) information and public relations, (5) networking and cooperation, and (6) evaluation and research. The application of theoretically based and evaluated prevention programs was specifically considered. For each activity domain specific goals and projects were defined and the agents responsible for realization were specified (for details see Spiel and Strohmeier 2007).

In December 2007, the Federal Minister of Education, the Arts, and Culture decided to implement the national strategy. For strategy management and implementation, a steering committee was established at the Federal Ministry with Christiane Spiel as an external member responsible for research issues. In 2008, the national strategy became part of the coalition agreement between the two governing parties and was planned through to the end of the legislative period in September 2013. Table 1 presents the projects implemented between 2008 and 2010.

Table 1: Projects implemented between 2008 and 2010

Activity domain	Projects
Policy and advocacy	The national strategy is an integral part of national government policy
Prevention and intervention	Increase the number of school psychologists Pilot projects by school social workers Implementation of behavior agreements Implementation of the “Faustlos” program Implementation of the Viennese Social Competence Training (ViSC) program Implementation of peer mediation
Knowledge transfer and education	Violence prevention and social competence promotion as obligatory components of basic teacher training Train-the-trainer course for teachers Train-the-trainer courses for ViSC coaches Information workshops for schools and kindergartens
Information and public relations	Establishment of a national internet platform Organization of events Press conferences Media reports Information material
Networking and cooperation	Establishment of a steering committee Conferences involving stakeholders (partners) Cooperation with national television networks
Evaluation and research	Documentation and evaluation of the implementation of the national strategy Evaluation of the prevention and intervention programs Development of online self-assessment instruments for classes and schools

2. The Importance of Strategy Implementation

An important factor for sustainable violence prevention in schools is the implementation quality of programs or strategies (Durlak and DuPree 2008; Fixsen et al. 2009; Berkel et al. 2011). According to Fixsen and Blasé (2009), implementation can be described as the missing link between research and practice. Shonkoff and Bales (2011) argue that the translation of research into policy and practice should be regarded as an important academic endeavor in its own right. In recent years, several research groups have formulated theoretical models of program implementation (Durlak and DuPree 2008; Fixsen et al. 2009; Berkel et al. 2011). In these models, fidelity and quality of implementation are considered important factors.

To date, most empirical research on implementation in general, and prevention programs in particular, has been conducted in Anglo-American countries (Elias et al. 2003). Datnow (2002, 2005) identifies the importance of understanding district and state contexts for the sustainability of comprehensive school reform models alongside teacher- and school-level factors (e.g., Beets et al. 2008; Cargo et al. 2006). According to Datnow's studies, the adoption, implementation, and sustainability of reform, and school change more generally, are the result of interrelations between and across groups in different contexts at various points in time (Datnow and Stringfield 2000). In other words, efforts to implement reforms are more likely to be effective when educators at various levels (e.g., state, district, reform design team, school) share goals and work together. Spoth and Greenberg (2011) show how practitioner-researcher partnerships and supporting infrastructures can support the local adoption of evidence-based interventions and produce community-level reductions in youth problem behaviors and concomitant positive youth development (see also Crowley et al. 2012). In Europe, Norway is a pioneer both in conducting violence prevention programs in schools and in evaluating their implementation on a national level (Roland 2011).

Development, implementation, and dissemination of strategies on a national or regional level involve intensive cooperation between researchers, politicians, and administrators (Roland 2000; Spiel and Strohmeier 2007, 2011) within a mutually respectful, collaborative process (Shonkoff and Bales 2011).

3. Aims of the Evaluation of Implementation

In 2010, Petra Wagner was commissioned by the Austrian Federal Ministry of Education, the Arts, and Culture to evaluate the implementation of the national strategy. Aims of the evaluation were (1) to analyze how the individual projects contribute to the six activity domains defined in the strategy plan by Spiel and Strohmeier (2007), (2) to analyze the national strategy at a general level, and (3) to provide recommendations for the individual projects and for the national strategy in its entirety.¹ The evaluation focuses on fidelity and quality of implementation (Carroll et al. 2007; Elias et al. 2003; Kalafat, Illback, and Sanders 2007) and on participant responsiveness (Dusenbury et al. 2003; for details see Wagner, 2011).

4. Method of the Evaluation

4.1. Expert Interviews

Expert interviews were conducted (Gläser and Laudel 2009). According to Schirmer (2009) interviewees are defined as experts if they have special knowledge related to the research interest. Expert interviews are based on a list of open questions (interview guideline).

The members of the steering committee were identified as experts for the projects within the activity domains (see Table 1) they were responsible for and for the national strategy at a general level. An interview guideline was developed consisting of the following topics: goals of the projects, schedules and application procedures of the projects, evaluation measures (where individual projects had been evaluated), and contribution to the national strategy. Concerning the national strategy at the general level, inter-

¹ The three aims of the strategy are formulated at a meta-level. Therefore, the evaluation described here focused on how the projects contribute to the activity domains as a necessary prerequisite for achieving the aims of the national strategy.

viewees were asked about internal changes (within the steering committee) and external changes (as consequences of the implementation of the national strategy) observed since the initiation of the strategy particularly concerning the recognition of the national strategy in the public. Additionally, they were asked to give recommendations for the future development of the strategy. The interviewees also assessed the quality of embedding at the political level (domain “Policy and advocacy”) and the quality of networking and cooperation (domain “Networking and cooperation”).

4.2. Sample and Procedure

The sample consisted of eight members of the steering committee; one was the general project manager and seven were responsible for particular projects within the activity domains. Six interviewees were female, two male. Seven interview partners were long-term employees of the Federal Ministry of Education, the Arts, and Culture, one was an external expert.

Sixteen interviews were conducted altogether. Each member of the steering committee was interviewed twice, in September 2009 and in November 2010. After the first interview phase, results relevant to improving the projects and the strategy in general were reported and discussed in the steering committee. The aim of the second interview was to identify any changes that had occurred. The interviews lasted between 60 and 90 minutes. All interviewees cooperated with the evaluation and were motivated to provide useful information.

4.3. Data Analysis

All interviews were transcribed and coded according to the interview guideline (Mayring 2002). This analytical procedure produced thematically classified protocols of each interview in tabular form. To ensure the reliability and validity of the protocols the interviewees were asked to review, revise, and authorize them. All interviewees cooperated in this procedure.

As each steering committee member was responsible for specific projects the evaluation results were based on the judgements of these single responsible members. The

protocols of the interviews served as a data base to evaluate how the projects contributed to the six activity domains.

Data about the national strategy at the general level were available from all respondents. The primary aim of the qualitative data analysis in this area was to elaborate similarities and differences between the interviews (Mayer 2008). Therefore, these parts of the protocols were summarized and correlated (Mayring 2002) and compared using the Delphi method (Linstone and Turoff 1975).

5. Results of the Evaluation

In the following, the most important results concerning the contribution of the projects to the activity domains of the national strategy are presented separately for each domain. Finally, evaluation results concerning the national strategy at a general level are presented. As no specific project is devoted to the activity domain “Policy and advocacy,” evaluation results concerning this domain are presented in the context of the results on the general level. The results are presented from the perspective of the second interview and changes over time are included. Evaluation results of individual projects are not presented here. If they have been published elsewhere references are given.

5.1. Prevention and Intervention

The projects “Increase the number of school psychologists” and “Pilot projects by school social workers” (see Table 1) aim at an Austria-wide support of teachers and students with the main focus on violence prevention. Concretely, both projects are designed to foster the social competence and social responsibility of students, their ability to deal with diversity, and their learning motivation directly (e.g., by advice, treatment, and mentoring) and indirectly (e.g., by advanced training and professional support of teachers). This is expected to improve the school and class climate and reduce aggression and violence in Austrian schools.

At the beginning (2008) there were about 140 school psychologists working in Austria. To achieve nationwide support for schools the Federal Ministry of Education, the Arts, and Culture financed 40 new part-time school psychologist posts. According to the interviewed expert, long-term funding for the new school psychologists is secure.

The evaluation of this project was conducted using internal reports.

The pilot projects by school social workers were prepared from 2008 to 2010 and their implementation was scheduled to run from September 2010 until August 2012 in six Austrian provinces. The projects are co-financed by the European Social Fund (ESF) and are being evaluated by an external research institute.

The behavior agreements project aims to establish Austria-wide support for the school partners (students, teachers, parents) to improve social interactions in schools. The medium-term goal is not only to increase the number of behavior agreements in schools, but also their quality. This project was established some years before the national strategy and later incorporated into it. In 2009, the interviewed expert conducted an Austria-wide survey to record the numbers of behavior agreements in schools and evaluate the current guidelines (<http://www.gemeinsam-gegen-gewalt.at/materialien-links/>). The results showed that these guidelines suffer several limitations. Based on the findings of the survey, the Federal Ministry planned to develop new guidelines for behavior agreements including recommendations on how to design the process to develop such behavioral arrangements.

The Faustlos (“no fists”) and ViSC programs aim to promote social competence and responsibility in students and encourage their participation to enhance the school community, to reduce aggression and violence in school, and to improve the school climate. Both programs are evidence-based and are primarily oriented towards prevention.

The Faustlos program is based on the Second Step program developed in the United States (Beland 1988) but translated, adapted, and evaluated in a German context for kindergarten and primary schools (Cierpka 2005). The Faustlos material comes in the form of a toolkit and has been delivered to approximately one third of Austrian primary schools. According to the interviewed expert, all Austrian primary schools had the opportunity to request a Faustlos toolkit. All teachers who received the Faustlos toolkit were required to complete Faustlos training. The

initiative ended with the last Faustlos training event in May 2010. To analyze the quality of the implementation the expert conducted an Austria-wide online survey developed in cooperation with the author of the Faustlos program (Cierpka 2005). In sum, more than four hundred teachers participated in the study.

The ViSC Social Competence Program (Atria and Spiel 2007; Spiel and Strohmeier 2011, Strohmeier et al 2012) is a school development project to prevent violence and foster social competencies in secondary schools. Activities are geared to operate on three different levels: the school as a whole, the classroom, and the individual level. A cascaded train-the-trainer model was developed and applied to implement the ViSC program sustainably in the school system: Scientists train ViSC coaches, ViSC coaches train teachers, and teachers train their students. The immediate target groups of this project (ViSC coaches) are teacher-training lecturers and psychologists. Between academic year 2008/09 and 2010/11 thirty-six coaches were trained.

The implementation quality and effectiveness of the ViSC program was intensively evaluated in 2009/10. Evaluation results showed that the program had very positive effects on students in comparison to a randomized control group (Strohmeier et al. 2012). Furthermore, to support schools in implementing the ViSC program a manual for teachers has been prepared.

The peer mediation project aims to provide Austria-wide support for teachers seeking to improve conflict resolution in schools. Within this project, students were trained to mediate in conflicts involving their peers. These peer mediators are supported by teachers trained to coach them. According to the interviewed expert, peer mediation was established in many Austrian schools within the framework of social learning even before the development and implementation of the national strategy. In other words, this project was incorporated into the national strategy as an existing measure. In 2009, the interviewed expert conducted an Austria-wide survey on peer mediation activities in schools to document and evaluate the project. In addition, quality standards for training coaches for peer mediation have been developed.

5.2. Knowledge Transfer and Education

The projects implemented within this activity domain aim to optimize basic and advanced teacher training regarding teaching of social skills and competencies for dealing with violence. The pedagogical universities play a crucial role here. One project in this domain was therefore to examine the curricula of the pedagogical universities to identify whether and to what degree they include obligatory components of violence prevention and social competence promotion. According to the interviewed expert it was a challenge to achieve adequate commitment of the pedagogical universities, which have only recently been established in Austria. The analysis of the curricula resulted in concrete recommendations for basic and advanced teacher training for dealing with violence and aggressive behavior. Subsequently, the actual implementation of these recommendations will be examined. The aim is to create a framework for a violence prevention curriculum for teacher training and to develop corresponding modules.

Another project in this domain was a two-day train-the-trainer seminar conducted in spring 2008 to provide teachers from the pedagogic universities with evidence-based knowledge for violence prevention. According to the interviewed expert, feedback reports from participants showed high acceptance ratings.

Furthermore, thirty-six ViSC coaches have been trained (see domain “Prevention and Intervention”). ViSC coaches are working with schools applying the ViSC program and also function as trainers for teachers.

In addition, twelve Austria-wide information workshops for schools and kindergartens were organized (“Joining forces against bullying and violence”). Each workshop was designed for about thirty participants. According to the interviewed expert parents, teachers, and other interested professional groups participated.

5.3. Information and Public Relations

The aim of this activity domain is to publicize the national strategy and make information material available. Its heart is the internet platform <http://www.gemeinsam-gegen-gewalt.at/>, which supplies information about the national

strategy and research on violence prevention in schools and provides school partners with access to materials and targeted information on violence in schools. In addition, the internet platform serves as a networking space for all partners contributing to the prevention of violence at school.

In addition, various events and press conferences were conducted, all related to different specific initiatives within the national strategy. For example, the national strategy was launched at a major press conference in December 2007 where the minister herself presented the “Together against Violence” initiative and described the initiative’s first plans and projects.

5.4. Networking and Cooperation

Networking and cooperation among the initiative’s partners is an important aim of the national strategy (Shonkoff and Bales 2011; Spoth and Greenberg 2011). To achieve this aim, annual conferences (see Table 1) have been conducted since 2008 to provide a platform for exchange between the different stakeholders in the field of school violence prevention. In these conferences, a common knowledge base for implementation of the national strategy should be created. In addition, the respective responsibility of the stakeholders (partners) in violence prevention should be discussed and clarified with the aim to create a platform for the schools on national and regional level and the public. These objectives are supported by the integration of national and international experts.

All members of the steering committee were interviewed about this activity domain. They agreed that the networking activities in general and the annual meetings in particular have a high priority for the national strategy. The steering committee has therefore taken a greater role in coordinating the planning process. There is also agreement among the interviewees that the planning and design of the networking meetings has developed very positively. The network meetings have been consecutively optimized on the basis of the experience and the evaluation results (participants’s assessments) of the previous meetings. As a consequence, representatives of the partner groups were involved in the preparation of the third networking meeting.

5.5. Evaluation and Research

This activity domain emphasizes the importance of evaluation and supporting research for the national strategy. Both the evaluation described here and the evaluations of the Faustlos and ViSC programs are concrete projects of this domain (see domain “Prevention and Intervention”).

In addition, this domain involves the development of on-line self-assessment instruments for schools. These tools enable principals and teachers to assess and interpret violence rates in their schools and classrooms, as well as to evaluate the effectiveness of interventions against violence. Consequently, these tools also support the sustainable implementation of violence prevention in Austria, as the presence of researchers is not needed for data collection, analysis, and interpretation. The AVEO self-evaluation tool (Austrian Violence Evaluation Online-Tool) provides information about violence rates from the perspective of students and is already operating (Spiel et al. 2011). The teacher and school perspective was systematically integrated into the development of the self-evaluation tool and the development carefully evaluated (Spiel et al. 2011). An analogous tool collecting data from teachers is in preparation.

5.6. Analysis of the National Strategy at General Level

All interview partners were asked about the national strategy at a general level. Their statements on the question of what changes they have observed since the initiation of the national strategy were very homogeneous. All of them pointed to the enhanced cooperation in the steering committee and the significantly improved project management compared to the starting phase. Cooperation in the steering committee was described as well developed, constructive, open concerning communication, and conducive to the exchange of knowledge between the individual projects and to overall coordination. Synergies have been increasingly identified and used. According to the interview partners, this positive trend is also reflected in increased networking between the projects.

The external development of the strategy was more cautiously assessed by the interviewees. They agree that the initiative has not actually been recognized as intended in

public and made several recommendations for improvement. One of them suggested a survey of schools to investigate awareness of the initiative, in order to acquire reliable data. In addition, some interviewees recommended optimizing and intensifying public relations (e.g., active involvement of Austrian broadcast media in the initiative). Some positive developments have been observed but further work needs to be invested. In particular, an overall public relations strategy was requested by interviewees at multiple levels (school, parents, and public) and in media with different levels of coverage (nationwide, state, and regional), as recommended in the strategy plan (Spiel and Strohmeier 2007).

In addition, strengthening the projects in the regions, establishing or strengthening local networks (schools), and raising teachers’ awareness were identified as future tasks of the national strategy. Here, the increased involvement of the pedagogical universities was seen as the key by all interviewees.

Furthermore, the interviewees agreed that the fact that the political declaration has not been realized at the national and state levels as recommended in the original strategy plan (Spiel and Strohmeier 2007) has been a limiting factor for the strategy’s success and for the commitment of certain stakeholder groups. According to the strategy plan there should have been a national declaration level signed by the president and the chancellor, as well as by representatives of the teaching unions, the national parents’ committee, and the students’ unions. Similar declarations at the state and the local level are suggested in the strategy plan. However, for political reasons the Federal Minister of Education, the Arts, and Culture did not support these declarations. This makes it clear that Austria has yet to achieve the national political commitment to violence prevention by all parties and the whole government that Spiel and Strohmeier (2007) identify as the central basis for the success of a national strategy as exemplified by the case of Norway (Roland 2011).

6. Conclusions and Lessons Learned

The lessons for the development of national strategies and the promotion of evidence-based policy and practice have

recently been discussed by Spiel and Strohmeier (2012). Therefore, we focus here on the implementation of the national strategy and the results of the implementation evaluation. However, the results are only preliminary as the strategy has not been fully implemented and the implementation evaluation had a formative rather than summative focus.

Nevertheless, the results of the evaluation should be discussed concerning fidelity, quality of implementation, and adaptation to local political and social circumstances. The starting point was the strategy plan recommended by Spiel and Strohmeier (2007). However, for political reasons the Federal Minister did not realize all parts of the plan (e.g., the declarations at different political levels). Furthermore, the Federal Minister has extended the national strategy by adding some pre-existing projects (e.g. the peer mediation and behavior agreements projects). Consequently, the strategy and also the steering committee itself became less focused and more heterogeneous.

If fidelity and quality of implementation were to be assessed in terms of the original strategy plan, the results would not be completely positive. However, if the politically modified strategy plan is used as the reference, taking into account the challenges caused by the modification, the results are satisfactory – in particular if it is borne in mind that implementation continues until the end of 2013. According to the implementation evaluation the projects are proceeding well. However, whether they together achieve the goals of the national strategy remains to be proven by a summative evaluation after full implementation of the national strategy, by collecting data from teachers and students. Attainment of the third goal (Increased responsibility and civil engagement) has so far only been successfully proven at project level (by the ViSC program evaluation; e.g., Strohmeier et al. 2012).

According to previous research (Datnow and Stringfield 2000; Shonkoff and Bales 2011; Spoth and Greenberg 2011) the views of the stakeholder groups actively engaged in the field of violence prevention were already considered in the development of the strategy plan. Furthermore, one out of six activity domains explicitly focuses on networking and cooperation. However, while internal cooperation (the steering committee) has been satisfactory improved it turns out that the engagement of stakeholder groups, in particular at state and local levels, plainly needs more time (Datnow 2002, 2005). Effective steps have been already set in motion.

The main future challenges are the systematic engagement of the pedagogical universities and the public visibility of the national strategy. Responsiveness at all levels is considered an important mediator of fidelity and quality of implementation and therefore of program outcome (Berkel et al. 2011). In agreement with Shonkoff (2000) we must acknowledge that science, policy, and practice reflect different ways of thinking about violence prevention. However, we also agree with him that success in the long run is best addressed as continuous work in progress (Shonkoff 2000). Aside from this, the development of the national strategy and its implementation have already had several positive effects on a more general level. The usual practice was changed from supporting single initiatives lacking standards of evidence (Atria and Spiel 2003) to promoting evidence-based programs. Moreover, a rigorous evaluation of the ViSC program was applied using randomized trials under real-world conditions. To our knowledge, this was the first time that the Austrian Federal Ministry financed such a procedure. Last but not least, the Federal Minister and the members of the steering committee were persuaded to commission an evaluation of the implementation of the national strategy and to use the evaluation results for improvement.

References

- Atria, Moira, and Christiane Spiel. 2003. The Austrian Situation: Many Initiatives Against Violence, Few Evaluations. In *Violence in Schools: The Response in Europe*, ed. Peter K. Smith, 83–99. London: RoutledgeFalmer.
- Atria, Moira, and Christiane Spiel. 2007. Viennese Social Competence (ViSC) Training for Students: Program and Evaluation. In *Bullying, Victimization, and Peer Harassment: A Handbook of Prevention and Intervention*, ed. Joseph E. Zins, Maurice J. Elias, and Charles A. Maher, 179–97. New York: Haworth.
- Beland, Kathy. 1988. *Second Step: A Violence-Prevention Curriculum: Grades 1–3*. Seattle: Committee for Children.
- Berkel, Cady, Anne M. Mauricio, Erin Schoenfelder, and Irwin N. Sandler. 2011. Putting the Pieces Together: An Integrated Model of Program Implementation. *Prevention Science* 12 (1): 23–33.
- Bundesministerium für Unterricht, Kunst und Kultur. 2008. Weiße Feder – Gemeinsam gegen Gewalt. Website. www.gemeinsam-gegen-gewalt.at.
- Beets, Michael W., Brian R. Flay, Samuel Vuchinich, Alan A. Acock, Kin-Kit Li, and Carol Allred. 2008. School Climate and Teachers' Beliefs and Attitudes Associated with Implementation of the Positive Action Program: A Diffusion of Innovations Model. *Prevention Science* 9 (4): 264–75. doi:10.1007/s1121-008-0100-2
- Cargo, Margaret, Jon Salsberg, Treena Delormier, Serge Desrosiers, and Ann C. Macaulay. 2006. Understanding the Social Context of School Health Promotion Program Implementation. *Health Education* 106 (2): 85–97. doi:10.1108/09654280610650936
- Carroll, Christopher, Malcolm Patterson, Stephen Wood, Andrew Booth, Jo Rick, and Shashi Balain. 2007. A Conceptual Framework for Implementation Fidelity. *Implementation Science* 2 (40): 1–9.
- Cierpka, Manfred. 2005. *FAUSTLOS – Wie Kinder Konflikte gewaltfrei lösen lernen*. Freiburg: Herder.
- Cross, Donna, Melanie Epstein, Lydia Hearn, Phillip Slee, Therese Shaw, and Helen Monks. 2011. National Safe Schools Framework: Policy and Practice to Reduce Bullying in Australian Schools. *International Journal of Behavioral Development* 35 (5): 398–404. doi:10.1177/0165025411407456.
- Crowley, D. Max, Mark T. Greenberg, Mark E. Feinberg, Richard L. Spoth, and Cleve R. Redmond. 2012. The Effect of the PROSPER Partnership Model on Cultivating Local Stakeholder Knowledge of Evidence-Based Programs: A Five-Year Longitudinal Study of 28 Communities. *Prevention Science* 13 (1): 96–105. doi:10.1007/s1121-011-0250-5
- Currie, Candace, Cara Zanotti, Antony Morgan, Dorothy Currie, Margaretha de Looze, Chris Roberts, Oddrun Samdal, Otto R. F. Smith, and Vivian Barnekow, eds. 2012. *Social Determinants of Health and Well-Being among Young People: Health Behaviour in School-Aged Children (HBSC) Study: International Report from the 2009/2010 Survey*, Health Policy for Children and Adolescents 6. Copenhagen: WHO Regional Office for Europe.
- Datnow, Amanda. 2002. Can We Transplant Educational Reform, and Does It Last? *Journal of Educational Change* 3 (3): 215–39. doi:10.1023/A:1021221627854
- Datnow, Amanda. 2005. The Sustainability of Comprehensive School Reform Models in Changing District and State Contexts. *Educational Administration Quarterly* 41 (1): 121–153. doi:10.1177/0013161X04269578
- Datnow, Amanda, and Sam Stringfield. 2000. Working Together for Reliable School Reform. *Journal of Education for Students Placed at Risk (JESPAR)* 5 (1–2): 183–204. doi:10.1080/10824669.2000.9671386
- Durlak, Joseph A., and Emily P. DuPre. 2008. Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology* 41 (3–4): 327–50.
- Dusenbury, Linda, Rosalind Brannigan, Mathea Falco, and William B. Hansen. 2003. A Review of Research on Fidelity of Implementation: Implications for Drug Abuse Prevention in School Settings. *Health Education Research* 18 (2): 237–56.
- Elias, Maurice J., Joseph E. Zins, Patricia A. Craczyk, and Roger P. Weissberg. 2003. Implementation, Sustainability, and Scaling Up of Social-Emotional and Academic Innovations in Public Schools. *School Psychology Review* 32 (3): 303–19.
- Ferguson, Christopher J., Claudia San Miguel, John C. Kilburn, and Patricia Sanchez. 2007. The Effectiveness of School-Based Anti-Bullying Programs. *Criminal Justice Review* 32 (4): 401–14.
- Fixsen, Dean L., Karen A. Blase, Sandra F. Naoom, and Frances Wallace. 2009. Core Implementation Components. *Research on Social Work Practice* 19 (5): 531–40.
- Fixsen, Dean L., and Karen A. Blase. 2009. *Implementation: The Missing Link Between Research and Practice*, NIRN Implementation Brief 1. Chapel Hill: The University of North Carolina, FPG, NIRN.
- Gläser, Jochen, and Grit Laudel. 2009. *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*. Wiesbaden: Verlag für Sozialwissenschaften
- Kalafat, John, Rober J. Illback, and Daniel Sanders Jr. 2007. The Relationship Between Implementation Fidelity and Educational Outcomes in a School-Based Family Support Program: Development of a Model for Evaluating Multidimensional Full-Service Programs. *Evaluation and Program Planning* 30 (2): 136–48.
- Linstone, Harold A., and Murray Turoff. 1975. *The Delphi Method: Techniques and Applications*. Reading, MA: Addison Wesley.
- Mayer, Horst. 2008. *Interview und schriftliche Befragung: Entwicklung Durchführung Auswertung*. Munich: Oldenbourg.
- Mayingr, Phillip. 2002. *Einführung in die qualitative Sozialforschung*. Weinheim: Beltz.
- Ogden, Terje, Freja Ulvestad Karki, and Katrine Stegenborg Teigen. 2010. Linking Research, Policy and Practice in Welfare Services and Education in Norway. *Evidence and Policy* 6 (2): 161–77.
- Olweus, Dan. 2004. The Olweus Bullying Prevention Programme: Design and Implementation Issues and a New National Initiative in Norway. In *Bullying in Schools: How Successful Can Interventions Be?* ed. Peter K. Smith, Debra J. Pepler, and Ken Rigby, 13–36. Cambridge: Cambridge University Press.
- Pepler, Debra, and Wendy Craig. 2011. Promoting Relationships and Eliminating Violence in Canada. *International Journal of Behavioral Development* 35 (5): 389–97. doi:10.1177/0165025411407455.
- Roland, Erling. 2000. Bullying in Schools: Three National Innovations in Norwegian Schools in 15 years. *Aggressive Behavior* 26 (1): 135–43.
- Roland, Erling. 2011. The Broken Curve: Effects of the Norwegian Manifesto Against Bullying. *International Journal of Behavioral Development* 35 (5): 383–88.
- Salmivalli, Christina, Antti Kärnä, and Elisa Poskiparta. 2011. Counteracting Bullying in Finland: The KiVa Program and Its Effects on Different Forms of Being Bullied. *International Journal of Behavioral Development* 35 (5): 405–11. doi:10.1177/0165025411407457.
- Schirmer, Dominique. 2009. *Empirische Methoden der Sozialforschung*. Paderborn: W. Fink.
- Shonkoff, Jack P. 2000. Science, Policy, and Practice: Three Cultures in Search of a Shared Mission. *Child Development* 71 (1): 181–87. doi:10.1111/1467-8624.00132.
- Shonkoff, Jack P., and Susan Nall Bales. 2011. Science Does Not Speak for Itself: Translating Child Development Research for the Public and Its Policymakers. *Child Development* 82 (1): 17–32. doi:10.1111/j.1467-8624.2010.01538.x.
- Spiel, Christiane, Christina Salmivalli, and Peter K. Smith. 2011. Translational Research: National Strategies for Violence Prevention in School. *International Journal of Behavioral Development* 35 (5): 381–82. doi:10.1177/0165025411407556.
- Spiel, Christiane, and Dagmar Strohmeier. 2007. *Generalstrategie zur Gewaltprävention an österreichischen Schulen und Kindergärten „Gemeinsam gegen Gewalt“*, report to the Federal Ministry of Education, the Arts, and Cultural Affairs. Vienna, Austria: University of Vienna. <http://www.gemeinsam-gegen-gewalt.at/resources/files/43/bericht-generalstrategie-29102007-ohne-anhang.pdf>

- Spiel, Christiane, and Dagmar Strohmeier. 2011. National Strategy for Violence Prevention in the Austrian Public School System: Development and Implementation. *International Journal of Behavioral Development* 35 (5): 412–18.
- Spiel, Christiane, and Dagmar Strohmeier. 2012. Evidence-Based Practice and Policy: When Researchers, Policy Makers, and Practitioners Learn How to Work Together. *European Journal of Developmental Psychology* 9 (1): 150–62.
- Spiel, Christiane, Dagmar Strohmeier, Marie-Therese Schultes, and Christoph Burger. 2011. *Nachhaltigkeit von Gewaltprävention in Schulen: Erstellung und Erprobung eines Selbstevaluationsinstruments*, report for the Austrian Federal Ministry for Education, the Arts, and Culture. Vienna, Austria: University of Vienna.
- Spoth, Richard, and Mark Greenberg. 2011. Impact Challenges in Community Science-with-Practice: Lessons from PROSPER on Transformative Practitioner-Scientist Partnerships and Prevention Infrastructure Development. *American Journal of Community Psychology* 48 (1):106–19. doi:10.1007/s10464-010-9417-7
- Strohmeier, Dagmar, Christine Hoffmann, Eva-Maria Schiller, Elisabeth Stefanek, and Christiane Spiel. 2012. ViSC Social Competence Program. *New Directions for Youth Development* 133:71–80.
- Ttofi, Maria M., and David P. Farrington. 2009. What Works in Preventing Bullying: Effective Elements of Anti-Bullying Programmes. *Journal of Aggression, Conflict and Peace Research* 1 (1): 13–24.
- Wagner, Petra. 2011. *Evaluation der Generalstrategie zur Gewaltprävention and Österreichischen Schulen und Kindergärten „Gemeinsam gegen Gewalt“*, unpublished report to the Federal Ministry of Education, the Arts, and Cultural Affairs. Vienna, Austria.

Christiane Spiel

christiane.spiel@univie.ac.at

Petra Wagner

petra.wagner@fh-linz.at

Dagmar Strohmeier

dagmar.strohmeier@fh-linz.at

Clinical Significance of Parent Training for Children with Conduct Problems

Martin Forster, Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Åsa Kling, Department of Psychology, Uppsala University, Sweden

Knut Sundell, The Swedish National Board of Health and Welfare (Socialstyrelsen), Stockholm, Sweden

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

► **Clinical Significance of Parent Training for Children with Conduct Problems** Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Clinical Significance of Parent Training for Children with Conduct Problems

Martin Forster, Division of Psychology, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Åsa Kling, Department of Psychology, Uppsala University, Sweden

Knut Sundell, The Swedish National Board of Health and Welfare (Socialstyrelsen), Stockholm, Sweden

While there is a strong evidence base for behavioral parent training in the treatment of child conduct problems, the clinical impact is less well known. Meta-analyses report effect sizes in the medium range, but the common practice of reporting “small,” “medium,” and “large” effects can be misleading and difficult to understand for practitioners and clients. There is a need for more research addressing the clinical significance of behavioral parent training, which would help to bridge the gap between research and practice. In the first part of this report, a reanalysis in terms of clinical significance of two outcome studies published by the authors was conducted. In the second part, the results from the first part were compared to six outcome studies published by other authors. The median number needed to treat across studies was five, which means that for every five treated children, one shows reliable change and moves from the dysfunctional to the functional population.

Evidence for the efficacy of psychological interventions generally relies on reports of statistical and practical significance (e.g., Chambless and Hollon 1998). Although statistical significance testing provides information as to the reliability of outcomes, it tells us little about the importance of such outcomes. Practical significance (i.e., effect sizes) provides information as to the magnitude of treatment effects at a group level and has the advantage that effects can be compared across studies. It is however not easily understood by clinicians and can be influenced by factors such as within-group variance and baseline levels of outcome measures. Two studies with equal effect sizes can, for example, differ considerably in proportions of participants who recover or improve. This points to the inherent problem in using the conventional definitions of “small” (.20), “medium” (.50), and “large” (.80) effect sizes to classify treatment effects (Cohen 1988). A small effect size may be clinically meaningful in one context, while close to meaningless in others. Therefore, several researchers have stressed the importance of including clinical significance, in addition to statistical and practical significance, in reports of treatment effects (Campbell 2005; La Greca 2005). The primary aim of the current work was

to investigate the clinical significance of behavioral parent training in the treatment of child conduct problems by synthesizing the results of published outcome studies.

Clinical significance refers to the importance or practical meaning of treatment effects – that is, proportions of clients who recover or improve and whether the changes make a real difference in the everyday life of the clients, besides reduction of the specific clinical symptoms being measured (Kazdin 1999). Despite the apparent benefits of considering clinical significance in syntheses of intervention research, it is seldom included in research reviews and clinical guidelines. The term was, for example, mentioned only once in an entire special issue of the *Journal of Clinical Child and Adolescent Psychology* on empirically supported treatments for children (initiated by American Psychological Association) (Silverman and Hinshaw 2008). There are at least two reasons for this. First, even if influential scholars and journals have called for the inclusion of clinical significance in outcome studies, most studies still only report results in terms of statistical significance and effect sizes (Ogles, Lunnen, and Bonesteel 2001; Campbell 2005). Second, there is no consensus as to how clinical significance

should be operationalized and measured (Ogles et al. 2001; Campbell 2005). Before to giving clinical significance the important role it deserves and including it in the guidelines for establishing empirically supported treatments as effective, it is essential to first agree upon its common operationalization and the analytical approach to its assessment.

The closest there is to a common standard is the procedure described by Jacobson and Truax (1991) – the “JT method” – which is the most widespread standardized method for assessment of clinical significance (Ogles et al. 2001). The JT method is based on two criteria that are used to classify participants in outcome studies. To satisfy the first criterion, an individual identified as member of the dysfunctional distribution on a given outcome measure must move to the functional distribution after treatment. This establishes *clinical change*. To satisfy the second criterion, the change in the individual has to be of sufficient magnitude to determine that it is significant rather than simply an artifact of measurement error. This establishes *reliable change*. Participants who satisfy both criteria are classified as *recovered*, whereas those who experience reliable change without passing the clinical cutoff are classified as *improved*.

The JT method has been shown to be as reliable and valid as more advanced statistical methods used to assess clinical significance (e.g., Bauer, Lambert, and Nielsen 2004). Therefore, and because the it is relatively easy to apply and understand, it is generally recommended over other methods (Bauer et al. 2004). It is important to point out that the JT method is not applicable to every type of clinical problem or context (Campbell 2005). The method does for example require that there is a clinically relevant cutoff point between dysfunctional (e.g., a diagnosis or defined risk group) and functional populations. In studies of problems without clinically relevant cutoff points (e.g., cigarette smoking), effect sizes may be a better way to operationalize meaningful change. On the other hand, for treatments targeting clinically defined groups, such as children with conduct problems, there is seldom an excuse not to use some variation of the JT method.

A strong argument for including clinical significance in research reports is that policymakers, practitioners, and con-

sumers can more easily understand the magnitude of treatment effects. Some authors also argue that results obtained in analyses of clinical significance should preferably be reported as numbers needed to treat – NNT (Marrs-Garcia 2010). The NNT is the number of individuals who would need to be exposed to a particular treatment before one individual would recover. Hence, a NNT close to 1 suggests that nearly all study participants recovered. In controlled studies, the NNT represents the relative advantage of the treatment group over the control group. For example, if every other participant (50 percent) recovered in a treatment group and every fourth (25 percent) spontaneously recovered in a no-treatment control group, the “net” gain of the treatment is 25 percent (50 percent minus 25 percent). For every four treated patients one would recover as a result of the treatment, which translates to a NNT of four.

Like intervention research in general, research on behavioral parent training for children with conduct problems suffers from a lack of standardized analyses of clinical significance. In most outcome studies of behavioral parent training reports of clinical significance are not included at all. Some studies use procedures that prevent comparison across studies, such as defining clinical significance as participants who show at least 30 percent improvement on a given outcome measure (e.g., Reid, Webster-Stratton, and Hammond 2007). Only a handful of studies use standardized methods (e.g., the JT method) that allow for comparison across studies and synthesis of data. The omission of clinical significance in published reviews and meta-analyses of behavioral parent training in the treatment of child conduct problems therefore comes as no surprise. This stands in sharp contrast to the impressive body of research supporting the statistical and practical significance of behavioral parent training, which has been reported in numerous reviews and meta-analyses (Eyberg, Nelson, and Boggs 2008; Furlong et al. 2012; Dretzke et al. 2009).

In the first part of this article, the aim is to contribute to the small body of research that properly reports clinical significance in studies of behavioral parent training for children with conduct problems by reanalyzing two studies previously published by some of the authors of the present

report in terms of clinical significance. In the second part, we investigate the clinical significance of behavioral parent training by synthesizing the results from the handful of published studies that have used standardized procedures to assess the clinical significance of behavioral parent training.

1. Part I: Reanalysis of Two Published Parent Training Studies

In the first part of this contribution, we report an analysis of clinical significance performed with data from two previously published studies by some of the authors of the present report. Both studies were randomized trials of a Swedish parent training program called Comet (Kling et al. 2010; Enebrink et al. 2012). In both studies there were statistically significant differences between treatment groups and waitlist control groups, with effect sizes in the medium to large range. The program has been implemented on a wide scale in Sweden through different methods of delivery. The standard method of delivery (Comet-S) consists of eleven 2.5-hour workshops, in which two practitioners, usually from the social services, provide guidance in effective parenting practices to groups of parents. The program is based on a manual, which contains theory and practice in sensitive play, praise, incentives, ignoring of misconduct, and rules and expectations. Video modeling, role-play, and homework assignments are key ingredients in the process of delivery. Parents participating in the self-directed version of the program (Comet-SD) receive exactly the same written material as parents in Comet-S, but the material is introduced at a single workshop without further practitioner support. The internet-based delivery format (Comet-I) also contains the same material as Comet-S, including instructional text and video vignettes, but also offers several interactive features such as participant support forums and minimal e-mail contact with a practitioner. For further description of the content and evaluations of the Comet program, see Kling et al. (2010) and Enebrink et al. (2012).

1.1. Method

1.1.1. Analysis of Clinical Change

The JT method was used in the reanalysis of the two studies. The first step of the JT method is to determine whether participants experience a clinical change, i.e., move from the dysfunctional to the functional distribution

on a given outcome measure. To make such an analysis, a cutoff point that divides the two distributions has to be determined. Cutoff C, which is defined as the weighted midpoint between the means of functional and dysfunctional populations, is generally the recommended method (Bauer et al. 2004; Evans, Margison, and Barkham 1998). Computation of Cutoff C requires that normative data is available for the selected outcome measure(s). While several outcome measures were used in the studies, norms were only available for the Eyberg Child Behavior Inventory (ECBI). Therefore, that measure was used to assess clinical significance in the two studies.

The ECBI (Eyberg and Pincus 1999) consists of thirty-six items describing disruptive and aggressive behaviors (e.g., “Hits parents” and “Does not obey house rules”), which are each rated in terms of their frequency on a seven-point likert scale (1 = never happens, 7 = always happens). The sum of these items is called the “intensity scale” (ECBI-IS) with a range of 36–252. The same items are also rated on a “problem scale” (ECBI-PS), which measures whether the parents experience the occurring behaviors as problematic (1 = yes) or not (0 = no). That scale thus has a range of 0–36. The ECBI is probably the most common outcome measure in studies of behavioral parent training for children with conduct problems (Dretzke et al. 2009) and numerous studies have investigated and confirmed its psychometric properties (Plake, Impara, and Spies 2003). In the study by Kling et al. (2010), mothers alone were the respondents for 84 percent of the participants. Fathers alone were respondents for 10 percent of the participants and both parents responded for the final 6 percent. In the study by Enebrink et al. (2012), it is unknown who the responding parent was.

The means and standard deviations for the dysfunctional population (i.e., children with conduct problems) and functional population (i.e., normal children) are required to compute the cutoff C. The pretest means and standard deviations in each study were used to represent the dysfunctional population in the present analysis, which is recommended instead of using published normative/clinical data (Jacobson and Truax 1991). On the other hand, published normative data is required to obtain means and

standard deviations for functional populations. In the present analysis, Swedish normative data was used as means and standard distributions for the functional distributions (Axberg, Johansson Hanse, and Broberg 2008). Normative data for six-year olds were used for the analysis of Kling et al. (2010); $M = 90.7$, $SD = 23.6$ for ECBI-IS and $M = 2.95$, $SD = 4.10$ for ECBI-PS. For Enebrink et al. (2012), normative data for seven-year olds was used; $M = 85.2$, $SD = 23.5$ for ECBI-IS and $M = 2.46$, $SD = 4.08$ for ECBI-PS. The weighted midpoint (Cutoff C) between functional and dysfunctional distributions was 113 (ECBI-IS) and 8 (ECBI-PS) for participants in Kling et al. (2010). In Enebrink et al. (2012), the corresponding cutoff points were 122 (ECBI-IS) and 9 (ECBI-PS). Participants who scored above these cutoff points at pretest, and below at posttest, satisfied the criterion for clinical change.

1.1.2. Analysis of Reliable Change

In the second step of the JT method, a reliable change index (RCI) is computed for each participant, representing the change between pretest and posttest divided by the standard error of difference between the two scores (Jacobson and Truax 1991). The standard error of difference is dependent on the variability in the studied sample (i.e., the standard deviation at pretest), but also the reliability of the measurement. The internal consistency of the ECBI, which is recommended over other types of reliability measures, was used as the reliability coefficient in the present analyses (Bauer et al. 2004; Evans et al. 1998). Furthermore, the reliability coefficient should be obtained from the studied sample, rather than published test data (Campbell 2005). In Kling et al. (2010), the internal consistency (Cronbach's alpha) was $\alpha = .92$ (ECBI-IS) and $\alpha = .89$ (ECBI-PS), while the corresponding coefficients were $\alpha = .81$ (ECBI-IS) and $\alpha = .79$ in Enebrink et al. (2012).

For individuals with a reliable change index larger than 1.96 change is unlikely to be due to measurement error ($p < .05$), which means that they satisfy the criterion for reliable change. It is also possible to calculate how much an individual must change on a given outcome. For participants in Kling et al. (2010), the minimum difference between pretest and posttest that constituted a reliable change was 20.4 points on the ECBI-IS and 5.6 points on the ECBI-PS. The

corresponding thresholds for participants in Enebrink et al. (2012) were 22.4 (ECBI-IS) and 7.0 (ECBI-PS).

1.1.3. Classification of Participants

The participants in the current analysis were classified as *recovered* if they made both a reliable and a clinical change (satisfied both criteria in the JT method). They were classified as *improved* if they satisfied the criterion of reliable change, but not that of clinical change. If they made a reliable change in the undesired direction, they were classified as *deteriorated*. Finally, participants who made no reliable change in any direction were classified as *unchanged*. Sometimes the unchanged category is defined as participants who "pass neither criteria" (e.g., Campbell, 2005), but McGlinchey, Atkins, and Jacobson (2002) recommend the definition used here. Finally, chi-square analyses (Fisher's exact test) were performed to assess whether the clinical significance differed significantly between the treatment and control groups, as recommended by Kendall et al. (1999).

Using the intention to treat principle in the analyses of the two studies makes particular sense with regard to clinical significance. Early termination of treatment may be an even bigger problem in clinical practice than in research settings (Kazdin 2008) and it is therefore reasonable to assess clinical significance including the total sample rather than just study completers. To obtain complete data for every participant, the last observed score was carried forward in cases of missing data at posttest and/or follow-up. This implies that every participant who dropped out or had a missing score was classified as unchanged.

1.2. Results

Table 1 shows the effect sizes and clinical significance for the three study conditions in Kling et al. (2010). After subtracting the proportions of the control group from the treatment groups, the recovery rates for the ECBI-IS were 28 percent (Comet-S), 13 percent (Comet-SD), and 26 percent (Comet-I), which translates to NNTs of four, eight, and four. The corresponding rates for the ECBI-PS were 28 percent, 11 percent, and 39 percent, with NNTs of four, nine, and 13. The recovery rates were statistically significantly larger in all treatment groups compared to the waitlist control groups.

Table 1: Clinical significance of the Comet program at posttest

Program/outcome	Cohen's <i>d</i>	Recovered n (%)	Improved n (%)	Unchanged n (%)	Deteriorated n (%)
Comet-S					
ECBI-IS	.71	16 (28)***	6 (10)	35 (60)*	1 (2)
ECBI-PS	.90	16 (28)***	8 (14)	34 (58)*	1 (2)
Comet-SD					
ECBI-IS	.56	8 (13)*	7 (11)	45 (74)	1 (2)
ECBI-PS	.52	7 (11)*	7 (11)	45 (74)	2 (4)
Waitlist					
ECBI-IS	.01	0 (0)	4 (10)	33 (82)	3 (8)
ECBI-PS	.00	0 (0)	3 (7)	32 (80)	5 (13)
Comet-I					
ECBI-IS	1.62	30 (52)*	3 (5)	20 (34)**	5 (9)
ECBI-PS	1.53	30 (52)***	6 (10)	21 (36)***	1 (2)
Waitlist					
ECBI-IS	.83	12 (26)	3 (7)	31 (67)	0 (0)
ECBI-PS	.72	6 (13)	3 (7)	37 (80)	0 (0)

Notes: Cohen's *ds* are within-group effect sizes.

Proportion significantly different from corresponding proportion in waitlist control group at * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 2 shows the results at six-month follow-up for Comet-S and Comet-SD. The proportions of recovered or improved participants were larger or similar to the corresponding proportions at posttest in both groups. The recovery rates were about twice as large for Comet-S (29

percent and 43 percent) as for Comet-SD (15 percent and 18 percent), but only the advantage pertaining to ECBI-PS was statistically significant. No follow-up data on clinical significance for Comet-I is reported in Enebrink et al. (2012).

Table 2: Clinical significance of Comet-S versus Comet-SD at follow-up

Program/outcome	Cohen's <i>d</i> ^a	Recovered n (%)	Improved n (%)	Unchanged n (%)	Deteriorated n (%)
Comet-S					
ECBI-IS	.85	17 (29)	9 (16)	30 (52)	2 (3)
ECBI-PS	1.20	25 (43)**	8 (14)	24 (41)*	1 (2)
Comet-SD					
ECBI-IS	.89	9 (15)	15 (25)	36 (59)	1 (2)
ECBI-PS	.82	11 (18)	10 (16)	37 (61)	3 (5)

Notes: Cohen's *ds* are within-group effect sizes (pretest/follow-up).

Proportion significantly different from corresponding proportion in Comet-SD group at * $p < .05$; ** $p < .01$.

2. Part II: Synthesis of Results across Studies

2.1. Method

2.1.1. Inclusion and Exclusion of Studies

The second part of this contribution compares published studies of behavioral parent training that include reports of clinical significance. The databases of PsychInfo and PubMed were searched up to July 2012. In addition, citations from a recent meta-analysis of behavioral parent training (Dretzke 2009) were also investigated. Marrs-Garcia (2010) specifies three conditions that have to be fulfilled to enable meaningful comparisons of NNTs across studies, which also apply to comparisons of clinical significance in general: (a) clinical significance has to be operationalized the same way across studies, (b) the control or comparison groups to which treated groups were compared have to be equivalent, and (c) the same outcome measure has to be used across studies. With these guidelines in mind, a set of criteria was developed to select studies for inclusion. First, only studies that based the analysis of clinical significance on the JT method, including analysis of both reliable and clinical change, were included. Second, only studies that included a waitlist/no-treatment control group were included, because this was the only type of comparison group that occurred in several studies. Third, only studies that based the analysis of clinical significance on the ECBI were included, because that was the only measure that occurred with sufficient frequency to allow proper comparisons across studies. Fourth, only studies published in peer-reviewed journals were included.

The search found in twenty-one studies of behavioral parent training for children with conduct problems that compared the treatment to a waitlist/no-treatment control group and employed the JT method to assess clinical significance. Five studies were excluded for reporting only reliable, but not clinical change, and another five studies were excluded for the opposite reason. Finally, three studies were excluded for basing the analysis of clinical significance on measures other than the ECBI. No authors of the excluded studies were contacted, because it was considered difficult or impossible for them to perform the necessary analyses to make the studies eligible for inclusion. The final sample therefore consisted of eight studies (including the two from

the first part of this report), altogether including 13 treatment conditions (Table 3).

In four of the studies in Table 3, different versions of the Triple-P program (Sanders 1999) were evaluated. Triple-P is a multilevel behavioral parent training program that targets different risk groups of children with conduct problems. In the self-directed version of the program (Triple-P-SD), parents receive training material (video and text) that they implement without any practitioner support. Triple-P-SD has also been enhanced in some studies with limited telephone support and/or a single session led by practitioners (Triple-P-SD+). In the standard version of the program (Triple-P-S), parents take part in ten individual one-hour sessions with a practitioner. Finally, the program has also been offered as an enhanced version (Triple-P-E). In addition to the ten sessions offered in Triple-P-S, parents in Triple-P-E also receive ten to fifteen sessions involving strategies to increase support from partners and friends as well as methods to manage stress, anxiety, and depression.

In the study by Nixon et al. (2003), two versions of the Parent-Child Interaction Therapy (PCIT) (Eyberg 1988) were evaluated. In the standard version of the program (PCIT-S), parents take part in twelve sessions (one to two hours) with a practitioner. In an abbreviated version of the program (PCIT-ABB), parents receive videotapes to learn the skills that are taught in PCIT-S. They also attend five face-to-face sessions with a practitioner, alternated with brief telephone sessions.

In the last study in Table 3, The Incredible Years program (IY) (Webster-Stratton 2000) was evaluated. In that program, parents of six to eight children meet for twelve to fourteen weekly two-hour sessions. Several video vignettes on specific parenting skills are shown and discussed during the sessions. Skills are role-played in the group and the parents get a weekly assignment to practice their newly acquired skills at home between sessions.

In all studies, with the exception of Enebrink et al. (2012), it was possible to conclude that the analysis of clinical significance was mainly or completely based on mothers' responses on the ECBI.

Table 3: Characteristics of synthesized studies

Study	Study conditions	n	Child age <i>M (SD)</i>	ECBI-IS pretest <i>M (SD)</i>	Cutoff
Kling et al. (2010)	Comet-S	58	6.0 (2.3)	138.0 (26.0)	113 ^a
	Comet-SD	61			
	Waitlist	40			
Enebrink et al. (2012)	Comet-I	58	6.8 (2.3)	150.7 (18.5)	122 ^a
	Waitlist	46			
Sanders et al. (2000)	TripleP-E	76	3.4 (0.3)	152.8 (26.0)	Not specified
	TripleP-S	77			
	TripleP-SD	75			
	Waitlist	77			
Morawska and Sanders (2006)	TripleP-SD+	43	2.2 (0.4)	119.1 (26.4)	131 ^b
	TripleP-SD	42			
	Waitlist	41			
Morawska et al. (2011)	TripleP-ABB	33	3.6 (0.9)	146.6 (28.0)	131 ^b
	Waitlist	34			
Joachim, Sanders, & Turner (2010)	TripleP-ABB	26	3.3 (1.1)	129.4 (25.8)	131 ^b
	Waitlist	20			
Nixon et al. (2003)	PCIT-S	22	3.9 (0.6)	164.9 (19.4)	131 ^b
	PCIT-ABB	23			
	Waitlist	18			
Axberg and Broberg (2012)	IY	38	6.0 (1.3)	156.4 (21.4)	121 ^b
	Waitlist	24			

^a The cut-off point was the weighted midpoint between the study sample mean and the mean of a normative population, as recommended by Jacobson and Truax (1991).

^b Cut-off point based on normative data only.

2.1.2. Analytic Strategy

Several different outcomes were included in the comparison of clinical significance across studies. First, the effect sizes were computed (Cohen's *d*), to enable comparison between practical and clinical significance. To make the effect sizes comparable across studies, they were not retrieved from the original articles, but re-computed from reported means and standard deviations. First, within-group effect sizes were computed separately for the treatment and control groups. The pooled standard deviation at pretest was used as denominator, with correction for small samples, and pre-post change scores were used as numerator. Second, the be-

tween-group effect sizes were computed by subtracting the within-group effect size in each control group from the corresponding treatment group.

Second, the proportions of participants experiencing reliable and clinical change were computed. When results are to be compared across studies, it is necessary to analyze between-group effects (Marrs-Garcia 2010), which therefore were computed by subtracting proportions of reliable and clinical change in the control groups from the corresponding proportions in the intervention groups. For example, in Kling et al. (2010), 38 percent of the participants in Comet-S and 10

percent in the waitlist control group experienced reliable change. The reliable change in terms of between-group effects therefore was 28 percent (38 minus 10).

Third, NNTs were operationalized and computed in two different ways. Some of the included studies used the classification of participants, as suggested in the JT method (Enebrink et al. 2012; Kling et al. 2010; Nixon et al. 2003). For those studies, NNTs based on proportions of recovered participants were computed. In the other studies, the JT method was used to compute proportions of participants experiencing reliable and clinical change, but the proportions were not combined to classify participants as recovered, improved, unchanged, or deteriorated, as suggested by Jacobson and Truax (1991). For those studies, the NNTs were based on either the reliable change or clinical change, whichever proportion was the smallest.

In several of the included studies the reported clinical significance was based on participants who completed the study, with no account of dropouts. In this synthesis, all results were instead analyzed as intention to treat. Dropouts were consequently counted as unchanged, which corresponds to the last observation carried forward method of handling missing data. A few studies also included follow-up measurements, but the dropout rates were generally high. Therefore, this report only include results at posttest in the synthesis across studies (Figure 1).

2.2. Results

The effect size and clinical significance based on the ECBI-IS for each of the thirteen treatments from the eight included studies are presented in Figure 1. The median was computed, instead of the mean, due to large variability across studies and treatments. Most effect sizes were in the medium to large range and all were of sufficient magnitude to be statistically significant. The median effect size across the thirteen treatments was $d = 1.31$ based on within-group effect sizes, and $d = .59$ based on between-group effect sizes (treatment vs. control). The median proportion of participants who made reliable change was 38 percent (within-group) and 19 percent (between-group). The corresponding proportions for clinical change were 37 percent (within-group) and 23

percent (between-group). For six treatments in Figure 1, proper recovery rates combining the criteria of reliable and clinical change were available. Across treatments, the median proportion of participants who experienced recovery was 31 percent (within-group) and 26 percent (between-group). Because the NNT represents a comparison between treatment and control conditions in controlled studies, all NNTs represent between-group effects. The median NNT across all treatments was five, which means that for every five treated children, one recovers. The median NNT for the treatments that involved full practitioner support (Comet S, TripleP-E, TripleP-S, PCIT-S, and IY) was four, while the median for the other treatments that involved no or minimal support was seven. Five of the included studies also reported the clinical significance based on the ECBI-PS (Kling et al. 2010; Enebrink et al. 2012; Morawska and Sanders 2006; Morawska et al. 2011, Joachim et al. 2010). Compared to the ECBI-IS, the NNTs based on the ECBI-PS were similar or slightly lower, with a median across treatments of four.

Figure 1: Effect size, reliable change, clinical change, and NNT at posttest (based on ECBI-IS)

		Cohen's <i>d</i>	Reliable change	Clinical change	Recovery	NNT
Kling et al. (2010)	Comet-S	*** 0.69	28 %	26 %	28 %	4 ^a
	Comet-SD	*** 0.54	15 %	10 %	13 %	8 ^a
Enebrink et al. (2012)	Comet-I	** 0.76	24 %	23 %	26 %	4 ^a
Sanders et al. (2000)		*** 0.98	27 %	25 %		4 ^b
		*** 0.65	19 %	18 %		6 ^b
		* 0.58	15 %	15 %		7 ^b
Morawska and Sanders (2006)	TripleP-SD+	*** 0.57	25 %	18 %		5 ^b
	TripleP-SD	* 0.52	2 %	7 %		48 ^b
Morawska et al. (2011)	TripleP-SD+	** 0.59	24 %	34 %		4 ^b
Joachim et al. (2010)	TripleP-SD+	* 0.57	11 %	31 %		9 ^b
Nixon et al. (2003)	PCIT-S	** 0.81	5 %	38 %	34 %	3 ^a
	PCIT-ABB	*** 0.21	-2 %	18 %	15 %	7 ^a
Axberg and Broberg (2012)	IY	*** 1.18	38 %	25 %	26 %	4 ^a
<i>Median</i>		0.59	19 %	23 %	26 %	5

Notes:

All results are between-group effects (i.e., the effects or proportions of the waitlist control groups are subtracted from those for the treatment groups). The significance levels of the effect sizes (**p* < .05, ***p* < .01, ****p* < .001) were retrieved from the original articles. NNT = Numbers Needed to Treat.

^a The NNT is based on the proportion of participants who recovered according to the JT method.

^b The NNT is based on whichever of the reliable or clinical change proportions was smallest.

3. Discussion

Part I investigated the clinical significance of behavioral parent training for children in two studies previously published by the authors. In Kling et al. (2010), about one fourth of the participants recovered in the practitioner-assisted version of the program (Comet-S), while only about one participant in eight made a recovery in the self-directed version of the program (Comet-SD). The relative advantage of Comet-S over Comet-SD was preserved at follow-up. This result is worth noting, considering that the

advantage of Comet-S was less apparent in the original article where results were reported in terms of statistical and practical significance (Kling et al. 2010). One interpretation of this result is that practitioners played an important role in helping clients making an actual recovery. While many participants in the self-directed version did improve, as shown by the effect size, most of the changes at the individual level were too small to be clinically significant.

In Enebrink et al. (2012) the within-group recovery rate in the treatment group was considerably higher than in Kling et al. (2010). However, after taking the control group into account, the recovery rates were similar to Comet-S. It is still striking that an internet-based version of the program with minimal practitioner assistance (Comet-I) was as effective as Comet-S in terms of clinical significance. One possible bias may be that the samples in the two studies were quite different. For example, only one third of the participating parents in the Comet-S study were well educated, compared to two thirds in the Comet-I study. It is well known that the social characteristics of families that take part in behavioral parent training can impact the treatment effects (Reyno and McGrath 2006). It is also worth noting that almost one in every ten parents in the Comet-I condition deteriorated. Even if this number was non-significant compared to the waitlist, it may warrant further investigation.

Part II investigated the clinical significance of behavioral parent training for children with conduct problems by synthesizing results from published studies. The median NNT was five across all treatments, four for the treatments offering full practitioner support, and seven for treatments offering no or minimal practitioner support. Even in the most effective programs, the NNTs were not lower than three. This means that, at best, one third of children with conduct problems actually recover as a result of treatment with behavioral parent training, while the rest only improve to some degree, show no change, or even deteriorate. This result is important for at least two reasons. First, the between-group recovery rates presented in Figure 1 are considerably lower than several of the within-group rates reported in the original articles. We argue that the former rates are of greater interest to practitioners and clients, because they account for bias such as spontaneous recovery and instead reflect the proportion of participants who recover *as a result of* the treatment. Second, many practitioners and clients may not realize that programs that are characterized as evidence-based and have reported “large” effect sizes, will only “cure” a minority of the treated children. However, from a researcher or policymaker point of view, curing one out of three patients may be of tremendous importance. The meaning of treatment effects is con-

text-dependent and factors such as severity of the treated problems and cost-benefit analyses have to be considered (Campbell 2005). For example, the similar effect sizes between Comet-S and Comet-SD at follow-up in Kling et al. (2010) mask the fact that the proportion of children making an actual recovery was much larger in Comet-S. The higher cost of implementing Comet-S compared to Comet-SD would be returned many times if twice as many children fully recovered from conduct problems.

The fact that only eight studies were included in the second part of this report limited the possibility to draw conclusions about the effects of moderators. It made little sense, for example, to compare the effects of different programs or the effect of child age, when there were only one or two studies representing a certain program or age group. It was however less problematic to investigate the effects of different methods of delivery on clinical significance, because such comparisons could be made both within and across studies. Therefore, method of delivery is the only moderator that can be discussed in any depth. The effects of different methods of delivery within the same trial were investigated in four studies (Kling et al. 2010; Morawska et al. 2006; Nixon et al. 2003; Sanders et al. 2000). As in Kling et al. (2010), the analyses of the other studies also showed that the clinical significance was larger for treatments involving more practitioner support (Figure 1). In some studies, advantages for treatment conditions involving more practitioner support were apparent in terms of clinical significance, but not in terms of statistical significance or effect sizes. For example, in Morawska and Sanders (2006) the effect sizes were similar for the two compared treatments, but there was a large difference between NNTs. The NNTs of self-directed treatments in Enebrink et al. (2012), Morawska et al. (2006) and Morawska et al. (2011) were similar to the most effective of programs offering full practitioner support in Figure 1. However, as in Enebrink et al. (2012), the participants in the two studies by Morawska and colleagues were particularly well educated and had few social problems. The other studies in Figure 1 all recruited average or at-risk samples. In conclusion, practitioner support seems to have a greater impact on clinical significance than on statistical significance and effect sizes. This conclusion at least holds based on comparisons of dif-

ferent levels of practitioner support within studies, which compared to comparisons across studies are less influenced by possible confounding variables such as characteristics of the study sample.

Two results that were found in the synthesis (Figure 1) especially warrant discussion. In Morawska et al. (2006), very few participants made a reliable and clinical change in the Triple-P-SD condition compared to the Triple-P-SD+, despite the fact that the effect sizes were similar in the two treatment groups. This means that many of the participants in Triple-P-SD must have improved, but not by enough to satisfy the criteria for reliable and clinical change. In Nixon et al. (2003), an unusually large proportion of participants in the waitlist control group made a reliable change (50 percent). The between-group proportions of reliable change thus turned out to be very small or even negative. Still, the NNTs were quite small for the treatment groups in that study. The reason for this was that they were based on recovery rates. Only 11 percent of the participants actually recovered in the control group, while 45 percent recovered in PCIT-S and 26 percent recovered in PCIT-ABB.

Several limitations of this report are recognized. First of all, the small number of studies included in the synthesis limited the possibility for more detailed analyses and generalization of the results. However, more liberal inclusion criteria would have made comparison across studies impossible (Marrs-Garcia 2010). It is therefore imperative that future studies of behavioral parent training include standardized analysis of clinical significance, preferably using the JT method (e.g., Bauer et al. 2004; McGlinchey et al. 2002). Another possibility would be to reanalyze the original data from a larger number of published outcome studies of behavioral parent training that lack reports of clinical significance.

A second limitation is that the analyses of clinical significance in this article were based on only one outcome measure. In Kling et al. (2010) and Sanders et al. (2000), a structured telephone interview measuring child conduct problems was also used as a basis for analyses of clinical significance (Parent Daily Report or PDR; Chamberlain and Reid 1987). Due to lack of proper normative data, the

PDR could not be used to analyze clinical change. The reliable change proportions were, however, computed. Based on the PDR, the reliable change proportions were 6 percent for Comet-SD and 5 percent for Triple-P-SD, as compared to 15 percent in both programs when the analyses were based on the ECBI-IS. There are a number of possible explanations for this difference, which also apply to differences in reliable change across measures in general. It could reflect general characteristics of the measures, such as sensitivity to change, reactivity, and different forms of reliability and validity. Further, the parameters that are used to compute the reliable change index are obvious sources of variation. For example, the internal consistency was .92 for the ECBI-IS as compared to .79 for the PDR in Kling et al. (2010), which strongly impacts the resulting thresholds for reliable change. To conclude, the choice of measures will often have an impact on results in analyses of clinical significance. Instead of relying on single measures, a compound of measures that target the construct of interest should ideally be used in analyses of clinical significance (Ogles et al. 2001). This is, however, often impossible due to lack of published norms for many outcome measures. Furthermore, because clinical significance refers to meaningful changes in the everyday life of clients (Kazdin 1999), it would also often be necessary to base analyses of clinical significance on several outcome domains (e.g., quality of life), besides the treated symptoms (e.g., child conduct problems). It is however by no means certain that a narrow focus on measuring symptoms will overestimate the clinical significance in terms of everyday functioning, as there are examples of the opposite (Karpenko et al. 2009).

A third limitation pertains to the application of the JT method in the six studies located by database search (Axberg and Broberg 2012; Joachim et al. 2010; Morawska et al. 2006; Morawska et al. 2011; Nixon et al. 2003; Sanders et al. 2000). First, none of the Triple-P-studies combined the criteria of reliable and clinical change. The discrepant results found in the analyses of Nixon et al. (2003) illustrate that it can be quite misleading to report reliable and clinical change separately, as opposed to combining the two criteria according to the JT method. This also justifies the exclusion of ten studies reporting only reliable or only clinical change in the process of finding eligible studies for the synthesis.

Second, none of the studies that were included from the database search reported which reliability coefficient were used to compute reliable change (e.g., internal consistency or test-retest). This limits the transparency and accuracy of the comparison of results across studies. Third, none of the six studies seem to have applied the recommended cutoff for clinical change, which is the weighted midpoint between the functional and dysfunctional populations (Jacobson and Truax 1991). The theory underlying the recommended cutoff point is that each study in essence investigates a unique population, whose cutoff in relation to the normative population also will be unique. It is generally difficult to compare results across studies with different populations, and using a cutoff point that partly is based on the study sample is probably more accurate than imposing an absolute cutoff (Wise 2004). With an absolute cutoff point, there is a risk that a substantial number of participants happen to score just above (or just below) the cutoff at pretest, which will result in misleading proportions for clinical change. This may have been the case in Joachim et al. (2010), where the pretest mean was close to the selected cutoff point (Table 3). Fourth, the potential for clinical change was limited in several studies because a substantial number of participants already scored below the cutoff at pretest. This was not the case in Enebrink et al. (2012) and Axberg and Broberg (2012), where only 5–10 percent of participants were below the cutoff at pretest, and probably not in Sanders et al. (2000), which used an elevated ECBI score as inclusion criterion. It was more of a problem in Kling et al. (2010) and Morawska et al. (2011) with 21 percent and 24 percent respectively scoring below the cutoff at pretest. In Joachim et al. (2010), 50 percent of the participants scored below the cutoff at pretest and in Morawska and Sanders (2006) the proportion was as large as 62 percent. Sometimes analyses of subsamples of participants

who score above the cutoff at pretest are used in such cases, but in randomized trials that strategy can result in selection bias. Instead, if a study intends to investigate treatment effects for a defined population, proper screening should be used to ensure that included participants actually belong to the dysfunctional/clinical population. Alternatively, several cutoff points could be used to represent different levels of severity of a given problem or condition (Ogles et al. 2001). Such a procedure could also be warranted, given that some authors suggest that the JT method sometimes may be too conservative (Tingey et al. 1996).

4. Conclusion

The results of this synthesis shows that the effects of behavioral parent training in the treatment of child conduct problems are generally clinically significant, but maybe to a lesser degree than would have been expected. The median recovery rate across studies showed that only one out of five children recovered. It was also evident that results in terms of clinical significance may lead to different conclusions than where conclusions are based solely on results in terms of statistical significance and effect sizes. The effects of practitioner support were considerably stronger in terms of clinical significance than in terms of effect sizes. The results further support the importance of including reports of clinical significance in outcome studies, which has called for by influential scholars and journals (Ogles et al. 2001; Campbell 2005). The fact that only eight studies were found to be eligible for inclusion in this synthesis points to the need for future research to adopt similar standards for the analysis of clinical significance, such as the JT method. A study that states that one out of five patients recover, rather than saying that the effect was $d = .59$, would probably be more effective in bridging the gap between science and practice.

References

- Axberg, Ulf, and Anders G. Broberg. 2012. Evaluation of the “Incredible Years” in Sweden: The Transferability of an American Parent-Training Program to Sweden. *Scandinavian Journal of Psychology* 53:224–32.
- Axberg, Ulf, Jan Johansson Hanse, and Anders G. Broberg. 2008. Parents’ Description of Conduct Problems in their Children – A Test of the Eyberg Child Behavior Inventory (ECBI) in a Swedish Sample Aged 3–10. *Scandinavian Journal of Psychology* 49:497–505.
- Bauer, Stephanie, Michael J. Lambert, and Steven L. Nielsen. 2004. Clinical Significance Methods: A Comparison of Statistical Techniques. *Journal of Personality Assessment* 82:60–70.
- Campbell, Todd C. 2005. An Introduction to Clinical Significance: An Alternative Index of Intervention Effects for Group Experimental Designs. *Journal of Early Intervention* 27:210–27.
- Chamberlain, Patricia., and John B. Reid. 1987. Parent Observation and Report of Child Symptoms. *Behavioral Assessment* 9:97–109.
- Chambless, Dianne L., and Steven D. Hollon. 1998. Defining Empirically Supported Therapies. *Journal of Consulting and Clinical Psychology* 66:7–18.
- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2d ed. Hillsdale, NJ: Erlbaum.
- Dretzke, Janine, Clare Davenport, Emma Frew, Jane Barlow, Sarah Stewart-Brown, Sue Bayliss, Rod S. Taylor, Josie Sandercock, and Chris Hyde. 2009. The Clinical Effectiveness of Different Parenting Programmes for Children with Conduct Problems: A Systematic Review of Randomized Controlled Trials. *Child and Adolescent Psychiatry and Mental Health* 3:7.
- Enebrink, Pia, Jens Högstrom, Martin Forster, and Ata Ghaderi. 2012. Internet-Based Parent Management Training: A Randomized Controlled Study. *Behaviour Research and Therapy* 50: 240–49.
- Evans, Chris, Frank Margison, and Michael Barkham. 1998. The Contribution of Reliable and Clinically Significant Change Methods to Evidence-Based Mental Health. *Evidence-Based Mental Health* 1:70–72.
- Eyberg, Sheila M. 1988. Parent–Child Interaction Therapy: Integration of Traditional and Behavioral Concerns. *Child and Family Behavior Therapy* 10:33–46.
- Eyberg, Sheila M., Melanie M. Nelson, and Stephen R. Boggs. 2008. Evidence-Based Psychosocial Treatments for Children and Adolescents with Disruptive Behavior. *Journal of Clinical Child and Adolescent Psychology* 37:215–37.
- Eyberg, Sheila M., and Donna Pincus. 1999. *Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory – Revised, Professional Manual*. Odesa, FL: Psychological Assessment Resources.
- Furlong, Mairead., Sinead McGilloway, Tracey Bywater, Judy Hutchings, Susan M. Smith, and Michael Donnelly. 2012. Behavioural and Cognitive-Behavioural Group-Based Parenting Programmes for Early-Onset Conduct Problems in Children Aged 3 to 12 Years. *Cochrane Database of Systematic Reviews* 2 CD008225. doi:10.1002/14651858.CD008225.pub2
- Jacobson, Neil S., and Paula Truax. 1991. Clinical Significance: A Statistical Approach to Defining Meaningful Change in Psychotherapy Research. *Journal of Consulting and Clinical Psychology* 59:12–19.
- Joachim, Sabine, Matthew R. Sanders, and Karen M. T. Turner. 2010. Reducing Preschoolers’ Disruptive Behavior in Public with a Brief Parent Discussion Group. *Child Psychiatry and Human Development* 41:47–60.
- Karpenko, Veronika, Julie S. Owens, Nicole M. Evangelista, and Cassandra Dodds. 2009. Clinically Significant Symptom Change in Children with Attention-Deficit/Hyperactivity Disorder: Does It Correspond with Reliable Improvement in Functioning? *Journal of Clinical Psychology* 65:76–93.
- Kazdin, Alan E. 1999. The Meanings and Measurement of Clinical Significance. *Journal of Consulting and Clinical Psychology* 67:332–39.
- Kazdin, Alan E. 2008. Evidence-Based Treatments and Delivery of Psychological Services: Shifting our Emphases to Increase Impact. *Psychological Services* 5:201–15.
- Kendall, Philip C., Abbe Marrs-Garcia, Sanjay R. Nath, and Radley C. Sheldrick. 1999. Normative Comparisons for the Evaluation of Clinical Significance. *Journal of Consulting and Clinical Psychology* 67:285–99.
- Kling, Åsa, Martin Forster, Knut Sundell, and Lennart Melin. 2010. A Randomized Controlled Effectiveness Trial of Parent Management Training with Various Degrees of Therapist Support. *Behaviour Therapy* 41:530–42.
- La Greca, Annette M. (2005). Editorial. *Journal of Consulting and Clinical Psychology*, 73:3–5.
- Marrs-Garcia, Abbe. 2010. What Does “Work” Mean? Reopening the Debate About Clinical Significance. *Clinical Psychology Science and Practice* 17:48–51.
- McGlinchey, Joseph B., David C. Atkins, and Neil S. Jacobson. 2002. Clinical Significance Methods: Which One to Use and How Useful Are They? *Behavior Therapy* 33:529–50.
- Morawska, Alina, and Matthew R. Sanders. 2006. Self-Administered Behavioural Family Intervention for Parents of Toddlers: Effectiveness and Dissemination. *Behaviour Research and Therapy* 44:1839–48.
- Morawska, Alina, Divna Haslam, Danielle Milne, Matthew R. Sanders. 2011. Evaluation of a Brief Parenting Discussion Group for Parents of Young Children. *Journal of Developmental and Behavioral Pediatrics* 32:136–45.
- Nixon, Reginald D. V., Lynne Sweeney, Deborah B. Erickson, and Stephen W. Touyz. 2003. Parent–Child Interaction Therapy: A Comparison of Standard and Abbreviated Treatments for Oppositional Defiant Preschoolers. *Journal of Consulting and Clinical Psychology* 71:251–60.
- Ogles, Benjamin M., Kirk M. Lunnen, and Kyle Bonesteel. 2001. Clinical Significance: History, Application, and Current Practice. *Clinical Psychology Review* 21:421–46.
- Plake, Barbara S., James C. Impara, and Robert A. Spies. 2003. *The Fifteenth Mental Measurements Yearbook*. Lincoln, NE: Buros Institute of Mental Measurements.
- Reid, M. J., Maryam Webster-Stratton, and Mary Hammond. 2007. Enhancing a Classroom Social Competence and Problem-Solving Curriculum by Offering Parent Training to Families of Moderate- to High-Risk Elementary School Children. *Journal of Clinical Child and Adolescent Psychology* 36:605–20.
- Reyno, Sandra M., and Patrick J. McGrath. 2006. Predictors of Parent Training Efficacy for Child Externalizing Behavior Problems – a Meta-Analytic Review. *Journal of Child Psychology and Psychiatry* 47:99–111.
- Sanders, Matthew R. 1999. Triple P-Positive Parenting Program: Towards an Empirically Validated Multilevel Parenting and Family Support Strategy for the Prevention of Behavior and Emotional Problems in Children. *Clinical Child and Family Psychology Review* 2:71–90.
- Sanders, Matthew R., Carol Markie-Dadds, Lucy A. Tully, and William Bor. 2000. The Triple P – Positive Parenting Program: A Comparison of Enhanced, Standard, and Self-Directed Behavioral Family Intervention for Parents of Children with Early Onset Conduct Problems. *Journal of Consulting and Clinical Psychology* 68:624–40.
- Silverman, Wendy K., and Stephen P. Hinshaw. 2008. The Second Special Issue on Evidence-Based Psychosocial Treatments for Children and Adolescents: A 10-Year Update. *Journal of Clinical Child and Adolescent Psychology* 37:1–7.
- Tingey, Richard C., Michael L. Lambert, Gary M. Burlingame, and Nathan B. Hansen. 1996. Assessing Clinical Significance: Proposed Extensions to the Method. *Psychotherapy Research* 6:109–23.
- Webster-Stratton, Carolyn. 2000. The Incredible Years Training Series. *OJJDP Bulletin*, June, 1–24.
- Wise, Edward A. 2004. Methods for Analyzing Psychotherapy Outcomes: A Review of Clinical Significance, Reliable Change and Recommendations for Future Directions. *Journal of Personality Assessment* 82:50–59.

Martin Forster
mrmartinforster@gmail.com

Åsa Kling
asa.kling@psyk.uu.se

Knut Sundell
knut.sundell@socialstyrelsen.se

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool

Gil Noam, Program in Education, Afterschool and Resiliency (PEAR), Harvard Medical School and McLean Hospital, Belmont, United States

Tina Malti, Department of Psychology, University of Toronto, Canada

Martin Guhn, Human Early Learning Partnership, University of British Columbia, Vancouver, Canada

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

► **From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool** Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool

Gil Noam, Program in Education, Afterschool and Resiliency (PEAR), Harvard Medical School and McLean Hospital, Belmont, United States

Tina Malti, Department of Psychology, University of Toronto, Canada

Martin Guhn, Human Early Learning Partnership, University of British Columbia, Vancouver, Canada

A description and test of the Holistic Student Assessment Tool (HSA), an assessment tool to measure children's and adolescents' resiliencies in relation to externalizing and internalizing problem behaviors. The HSA is based on the authors' research-based clinical-developmental Clover Leaf Model of resilience and psychopathology, and is one of the first attempts at closing the gap between risk and resilience approaches in developmental assessment. The HSA was tested in a cross-sectional sample of 423 children and adolescents.

The results lend support to the HSA as a valid measure of children's and adolescents' resiliencies. Furthermore, the resilience scales mostly exhibited the theoretically expected convergent and divergent relationships with the psychopathology scales. In addition, we show how the resilience scales predict adolescents' externalizing and internalizing symptoms. We contend that evidence-based intervention to address youth aggression needs to be based on sound developmental assessment.

It has been estimated that more than 20 percent of U.S. children and youth aged nine to seventeen suffer from significant behavioral and emotional problems and are at risk for school failure (e.g., Costello, Egger, and Angold 2005). Furthermore, aggression, violence, and bullying can seriously impact children's and adolescents' mental health (Farrington 2005). These types of externalizing behavior also interfere with children's ability to develop resilience (Masten and Wright 2009). The early identification of aggressive behavior and the precursors of psychopathology is a priority as it can reduce the individual burden and societal costs of related problems throughout life as well as promote social-emotional development and well-being (Jones et al. 2002; Powell, Lochman, and Boxmeyer 2007).

But why is it important to think about developmental theory and assessment in the prevention of bullying, aggression, and violence among children and youth? Researchers have argued that any evidence-based approach to violence and bullying prevention and intervention needs to be based in a sound developmental theory that identifies important risk- and resilience factors that contribute to, or

impede problem behavior, such as violence and antisocial conduct (Beelmann 2011; Eisner and Malti 2012; Lösel and Farrington 2012; Rutter 2012). In addition to sound developmental models that emphasize a strengths-oriented approach that relies on resiliencies and protective factors (Luthar 2006; Masten 2009, 2011), the use of early developmental screening tools is an important step to ensure that these risk and resilience factors are identified in practice. Assessment results, in turn, can help to inform intervention practice. For example, they can help in the decision-making process associated with the kind of services and the intensity of the intervention that a child may need. For example, children who are at-risk for, or already show elevated levels of aggression and antisocial conduct, may benefit from targeted interventions that utilize resilience factors in the treatment to reduce behavioral outcomes (Malti, Liu, and Noam 2010). Thus, early assessments that systematically integrate developmental research and risk and resilience factors are likely to facilitate the delivery of treatment methods that are sensitive to the developmental needs of the child (Liu, Malti and Noam 2008).

In line with this argument, the importance of developing school-based early assessment tools for identifying children's and adolescents' mental health problems, including bullying and antisocial conduct, has been underscored (see Malti and Noam 2009). Accordingly, several assessment tools for use in school and afterschool contexts have been developed. Developmental studies provide ample evidence for the role of resiliencies and social-emotional development in the prevention of children's problem behaviors, such as bullying and antisocial conduct (e.g., Lansford et al. 2006; Orobio de Castro et al. 2002).

Despite these findings and an increasing emphasis on social-emotional development and resiliencies in assessment and intervention research (see also WHO 2003, Guhn et al. 2012; Schonert-Reichl et al. 2012), the great majority of existing school-based assessments typically include only questions about risks and symptomatology. From both a developmental and clinical perspective, however, holistic measures that address both risk and resiliencies can be more effective in engaging students in high-quality in-school and out-of-school-time activities that fit their developmental strengths and clinical needs (Malti, Liu, and Noam 2009). In addition, symptom checklists yield clinical and sub-clinical results that far exceed the treatment capacity of schools and associated institutions. Thus, simply from a pragmatic point of view it is important to understand the balance between risk and protective factors and to evaluate the vulnerabilities, problems, strengths, and assets to develop appropriate referral systems.

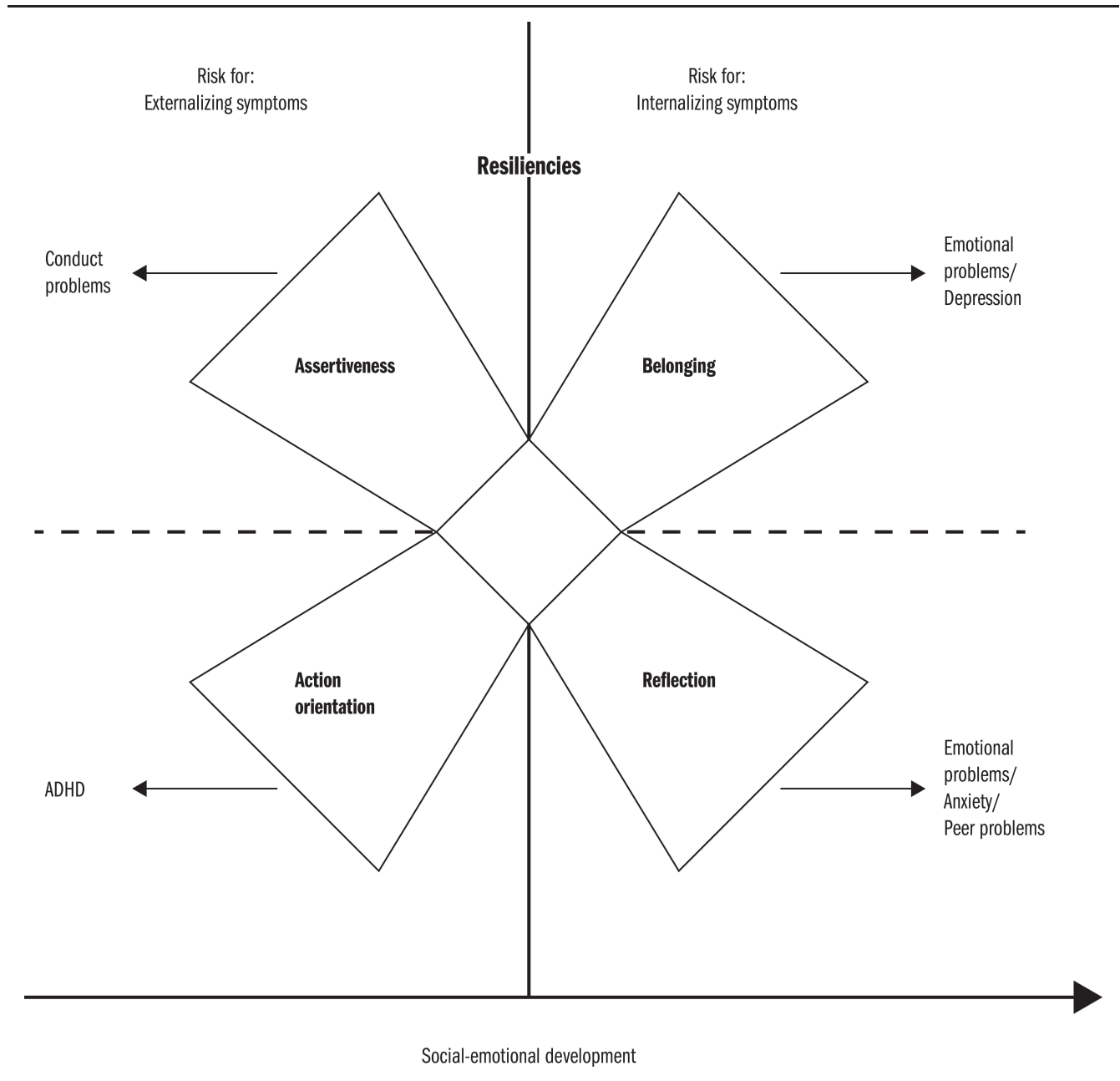
Here we present a new assessment tool, the Holistic Student Assessment (HSA), in which children and adolescents report their resiliencies and socio-emotional strengths. The

aim is to measure key dimensions of resilience and strength in order to complement existing school-based assessments of risk factors and psychopathology, including aggression and antisocial conduct. We sought to evaluate the psychometric properties of the HSA and test the theoretical assumptions between social-emotional skills, resiliencies, and externalizing and internalizing psychopathology.

1. Theoretical Background: The Clover Leaf Model

The theoretical model underlying the HSA tool – the *Clover Leaf Model* – is a research-based clinical-developmental model of resilience and psychopathology (Noam and Malti 2008; for a comparison with other developmental resilience models, see Noam, Malti, and Karcher, forthcoming). The model interconnects adolescent psychopathology with social-emotional development and resilience; problem behaviors emerge as developmental difficulties, and adaptation emerges from social-emotional development and resiliencies (Noam 1996). Hence, young peoples' socio-emotional development and resiliencies may help determine whether early signs of a problem will evolve into a clinically relevant disorder or resolve into healthy development. In our research and theory on developmental psychopathology, we have systematically linked social-emotional development to resiliency and to the risk of psychopathology (Noam 1999). In the Clover Leaf Model, development in adolescence is described as the leaves of a clover, with each leaf reflecting a particular form of social-emotional development (Noam and Malti, 2008): need for action, assertiveness, interpersonal sensitivity/belonging, and reflection (Figure 1). Each leaf represents particular resilience factors. However, each of also has its own risks, which represent behavioral and emotional problems (for a detailed description of the Clover Leaf Model, see Noam et al., forthcoming).

Figure 1: The Clover Leaf Model



It follows logically from this perspective that socio-emotional development is inevitably linked to specific risks and resiliencies. There are also different windows of risks and psychopathology in each developmental leaf (Noam, Chandler, and LaLonde 1995). For example, the assertive-

ness leaf includes the risk of aggressive behavior problems and is associated with high assertiveness (resilience) but low interpersonal sensitivity (risk; see Malti and Keller 2009). Thus, the Clover Leaf Model not only distinguishes the pathways for growth which may be used to advance

mental health, but its application may also reduce problem behavior and the risks inherent in the developmental process. The HSA empirically captures the resilience dimensions conceptualized in the Clover Leaf Model and, therefore, enables researchers to test the strengths associated with risk and behavioral problems. The HSA also includes additional dimensions of social-emotional development that have shown to be of significance for behavioral and emotional problems, such as empathy. However, here we focus on the resilience dimensions that directly capture the clover leaves because our interest is in relations between these resilience factors with aggression and antisocial behavior outcomes on the one hand, and internalizing symptoms on the other.

In summary, this study set out to investigate the psychometric properties of the Holistic Student Assessment (HSA). We tested the unidimensionality of the HSA scales, as well as the overall factor structure of the HSA. Unidimensionality of the individual scales was examined via factor analysis, in order to determine the ratio of the first to the second eigenvalue. In addition, the overall factor structure was examined via exploratory factor analysis as well as via exploratory bi-factor analysis (Jennings and Bentler 2011).

We also examined the convergent validity of the HSA by studying links between the resiliency scales and psychopathology. Previous research has shown associations between social-emotional development and psychopathology (Noam, Young, and Jilnina 2006). Based on this research, we examined the HSA in relation to externalizing (i.e., aggression, ADHD) and internalizing (i.e., emotional symptoms, peer relationship problems) symptomatology. In order to examine the extent to which the HSA resilience scales (Clover Leaf constructs) are jointly related to the scales of the Strength and Difficulties Questionnaire (SDQ; Goodman 1997), we conducted multiple regression analyses which contained all Clover Leaf scales as predictors and each of the SDQ scales as dependent variables. In addition to the beta coefficients of the multiple regressions, we report Pratt's measure of variable importance for each predictor variable, because this identifies whether specific predictor variables func-

tion as suppressor variables in the multiple regression context (Thomas, Hughes, and Zumbo 1998).

Our hypotheses were that children who had high resiliencies related to externalizing problems (action orientation and assertiveness) would be more likely to report externalizing problems if their resiliencies related to internalizing symptoms (interpersonal sensitivity/belonging and reflection) were low. In addition, children with a relatively balanced profile on all the four clover leaves – i.e., a combination of resiliencies – would report low levels of symptoms. In other words, we expected the beta coefficients of the predictor variables to be generally larger than their respective zero-order correlations (i.e., represent suppression effects). This hypothesis is based on the theoretical assumption that high action orientation, assertiveness, interpersonal sensitivity, and reflection are only associated with externalizing and internalizing problems respectively, to the extent that they are not balanced by competencies in the other three domains (see Noam, Malti, and Karcher, forthcoming).

2. Method

2.1. Participants

The sample comprised 423 children and adolescents (grades 4 to 9; $M = 12.7$ years, $SD = 1.1$; 52 percent girls) attending ten public elementary, middle, and junior high schools in Boston, Massachusetts. We collected data from schools that had high proportions of at-risk youth and low-income backgrounds according to public school district records. The student populations reflected the ethnic diversity of the Boston public school system.

2.2. Measures

Holistic Student Assessment (HSA). The HSA is a newly-developed eighty-four-item measure designed to assess the resiliencies and social-emotional development of children and adolescents aged 10 to 18 years (grade 4 to grade 12). It is based on our previous research and on the Resilience Inventory developed by Noam and Goldstein (1998) and Song (2003). All HSA items have a four-point Likert response format (not at all = 0, sometimes = 1, often = 2, almost always = 3).

The HSA contains nine subscales. For the present study, we used the four scales that represent the four dimensions of the Clover Leaf Model directly: action orientation (five items; e.g., “I like being active,” Cronbach’s $\alpha = .72$); assertiveness (six items; e.g., “I defend myself against unfair rules,” Cronbach’s $\alpha = .69$); interpersonal sensitivity/belonging (eight items; e.g., “I try to understand how other people think and feel about things,” Cronbach’s $\alpha = .81$); and reflection (nine items; e.g., “I think about the problems of the world,” Cronbach’s $\alpha = .86$).

The HSA is filled out by the students in a group setting with careful adult supervision, and its administration takes approximately 20 minutes.

Strength and Difficulties Questionnaire (SDQ). Children evaluated their social behavior on a three-point Likert scale using the twenty-five items from the SDQ (Goodman 1997). The SDQ contains five subscales, each with five items: hyperactivity/inattention; conduct problems; peer relationship problems; emotional symptoms; and prosocial behaviour. It is a validated and widely used measure of psychopathology and prosocial behavior (e.g., van Roy, Veenstra, and Clench-Aas 2008). In our sample, Cronbach’s α was .68 for hyperactivity/inattention (ADHD), .53 for conduct problems, .56 for peer relationship problems, and .70 for emotional symptoms.

2.3. Procedure

Participation in the survey was voluntary.

Data analysis procedure. In step 1, separate factor analyses were conducted for all individual Clover Leaf scales to examine their unidimensionality (i.e., according to the ratio of first to second eigenvalue). In step 2, the overall factor structure of the HSA items was examined via ex-

ploratory factor analysis and exploratory bi-factor analysis (Jennings and Bentler 2011). Exploratory bi-factor analysis allows examination of the nature of second-order factors, in cases in which factors are correlated, and most or all items load one general factor – similar to how one may find a g-factor for intelligence, that accounts for the correlation among subscales of intelligence.

To test convergent and divergent validity, we explored the relationship between the Clover Leaf and SDQ subscales using correlation matrices, Fisher’s Z-test to compare pairs of correlations, and multiple regressions. In the multiple regression analyses, gender was included as a covariate, to control for gender differences. Finally, taking into account the zero-order correlations and the beta coefficients from the multiple regression analyses, we calculated Pratt’s measure of variable importance for each predictor variable (Thomas et al. 1998). Pratt’s measure helps to interpret the importance of predictor variables in the presence of suppression effects, as well as multicollinearity. In all correlational analyses, we used the continuous mean score across all scale items. Analyses were conducted in SPSS (version 17).

3. Results

3.1. Unidimensionality of the Clover Leaf Scales

The factor analyses and examinations of the scree plots indicate essential unidimensionality for all the Clover scales. For action orientation, interpersonal sensitivity/belonging, and reflection, only one eigenvalue was greater than 1, and all items had loadings of .4 or higher (ranging from .45 to .73). For assertiveness, the first eigenvalue was 3.5 and the second eigenvalue was 1.0; hence, the ratio of first to second eigenvalue indicated essential unidimensionality, as did the scree plot and the item loadings (ranging from .48 to .67).

3.2. Overall Factor Structure

Table 1: Higher order factor analysis

Items	g	Factor 1	Factor 2	Factor 3	h2
Reflection 1 (65)	.69	.28			.56
Reflection 2 (48)	.57	.25			.39
Reflection 3 (79)	.65	.25			.49
Reflection 4 (58)	.70	.31			.60
Reflection 5 (80)	.62	.23			.44
Reflection 6 (81)	.60				.43
Reflection 7 (76)	.52				.33
Reflection 8 (77)	.47				.29
Interpersonal 1 (5)	.36		.59		.48
Interpersonal 2 (22)	.42		.58		.51
Interpersonal 3 (35)	.45		.61		.57
Interpersonal 4 (4)	.28		.36		.21
Interpersonal 5 (8)	.40		.41		.34
Interpersonal 6 (23)	.45		.23		.27
Interpersonal 7 (34)	.47		.33		.33
Interpersonal 8 (38)	.52				.33
Action 1 (89)	.29			.61	.46
Action 2 (90)				.49	.25
Action 3 (91)				.67	.50
Action 4 (92)	.27			.37	.24
Action 5 (93)	.35			.51	.39
Assertiveness 1 (3)	.38		.24		.21
Assertiveness 2 (63)	.32			.25	.18
Assertiveness 3 (68)	.53			.32	.41
Assertiveness 4 (46)	.40			.23	.23
Assertiveness 5 (56)	.22				.09
Assertiveness 6 (61)	.50				.33

An exploratory factor analysis across all HSA items identified one dominant first factor (eigenvalue of 8.1), explaining 29 percent of the total variance. The second largest eigenvalue was 2.3. The ratio of first to second eigenvalue is thus larger than 3:1, indicating the presence of an overall “resiliency” factor. We proceeded by conducting an exploratory bi-factor analysis (Jennings and Bentler 2011). The results (see Table 1) suggest the presence of three secondary factors, in addition to one primary factor. The fit for the higher-order factor model was good (RMSEA=0.063; Chi-square=769; df=297), and significantly better ($p < .001$) than for the one-factor solution (RMSEA=0.091; Chi-square=1728; df=350). Five out of eight items from the Clover Leaf reflection scale loaded primarily on the (second-order) factor 1, and the remaining three reflection items loaded on the general factor (g), but not on any of the three second-order factors. Seven of the eight items from the interpersonal sensitivity scale loaded primarily on factor 2, and the remaining interpersonal sensitivity item loaded only on the general factor. All five items from the action scale loaded highly on factor 3. In addition, three of the eight items from the assertiveness scale loaded also on factor 3. One assertiveness item had its highest loading on factor 2, and the remaining assertiveness items did not load significantly on any second-order factor. In sum, three of the four clover leaves were relatively closely reproduced by the three second-order factors – the exception being the assertiveness scale. The fact that three assertiveness items loaded on the same second-order factor as all the action orientation items indicates that the two “externalizing scales” are relatively closely associated with each other. The implications of these findings will be discussed below.

3.3. Convergent and Divergent Validity

Table 2 shows the Pearson zero-order correlations among the Clover Leaf scales and the SDQ scales. To test the convergent and discriminant validity, we compared the correlations of the Clover Leaf constructs with the SDQ scales. We used Fisher’s Z-test to test the statistical significance between two correlation coefficients from one sample (using an online tool described by Uitenbroek [1997]; <http://www.quantitativeskills.com/sisa/statistics/correl.htm>).

Table 2: Pearson correlations (p-values) between Clover Leaf and SDQ scales

	1	2	3	4	5	6	7
1 ADHD							
2 Conduct problems	0.52 (0.00)						
3 Peer problems	0.18 (0.00)	0.29 (0.00)					
4 Emotional problems	0.35 (0.00)	0.26 (0.00)	0.39 (0.00)				
5 Action orientation	0.07 (0.15)	0.04 (0.41)	-0.23 (0.00)	-0.10 (0.05)			
6 Assertiveness	-0.07 (0.14)	0.04 (0.47)	-0.17 (0.00)	-0.07 (0.18)	0.41 (0.00)		
7 Interpersonal sensitivity/belonging	-0.31 (0.00)	-0.27 (0.00)	-0.10 (0.04)	0.08 (0.12)	0.25 (0.00)	0.44 (0.00)	
8 Reflection	-0.24 (0.00)	-0.11 (0.03)	-0.05 (0.27)	0.00 (0.95)	0.37 (0.00)	0.60 (0.00)	0.63 (0.00)

The primary hypothesis was that the Clover Leaf action orientation and assertiveness scale items would be significantly more protective against (i.e., negatively correlated with) internalizing problems as indicated by the SDQ scale (emotion symptoms) than against the corresponding externalizing items on the SDQ scales (ADHD and conduct problems). Interpersonal sensitivity/belonging and reflection, on the other hand, were expected to be significantly more protective against the SDQ scales indicative of externalizing

problems (ADHD and conduct problems) than against the corresponding internalizing items on the SDQ scales (peer problems and emotion symptoms). As can be seen in Table 2, action orientation correlated negatively with peer problems and emotional problems. Similarly, assertiveness correlated negatively with peer problems. Moreover, interpersonal sensitivity/belonging correlated negatively with ADHD, conduct problems, and peer problems. Reflection was negatively associated with ADHD and conduct problems.

Table 3: Gender differences on predictor and outcome variables

Scale (score range)	Girls, mean (SD)	Boys, mean (SD)	t-value (p-value)
ADHD (0–2)	0.8 (0.4)	0.8 (0.4)	0.1 (.91)
Conduct problems (0–2)	0.5 (0.4)	0.5 (0.3)	1.1 (.26)
Peer problems (0–2)	0.5 (0.4)	0.5 (0.4)	1.5 (.14)
Emotional symptoms (0–2)	0.8 (0.5)	0.6 (0.4)	-3.8 (.00)
Action orientation (0–3)	2.0 (0.6)	2.3 (0.6)	3.6 (.00)
Assertiveness (0–3)	1.9 (0.6)	1.8 (0.6)	-0.2 (.84)
Interpersonal sensitivity (0–3)	2.0 (0.6)	1.9 (0.6)	-2.4 (.02)
Reflection (0–3)	1.8 (0.7)	1.8 (0.6)	0.3 (.78)

Gender differences are shown in Table 3. As expected, girls reported significantly higher levels of emotional symptoms and interpersonal sensitivity, and boys reported higher levels of action orientation. All other scales showed no significant gender difference. Given

that gender differences were observed on two of the Clover Leaf scales – which are used as the predictors in the multiple regression analyses – and on one SDQ scale, gender was included as a covariate in all regression analyses.

3.4. Multiple Regressions: Predicting Psychopathology by Resiliencies

Table 4: Standardized multiple regression coefficients (p-values) for Clover Leaf scales predicting SDQ scales

Dependent variable	Externalizing symptoms				Internalizing symptoms			
	ADHD		Conduct problems		Peer problems		Emotional problems	
Predictor variable	β (p)	Pratt score	β (p)	Pratt score	β (p)	Pratt score	β (p)	Pratt score
Gender	0.06 (0.21)	0%	0.00 (0.98)	0%	-0.11 (0.03)	10%	0.15 (0.00)	56%
Action orientation	0.18 (0.00)	9%	0.05 (0.36)	2%	-0.24 (0.00)	68%	-0.07 (0.20)	14%
Assertiveness	0.08 (0.18)	-4%	0.19 (0.00)	8%	-0.12 (0.05)	25%	-0.10 (0.12)	14%
Interpersonal sensitivity	-0.29 (0.00)	65%	-0.34 (0.00)	91%	-0.06 (0.30)	7%	0.10 (0.10)	16%
Reflection	-0.18 (0.01)	31%	-0.02 (0.73)	2%	0.15 (0.04)	-9%	0.02 (0.73)	0%
R^2	0.14 (0.00)		0.10 (0.00)		0.08 (0.00)		0.05 (0.01)	

Next, we predicted externalizing and internalizing symptoms by the Clover Leaf resiliency scales. Table 4 shows the results for the multiple regression analyses. The regression model’s overall predictive power for hyperactivity ($R^2 = .14$) was equivalent to a medium/large effect size. For conduct problems ($R^2 = .10$) and peer problems ($R^2 = .08$), the R^2 was equivalent to a medium effect size, and for emotional symptoms ($R^2 = .05$), it was equivalent to a small effect size. Overall, the coefficients were similar in pattern to the zero-order correlations, but the findings also indicated the presence of several suppression effects.

3.4.1. Externalizing Problems: ADHD and Conduct Problems

Interestingly, the zero-order correlations between action orientation and ADHD, as well as between assertiveness and conduct problems were statistically not different from zero. However, in the multiple regressions, the beta coefficients for action orientation in relation to hyperactivity ($r = .18, p < .001$) and assertiveness in relation to conduct

problems ($r = .19, p < .001$) were larger, and statistically significant. In other words, action orientation only predicted ADHD, and assertiveness only predicted conduct problems, in the presence of the other Clover Leaf predictor variables. These findings suggest that interpersonal sensitivity and reflection are protective factors with regard to ADHD and conduct problems; this was indicated by the zero-order correlations and the negative (and statistically significant) beta coefficients. In fact, interpersonal sensitivity was the most important variable with regard to ADHD (Pratt = 65 percent) and conduct problems (Pratt = 91 percent). Also, the Pratt indices, zero-order correlations, and beta coefficients indicate that action orientation contributes to the regression (with ADHD as the outcome variable) as a suppressor: The beta coefficient for action orientation (.18) is of equal size to the beta coefficient for reflection (-.18), but its Pratt score is less than one third (9 percent versus 31 percent). The same pattern is found for assertiveness in the regression with conduct problems as the outcome variable (cf. Thomas et al. 1998).

3.4.2. Internalizing Problems: Peer problems and Emotional Symptoms

In relation to peer problems and emotional symptoms, the zero-order correlations for action orientation and assertiveness were significantly negative, and of almost identical size as the corresponding beta coefficients in the regression analyses. Action orientation was the most important variable with regard to peer problems, according to its Pratt measure (Pratt = 65 percent). In addition, the multiple regression analysis with peer problems as the dependent variable indicated a suppression effect with respect to the reflection scale. The zero-order correlation between reflection and peer problems was not significant; however, in the multiple regression, reflection was significantly associated with peer problems ($r = .15$, $p < .04$).

Again, the beta coefficients and Pratt scores indicate that a suppression effect is present: In relation to peer problems reflection has a larger beta coefficient than assertiveness; however, the Pratt score is considerably smaller (9 percent versus 25 percent). In other words, reflection seems to predict peer problems to a larger than extent once action orientation, assertiveness, and interpersonal sensitivity are taken into account.

In relation to emotional symptoms, gender was the only significant predictor. The beta coefficients of the four Clover Leaf scales were all not significant. Among all four multiple regression analyses that were conducted, the one on emotional symptoms was the only one that did not show the pattern of a suppression effect.

4. Discussion

The Holistic Student Assessment (HSA) is a new assessment tool designed to measure children's and adolescents' resiliencies and social-emotional development. It complements existing assessments of risk and psychopathology, such as the Strength and Difficulties Questionnaire, in both school and out-of-school-time settings. Conceptually, the HSA is based on the Clover Leaf Model which combines psychopathology and risk with resiliencies and social-emotional development (Malti and Noam 2009). Despite a strong emphasis on resilience and protective factors in the literature (see Rutter 2012), early as-

sessments that integrate social-emotional development and resilience factors are still scarce. However, we argue here that their development is important because they can help tailor intervention strategies for the prevention of bullying, violence, and antisocial conduct.

Our results are the first to lend empirical support to the HSA as a valid measure of children's and adolescents' resiliencies. Factor analyses show that the unidimensionality of the Clover Leaf resilience scales was mostly plausible. Furthermore, the resilience scales mostly exhibited the theoretically expected convergent and divergent relationships. More specifically, action orientation was negatively associated with internalizing symptoms (peer problems and emotional symptoms), while assertiveness correlated negatively with internalizing symptoms (peer problems). In addition, as expected, interpersonal sensitivity/belonging correlated negatively with externalizing symptoms, but positively with internalizing symptoms (peer problems). Reflection was negatively related to externalizing symptoms. Taken together, these findings confirm the Clover Leaf Model's assumptions regarding the interrelations between resiliencies and different types of externalizing and internalizing psychopathology. The negative relationship between interpersonal sensitivity/belonging and externalizing symptoms is also consistent with previous research reporting a negative relationship between empathy and antisocial conduct (Hastings et al. 2000). In contrast, the negative association between assertiveness with internalizing symptoms is in line with related research on social skill deficits in children with depressive symptoms (Perren and Alsaker 2009).

Interestingly, none of the Clover Leaf resiliency constructs was consistently either a protective factor or a risk factor for all psychopathology scales; rather, each resilience construct was significantly correlated with at least one of the four psychopathology scales. This is in line with our hypothesis that the different resiliency scales are associated differentially with different externalizing and internalizing symptoms, suggesting that each resilience scale has a specific function as a measure of social-emotional development and in relation to risk for psychopathology.

In addition to the bivariate relationships between the Clover Leaf resiliency scales and the psychopathology scales, our regression analyses show that action orientation predict ADHD positively, whereas interpersonal sensitivity/belonging and reflection predict ADHD negatively. As theoretically expected, high assertiveness and low interpersonal sensitivity predicted conduct problems. Given the positive association between assertiveness and interpersonal sensitivity/belonging, this pattern indicates a suppression effect which may indicate that assertiveness alone is indeed not a resilience factor, but a risk factor for antisocial conduct. However, this is only the case if interpersonal sensitivity/belonging is missing. In contrast, being both assertive and interpersonally sensitive is a resilience factor and contribute to developmentally adaptive outcomes. This interpretation is in line with related research on the social and moral antecedents of bullying (e.g., Gasser and Keller 2009); children and adolescents with aggression and bullying behavior may not necessarily lack social skills, but may have deficiencies in the moral qualities of empathy and interpersonal sensitivity. The finding also points to the need to assess various resiliencies to fully understand individual risk and protective factors of psychopathology.

Regarding internalizing symptomatology, the results show that action orientation and assertiveness predict peer problems negatively, whereas reflection predicts them positively. In addition, emotional symptoms are positively predicted by interpersonal sensitivity/belonging. These findings are fully in line with the theoretical expectations and provide additional evidence for the notion of resiliencies as risk and protective factors for psychopathology. Specifically, the findings show that each psychopathology, such as aggressive behavior, is associated with a lack in specific resiliencies (e.g., belonging) and high levels in other resiliencies (e.g., assertiveness).

Thus, the HSA can help to detect specific resiliencies and risks in children and adolescents who are at risk for, or already show, elevated levels of aggression and antisocial conduct. Specifically, children with these problem behaviors show high levels on assertiveness but simultaneously low levels of interpersonal sensitivity. This

information can be used to prepare targeted intervention strategies for these children, for example using at-risk children's assertiveness (high resilience) to improve their low levels of interpersonal sensitivity and feelings of belonging and empathy, while they retain their assertiveness and reduce their aggression (see Malti and Noam 2009). Thus, this approach implies that it is important to target at-risk children or children with elevated problem behaviors, such as high levels of aggression, by looking at specific resiliencies (i.e., assertiveness) in order to tailor interventions and strengthen specific other resiliencies (interpersonal sensitivity/belonging). The Clover Leaf Model predicts that a balance between different resilience dimensions is most adaptive. From this perspective, it seems warranted to seek a balance between assertiveness and interpersonal sensitivity for children with elevated aggression levels. This differs from a perspective that focuses exclusively on the reduction of risk factors by emphasizing the child's strengths and using these strengths to overcome risks and vulnerabilities such as low levels of other resiliencies and related problem behaviors.

In summary, these findings provide support for the HSA as a psychometrically valid measure of resilience. However, all of the convergent and divergent relationships were based exclusively on self-report, thus a further examination of the psychometric properties of the HSA with self- and other-reports is warranted. In addition, Cronbach's alphas for the conduct disorder and peer problem scales were only moderate. However, these alphas compare to findings from other studies using the self-report version of the SDQ (e.g., Hawes and Dadds 2004). Furthermore, our data analysis approach was merely correlational and cross-sectional, and thus no causal implications can be drawn from the current research. Lastly, the resilience factor explained only a relatively small proportion of variance in predicting behavioral outcomes. Future research needs to take into account other well-known risk factors, such as contextual risk factors that contribute to problem behavior, in order to test the relative predictive power of the resilience scales when compared to these classical risk factors (Eisner and Malti 2012). However, previous research suggests that these and related resilience factors predict problem behavior longitudinally (Malti and Krettenauer 2012).

Nevertheless, the current results provide evidence that the HSA validly captures the four dimensions of the Clover Leaf Model. As such, the HSA has potential to help tailor interventions based on the developmental needs and resiliencies of adolescents at risk for, or already engaging in, ex-

ternalizing or internalizing psychopathology. Early screenings, such as the HSA, can help educators choose more effective strategies in order to reduce bullying and aggressive behavior through a three-tiered delivery system (i.e., promotion, prevention, intervention).

References

- Benson, Peter L., and Peter C. Scales. 2009. Positive Youth Development and the Prevention of Youth Aggression and Violence. *European Journal of Developmental Science* 3:218–34.
- Costello, E. Jane, Helen L. Egger, and Adrian Angold. 2005. Ten-Year Research Update Review: The Epidemiology of Child and Adolescent Psychiatric Disorders: I. Methods and Public Health Burden. *Journal of the American Academy of Child and Adolescent Psychiatry* 44:972–86.
- Eisner, Manuel, and Tina Malti. 2012. The Future of Research on Evidence-Based Developmental Violence Prevention in Europe – Introduction to the Focus Section. *International Journal of Conflict and Violence* 6 (2): ##–##.
- Farrington, David P. 2005. Childhood Origins of Antisocial Behaviour. *Clinical Psychology and Psychotherapy* 12:177–90.
- Gasser, Luciano, and Monika Keller. 2009. Are the Competent the Morally Good? Social-Cognitive Competencies and Moral Motivation of Children Involved in Bullying. *Social Development* 18:798–816.
- Goodman, Robert. 1997. The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology, Psychiatry, and Allied Disciplines* 38:581–86.
- Guhn, Martin, Kimberly A. Schonert-Reichl, Anne M. Gadermann, Dan Martriott, Lisa Pedrini, Shelley Hymel, and Clyde Hertzman. 2012. Well-Being in Middle Childhood: An Assets-Based Population-Level Research-to-Action Project. *Child Indicators Research* 5:393–418.
- Hastings, Paul D., Carolyn Zahn-Waxler, JoAnn Robinson, Barbara Usher, and Dana Bridges. 2000. The Development of Concern for Others in Children with Behavior Problems. *Developmental Psychology* 36:531–46.
- Hawes, David J., and Mark R. Dadds. 2004. Australian Data and Psychometric Properties of the Strengths and Difficulties Questionnaire. *Australian and New Zealand Journal of Psychiatry* 38 (8): 644–51.
- Jennrich, Robert I., and Peter M. Bentler. 2011. Exploratory Bi-Factor Analysis. *Psychometrika* 76:537–49.
- Jones, Damon, Kenneth A. Dodge, E. Michael Foster, Robert Nix, and the Conduct Problems Prevention Research Group. 2002. Early Identification of Children at Risk for Costly Mental Health Service Use. *Prevention Science* 3:247–56.
- Lansford, Jennifer E., Patrick S. Malone, Kenneth A. Dodge, Joseph C. Crozier, Gregory S. Pettit, and John E. Bates. 2006. A 12-Year Prospective Study of Patterns of Social Information Processing Problems and Externalizing Behaviors. *Journal of Abnormal Child Psychology* 34:709–718.
- Lösel, Friedrich, and David P. Farrington. 2012. Direct Protective and Buffering Protective Factors in the Development of Youth Violence. *American Journal of Preventive Medicine* 43 (2): S8–S23.
- Luthar, Suniya S. 2006. Resilience in Development: A Synthesis of Research Across Five Decades. In *Developmental Psychopathology: Risk, Disorder, and Adaptation*, ed. Dante Cicchetti and Donald J. Cohen (740–95). New York: Wiley.
- Malti, Tina, and Monika Keller. 2009. The Relation of Elementary School Children's Externalizing Behaviour to Emotion Attributions, Evaluation of Consequences, and Moral Reasoning. *European Journal of Developmental Psychology* 6:592–614.
- Malti, Tina, and Tobias Krettenauer. 2012. The Relation of Moral Emotion Attributions to Pro- and Antisocial Behavior: A Meta-Analysis. *Child Development*, early online publication, 24 September 2012. doi:10.1111/j.1467-8624.2012.01851.x
- Malti, Tina, Cindy H. Liu, and Gil G. Noam. 2010. Holistic Assessment in School-Based, Developmental Prevention. *Journal of Prevention and Intervention in the Community* 38:244–59.
- Malti, Tina, and Gil G. Noam. 2009. A Developmental Approach to the Prevention of Adolescents' Aggressive Behavior and the Promotion of Resilience. *European Journal of Developmental Science* 3:235–46.
- Masten, Ann S. 2009. Ordinary Magic: Lessons from Research on Resilience in Human Development. *Education Canada* 49 (3): 28–32.
- Masten, Ann S. 2011. Resilience in Children Threatened by Extreme Adversity: Frameworks for Research, Practice, and Translational Synergy. *Development and Psychopathology* 23:141–54.
- Masten, Ann S., and Margaret O'D. Wright. 2009. Resilience Over the Lifespan: Developmental Perspectives on Resistance, Recovery, and Transformation. In *Handbook of Adult Resilience*, ed. John W. Reich, Alex J. Zautra, and John S. Hall (213–37). New York: Guilford.
- Noam, Gil G. 1996. High-Risk Youth: Transforming Our Understanding of Human Development. *Human Development* 39:1–17.
- Noam, Gil G. 1999. The Psychology of Belonging: Reformulating Adolescent Development. In *Adolescent Psychiatry*, vol. 24, ed. Lois Flaherty and Harvey Horowitz, (49–68). Hillsdale, New Jersey: The Analytic Press.
- Noam, Gil G., Michael J. Chandler, and Chris E. Lalonde. 1995. Clinical-Developmental Psychology: Constructivism and Social Cognition in the Study of Psychological Dysfunction. In *Developmental Psychopathology*, ed. Dante Cicchetti and Donald J. Cohen (424–66). New York: John Wiley.
- Noam, Gil G., and Larry S. Goldstein. 1998. The Resilience Inventory. Unpublished protocol.
- Noam, Gil G., and Tina Malti. 2008. Responding to the Crisis: RALLY's Developmental and Relational Approach. *New Directions for Youth Development* 120: 31–55.
- Noam, Gil G., Tina Malti, and Michael J. Karcher. Forthcoming. Mentoring Relationships in Developmental Perspective. In *Handbook of Youth Mentoring*, ed. David DuBois and Michael J. Karcher. Thousand Oaks, CA: Sage.
- Noam, Gil G., Kendra Winner, Andrew Rhein, and B. Molad. 1996. The Harvard RALLY Program and the Prevention Practitioner: Comprehensive, School-Based Intervention to Support Resiliency in At-Risk Adolescents. *Journal of Child and Youth Care Work* 11:32–47.
- Noam, Gil G., Copeland H. Young, and Janna Jilnina. 2006. Social Cognition, Psychological Symptoms, and Mental Health: The Model, Evidence, and Contribution to Ego Development. In *Developmental Psychopathology*, volume 1, *Theory and Method*, ed. D. Cicchetti and D. J. Cohen (750–94). Hoboken, New Jersey: John Wiley.
- Orobio de Castro, Bram, Jan W. Veerman, Willem Koops, Joop D. Bosch, and Heidi J. Monshouwer. 2002. Hostile Attribution of Intent and Aggressive Behavior: A Meta-Analysis. *Child Development* 73:916–34.
- Perren, Sonja, and Francoise Alsaker. 2009. Depressive Symptoms from Kindergarten to Early School Age: Longitudinal Associations with Social Skills Deficits and Peer Victimization. *Child and Adolescent Psychiatry and Mental Health* 3, article 28.
- Powell, Nicole R., John E. Lochman, and Caroline L. Boxmeyer. 2007. Prevention of Conduct Problems. *International Review of Psychiatry* 19:597–605.

- Schonert-Reichl, Kimberly A., Martin Guhn, Shelley Hymel, Anne M. Gadermann, Lina Sweiss, and Clyde Hertzman. 2012. Development and Psychometric Properties of an Assessment Tool on Children's Well-Being and Contextual Assets. Paper presented at the Annual Meeting of the American Educational Research Association, April, Vancouver, Canada.
- Song, Mehyun. 2003. *Two Studies on the Resilience Inventory (RI): Toward the Goal of Creating a Culturally Sensitive Measure of Adolescence Resilience*. Unpublished PhD diss., Harvard University, Cambridge, MA.
- Thomas, D. Roland, Edward Hughes, and Bruno D. Zumbo. 1998. On Variable Importance in Linear Regression. *Social Indicators Research* 45:253–75. doi:10.1023/A:1006954016433
- Uitenbroek, Daan G.. 1997. *SISA Correlations*. Southampton: D. G. Uitenbroek. <http://www.quantitativeskills.com/sisa/statistics/correl.htm>
- Van Roy, Betty, Marijke Veenstra, and Jocelyne Clench-Aas. 2008. Construct Validity of the Five-Factors Strengths and Difficulties Questionnaire (SDQ) in Pre-, Early, and Late Adolescence. *Journal of Child Psychology and Psychiatry* 49:1304–12.
- World Health Organization. 2003. *Creating an Environment for Emotional and Social Well-Being: An Important Responsibility of a Health Promoting and Child-Friendly School*, Information Series on School Health 10. Geneva, Switzerland: Author. http://www.who.int/school_youth_health/media/en/sch_childfriendly_03_v2.pdf

Gil Noam

gil_noam@harvard.edu

Tina Malti

tina.malti@utoronto.ca

Martin Guhn

martin.guhn@ubc.ca

Preventing Child Behavior Problems in the Erlangen-Nuremberg Development and Prevention Study: Results from Preschool to Secondary School Age

Friedrich Lösel, Institute of Criminology, Cambridge University, United Kingdom, and Institute of Psychology,
University of Erlangen-Nuremberg, Germany

Mark Stemmler, Institute of Psychology, University of Erlangen-Nuremberg, Germany

Vol. 6 (2) 2012

Evidence-based Developmental Prevention of Bullying and Violence in Europe

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

► **Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study** Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Preventing Child Behavior Problems in the Erlangen-Nuremberg Development and Prevention Study: Results from Preschool to Secondary School Age

Friedrich Lösel, Institute of Criminology, Cambridge University, United Kingdom, and Institute of Psychology, University of Erlangen-Nuremberg, Germany

Mark Stemmler, Institute of Psychology, University of Erlangen-Nuremberg, Germany

A brief overview of the prevention part of the long-term Erlangen-Nuremberg Development and Prevention Study, which combines a prospective longitudinal and experimental design. Findings up to five years after intervention are reported. From a sample of 609 families with kindergarten children, subgroups participated in the universal prevention program EFFEKT (child social skills training, a parent training and a combination of both) or were assigned to equivalent control groups. The short-term evaluation showed significant effects in mediating constructs (social problem solving and parenting behavior) and in educators' ratings of children's social behavior. In a follow-up after two to three years, school report cards showed fewer children with multiple behavior problems. In a further follow up after four to five years program children reported fewer externalizing and internalizing problems than the control group. There were no significant effects in the mothers' reports on their children's behavior. Most significant effect sizes ranged between $d = 0.20$ and $d = 0.40$. The findings suggest various positive long-term effects of the intervention. However, one needs to be cautious with regard to over-generalizing the positive findings, because effect sizes vary over time and the positive findings could not be replicated in all investigated variables.

Introduction

In recent years development-oriented prevention of delinquency and violence has become a key topic of criminology and crime policy (Beelmann 2012; Farrington and Welsh 2007; Lösel and Bender 2012). Numerous programs have been created and implemented in families, schools, preschools, social services, clinics, and neighborhoods. The child age may range from pre-birth to adolescence and the prevention can be universal (for all members of a population), selected (for at-risk groups), or indicated (for children with pre-existing behavior problems). Many programs have broader targets such as preventing not only crime but also internalizing problems, substance misuse, and school dropout, or generally supporting a desirable child development.

The expansion of developmental prevention in criminology and related disciplines has sound reasons such as sub-

stantial prevalence rates of behavioral problems, problem stability in a small group of "early starters," frequent comorbidity of various disorders, difficulties of later treatment, and high monetary costs for society (Lösel 2012a). Early prevention programs are based on a range of theoretical concepts. For example, child social skills training programs address social information processing and problem-solving, which play an important role in aggressive behavior (e.g. Dodge and Pettit 2003). Parenting programs aim to reduce coercive interactions, inconsistency, and corporal punishment (e.g. Dishion and Patterson 1994). Parent- and child-oriented programs are based on social learning and parenting theories (e.g. Webster-Stratton, Reid, and Hammond 2004). Early home-visiting programs use theories of social learning and attachment (e.g. Olds et al. 2007). Other programs integrate various concepts in a systems- and development-oriented perspective (e.g. Haw-

The research reported in this article was funded by the German Federal Ministry of Family Affairs, Seniors, Women, and Youth; the German Red Cross;

the European Union; and other institutions. We wish to thank Andreas Beelmann, Sabine Beer, Doris Bender, Stefanie Hacker, Stefanie Jaurisch, Antje

Klindworth-Mohr, Daniela Runkel, and Susanne Wallner for their work on the project.

kins et al. 2008). Although the respective theoretical constructs form a sound basis for prevention programs, their empirical correlations with antisocial outcomes are mainly small to moderate (Hawkins et al. 1998), particularly with regard to protective effects (Lösel and Farrington 2012). Therefore, it would be unrealistic to expect very large effects even for sound developmental programs.

Overall, systematic reviews and meta-analyses have shown positive results (e.g. Beelmann 2012; Farrington and Welsh 2007; Lösel 2012a) and also desirable cost-benefit ratios (e.g. Aos et al. 2004). However, there is great heterogeneity in outcomes and the field is confronted with various problems: 1. Most programs are not based on empirical evaluation studies (e.g. Junger et al. 2007). 2. Even when randomized controlled trials or sound quasi-experimental evaluations are carried out, most studies have only short follow-up periods, so there is no information on the impact on youth crime (Farrington and Welsh 2012; Lösel and Beelmann 2003). 3. There are frequently difficulties reaching high-risk families, early dropout, and other implementation problems (Eisner and Meidert 2011; Lösel 2012a). 5. There is a particular lack of long-term systematic research outside the English-speaking world (Beelmann 2012; Lösel and Bender 2012).

Against this background we carried out the Erlangen-Nuremberg Development and Prevention Study (ENDPS). This project is one of the few examples that follow the recommendation of Farrington, Ohlin and Wilson (1996) to combine a prospective longitudinal design with an experimental program evaluation. The ENDPS started in 1999 and is to our knowledge currently the longest-running study with such a combined design in Europe. As requested by the editors of this special issue, the present article contains a brief overview of the evaluation part of the project. While we refer mainly to previously published empirical articles, hitherto unpublished findings will also be presented.

1. Method

1.1. Sample and Participation Rates

The ENDPS comprises a core study on universal prevention programs and various smaller evaluations of program adaptations for specific risk groups. The sample of

the core study consisted of 675 kindergarten children (336 boys, 339 girls) from 609 families at 61 kindergartens in the cities of Erlangen and Nuremberg in Bavaria, Germany. The average age of the children at the first measurement was $M = 4.7$ years ($SD = 9.3$ months). According to an index of socio-economic status (cf. Geißler 1994) the sample was very similar to the population of the area (Beelmann et al. 2006). In total, seven waves of data collection have been carried out to date (the fourth only in a small subsample). The first three waves took place annually, the others at longer intervals. The retention rates varied over time and with regard to type of data gathering (e.g., number of mothers' versus fathers' questionnaires). As usual, not all measures were completed by all participants. The attrition rate of families after the first three annual waves was 5 percent. In the sixth assessment (four to five years after the intervention) 85 percent of the families were retained. In the most recent seventh wave (nine to eleven years after the first one) the retention rate was about 90 percent. Because it sometimes took a long time to complete the assessment of all families, the later follow-up times have a bandwidth.

The sample for the program evaluation consisted of 282 children (age: $M = 4.6$ years, $SD = 8.8$ months). 9.4 percent of the families were lower class, 29.4 percent lower middle class, 42.0 percent middle class, 16.4 percent upper middle class, and 2.8 percent upper class (cf. Geißler 1994). Ten percent of all parents had foreign ethnicity and 11 percent were single mothers (there was one single father in the study). We grouped the children/families as follows: 1. child training; 2. parent training; 3. combination of child and parent training, and 4. control group. All training took place in the year after the first assessment and ended between two to three months before the second assessment. The assignment of families/children to the training groups and control groups followed both methodological and practical considerations. A random assignment on the individual level would have caused serious threats to validity, for example reactance of families not included in a program, experimental or compensatory rivalry of control group parents, and diffusion of treatment if training group and control group children/parents were in contact in the same kindergarten (Lösel 2007b). Furthermore, not all kin-

dergartens were suitable for the training with regard to distance to the families' homes, available space, and group size. To cope with such typical problems in prevention practice but achieve equivalence of training group and control group, our design followed several steps: First, the above-mentioned organizational criteria were used to select kindergartens suitable for training. Second, 21 training groups from training kindergartens were selected randomly from all suitable kindergarten groups. Third, matched pairs were recruited from the other kindergartens with regard to age, gender, socio-economic status, and pre-training behavioral problems as assessed by the educators (total score of the Social Behavior Questionnaire, SBQ; see instruments section). With regard to quantitative matching criteria we used the "untreated neighbor" with the closest score. Due to this procedure, pretest SBQ means in the training group and control group were clearly equivalent (0.02 *SD* difference; Lösel et al. 2009). To control other aspects for potential heterogeneity in group-wise randomization we applied not only ANCOVAs but also causal regression models that separate average, covariate, and conditional effects (Steyer et al. 2000).

The child skills program was offered to the parents of 190 children. Twelve children were not permitted to participate, leaving 178 children from 157 families who took part in the training. This represents a participation rate of 93.7 percent. Of the participating children 96 percent were present for at least half of the sessions. The parent training was offered to 255 families, of which 170 (67 percent) participated. In most cases the mothers represented the family ($n = 163$), but there were also 48 fathers in the courses (sometimes together with the mother). Three-quarters of the parents attended at least half of the program.

As there is no generally valid solution for dealing with dropouts in program evaluation, one component of our analyses focused on children and parents who had attended at least half of their program (Lösel et al. 2006). In addition, intent-to-treat analyses that allocated all dropouts to the TC were also carried out (Lösel et al. 2009). The respective results were rather similar. Although we observed no significant differences in matching variables between dropouts and completers, we nonetheless used equal n

comparisons in our evaluations; i.e. only individuals for whom the matched partner was available were entered in the analyses. This ensures equivalence between training group and control group.

1.2. Prevention Programs

After various pilot studies (e.g. Beelmann 2003) the following programs were chosen for the main evaluation in the ENDPS:

Child training: The training of children's social skills was based on "I Can Problem Solve" (Shure 1992), but updated and modified for the German context. It was delivered to twenty-one groups of six to ten children. The course is a manual-based group training in social problem-solving (Beelmann, Jausch, and Lösel 2004). The first part addresses verbal concepts, identification of emotions, and reflection on reasons for behavior. The second part contains training in problem-solving skills such as providing alternative solutions in conflicts, anticipation of actions and evaluation of consequences. The training uses a range of didactical methods. Each of the fifteen sessions lasted 30 to 60 minutes and there were three to five sessions per week (guided by two trained facilitators from the ENDPS).

Parent training: The parenting program was delivered in twelve courses in the afternoon or evening. Child-care was provided to enable parents to attend. The training aims to enhance positive parenting skills (Beelmann and Lösel 2004). It is partly based on the programs of the Oregon Social Learning Center (Dishion and Patterson 1996; Fisher et al. 1997) and was updated and adapted to the German context. Pilot studies suggested keeping the program short to increase participation and reduce dropout. The training consisted of five 90- to 120-minute sessions spread over five weeks. The courses were delivered by two experienced facilitators from our team (group sizes 6 to 15). The content included issues of positive parenting, requests and demands, setting limits, dealing with difficult parenting situations, coping with stress, and enhancing the family's social relationships. Structured presentations, group discussions, role-playing, self-awareness exercises, homework, and other didactic measures were used.

1.3. Instruments

The assessments of the children and families varied over time and according to the children's age (e.g., the problem-solving test for six-years olds contained more items than for three-years olds). They employed multiple methods and data sources, ranging from parent interviews and questionnaires through child assessments and kindergarten staff's ratings to school report cards and pediatric data (Lösel et al. 2005). As the focus of this article is on program evaluation, we mainly refer to the children's behavioral problems as reported by various informants. In addition, data on process evaluation and theoretically relevant mediating factors for program outcome are briefly reported (i.e. children's social problem-solving, and parenting behavior).

For *outcome evaluation* the children's behavioral problems were measured using our German adaptations (Lösel, Beelmann, and Stemmler 2002) of the Social Behavior Questionnaire (SBQ; Tremblay et al. 1992). We used the SBQ to gather independent information from kindergarten educators, mothers, and (when they were older) children's self-reports. Behavioral problems in the first and second grades of elementary school were assessed by a content analysis of the school report cards (Stemmler et al. 2005). Teachers' comments on behavioral and emotional aspects such as aggression, hyperactivity, emotional tone, and obeying rules were reliably categorized and used as indicators of behavioral problems. From the sixth wave onwards we also applied short versions of a German self-reported delinquent behavior scale (the Delinquenz-Belastungsskala, DBS; Lösel 1975).

Parenting behavior was measured by mothers' reports on our German adaptation of the Alabama Parenting Questionnaire (Shelton, Frick, and Wotton 1996). The children's social problem-solving competence was assessed via the German version of Spivak and Shure's (1982; Shure, 1990) Preschool Interpersonal Problem-Solving Test (PIPS; Döpfner, Lorch, and Reihl 1989). In this test the children are asked for interpretations, potential motives, and behavioral alternatives in response to conflict scenarios presented in pictures.

For *process evaluation* of the parent training we used anonymous ratings on five-point scales that addressed as-

pects such as the selection of topics, comprehensibility, usefulness for participants' own parenting, and overall user satisfaction (Lösel et al. 2005). The implementation of the child training was assessed via ratings of the children's on-task and off-task behavior in each session (Cangelosi 1996).

2. Results

Most of the following results are presented as effect sizes using Cohen's d coefficients. When the outcome measure had already been used in the pretest (e.g. the SBQ), we calculated the net difference between standardized effects in the training group and control group. When an outcome measure could only be measured at follow up (e.g. the DBS), d was based on the standardized difference at the respective measurement point.

2.1. Implementation Quality

The above-mentioned participation and completion rates suggest that the implementation of the programs was satisfactory. This was particularly the case for the child program, where both rates were clearly above 90 percent. More detailed process evaluations also indicate that the implementation was appropriate (Lösel et al. 2005). On average the participants of the parent training rated all aspects between "very good" (1) and "good" (2) with a variation from $M = 1.20$ to 1.74 . Overall satisfaction with the training was $M = 1.73$ ($SD = 0.43$), the quality of the facilitators was rated particularly positively ($M = 1.20$, $SD = 0.33$).

The children's behavior during the program was also satisfactory. Over all sessions the rates of relevant on-task behavior were between 75 and 80 percent ($M = 79.2$ percent, $SD = 16.7$). Disruptive off-task behavior was observed much less commonly ($M = 7.2$ percent, $SD = 8.3$). Only three children (2.9 percent) had on-task scores of less than 50 percent. These and other process data suggest that the programs were implemented at high quality. Therefore, a potential lack of effects in the outcome evaluation could not be attributed to poor implementation quality.

2.2. Effects on Mediating Factors

Our findings show significant program effects on theoretically mediating proximal factors. The child program had a positive influence on social problem-solving as measured

by the PIPS (Lösels and Beelmann 2005). Children in the training group produced a larger overall number of conflict solutions and a smaller proportion of aggressive solutions, and made fewer aggressive decisions ($d = 0.25\text{--}0.47$). Such effects were not observed when only the parents had participated in parent training. Although our project showed an overall improvement of social information processing and problem-solving with increased age (Beelmann, Lösels, and Stemmler 2010), the child social skills training seems to accelerate this process.

Similarly, there were specific effects of parent training on parenting behavior and attitudes (Stemmler et al. 2007). Shortly after the training, mothers from the training group reported significantly more positive parenting ($d = 0.30$) and less inconsistent discipline ($d = 0.29$) than the control group. The latter effect remained stable during the first year. There was no overspill of impact from mothers to fathers who did not participate in the program.

As expected, characteristics of the children's social information processing and of the parents' educational behavior were related to child behavior problems, although most correlations were small (Beelmann et al. 2010; Stemmler et al. 2007).

2.3. Short-Term Effects on Child Behavior (Two to Three Months)

The first outcome assessment of child behavior was carried out two to three months after the training. The kindergarten educators' ratings of the children showed a significant positive effect of the total EFPEKT program in the SBQ-total score ($d = 0.30$). The effects of the various program components were $d = 0.26$ for the child training, $d = 0.22$ for the parent training, and $d = 0.39$ for the combined parent and child training (Lösels, Beelmann, et al. 2006). These results remained consistent when we used causal regression models instead of covariance analyses (Lösels et al. 2009). The effects in the subscales on externalizing and internalizing problems were significant for the total program ($d = 0.17$ and 0.19), the child training (0.25 and 0.26), and the combined training (0.36 and 0.33), but not for the parent training (0.11 and 0.09). There were also some conditional effects showing that those children who had more behavioral problems before

the program benefitted most (Lösels et al. 2009). The effect sizes for this subgroup of those in greatest need ranged between $d = 0.25$ and 0.66 (partially due to a slight increase of problems in the respective control groups; Lösels, Beelmann et al. 2006).

In contrast to the educators' ratings, there was no outcome in the positive direction in the mothers' ratings of child behavior (Lösels et al., 2009). We even found a small negative effect ($d = -0.22$) on internalizing problems.

2.4. Long-Term Effects on Child Behavior (Two to Five Years)

Two years after the training we analyzed the content of the report cards at elementary school. This is a particularly valid outcome indicator because the teachers did not know who participated in the program at kindergarten age. Overall, there were significantly less behavioral problems in the training group than in the control group ($d = 0.17$; Lösels, Beelmann et al. 2006). The specific effects of the parent training and combined training were not significant, but the child skills training had a significant and substantial effect ($d = 0.35$). In further analyses we focused on those children for whom the teachers reported at least three behavioral problems at grade one and two (Lösels et al. 2009). Because these children showed relatively stable problems two to three years after the program they are at high risk for longer-term deviance. The results for the total training and for the combined training showed significant positive effects (1.2 percent vs. 4.3 percent and 0.0 percent vs. 6.1 percent in the training group vs. control group).

In contrast to this non-reactive information from the school teachers, the mothers' reports did not reveal any significant program effect (Lösels et al. 2009). The above-mentioned slightly negative short-term effect in the mothers' reports of internalizing child symptoms had also disappeared.

In the follow up four to five years after the program we were able to assess children's self-reports. Table 1 contains the findings on all trained children and their parents versus the control group in the SBQ scales and delinquency self-report (DBS).

Table 1: Results for children's self-reports of behavioral and emotional problems four to five years after the intervention (all trained EFFEKT children and their parents)

	Program	Control	Difference	Effect size
	<i>M</i> <i>SD</i>	<i>M</i> <i>SD</i>	<i>t-test</i>	<i>d</i>
SBQ - Total scale	9.26 (5.37)	11.29 (5.76)	3.07**	0.28
SBQ - Externalizing score	5.99 (5.37)	7.35 (4.99)	2.37*	0.28
SBQ - Internalizing score	3.26 (2.20)	3.94 (2.56)	2.57*	0.29
DBS - Delinquency scale	0.20 (0.47)	0.27 (0.67)	1.14	0.13

Notes:

d = Cohen's standardized mean difference*t*-tests controlled for variance heterogeneity

SBQ = Social Behavior Questionnaire

DBS = Delinquency Self Report Scale

total *n* = 282 (equal *n* in training group and control group)* *p* < 0.05; ** *p* < 0.01.

There were positive effects for the SBQ total problems score and the subscales on both externalizing and internalizing problems. The result in the DBS scale on self-reported delinquency went in the expected direction, but was not significant (perhaps due to the generally small numbers of offences at this age). The effects were similar for the subsamples rated above and below the SBQ median at Time 1. The outcomes for the different types of interventions varied. They were all positive for the combined parent and child training ($d = 0.21$ – 0.34), but due to the now smaller *n* they did not reach statistical significance. The results of the parent training in the SBQ were particularly positive and significant ($d = 0.49$ – 0.63), but not significant in the delinquency scale. The findings for the child training all went in the desirable direction ($d = 0.19$ – 0.45) and were significant in the DBS. The effects of the combined trainings in the delinquency scale were positive ($d = 0.17$ – 0.26), but not statistically significant.

We also recorded the mothers' evaluations of child behavior in the SBQ scales four to five years after the training. There

were no significant effects in the total score, nor in the subscales on externalizing and internalizing child problems.

3. Discussion

In comparison to the majority of evaluations on developmental prevention, follow-up periods of two to three and four to five years are rather long. Most studies comprise only one year or less (Beelmann 2012). This is particularly the case in Europe where well controlled long-term evaluations are very rare (e.g. Beelmann and Lösel 2007; Eisner et al. 2007).

A substantial proportion of our short- and long-term effects went in the desirable direction, i.e. less behavioral problems in the program groups. The significant effects were mainly small and occasionally moderate. This is in line with international meta-analyses that report lower effects for universal prevention than for risk-focused selective and indicated approaches (Beelmann and Raabe, 2009; Lösel 2012a). In principle, one should not expect larger long-term effects of universal programs because the majority of the children in the more or less "normal" intervention groups would not develop behavioral problems without receiving a program. This does not imply that universal programs for the whole population of a kindergarten, school, or neighborhood should not be carried out. Universal programs have the advantage of avoiding potential stigmatization, do not require risk assessment procedures and can more easily be implemented in routine practice (e.g. in schools). As in public health programs, universal approaches can also have a particular impact on those groups who are most in need or at highest risk (Coid 2003). This was confirmed in our short-term outcomes and in the school report cards two to three years after the training. After four to five years we did not find such a differential effect for those at highest risk in the children's self-reports. This may have been due to the overall small correlations between different informants and reduced stability over time (Lösel et al. 2005; see also below).

We conclude from these findings that a relatively short universal prevention program such as EFFEKT can reduce child behavior problems. This is in accordance with meta-analyses that found no strong relationship between pro-

gram intensity (dosage) and effect size (Beelmann 2008; Lösel and Beelmann, 2003). However, the universal approach should be seen as a “foot in the door” with regard to lasting influences in high-risk groups. Universal programs would become too expensive for the society if large numbers of low-risk children and families received high-dosage measures (Foster et al. 2008; Offord 2000).

Despite a number of theoretically plausible and practically desirable outcomes of the EFFEKT program one should be aware of various issues. Firstly there was some inconsistency of effect sizes over time. Although, as expected, a number of effects decreased in the follow-ups, some outcomes were stronger in the long-term than in the short-term evaluation. There was also partial inconsistency of effects in different outcome measures. For example, the positive effects in the kindergarten educators’ SBQ reports, elementary school report cards, and children’s self reports clearly differed from the results in the mothers’ reports where we found no positive effects. Similarly, we observed certain inconsistencies in the outcomes of the different parts of the program. For example, the child training exhibited the largest effect after two to three years, whereas the parent training seemed to be more effective in the follow up after four to five years.

Some of these variations may be due to random fluctuation. Others may be due to numerous program, individual, contextual, and methodological factors (Lösel 2012a). One must be aware that programs are part of the child’s “natural” development in which (causal) risk and protective factors vary over time, accumulate, and interact with each other (Dodge et al. 2008; Lösel and Bender 2003). This can lead to complex patterns of influences. For example, the small negative effects in the short-term evaluation of the mother’s reports on internalizing child problems may have been due to a temporal increase of the mother’s sensitivity. Larger effect sizes in the follow-up could be due to “sleeping effects” that may result from greater experience in the application of training content. Certain inconsistencies between the data from the educators/teachers and from the mothers reflect the generally small correlations between different informants on child behavior problems (Achenbach 2006) that was also observed in the ENDPS (Lösel et al. 2005). The reports of the pro-

fessionals may be more reliable because they contain comparisons between children (which is often lacking in small families). Mothers’ and teachers’ reports can be influenced by a relatively stable general image of the child. As Lösel reports (2002), the longitudinal correlations between behavior ratings by the same informants were larger than the cross-sectional correlations between different informants. In another study of the ENDPS we only found significant effects of a child skills training when the teachers who rated the child were not the same before and after the program (Hacker et al. 2007).

Variations and partial inconsistencies in findings across times, subprograms, and outcome measures are rather common in evaluations of developmental prevention programs (even in some of the best long-term studies). For example, the FAST Track project found substantial variation in different outcomes (Conduct Problems Prevention Research Group 2002, 2004, 2010); most recently there were positive effects in official indicators of offending but not in self-reports. The Seattle Development Project reported positive effects of “Communities That Care,” but with variations between parts of the program and different outcome measures (e.g. Hawkins et al. 1999, 2008). The Montreal Prevention Experiment had less effects in short-term than in long-term evaluations (Tremblay et al. 1995; Vitaro, Brendgen, and Tremblay 2001). These and other examples suggest that one should not expect too simple and uniform messages from rather complex and multimodal long-term evaluations of developmental prevention programs. One should also be aware of the risk of “fishing for significances” when many variables and measurement times are included.

Our evaluations of EFFEKT revealed not only various long-term effects, but also no lasting negative outcomes. As McCord (2003) has shown, some programs can harm in spite of best intentions. Having various positive and no harmful impact justified disseminating the EFFEKT program into routine practice. Meanwhile more than one thousand facilitators have been trained in all parts of Germany on a non-profit basis. We also learnt that the standard program needs to be enriched with modules for subpopulations with specific needs. Therefore, we devel-

oped adapted versions that addressed cultural differences in parenting (e.g. in Turkish families) and also used simpler language. Ethnic minority families in deprived neighborhoods benefitted from this intervention (Runkel 2009; Runkel et al., forthcoming). Another adaptation addressed emotionally burdened or depressed mothers (Kötter et al. 2010). This program was evaluated in clinic contexts and showed positive effects on mothers' parenting and children's behavior (Bühler et al. 2011; Stemmler et al., forthcoming). In principle, we recommend working with the core structure of evidence-based programs, but adding and evaluating modules for specific family needs where necessary. Such a more individualized approach is supported by a meta-analysis of family-oriented prevention programs in Germany (Lösel, Schmucker et al. 2006).

4. Conclusions

The findings and experiences from our primary evaluations and research syntheses within the ENDPS lead us

to the following conclusions: First, universal developmental prevention programs such as EFFEKT can have positive long-term effects. Second, it is realistic to expect mainly small effect sizes, particularly in routine practice (as opposed to demonstration projects). Third, because of the high costs of persistent criminality even small effects may well pay off if only a few cases become more resilient. Fourth, more well-controlled and replicated evaluations of the long-term outcomes of both universal and targeted prevention programs are needed, particularly outside North America. Fifth, there is a need for more research on the outreach and implementation of programs in routine practice. Sixth, specific evidence-based programs must be more closely integrated into the broad range of routine services in practice. And finally, developmental prevention on the individual and micro-social level should be accompanied by approaches designed to reduce risks at the macro-level (e.g. in order to avoid social segregation).

References

- Achenbach, Thomas. 2006. As Others See Us: Clinical and Research Implications of Cross-Informant Correlations for Psychopathology. *Current Directions in Psychological Science* 15:94–98.
- Aos, Steve, Roxanne Lieb, Jim Mayfield, Marna Miller and Annie Pennucci. 2004. *Benefits and Costs of Prevention and Early Intervention Programs for Children and Youth*. Olympia, WA: Washington State Institute of Public Policy.
- Beelmann, Andreas. 2003. Effektivität behavioraler Elternt Trainingsprogramme: Ergebnisse zweier Pilotstudien zur Prävention dissozialen Verhaltens. *Psychologie in Erziehung und Unterricht* 50:310–23.
- Beelmann, Andreas. 2008. *The Effect of Parent Training Programs in the Prevention and Treatment of Antisocial Behavior in Childhood and Adolescence*. Paper presented at the 18th European Conference of Psychology and Law, 2–5 July 2008, Maastricht, the Netherlands.
- Beelmann, Andreas. 2012. The Scientific Foundation of Prevention: The Status Quo and Future Challenges for Developmental Prevention. In *Antisocial Behavior and Crime: Contributions of Developmental and Evaluation Research to Prevention and Intervention*, ed. Thomas Bliesener, Andreas Beelmann, and Mark Stemmler, 137–63. New York: Hogrefe.
- Beelmann, Andreas, Stefanie Jaurisch, and Friedrich Lösel. 2004. *Ich kann Probleme lösen: Soziales Trainingsprogramm für Vorschulkinder*. Universität Erlangen-Nürnberg: Institut für Psychologie.
- Beelmann, Andreas, Stefanie Jaurisch, Friedrich Lösel, and Mark Stemmler. 2006. Frühe universelle Prävention von dissozialen Entwicklungsproblemen: Implementation und Wirksamkeit eines verhaltensorientierten Elternt Trainings. *Praxis der Rechtspsychologie* 16:120–43.
- Beelmann, Andreas, and Friedrich Lösel. 2004. *Elternt raining zur Förderung der Erziehungskompetenz*. Universität Erlangen-Nürnberg: Institut für Psychologie.
- Beelmann, Andreas, and Friedrich Lösel. 2007. Prävention von externalisierendem Problemverhalten. In *Prävention und Gesundheitsförderung*, vol. 3, Kinder und Jugendliche, ed. Bernd Röhrle, 557–95. Tübingen: Deutsche Gesellschaft für Verhaltenstherapie.
- Beelmann, Andreas, Friedrich Lösel, and Mark Stemmler. 2010. Die Entwicklung von sozialer Informationsverarbeitung und die Vorhersage physischer Aggression im Vorschulalter. *Psychologie in Erziehung und Unterricht* 57:119–31.
- Beelmann, Andreas, Friedrich Lösel, Mark Stemmler, and Stephanie Jaurisch. 2006. Beurteilung von sozialen Verhaltensproblemen und Erziehungsschwierigkeiten im Vorschulalter: Eine Untersuchung zur deutschen Adaptation des Eyberg Child Behavior Inventory (ECBI). *Diagnostica* 52:189–98.
- Beelmann, Andreas, and Thomas Raabe. 2009. The Effects of Preventing Antisocial Behavior and Crime in Childhood and Adolescence: Results and Implications of Research Reviews and Meta-analyses. *European Journal of Developmental Science* 3:260–81.
- Bender, Doris, and Friedrich Lösel. 2006. Working With Violent Children in German Youth Services: Results of a Survey. In *Children Who Commit Acts of Serious Interpersonal Violence: Messages for Practice*, ed. Ann Hagell and Renuka Jeyarajah-Dent, 167–85. London: Jessica Kingsley.
- Bühler, Anneke, Charlotte Kötter, Stefanie Jaurisch, and Friedrich Lösel. 2011. Prevention of Familial Transmission of Depression: EFFEKT-E, A Selective Program for Emotionally Burdened Families. *Journal of Public Health* 19:321–27.
- Cangelosi, James S. 1996. *Classroom Management Strategies: Gaining and Maintaining Students' Cooperation*. New York: Longman.
- Coid, Jeremy W. 2003. Formulating Strategies for the Primary Prevention of Adult Antisocial Behaviour: "High Risk" or "Population" Strategies? In *Early Prevention of Adult Antisocial Behaviour*, ed. David P. Farrington and Jeremy W. Coid, 32–78. Cambridge, UK: Cambridge University Press.
- Conduct Problems Prevention Research Group. 2002. Evaluation of the First 3 Years of the Fast Track Prevention Trial with Children at High Risk for Adolescent Conduct Problems. *Journal of Abnormal Child Psychology* 19:553–67.
- Conduct Problems Prevention Research Group. 2004. The Effects of the Fast Track Program on Serious Problem Outcomes at the End of Elementary School. *Journal of Clinical Child and Adolescent Psychology* 33:650–61.

- Conduct Problems Prevention Research Group. 2010. Fast Track Intervention Effects on Youth Arrests and Delinquency. *Journal of Experimental Criminology* 6:131–57.
- Dishion, Thomas J., and Gerald S. Patterson. 1996. *Preventive Parenting with Love, Encouragement and Limits: The Preschool Years*. Eugene, OR: Castalia.
- Dodge, Kenneth A., Mark T. Greenberg, Patrick S. Malone, and Conduct Problems Prevention Research Group. 2008. Testing an Idealized Dynamic Cascade Model of the Development of Serious Violence in Adolescence. *Child Development* 79:1907–27.
- Dodge, Kenneth A., and Gregory S. Pettit. 2003. A Biopsychosocial Model of the Development of Chronic Conduct Problems in Adolescence. *Developmental Psychology* 39:349–71.
- Döpfner, Manfred, Ruth Lorch, and Dieter Reihl. 1989. Soziale Informationsverarbeitung in Konfliktsituationen – eine empirische Studie an Vorschulkindern. *Zeitschrift für Pädagogische Psychologie* 3:239–48.
- Eisner, Manuel, and Ursula Meidert. 2011. Stages of Parental Engagement in a Universal Parent Training Program. *Journal of Primary Prevention* 32:83–93.
- Eisner, Manuel, Denis Ribeaud, Rahel Jünger, and Ursula Meidert. 2007. *Frühprävention von Gewalt und Aggression*. Zürich: Rüeegger.
- Farrington, David P., Lloyd E. Ohlin, and James Q. Wilson. 1986. *Understanding and Controlling Crime: Toward a New Research Strategy*. New York: Springer.
- Farrington, David P. and Brandon C. Welsh. 2007. *Saving Children From a Life of Crime*. Oxford, UK: Oxford University Press.
- Fisher, Philip A., Elizabeth Ramsay, Karla Antoine, Kate Kavanagh, Allen Winebarger, J. Mark Eddy, and John B. Reid. 1997. *Success in Parenting: A Curriculum for Parents with Challenging Children*. Eugene, OR: Oregon Social Learning Center.
- Foster, E. Michael, Ronald J. Prinz, Matthew R. Sanders, and Cheri J. Shapiro. 2008. The Costs of a Public Health Infrastructure for Delivering Parenting and Family Support. *Children and Youth Services Review* 30:493–501.
- Geißler, Rainer, ed. 1994. *Soziale Schichtung und Lebenschancen in Deutschland*. Stuttgart: Enke.
- Hacker, Stefanie, Friedrich Lösel, Mark Stemmler, Stefanie Jaurisch, Daniela Runkel, and Andreas Beelmann. 2007. Training im Problemlösen (TIP): Implementation und Evaluation eines sozial-kognitiven Kompetenztrainings für Kinder. *Heilpädagogische Forschung* 23:11–21.
- Hawkins, J. David, Eric C. Brown, Sabrina Oesterle, Michael W. Arthur, Robert D. Abbott, and Richard F. Catalano. 2008. Early Effects of Communities That Care on Targeted Risks and Initiation of Delinquent Behavior and Substance Use. *Journal of Adolescent Health* 43:15–22.
- Hawkins, J. David, Richard F. Catalano, Rick Kosterman, Robert D. Abbott, and Karl G. Hill. 1999. Preventing Adolescent Health-risk Behaviors by Strengthening Protection During Childhood. *Archives of Pediatric and Adolescent Medicine* 153:226–34.
- Junger, Marianne, Lynette Feder, Joy Clay, Sylvana M. Coté, David P. Farrington, Kate Freiberg, Vicente Garrido, et al. 2007. Preventing Violence in Seven Countries: Global Convergence in Policy. *European Journal on Criminal Policy and Research* 13:327–56.
- Kötter, Charlotte, Mark Stemmler, Anneke Bühler, Stefanie Jaurisch, and Friedrich Lösel. 2010. Mittelfristige Effekte des Präventionsprogramms EFFEKT-E für emotional belastete Mütter und ihre Kinder unter besonderer Berücksichtigung psychosozialer Risikofaktoren. *Zeitschrift für Gesundheitspsychologie* 19:122–33.
- Lösel, Friedrich. 1975. *Handlungskontrolle und Jugenddelinquenz: Theoretische Integration und empirische Prüfung*. Stuttgart: Enke.
- Lösel, Friedrich. 2002. Risk/Need Assessment and Prevention of Antisocial Development in Young People: Basic Issues from a Perspective of Cautionary Optimism. In *Multiproblem Violent Youth*, NATO SPS series, ed. Ray Corrado, Ron Roesch, Stephen D. Hart, and Josef Gierowski, 35–57. Amsterdam: IOS Press.
- Lösel, Friedrich. 2007a. It's Never Too Early and Never Too Late: Towards an Integrated Science of Developmental Intervention in Criminology. *Criminologist* 35 (2): 1–8.
- Lösel, Friedrich. 2007b. Doing Evaluation in Criminology: Balancing Scientific and Practical Demands. In *Doing Research on Crime and Justice*, 2d ed., ed. Roy D. King and Emma Wincup, 141–170. Oxford, UK: Oxford University Press.
- Lösel, Friedrich. 2012a. Entwicklungsbezogene Prävention von Gewalt und Kriminalität: Ansätze und Wirkungen. *Forensische Psychiatrie, Psychologie und Kriminologie* 6:71–84.
- Lösel, Friedrich. 2012b. Offender Treatment and Rehabilitation: What Works? In *The Oxford Handbook of Criminology*, 5th ed., ed. Mike Maguire, Rod Morgan, and Robert Reiner 986–1016. Oxford, UK: Oxford University Press.
- Lösel, Friedrich, and Andreas Beelmann. 2003. 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* 587:84–109.
- Lösel, Friedrich, and Andreas Beelmann. 2005. Social Problem-Solving Programme for Preventing Antisocial Behaviour in Children and Youth. In *Social Problem Solving and Offending: Evidence, Evaluation and Evolution*, ed. Mary McMurrin and James McGuire, 127–43. Chichester, UK: Wiley.
- Lösel, Friedrich, and Andreas Beelmann. 2006. Child Skills Training. In *Preventing Crime: What Works for Children, Offenders, Victims, and Places*, ed. Brandon C. Welsh and David P. Farrington, 33–54. Dordrecht, NL: Wadsworth.
- Lösel, Friedrich, Andreas Beelmann, Stefanie Jaurisch, Ute Koglin, and Mark Stemmler. 2005. Entwicklung und Prävention früher Probleme des Sozialverhaltens: Die Erlangen-Nürnberg Studie. In *Möglichkeiten der Gewaltprävention*, ed. Manfred Cierpka, 201–49. Göttingen: Vandenhoeck und Ruprecht.
- Lösel, Friedrich, Andreas Beelmann and Mark Stemmler. 2002. *Skalen zur Messung sozialen Problemverhaltens bei Vorschul- und Grundschulkindern*. Universität Erlangen-Nürnberg: Institut für Psychologie.
- Lösel, Friedrich, Andreas Beelmann, Mark Stemmler, and Stefanie Jaurisch. 2006. Prävention von Problemen des Sozialverhaltens im Vorschulalter: Evaluation des Eltern- und Kindertrainings EFFEKT. *Zeitschrift für Klinische Psychologie und Psychotherapie* 35:127–39.
- Lösel, Friedrich, and Doris Bender. 2003. Protective Factors and Resilience. In *Early Prevention of Adult Antisocial Behaviour*, ed. David P. Farrington and Jeremy W. Coid, 130–204. Cambridge: Cambridge University Press.
- Lösel, Friedrich, and Doris Bender. 2006. Risk Factors for Serious Juvenile Violence. In *Children Who Commit Acts of Serious Interpersonal Violence: Messages for Practice*, ed. Ann Hagell and Renuka Jeyarajah-Dent, 42–72. London: Jessica Kingsley.
- Lösel, Friedrich, and Doris Bender. 2012. Child Social Skills Training in the Prevention of Antisocial Development and Crime. In *Handbook of Crime Prevention*, ed. David P. Farrington and Brandon C. Welsh, 102–29. Oxford, UK: Oxford University Press.
- Lösel, Friedrich, and David P. Farrington. 2012. Direct Protective and Buffering Protective Factors in the Development of Youth Violence. *American Journal of Preventive Medicine* 43 (2, supplement): 8–23.
- Lösel, Friedrich, Martin Schmucker, Birgit Plankensteiner, and Maren Weiss. 2006. *Bestandsaufnahme und Evaluation der Elternbildung*. Berlin: Bundesministerium für Familie, Senioren, Frauen und Jugend.
- Lösel, Friedrich, Mark Stemmler, Andreas Beelmann, and Stefanie Jaurisch. 2005. Aggressives Verhalten im Vorschulalter: Eine Untersuchung zum Problem verschiedener Informanten. In *Aggressionsentwicklung zwischen Normalität und Pathologie*, ed. Inge Seiffge-Krenke, 141–67. Göttingen: Vandenhoeck und Ruprecht.
- Lösel, Friedrich, Mark Stemmler, Stefanie Jaurisch, and Andreas Beelmann. 2009. Universal Prevention of Antisocial Development: Short- and Long-Term Effects of a Child- and Parent-Oriented Program. *Monatsschrift für Kriminologie und Strafrechtsreform* 92: 289–308.
- McCord, Joan. 2003. Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs. *Annals of the American Academy of Political and Social Science* 587:16–30.
- Offord, David R. 2000. Selection of Levels of Prevention. *Addictive Behaviors* 25:833–42.

- Olds, David L., Harriet Kitzman, Carole Hanks, Robert Cole, Elizabeth Anson, Kimberly Sidora-Arcoleo, Dennis W. Luckey, et al. 2007. Effects of Nurse Home Visiting on Maternal and Child Functioning: Age-9 Follow-up of a Randomized Trial. *Pediatrics* 120:832–45.
- Runkel, Daniela. 2009. *EFFEKT-Interkulturell: Implementierung und Evaluation eines präventiven Eltern- und Kindertrainings an Grundschulen mit einem hohen Anteil von Kindern mit Migrationshintergrund*. PhD diss., Universität Erlangen-Nürnberg.
- Runkel, Daniela, Friedrich Lösel, Mark Stemmler, and Stefanie Jaursch. Forthcoming. Preventing Social Behaviour Problems in Children From Deprived Migrant Families: Evaluation of a Child and Parent Training.
- Shelton, Karen K., Paul J. Frick, and Jane Wootton. 1996. Assessment of Parenting Practices in Families of Elementary School-age Children. *Journal of Clinical Child Psychology* 25:317–29.
- Shure, Myrna B. 1990. *The PIPS Test Manual: A Cognitive Measure of Interpersonal Problem-Solving Ability for Four to Six Year Old Children*, 2d ed. Philadelphia: Hahneman University, Department of Mental Health Sciences.
- Shure, Myrna B. 1992. *I Can Problem Solve: An Interpersonal Cognitive Problem-Solving Program: Kindergarten & Primary Grades*. Champaign, IL: Research Press.
- Spivack, George, and Myrna B. Shure. 1982. Interpersonal cognitive problem solving and clinical theory. In *Advances in Child Clinical Psychology*, vol. 5, ed. B. Lahey and Alan E. Kazdin, 323–72). New York: Plenum.
- Stemmler, Mark, Andreas Beelmann, Stefanie Jaursch, and Friedrich Lösel. 2007. Improving Parenting Practices in Order to Prevent Child Behavior Problems: A Study on Parent Training as Part of the EFFEKT Program. *International Journal of Hygiene and Environmental Health* 210:563–70.
- Stemmler, Mark, Charlotte Kötter, Anneke Bühler, Stefanie Jaursch, Andreas Beelmann, and Friedrich Lösel. Forthcoming. Prevention of Familial Transmission of Depression Through a Family Oriented Programme Targeting Parenting as Well as the Child's Social Competence. *Journal of Children's Services*.
- Stemmler, Mark, Friedrich Lösel, Andreas Beelmann, Stefanie Jaursch, and Barbara Zenkert. 2005. Child Problem Behavior in Kindergarten and in Primary School: A Comparison Between Prediction Configural Frequency Analysis and Multiple Regression. *Psychology Science* 47 (3/4): 467–78.
- Steyer, Rolf, Siegfried Gabler, Alina A. von Davier, Christof Nachtigall, and Thomas Buhl. 2000. Causal Regression Models I: Individual and Average Causal Effects. *Methods of Psychological Research Online* 5:39–71.
- Tremblay, Richard E., Linda Pagani-Kurtz, Frank Vitaro, Louise C. Masse, and Robert O. Pihl. 1995. A Bimodal Preventive Intervention for Disruptive Kindergarten Boys: Its Impact Through Mid-Adolescence. *Journal of Consulting and Clinical Psychology* 63:560–68.
- Tremblay, Richard E., Frank Vitaro, Claude Gagnon, C. Piché, and N. Royer. 1992. A Prosocial Scale for the Preschool Behavior Questionnaire: Concurrent and Predictive Correlates. *International Journal of Behavioral Development* 15:227–45.
- Vitaro, Frank, Mara Brendgen, and Richard E. Tremblay. 2001. Preventive Intervention: Assessing its Effects on the Trajectories of Delinquency and Testing for Mediation Processes. *Applied Developmental Science* 5:201–13.
- Webster-Stratton, Carolyn, M., John Reid, and Mary Hammond. 2004. Treating Children with Early-Onset Conduct Problems: Intervention Outcomes for Parent, Child, and Teacher Training. *Journal of Clinical Child and Adolescent Psychology* 33:105–24.

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales

Judy Hutchings, Centre for Evidence Based Early Intervention, School of Psychology, Bangor University, United Kingdom

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

► **Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales** Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales

Judy Hutchings, Centre for Evidence Based Early Intervention, School of Psychology, Bangor University, United Kingdom

A case study reviewing the establishment of the evidence-based Incredible Years programme in Wales, describing the rationale for selecting the programme, the outcomes achieved in Wales, and the influence on policy leading to a Wales-wide dissemination strategy. The UK context features a growing trend towards evidence-based anti-violence services and significant increases in funding for early intervention. Factors that contributed to the success of this project included careful selection of a programme with evidence, establishing a local evidence base for it, ensuring that information was disseminated to government and service providers, and the need to build in a sustainability plan. The biggest challenge, lack of leader time and resources to deliver the programme effectively, is explored and solutions from Wales, including leader feedback surveys and manager training days are described.

Over the last twenty years there has been growing recognition that conduct disorder and antisocial behaviour in adolescence and subsequent life-long problems, including violence, criminality, substance misuse and enduring mental health problems, often have their origins in patterns of behaviour that are identifiable in early childhood (Hutchings and Gardner 2012). There is also good evidence that interventions for parents of younger children, who are at risk of poor long-term outcomes, are effective in reducing this risk (Furlong et al. 2012). As a result, over the last fifteen years, early intervention strategies targeting disadvantaged high-risk communities, such as the Sure Start programme, have become part of government strategy across the UK. These developments have coincided with a growing interest in the UK in ensuring that services make use of evidence-based interventions.

By the 1990s researchers in Britain were recognising the need for randomised controlled trials (RCTs) in the mental health field and the Cochrane Collaboration, founded in 1993 to provide systematic reviews of evidence-based health care interventions, began to publish reviews in the child mental health field (e.g. Barlow and Parsons 2003). In 1995 grant-funding was established for health research in

Wales and, by 1998, the National Institute for Health and Clinical Excellence was established to advise the UK National Health Service on appropriate interventions for a wide range of health problems. However, whilst these developments were taking place in relation to health provision, they were slower to filter through to early intervention and social care service providers. As a result, services such as Sure Start were initially set up without any guidance on suitable programmes (Belsky, Barnes, and Melhuish 2007).

Things have started to change in the UK and the demand for publicly funded early intervention services to deliver effective evidence-based programmes is growing (Hutchings and Gardner 2012). However choosing such programmes is only the first step and delivering them effectively at the local level presents a variety of challenges. Research trials of parenting interventions have demonstrated strong evidence with parents of pre-school children (Olds 2006; Hutchings, Bywater, Daley, Gardner et al. 2007) but although many programmes have been demonstrated to work in RCT trials conducted by their developers, few have rigorous evidence of effectiveness when delivered in real world settings. This leaves service providers with two important questions:

Which programme to choose? How to deliver it so as to achieve comparable outcomes to those reported in the research trials? These two issues, choice of programme and how to take interventions to scale and deliver them effectively in service settings, are the subject of the growing field of “implementation science” that has been informed by the Society for Prevention Research and set out in their guidance to service providers on how to ensure that evidence based programmes work in service settings (Flay et al. 2004) and by the work at the University of Colorado Center for Violence Prevention in identifying strongly evidence-based “blueprint” programmes (Mihalic et al. 2002).

This paper provides a case study in the implementation and dissemination of the strongly evidence-based Incredible Years parent, child and teacher programmes across Wales. It describes the author’s work in delivering, researching and supporting the dissemination of the parent programme, and subsequently the child and teacher programmes, with the support of the Welsh government. It sets out the reasons for choosing the programmes, the steps taken to test their effectiveness in service settings across Wales, the dissemination process and lessons learned.

1. Evidence for the Incredible Years Programme

The Incredible Years programme was developed by Webster-Stratton at the University of Washington, Seattle, and has over thirty years of research behind it. It has components for parents, children and teachers and is one of the best evidence-based programmes in the world for both the prevention and treatment of conduct disorder and related difficulties. It has been highlighted in many systematic reviews and is one of only eleven “Blueprint” programmes identified by the Center for Violence Prevention at the University of Colorado, from over nine hundred programmes that they have reviewed. To attain “Blueprint” status, programmes require high standards of evidence, independent replication (ideally in service settings), long-term follow-up and tools to enable effective dissemination (www.colorado.edu/cspv/blueprints/).

Evidence of the effectiveness of the Incredible Years programmes in service settings for both the prevention and treatment of conduct problems comes from several coun-

tries in different continents. Programmes cover a wide age range, from birth to 12 years of age, have demonstrated good outcomes with people from a variety of cultural groups and are effective when delivered in everyday service settings (Webster-Stratton et al. 2001; Webster-Stratton 2011). The parent programme achieves significant improvements in child problem behaviour as well as increases in positive parenting, parental mental health and parenting confidence (Hutchings, Bywater, Daley, Gardner et al. 2007). It also contains all of the components that characterise programmes that achieve better outcomes, particularly with families in disadvantaged circumstances (Gardner et al. 2010). It teaches participants to identify and use social learning theory principles, has a collaborative focus, uses role-play rehearsal of new skills, teaches the accurate observation skills needed for effective problem-solving and realistic goal-setting, emphasises home activities and incorporates strategies to ensure that access issues are addressed (Hutchings, Gardner, and Lane 2004). These three components – curriculum, collaborative delivery and access – contribute to the programme’s effectiveness with the most challenged and high-risk families that other programmes have traditionally failed to help (Hartman, Stage, and Webster-Stratton 2003; Gardner et al. 2010). It has comprehensive fidelity tools to enable effective replication including books, CDs and hand-outs for parents, manuals, standardised training, supervision and consultation, and an effective and rigorous accreditation process for leaders.

2. Establishing the Programme in Wales

2.1 Getting Started

Having worked in the National Health Service since 1976, by 1996 the author held a joint appointment as a Consultant Clinical Psychologist in the North-West Wales Child and Adolescent Mental Health Service (CAMHS) with responsibility for children with conduct disorder and as Director of a research team at Bangor University. In 1999 she started to trial the Incredible Years group parent programme (Webster-Stratton et al. 2001) as a CAMHS-based treatment for parents of children with conduct disorder and related difficulties. At the same time the Sure Start programme was launched in Wales to provide early intervention support for pre-school children in socially disadvantaged communities (Ball 2002). The publication of

Everybody's Business (National Assembly for Wales 2001) identified an expanded role for CAMHS staff, to include supporting primary care staff in the prevention of children's mental health problems, allowing the author to support Sure Start services alongside her work with children with severely challenging behaviour. The next step was for the author to attain programme leader accreditation from the Incredible Years programme developer and to progress to mentor status enabling her to undertake training and supervision of parent group leaders.

Having established the Incredible Years parent programme as a treatment intervention within the CAMHS service, evidence that the programme also functioned as an early preventive intervention provided the author with an opportunity to suggest that it be used in Sure Start services. Basic leader training was delivered by the author to Sure Start staff across North Wales and a peer-support group was established to support programme delivery. Sure Start staff were enthusiastic about the programme, which created a potential research opportunity.

2.2. Evaluating the Programmes

Despite the Incredible Years programme having "Blueprint" status, a number of the "Blueprint" programmes have failed to show good outcomes in service-based replication trials (Mihalic et al. 2002) and the issue of how to deliver evidence-based programmes effectively in service settings was becoming an important research topic (Flay et al. 2004).

Unlike Sure Start projects across England and the rest of Wales, where services were left to decide what to deliver and delivered a plethora of different programmes (Belsky, Barnes, and Melhuish 2007), eleven Sure Start services in North and Mid Wales were delivering the Incredible Years parent programme. This enabled funding to be obtained from the Health Foundation to undertake an RCT evaluation with identified high-risk three and four-year-old children living in these communities. All of the fidelity components were incorporated, including basic leader training, on-going weekly supervision of leaders by the author, provision of all materials for parents and leaders, funding of crèche facilities, lunches for families. Video recordings of sessions were reviewed at supervision and sub-

sequently processed for leader certification which was achieved by twenty-one of the twenty-two leaders in the trial (Hutchings, Bywater, and Daley 2007).

Short- and longer-term outcomes included significant improvements in child and parenting behaviour, parental stress and depression (Hutchings, Bywater, Daley, Gardner et al. 2007; Bywater et al. 2009), replicating those achieved by the programme developer, including similarly high retention rates and good outcomes with the most disadvantaged and hard-to-engage families (Gardner et al. 2010). In a study of a sub-sample of children at risk of ADHD, independent improvements were found in child hyperactivity and inattentivity (Jones et al. 2008). Research to establish the key leader behaviours associated with changes in parent behaviour demonstrated that leader praise and reflective statements, coded from the session videotapes, were associated with similar parental behaviours recorded in the home (Eames et al. 2010). A cost effectiveness analysis showed the programme to have achieved significant benefits at reasonable cost (Edwards et al. 2007). We had shown that a service based intervention, delivered across eleven Sure Start sites with leaders that were locally employed and had relatively low levels of qualifications, was as successful as similar trials by the programme developer.

2.3. Factors Contributing to the Trial Success

This topic is explored in more detail in "Early Prevention of Conduct Disorder: How and Why Did the North and Mid Wales Sure Start Study Work?" (Hutchings, Bywater and Daley 2007); space here permits only a summary. Firstly, there was locally available leader training, supervision and support from the author, a senior and experienced clinical psychologist. Sure Start staff members had already been trained, were running the programme prior to the research trial, and were enthusiastic about it. Sure Start services were newly established and managers were open to advice on how to deliver effective services and seeking evidence of outcomes achieved. They were willing to commit the levels of staff time and resources needed to deliver the programme effectively and to participate in a randomised controlled trial (RCT) in which some families were in a waiting list control condition and would be of-

ferred the intervention after the first six-month follow-up. There was a research team in place at Bangor University with expertise in conducting an RCT and early intervention provision was increasing and accompanied by a demand for evidence. This provided the conditions to enable a successful grant application to the Health Foundation, made collaboratively by the author and the participating services.

Considerable attention was given to the recruitment process to ensure that families of children at risk of longer-term conduct disorder were targeted and recruited. This included having a health visitor on the research team to work with local health visitors, training them in collaborative, non-judgemental recruitment strategies. Health visitors helped parents to recognise that their child presented challenges that made parenting more difficult and that the intervention could help them to support their children. This was crucial in ensuring a high take-up rate, with 93 percent of eligible families with a child scoring within the clinical range on the Eyberg Child Behaviour Inventory (ECBI; Eyberg and Ross 1978) signing up for the trial.

Service access issues were addressed by the provision of all resources needed to enable parents to attend the programme. Transport and crèche facilities were arranged where needed and meals were provided at each group session. In addition implementation fidelity in programme delivery was encouraged through weekly supervision that included viewing videotapes of group sessions and leaders completing session checklists detailing the components delivered during each session. Leaders followed up parents that missed sessions and made weekly phone calls to all participants, and parents completed weekly session evaluations that guided delivery. As a result 83 percent of participants attended seven or more sessions with an overall mean attendance of 9.2 of the 12 sessions. These results with identified high-risk families with high levels of social disadvantage were impressive, particularly since families with all of the characteristics that normally predict poorer outcomes (poverty, single parenthood, young parenthood, maternal depression, etc.) demonstrated outcomes as good as or in some cases better than the sample as a whole (Gardner et al. 2010).

2.4. Impact of the Trial

The success of this trial had an important impact on policy and service development in Wales. The Welsh government monitored the study and, as a result of the extremely positive outcomes, incorporated funding to develop the parent programmes into their Parenting Action Plan for Wales (Department for Training and Education 2005). From 2006 this funded parent group leader training for staff across the twenty-two Welsh local authorities. Services had to commit to both delivering the programme and providing adequate resources for it to be delivered with fidelity. Staff from every county in Wales accessed the training and services in all counties have since delivered the programme and Welsh government funding for leader training and supervision has continued since that time.

In the meantime the research has continued. The Incredible Years Toddler parent programme has been researched, in an RCT, with parents of one- and two-year-olds living in nine Flying Start communities across Wales (Griffith 2011; Griffith et al. 2011; Hutchings, Griffith et al. 2012) and with nursery staff (Bywater, Hutchings, Gridley and Jones 2011). The Incredible Years BASIC parent programme has been evaluated with Welsh foster carers (Bywater et al. 2010) and we have recruited local authorities from across Wales into two on-going research trials: one of the parents and babies programme and one of the school readiness parent programme. In all of these trials the author trained and supervised the service-based staff that delivered the programmes to ensure fidelity of delivery. A study in six authorities in England, using the extended Incredible Years school-aged programme with parents of high challenge 8 to 13 year olds, overseen by the first author also reported good outcomes (Hutchings, Bywater, Williams et al. 2011).

3. From Parent Programmes to Child and Teacher Programmes in Wales

In 2002 we started to deliver and evaluate the child and teacher programmes in Wales and again the author went through the programme developer's accreditation process enabling her to train and support staff delivering these programmes. Pilot trials achieved good outcomes (Hutchings, Lane, Owen and Gwyn 2004; Hutchings, Bywater, Daley and Lane 2007; Hutchings, Daley et al. 2007; Hutchings et al. 2008). An RCT of the Incredible Years teacher classroom

management programme showed positive outcomes in terms of changes in both teacher and child behaviour (Hutchings, Martin-Forbes et al. 2012) and a pilot of the Incredible Years therapeutic small-group Dinosaur school programme was effective (Hutchings, Bywater et al. 2012). This led to funding for a large-scale on-going RCT of this programme targeting high-risk children in twenty-two schools across North and Mid Wales during their early school years (Bywater, Hutchings, Whitaker et al. 2011). In this trial the intervention is delivered by teachers and classroom support staff that are trained and supervised by the first author.

As the results from studies of the child and teacher programmes in Wales filtered through, and because the Incredible Years philosophy matched the Welsh government's developing discovery-based Foundation Phase early years education approach, Welsh government funding for programme leader training incorporated support for the child and teacher programmes with training and resources. Welsh government funding is now in its seventh year and has supported training, supervision and resources for local authorities across Wales in the Incredible Years parent, child and teacher programmes.

4. The Successes and Challenges in Taking the Incredible Years Programme to Scale Across Wales

Many of the recognised challenges of going to scale were reduced as a result of this small-scale bottom-up development leading to the larger RCT of the Sure Start implementation across North Wales. This fed into the 2005 Welsh government parenting strategy. The excellent research outcomes, recognised as being of international significance, meant that the lessons learned about fidelity (Hutchings, Bywater, Daley, Gardner et al. 2007) were taken on board in the discussions with the Welsh government.

4.1. Successes

1. The development across Wales started with the author's identification of a programme with evidence supporting its applicability for the targeted clinical population. After trialling it herself and seeing positive results, she became the "necessary" local champion for the programme in Wales and in particular was able to em-

phasise the necessary fidelity components to ensure effective delivery.

2. Like many of the best evidence-based programmes, the Incredible Years programme is not British and focusing from the start on getting locally available trainers that are accredited by the programme developer is essential. The author worked with the programme developer and was accredited as a mentor within two years, enabling training and supervision to be quickly brought in-house in Wales making it accessible and economically viable.
3. The author was able, as part of her CAMHS clinical psychology post, to ring-fence the time needed to provide support for early implementers. These activities take time and a dedicated person needs to be resourced to do this.
4. Co-ordination of training and support activity within Bangor University, at the Centre for Evidence Based Early Intervention, which the author established and directs, has meant that there is a register of all the people within Wales that have been trained to deliver the various programmes. This has enabled the Centre to maintain contact with them, to keep them updated on training opportunities and research findings, and invite them to annual conferences held in both North and South Wales.
5. The central database of trained staff has made it possible to undertake surveys to obtain feedback from programme deliverers about successes and challenges.
6. The research team at Bangor University, led by the author, were able to demonstrate that the programmes were effective in Wales, overcoming the common criticism that programmes developed overseas are not culturally relevant. Wales is a small country (population three million), making access to government possible and enabling positive results to become known and to have an impact on government policy.
7. The Welsh government was developing a series of policy initiatives (Sure Start, Everybody's Business, the Parenting Action Plan for Wales, Flying Start, the Child Poverty Strategy and others) and was looking for suitable coal-face intervention to support these initiatives and to deliver good outcomes.
8. From small beginnings in one CAMHS service twelve years ago, the Incredible Years parent programme is now

delivered across Wales, with over three hundred groups in 2011 (Hutchings, Williams and Morgan-Lee 2011) and the child and teacher programmes are also developing with supporting research. To date twenty-one of the twenty-two local authorities in Wales have accessed Welsh government funded training in the Incredible Years Teacher Classroom Management programme and twenty have accessed training in the Incredible Years child programmes. There is continued Welsh government funding to support leaders to deliver the programme well and achieve accreditation.

4.2. Challenges

Every local authority in Wales has delivered the parent programmes, mainly within early intervention services, using the parenting programmes targeting parents of babies, toddlers and young children. Despite specification by the Welsh Government of what authorities needed to provide, after training, for their staff to deliver programmes effectively (time, access to supervision and resources), informal feedback from group leaders suggested that not all programme leaders were provided with adequate time and support. This information was collected more formally in three Wales-wide leader surveys to learn about challenges and needs, conducted in 2007, 2009 and 2011 (Hutchings, Williams and Morgan-Lee 2011). The key issue was lack of time for supervision, preparation for sessions and wrap around time for additional scaffolding for high-need families. These surveys record growing understanding of what is required to deliver programmes effectively in terms of supervision, time and other resources. However, with more widespread programme delivery, ensuring that managers know what is needed for effective delivery becomes an ever-present challenge.

Although the survey results were circulated to all managers, this channel proved not to be entirely effective. A more effective strategy has been the development of manager workshops, delivered annually in North and South Wales since 2008, to enable service managers to understand the requirements of effective programme delivery. In these workshops managers learn about the international and local research outcomes, fidelity requirements and strategies for outcome evaluation. In addition they are provided

with information on the resources and activities required to deliver the programmes effectively. Within the workshops managers undertake goal planning in relation to programme delivery that helps them to identify assets and goals in relation to their own service plans.

Any programme that supports families facing significant challenges necessitates skilled staff and significant resources. It is an ever-present challenge to ensure that staff supervision is available from appropriately trained and experienced supervisors. With more people trained across Wales and in particular in South Wales, where the majority of the population live, providing adequate levels of supervision is an on-going challenge. The solution lies in our longer-term dissemination goal, that the twenty-two local authorities in Wales will each establish in-house training and supervision. To date, thirty programme leaders across Wales are either accredited or have completed part of the accreditation process. Six authorities have trained accredited group leaders as peer coaches, to provide in-house supervision, and three authorities have in-house mentors providing leader training. However, as these figures make clear, this leaves the majority of Wales at the very beginning of the road to ensuring fidelity and sustainability for the programmes.

5. Final Comments

What must be done to deliver effective services is known. The Society for Prevention Research guidelines (Flay et al. 2004) and the NICE guidance on how to overcome barriers (NICE 2006) together provide useful information on how to achieve effective programme delivery in service settings. However there are many challenges ahead, not least funding restrictions arising from the present severe cuts in government funds. Early intervention requires sustained support from politicians (Allen 2011) and it is difficult to get sustained commitment when the longer-term goals of reductions in lifelong problems fall outside the political time frame of elected governments.

We have benefitted from being a small country, which made gaining access to political decision-makers easier, and also from having undertaken rigorous research at a time when there was growing recognition of the importance of

outcomes as opposed to outputs. Ten authorities have partnered with us in one or more of our RCTs and our dissemination activity has been successful, with many published research and discussion articles (www.centreforearlyinterventionwales.co.uk).

Across the UK as a whole, there has been a poor take-up of evidence-based interventions for children (Little 2010). One of the recognised challenges of scaling up evidence-based interventions (Little 2010; Bumbarger, Perkins, and Greenberg 2010) is persuading politicians to incorporate

such interventions in their children's services. Many of these challenges were reduced in Wales as a result of this small-scale bottom-up development leading to the larger RCT of the Sure Start implementation across North Wales.

There is still a long way to go to ensure that all of the Welsh families that need help get effective evidence-based services that achieve the outcomes shown in both efficacy and effectiveness trials of the Incredible Years programmes but, with the support of the Welsh government, we have made a great start.

References

- Allen, Graham. 2011. *Early Intervention: The Next Steps – An Independent Report to Her Majesty's Government*. London: HM Government.
- Ball, Mog. 2002. *Getting Sure Start Started*. London: DfES Publications.
- Barlow, Jane, and Jacqueline Parsons. 2003. Group-Based Parent Training Programmes for Improving Emotional and Behavioural Adjustment in 0-3 Year Old Children (Cochrane review). *The Cochrane Library* Issue 1: Oxford, UK: Update Software.
- Belsky, Jay, Jacqueline Barnes, and Edward Melhuish. 2007. *The National Evaluation of Sure Start: Does Area-Based Early Intervention Work?* Bristol: Policy Press.
- Bumbarger, Brian K., Daniel F. Perkins, and Mark T. Greenberg. 2010. Taking Effective Prevention to Scale. In *Handbook of Youth Prevention Science*, ed. Beth Doll, William Pfohl, and Jina Yoon, 433–43. New York: Routledge.
- Bywater, Tracey, Judy Hutchings, David Daley, Christopher J. Whitaker, Karen Jones, and Catrin Eames. 2009. Long-Term Effectiveness of the Incredible Years Parenting Programme in Sure Start Services in Wales with Children at Risk of Developing Conduct Disorder. *British Journal of Psychiatry* 195:1–7.
- Bywater, Tracey, Judy Hutchings, Christopher J. Whitaker, Ceri Evans, and Laura Parry. 2011. Research Protocol: Building Social and Emotional Competence in Young High-Risk School Children: A Pragmatic Randomised Controlled Trial of the Incredible Years Therapeutic (Small Group) Dinosaur Curriculum in Gwynedd Primary Schools, Wales. *Trials* 12 (1): 39–46. doi:10.1186/1745-6215-12-39
- Bywater, Tracey, Judy Hutchings, Nicole Gridley, and Karen Jones. 2011. Delivering the Incredible Years Toddler Parent Programme to Nursery Staff. *Child Care in Practice* 17 (3): 285–302.
- Bywater, Tracey, Judy Hutchings, Pat Linck, Christopher J. Whitaker, David Daley, Seow Tien Yeo, and Rhiannon Tudor-Edwards. 2010. Incredible Years Parent Training Support for Foster Carers in Wales: A Multi-Centre Feasibility Study. *Child Care, Health and Development* 37 (2): 233–43.
- Department for Training and Education. 2005. *Parenting Action Plan: Supporting Mothers, Fathers and Carers with Raising Children in Wales*. Cardiff: Welsh Assembly Government (Information Document No. 054-05). <http://www.learning.wales.gov.uk/pdfs/parenting-action-plan-e.pdf>
- Eames, Catrin, David Daley, Judy Hutchings, Christopher J. Whitaker, Tracey Bywater, Karen Jones, and J. Carl Hughes (2010). The Impact of Group Leaders' Behaviour on Parents' Acquisition of Key Parenting Skills During Parent Training. *Behaviour Research and Therapy* 48 (12): 1221–26. doi:10.1016/j.brat.2010.07.011
- Edwards, Rhiannon Tudor, Alan Ó Céilleachair, Tracey Bywater, Dyfrig A. Hughes, and Judy Hutchings. 2007. Parenting Programme for Parents of Children at Risk of Developing Conduct Disorder: Cost-Effectiveness Analysis. *British Medical Journal* 334: 682–87.
- Eyberg, Sheila M. and A. William Ross, (1978). Assessment of Child Behaviour Problems: The Validation of a new Inventory. *Journal of Clinical Child Psychology* 7 (2): 113–16.
- Flay, Brian R., Anthony Biglan, Robert F. Boruch, Felipe Gonzalez Castro, Denise Gottfredson, Sheppard Kellam, Eve K. Moscicki, Steven Schinke, Jeffrey C. Valentine and Peter Ji. 2004. Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science*, 6 (3): 151–175.
- Furlong, Mairead, Sinead McGilloway, Tracey Bywater, Judy Hutchings, Susan M. Smith, and Michael Donnelly. 2012. Behavioural and Cognitive-Behavioural Group-Based Parenting Programmes for Early-Onset Conduct Problems in Children Aged 3 to 12 Years (Cochrane review). *The Cochrane Library*, Issue 2. Oxford: UK Software.
- Gardner, Frances, Judy Hutchings, Tracey Bywater, and Christopher J. Whitaker. 2010. Who Benefits and How Does it Work? Moderators and Mediators of Outcome in a Randomised Trial of Parenting Interventions in Multiple "Sure Start" Services. *Journal of Clinical Child and Adolescent Psychology* 39 (4): 568–80.
- Griffith, Nia. 2011. *Evaluating the Incredible Years Toddler Programme in Disadvantaged Areas of Wales*. Unpublished doctoral dissertation, Bangor University, UK.
- Griffith, Nia, Judy Hutchings, Karen Jones, and Margiad Williams. 2011. *The Evaluation of the Incredible Years Toddler Parent Programme: Project Protocol*. Bangor, UK: Bangor University. ISBN 978-1-84220-126-8
- Hartman, Renee R., Scott A. Stage, and Carolyn Webster-Stratton. 2003. A Growth Curve Analysis of Parent Training Outcomes: Examining the Influence of Child Risk Factors (Inattention, Impulsivity, and Hyperactivity Problems), Parental and Family Risk Factors. *Journal of Child Psychology and Psychiatry* 44:388–98.
- Hutchings, Judy, Tracey Bywater, and David Daley. 2007. Early Prevention of Conduct Disorder: How and Why Did the North and Mid Wales Sure Start Study Work? *Journal of Children's Services* 2 (1): 4–14.
- Hutchings, Judy, Tracey Bywater, David Daley, Frances Gardner, Christopher J. Whitaker, Karen Jones, Catrin Eames, and Rhiannon Tudor Edwards. 2007. Parenting Intervention in Sure Start Services for Children at Risk of Developing Conduct Disorder: Pragmatic Randomised Controlled Trial. *British Medical Journal* 334 (7595): 678–85.
- Hutchings, Judy, Tracey Bywater, David Daley, and Eleanor Lane. 2007. A Pilot Study of the Webster-Stratton Incredible Years Therapeutic Dinosaur School Programme. *Clinical Psychology Forum* 170:21–24.
- Hutchings, Judy, Tracey Bywater, Catrin Eames, and Pam Martin. 2008. Implementing Child Mental Health Interventions in Service Settings: Lessons from Three Pragmatic Randomised Controlled Trials in Wales. *Journal of Children's Services* 3 (2): 17–27.

- Hutchings, Judy, Tracey Bywater, Nicole Gridley, Christopher J. Whitaker, Pam Martin-Forbes, and Stella Gruffydd. 2012. The Incredible Years Therapeutic Social and Emotional Skills Programme: A Pilot Study. *School Psychology International* 33 (3): 285-293. doi:10.1177/0143034311415899
- Hutchings, Judy, Tracey Bywater, Margiad Elen Williams, Christopher J. Whitaker, and Eleanor Lane. 2011. The Extended School Aged Incredible Years Parent Programme. *Child and Adolescent Mental Health* 16 (3): 136-43.
- Hutchings, Judy, David Daley, Karen Jones, Pam Martin, Tracey Bywater, and Rhaiain Gwyn. 2007. Early Results from Developing and Researching the Webster-Stratton Incredible Years Teacher Classroom Management Training Programme in North Wales. *Journal of Children's Services* 2 (3): 15-26.
- Hutchings, Judy, and Frances Gardner. 2012. Support from the Start: Effective Programmes for Three to Eight Year-Olds. *Journal of Children's Services* 7 (1): 29-41.
- Hutchings, Judy, Frances Gardner, and Eleanor Lane. 2004. Making Evidence-Based Interventions Work. In *Support from the Start: Working with Young Children and Their Families to Reduce the Risks of Crime and Anti-Social Behaviour*, ed. Carole Sutton, David Utting, and David Farrington, 69-79. London: Department for Education and Skills.
- Hutchings, Judy, Nia Griffith, Tracey Bywater, Nicole Gridley, and Christopher J. Whitaker. 2012. *Outcomes from the Incredible Years Toddler Programme with Parents of One and Two Year Old Children Living in Flying Start Areas in Wales*. Unpublished manuscript.
- Hutchings, Judy, Eleanor Lane, Richard Ellis Owen, and Rhaiain Gwyn. 2004. The Introduction of the Webster-Stratton Classroom Dinosaur School Programme in Gwynedd, North Wales. *Education and Child Psychology* 21 (4): 4-15.
- Hutchings, Judy, Pam Martin-Forbes, David Daley, Christopher J. Whitaker, and Margiad Elen Williams. 2012. *The Impact of a Teacher Classroom Management Programme on the Classroom Behaviour of Children With and Without Behaviour Problems*. Unpublished manuscript.
- Hutchings, Judy, Margiad Elen Williams, and Catrin Morgan-Lee. 2011. *Delivering the Incredible Years Programmes in Wales: Background, Research and Feedback from Leaders and Managers*. Bangor University: Centre for Evidence-Based Early Intervention. ISBN: 978-1-84220-129-9
- Jones, Karen, David Daley, Judy Hutchings, Tracey Bywater, and Catrin Eames. 2008. Efficacy of the Incredible Years Basic Parent Training Programme as an Early Intervention for Children with Conduct Disorder and ADHD: Long-Term Follow-Up. *Child: Care Health and Development* 34 (3): 380-90.
- Little, Michael. 2010. Proof Positive. London: Demos. www.demos.co.uk
- Mihalic, Sharon, Abigail Fagan, Katie Irwin, Diane Ballard, and Delbert Elliot. 2002. *Blueprints for Violence Prevention Replications: Factors for Implementation Success*. Boulder, CO: Centre for Study and Prevention of Violence, University of Colorado.
- National Assembly for Wales. 2001. *Everybody's Business: Improving Mental Health Services in Wales: Child and Adolescent Mental Health Strategy*. Cardiff: National Assembly for Wales. Available at <http://www.wales.gov.uk>
- National Institute for Health and Clinical Excellence. 2006. *Parent Training/Education Programmes in the Management of Children with Conduct Disorders*, NICE Technology Appraisal Guidance 102. London: SCIE, NHS. <http://www.nice.org.uk/TA102>
- Olds, David L. 2006. The Nurse-Family Partnership: An Evidence-Based Preventive Intervention. *Infant Mental Health Journal* 27:5-25. doi:10.1002/imhj.20077
- Webster-Stratton, Carolyn, Sharon Mihalic, Abigail, Fagan, David Arnold, Ted Taylor, and Charles Tingley. 2001. *Blueprints for violence prevention, book eleven: The Incredible Years – Parent, Teacher, and Child Training Series*. Boulder, CO: Center for the Study and Prevention of Violence.
- Webster-Stratton, Carolyn. 2011. *The Incredible Years: Parents, Teachers, and Children's Training Series*. Seattle: Incredible Years Inc.

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment

Ferry X. Goossens, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands

Evelien M. J. C. Gooren, Faculty of Psychology and Education, VU University Amsterdam, the Netherlands

Bram Orobio de Castro, Department of Developmental Psychology, Utrecht University, the Netherlands

Kees W. van Overveld, Seminary for Remedial Education, Utrecht, the Netherlands

Goof J. Buijs, The Netherlands Institute for Health Promotion (NIGZ), Woerden, the Netherlands

Karin Monshouwer, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands, and Department of Interdisciplinary Social Science, Utrecht University, the Netherlands

Simone A. Onrust, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands

Theo G. W. M. Paulussen, TNO Quality of Life (Netherlands Organization of Applied Scientific Research), Leiden, the Netherlands

Vol. 6 (2) 2012

Focus:

Evidence-based Developmental Prevention of Bullying and Violence in Europe

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

- ▶ **Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment** Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Implementation of PATHS through Dutch Municipal Health Services: A Quasi-Experiment

Ferry X. Goossens, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands

Evelien M. J. C. Gooren, Faculty of Psychology and Education, VU University Amsterdam, the Netherlands

Bram Orobio de Castro, Department of Developmental Psychology, Utrecht University, the Netherlands

Kees W. van Overveld, Seminary for Remedial Education, Utrecht, the Netherlands

Goof J. Buijs, The Netherlands Institute for Health Promotion (NIGZ), Woerden, the Netherlands

Karin Monshouwer, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands, and Department of Interdisciplinary Social Science, Utrecht University, the Netherlands

Simone A. Onrust, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, the Netherlands

Theo G. W. M. Paulussen, TNO Quality of Life (Netherlands Organization of Applied Scientific Research), Leiden, the Netherlands

Only a limited number of effectiveness studies have been performed to study the benefits of efficacious behavior problems prevention programs for children when implemented through national health service systems. This study uses a quasi-experimental design to test the effectiveness of the school-based PATHS prevention program (Providing Alternative Thinking Strategies) when implemented through Dutch municipal health services by health promotion professionals. A sample of 1,294 children was followed for two years: 674 children attending nine schools providing PATHS and 620 children in nine comparison schools. We hypothesized finding an intervention effect of PATHS in terms of a significant reduction in teacher- and student-rated externalizing and internalizing problem behaviors, and a significant improvement in teacher-, student-, and peer-rated social skills and emotional skills. In fact, the results show low levels of program implementation and no intervention effects on problem behavior or social and emotional skills, suggesting that it is hard to reproduce positive intervention effects where an efficacious social-emotional prevention program is implemented through a national health service. More research is needed on the specific conditions required to implement efficacious programs effectively.

It has been shown that school-based prevention programs aimed at social and emotional learning can be efficacious (Durlak et al. 2011). Usually these programs are evaluated in efficacy trials characterized by optimal conditions, such as well-trained and carefully supervised intervention personnel, and ample resources, and by involvement of the program developers in the implementation process and research (Flay et al. 2005). The better the program im-

plementation and the greater the program fidelity, the stronger the effects are (Beelmann and Raabe 2009; Eisner 2009; Salmivalli, Kaukiainen, and Voeten 2005; Wilson and Lipsey 2007). However, efficacy trials have high internal validity but weaker external validity, for example actual utilization under normal community conditions (Weisz et al. 1995). Very few effectiveness studies have examined the benefits of efficacious prevention programs on a large scale

Acknowledgements: This study was funded by grant number 2430.0031 from the Dutch Health Care Research Organization (ZonMw). Furthermore, we are grateful to all who participated in this study (health promotion professionals, teachers, children, and their parents).

Authors' contributions: F. Goossens carried out the study, analyzed the data, and wrote the manuscript. E. Gooren coordinated the acquisition of data, cleaned the data, and revised the manuscript. B. Orobio de Castro revised the manuscript thoroughly. C. van Overveld (who is employed at the Dutch licensee of PATHS) trained and supported the municipal health

services and revised the manuscript. G. Buijs contributed to the design of the study, developed the implementation materials, and revised the manuscript. K. Monshouwer and S. Onrust revised the manuscript. T. Paulussen designed the study and revised the manuscript thoroughly. All authors have given final approval of the version to be published.

(Dodge 2009). Given that to date few efficacious prevention programs have been successfully implemented on a large scale (Elliot and Mihalic 2004), and their population impact on the prevention of problem behavior is very modest (Dodge 2009), more research on this topic is needed.

1. The PATHS Program

PATHS is a comprehensive school-based prevention program that aims to enhance social and emotional competence and reduce behavioral and mental problems of children from kindergarten to sixth grade. PATHS is based on five conceptual models that form a theoretical framework focusing on a broad range of protective and risk factors for internalizing and externalizing problem behavior. The curriculum provides one to three sessions each week focusing on self-control, emotional understanding, positive self-esteem, relationships, and interpersonal problem-solving skills. Lessons include didactic instruction, role playing, class discussion, modeling by teachers and peers, social and self-reinforcement, and worksheets. The lessons are grouped in three major units: readiness and self-control, feelings and relationships, and interpersonal cognitive problem-solving (Greenberg and Kusché 1993).

1.1. PATHS in the Netherlands

For a couple of decades now, the efficacious school-based prevention program PATHS (Providing Alternative THinking Strategies) (Domitrovich 1999; Kusché and Greenberg 1994) has been implemented in the Netherlands by a small group of dedicated professionals employed at the Dutch national licensee, which has directly trained staff at three hundred schools over the last twenty-five years. However, with approximately seven thousand elementary schools in the Netherlands this is no more than 5 percent of the potential, and its impact on society is therefore probably still limited. The present study examines the effectiveness of PATHS when implemented through health promotion professionals from Dutch municipal health services (MHSs). In this implementation strategy, the national licensee trained MHS health promotion professionals who in turn trained teachers. As it is the daily job of such professionals to implement school-based prevention programs on topics like bullying, alcohol, obesity, sexually transmitted diseases,

and social emotional competence, it was assumed that they could also be a useful link in implementing PATHS.

1.2. PATHS Efficacy and Effectiveness

PATHS has been studied in a number of trials within a variety of populations, including children in regular education, children at risk of behavior problems, children in special education, and children with hearing impairments. As the current study concerns the implementation of PATHS for children in regular education, we concentrate on studies focusing on that target group, as well as studies that examine the implementation process.

The results of the first PATHS study in regular education indicated that the intervention was effective in improving grade 2 and 3 children's range of vocabulary and fluency in discussing emotional experiences, their efficacy beliefs regarding the management of emotions, and their developmental understanding of some aspects of emotions (Greenberg et al. 1995). A second, larger study of 329 second and third graders showed that the intervention promoted inhibitory control, verbal fluency, and diminished internalizing and externalizing problem behaviors (Riggs et al. 2006). A third study, with 246 pre-school children, showed that children exposed to PATHS intervention had higher emotion knowledge skills and were more socially competent than peers (Domitrovich, Cortes, and Greenberg 2007). In all these studies, the program developers were involved in the research, and the level of support was high (i.e. teachers received monthly or even weekly consultation from the project staff to enhance the quality of implementation).

These intensive procedures may have contributed significantly to the implementation quality and the effects found. This assumption is supported by the results of an effectiveness trial involving 350 first graders in six inner-city public schools in a high-risk urban community in the United States (Kam, Greenberg, and Walls 2003). In this study, significant intervention effects were found only in schools where both implementation quality and support from the principal were high. Another study showed that it was not the number of PATHS sessions received, but the quality of these sessions that positively influenced the outcomes (CPPRG 1999).

Interestingly, PATHS with less intense support was also studied. A Dutch study concerning the effectiveness of PATHS for boys with severe aggressive behavior problems showed positive effects on proactive and reactive aggression (Louwe et al. 2007), in a context where training and support were provided directly by the national licensee. However, a negative trend was found for children in special needs schools, where implementation quality was substantially lower than in other schools (Louwe et al. 2007b). A recent study in Zurich showed intervention effects on teacher- and parent-rated aggressive behavior, and teacher-rated ADHD, but no significant positive effects for nine other externalizing and internalizing outcomes. The teachers who implemented PATHS received two days' training, and local coaches were trained to visit the classes and provide feedback to the teachers (Malti, Ribeaud, and Eisner 2011). Finally, the intervention was studied in ten US public elementary schools (SCDRC 2010). Of the twenty child-level outcomes, none was significant. The authors suggest that the lack of positive findings was probably caused by the control condition not being a non-treatment condition, but a standard practice condition including schools that use other social and character development activities. In sum, PATHS has been shown to be efficacious and potentially effective, but effectiveness depends to a large extent on the implementation conditions.

1.3. Aims and Hypotheses

The present study explored the effectiveness of PATHS when implemented through a regular health service system, i.e. Dutch municipal health services. We hypothesized finding an intervention effect of PATHS in terms of a significant reduction in teacher- and student-rated externalizing and internalizing problem behaviors, and a significant improvement in teacher-, student-, and peer-rated social and emotional skills. We further hypothesized finding a larger effect in classes with higher implementation quality, and in classes with higher implementation quantity.

2. Method

The effectiveness of PATHS was assessed using a quasi-experimental design with an intervention ($n = 674$) and a waiting list comparison ($n = 620$) condition, each containing nine regular elementary schools. The intervention

effects were measured during the first two years of PATHS implementation.

2.1. Implementation Strategy

First, the national project team recruited MHSs to participate in the study. All thirty Dutch MHSs were approached by e-mail, and three were willing and able to participate in the study. At each participating MHS, three health promotion professionals were recruited. These nine professionals were facilitated in their task of training and supporting teachers by means of 1) a two-day PATHS course run by the Dutch PATHS licensee, 2) a manual, a model recruitment letter, and a PATHS brochure for recruiting schools, 3) a pre-implementation teacher training protocol, and 4) follow-up support by the Dutch PATHS licensee.

Second, the health promotion professionals sent all principals of elementary schools in their region the PATHS model recruitment letter and the PATHS brochure. If principals and teachers expressed the intention to participate in the study, an implementation agreement was signed between the school and the MHS.

Third, schools were allocated randomly to either the intervention condition or the waiting list comparison condition, with four exceptions. In order to maximize ecological validity, we refrained from randomizing four of the schools. All the schools were very willing to implement PATHS, but two preferred to start two years later for organizational reasons and two were willing to participate only if they could start directly. These requests were complied with.

Fourth, under supervision of the national licensee, the health promotion professionals provided a two-day pre-implementation training course for the teachers in the intervention condition. Directly after this course, all teachers in the intervention condition started giving their PATHS lessons.

Fifth, in the course of each school year the teachers received a half-day booster session. Additionally, the PATHS coordinators received a half-day support session. Each school designated a staff member as PATHS coordinator with the task of supporting the implementation in their school by delivering supervision and feedback to the teachers.

Sixth, each school in the experimental condition organized, in co-operation with the MHS, an information meeting for parents. Parents also received frequent written information about the content of the program.

2.2. The Health Promotion Context

The PATHS strategy had previously been applied successfully to the implementation of a school-based sex education program by MHSs in the Netherlands (Wiefferink et al. 2005). The strategy had a positive impact on teachers' extent of use, as well as their curriculum-related beliefs. Moreover, implementing school-based prevention programs through MHS health promotion professionals is a common strategy for the implementation of school-based prevention programs in the Netherlands. So although this strategy clearly differs from the more intensive procedures advised by the developers – and as generally used in PATHS trials – it is a good example of how prevention programs are implemented in the Dutch context of school-based health promotion and thus complies with our aim of exploring the effectiveness of PATHS when implemented through a regular health service system.

2.3. The Intervention

The version of PATHS used in this study (SvO 2005) consisted of 161 lessons of 20 to 30 minutes, spread over the eight years of elementary school. It was an update of a Dutch translation of the US curriculum for regular schools that had already been in use for several years in the Netherlands, supplemented by a translation of the pre-school PATHS program (Domitrovich et al. 1999). During the study, all children in the intervention condition received the PATHS program for two years, consisting of approximately forty PATHS lessons. As children in the higher grades did not start the PATHS lessons from kindergarten, they received extra lessons to inform them about the story line and basic principles of the PATHS program. Nevertheless, as PATHS is a cyclical program, all major units are discussed each school year. The study proposal was approved by the medical ethical committee, which stated that passive informed consent was adequate.

2.4. Data Collection

Teacher assessments and student assessments were conducted in both conditions: at the start (T0) and the end

(T1) of the first year, and at the start (T2) and the end (T3) of the second year. Each teacher received his/her questionnaires by post and was asked to fill out ten to thirty questionnaires each wave. Child questionnaires were completed in face-to-face interviews with the three youngest cohorts that were followed in this study (kindergarten, grade 1 and grade 3) and by means of a self-report questionnaire for the oldest cohort (grade 5). The student assessments lasted 20 to 30 minutes per child per wave and were conducted by approximately seventy-five trained graduate psychology students who did not know the intervention condition of the school. Questionnaires concerning implementation quantity were sent to all teachers in the intervention condition four times a year. Implementation quality was measured twice, at the end of the first year and at the end of the second year. Children received a gift worth +/- €0.75 at each measurement point. To incentivize school participation, training and materials were provided free of charge (+/- €4,000).

2.5. Measures

2.5.1. Problem Behavior

The Problem Behavior at School Interview (PBSI) (ErasmusMC 2000) is a forty-two-item questionnaire enquiring about externalizing problems: attention deficit and hyperactivity (ADH), oppositional defiant disorder (ODD), conduct problems, and relational aggression, and internalizing problems: anxiety and depression. Teachers rated the child's behavior on a five-point Likert scale ranging from 0 (never applicable) to 4 (often applicable). Cronbach's α in this study varied between .78 and .92. The convergent validity of the PBSI was found to be good as indicated by the correlations between the PBSI and the Teacher's Report Form (Achenbach 1991), which were .75 for externalizing behavior and .55 for internalizing behavior (Witvliet et al. 2010). This measure was used for all cohorts.

The Social Experience Questionnaire – Teacher Report (SEQ-T) (Crick and Grotpeter 1996) was used to measure relational victimization, physical victimization, and prosocial behavior on a 5-point Likert scale ranging from 0 (never applicable) to 4 (often applicable). The relational victimization and physical victimization scales included three items and the prosocial behavior scale four items. Cronbach's α was .87 (relational victimization), .85 (physi-

cal victimization), and .75 (prosocial behavior). This measure was used for all four cohorts.

The twenty-four-item Dimensions of Depression Profile for Children (DDPC) (four-point Likert format) was used to measure the level of depression in the oldest cohort (Harter and Nowakowski 1987). It contained four subscales: depressed mood ($\alpha = .69$), self-blame ($\alpha = .59$), low energy/interest ($\alpha = .75$), low global self-worth ($\alpha = .77$), and one total score ($\alpha = .85$).

2.5.2. Social and Emotional Skills

Peer social preference was obtained using peer nominations of like most and like least, as described by Coie, Dodge, and Coppotelli (1982). Children (cohorts 1 and 4) were asked to nominate an unlimited number of classmates that they liked most and that they liked least. Each child could therefore be nominated by each classmate as “liked most,” “liked least,” or gain no score. For each child, the liked most as well as the liked least nominations were summed and divided by the number of children in the class minus one (self-nomination was not allowed). The standardized liked least score was then subtracted from the standardized liked most score to generate a social preference score. Social preference is generally regarded as a reliable and valid measure of sociometric status (Cillessen and Mayeux 2004; Rubin et al. 2006).

The teacher-based Preschool and Kindergarten Behavior Scale (PKBS) (Merrel 1996) used for cohorts 1 and 2 is designed to assess social skills and problem behaviors in children aged three to six years. In this study, the social skills scale, which has three subscales, was used. The social cooperation scale included twelve items reflecting behaviors and characteristics deemed important in following instructions from adults, cooperating and compromising with peers. The social interaction scale included 11 items reflecting behaviors and characteristics deemed important in gaining and maintaining the acceptance and friendship of others. The social independence scale included eleven items reflecting behaviors and characteristics deemed important in achieving social independence within the domain of the peer group. For all items, responses were based on a four-point scale. The internal consistency of all three subscales was high (respectively $\alpha = .89$, $\alpha = .87$, $\alpha = .86$).

The Head Start Competence Scale (HSCS) (Domitrovich, Cortes, and Greenberg 2001) used for cohorts 1 and 2 is a twelve-item measure of children’s social and emotional skills reflecting interpersonal relationships and emotion regulation. Teachers were asked to indicate on a four-point scale how well each item on the scale described the child. Internal consistency was high ($\alpha = .95$).

Children’s emotional awareness scores were obtained for cohort 1 and 4 using the Levels of Emotional Awareness Scale for Children (LEAS-C), which assesses the complexity of children’s emotional awareness (Bajgar et al. 2005). It contains twelve interpersonal scenarios featuring the child and another person. After each scenario had been presented, the children were asked to describe their own feelings and those of the other person. For each scenario, the self and other response was rated on a scale ranging from 0 to 4. The higher of the scores for the self and the other was taken as the total score for each scenario. In cases where both the self and the other score were 4, a total score of 5 was awarded. The total scores were summed across the scenarios. Cronbach’s α ranged from .89 to .92 over the assessments.

The child-based Difficulties in Emotion Regulation Scale (DERS) (Gratz and Roemer 2004) was used to measure deficits of emotion regulation. It contains thirty-six questions (five-point Likert scale) in six scales: non-acceptance of emotional responses ($\alpha = .73$), difficulties engaging in goal-directed behavior ($\alpha = .82$), impulse control difficulties ($\alpha = .80$), lack of emotional awareness ($\alpha = .78$), limited access to emotion regulation strategies ($\alpha = .74$), and lack of emotional clarity ($\alpha = .61$). This measure was used for the oldest cohort only.

To measure the children’s affective sharing of others’ emotions, a short ten-item version of Bryant’s Empathy Index was used (Bryant 1982; De Wied et al. 2007). This child-based measure was used for the youngest and oldest cohort only. Cronbach’s α was .68.

2.5.3. Implementation Quality

Implementation quality was operationalized as “conceptual use of the program” i.e. to what extent do teachers act according to the PATHS basic principles. Teachers received a

list of ten questions describing daily classroom situations. For each situation, teachers could choose one of four answers that described how they would react in this specific situation. The answers varied from most desired reaction according to the PATHS basic principles (score = 4) to least desired reaction (score = 1). All scores were averaged, resulting in a mean score for the first year and a mean score for the second year. For each class, we calculated one mean score (range 1 – 4) from these two scores.

2.5.4. Implementation Quantity

To measure implementation quantity (i.e. completeness), teachers completed a monthly log describing all the required lessons and elements thereof, recording whether they completed each specific element of each lesson. For both the first year and the second year, intervention completeness was assessed as the completed proportion of all prescribed activities for that year. We summed these two proportions to obtain one total score (range 0 – 2).

2.5.5. Covariates

The Peabody Picture Vocabulary Test – Third Edition (PPVT-III) (Dunn and Dunn 1997) was used to measure verbal ability. This assessment is a well-known and widely used measure of children's receptive vocabulary. A Word Comprehension Quotient (WCQ) score, using age-appropriate norms, was calculated from the raw total number of correct answers. The internal consistency of the PPVT-III standard scores ranged from .92 to .98. Child verbal ability was included as a covariate because of its potential to affect children's performance in the testing situation.

2.6. Statistical Analyses

The data were analyzed using Stata version 11.1 (StataCorp 2009) over all 1,294 students in accordance with the intention-to-treat principle. Missing data were handled through Full Information Maximum Likelihood estimation (FIML). ANOVAs and Chi-square analyses were conducted to check whether there was a balanced distribution of student characteristics and outcome variable values across the two conditions at baseline ($p < .05$).

The data in this study is hierarchically structured, i.e. measurement waves (T0, T1, T2, and T3) are nested within

students, students are nested within classes, and classes are nested within schools. In such a hierarchical structure, standard statistical formulas will underestimate the sampling variance, and consequently lead to significance tests with an inflated alpha level (type 1 error rate). Multilevel models are specifically geared toward the statistical analyses of nested or clustered data (Hox 2010). In the present study we use multilevel mixed-effects linear regression analyses (Twisk 2006) to test, for each of the outcome variables, whether changes over time varied as a function of the intervention.

For each outcome variable we analyzed the whole development of each outcome variable over time in one analysis by calculating the three change scores (T0 – T1, T1 – T2, and T2 – T3) for each variable. However, as there were baseline differences between the conditions we had to correct the change scores and used the analysis of covariance combination approach as described by Twisk and De Vente (2008) for that purpose. In this analytical approach, the change between baseline and the first reassessment is corrected for the baseline value by subtracting the individual baseline value from the first individual change score. For computational reasons only, a correction was also made for the remaining two change scores. These three adjusted change scores per variable were included in a longitudinal multilevel mixed-effects linear regression model. The resulting β coefficient represents the intervention effect over the whole period, i.e. from T0 to T3.

In order to assess short-term effects, the analyses were also performed for the T0 – T1 period alone (i.e. the models included only the T0 – T1 adjusted change score). In addition, we tested for possible moderating effects of group, gender, program completeness, and conceptual use by including interaction terms between these variables and condition in the models (each interaction term was tested in separate models). Because of multiple testing (27 outcomes), the level of statistical significance was set at $p < .01$ in all tests.

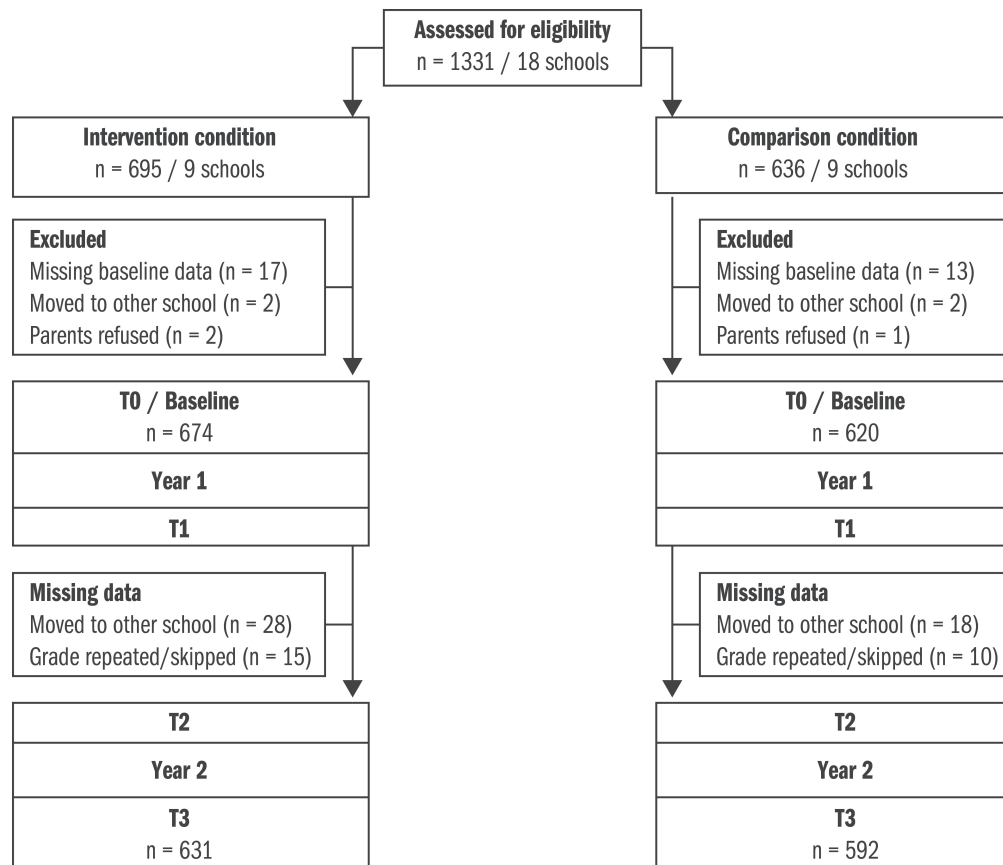
2.7. The Sample

In total, 1,331 children (five to 11 years old) from kindergarten and elementary school grades 1, 3, and 5 were eligible for inclusion (Figure 1). The 18 participating schools were located in rural areas and provincial towns in the

western (Noord-Holland) and eastern part (Gelderland and Overijssel) of the Netherlands. The baseline data for 30 children were missing or incomplete. Four children were excluded because they moved to another school soon after the baseline assessment, and three children's parents refused to allow their children to participate in the study. Therefore the baseline sample included 1,294 children. Of these, 65 changed school during the study. In accordance

with the intention-to-treat principle, we sought to collect data from these children by sending a questionnaire to their home address, asking the parents to fill out the forms. We were able to collect data from nineteen of them. We were also able to collect reassessment data for ten of the thirty-five children who repeated or skipped a grade. Participation in our study was high (97 percent), and the attrition rate low (5 percent at last assessment).

Figure 1: Participation flowchart



Age, gender, ethnicity, and verbal ability did not differ between the intervention and the comparison group (Table 1). Significant baseline differences ($p < .05$) were present with respect to the level of ADHD ($F(1,1292) = 10.443, p < .001$), ODD ($F(1,1292) = 20.896, p < .001$), conduct problems ($F(1,1292) = 30.338, p < .001$), relational aggression ($F(1,1292) = 60.891, p < .001$), anxiety ($F(1,1293) = 11.400, p < .001$), depression ($F(1,1292) = 10.161, p < .001$), relational victimization ($F(1,1292) = 45.082, p < .001$),

physical victimization ($F(1,1292) = 45.594, p < .001$), prosocial behavior ($F(1,1292) = 5.098, p < .05$), low energy ($F(1,1292) = 4.924, p < .05$), social interaction ($F(1,1292) = 4.722, p < .05$), social independence ($F(1,1292) = 4.725, p < .05$), and social and emotional skills ($F(1,1292) = 10.691, p < .01$). In general, the levels of problem behavior were higher in the intervention condition, and the levels of social and emotional skills were lower (Table 2). These differences were statistically corrected for in the analyses.

Table 1: Socio-demographic variables at baseline by cohort

		Intervention condition	Comparison condition
Cohort 1	n	158	166
(Kindergarten)	Male	56.3 percent	52.4 percent
	Age	5.4 years (SD = .36)	5.5 years (SD = .31)
	Dutch	98.6 percent	92.4 percent
Cohort 2	n	159	151
(Grade 1)	Male	52.2 percent	46.4 percent
	Age	6.5 years (SD = .38)	6.5 years (SD = .42)
	Dutch	96.7 percent	95.1 percent
Cohort 3	n	173	152
(Grade 3)	Male	49.1 percent	56.6 percent
	Age	8.6 years (SD = .53)	8.5 years (SD = .50)
	Dutch	92.1 percent	93.2 percent
Cohort 4	n	184	151
(Grade 5)	Male	44 percent	47 percent
	Age	10.5 years (SD = .42)	10.6 years (SD = .45)
	Dutch	94.4 percent	94.5 percent

3. Results

Table 2 shows the outcome variables over time for both conditions. A decline in the PBSI, SEQ-T (except prosocial behavior scale), DDPC, and DERS scores represents a decrease in these problem behaviors/skills. An increase in the PEER, PKBS, HSCS, LEAS, EMPATHY, and prosocial behavior scale (SEQ-T) scores represents an improvement in these skills.

Table 2: Mean scores (M) and standard deviations (SD) at baseline and reassessments for all four cohorts combined

Measure	Scale	Intervention condition					Comparison condition				
		T0		T1	T2	T3	T0		T1	T2	T3
		M	SD	M	M	M	M	SD	M	M	M
PBSI	ADH*	1.26	(.81)	1.00	1.07	1.11	1.11	(.84)	.97	1.10	1.07
	ODD*	1.11	(.75)	.96	.98	.98	.91	(.79)	.88	.88	.89
	Conduct problems*	.62	(.55)	.51	.54	.56	.45	(.52)	.45	.41	.42
	Relational aggression*	1.12	(.73)	.94	1.05	1.04	.80	(.76)	.82	.89	.86
	Anxiety*	1.29	(.71)	1.10	1.20	1.15	1.16	(.72)	1.09	1.13	1.01
	Depression*	.92	(.64)	.86	.95	.92	.80	(.69)	.81	.84	.82
SEQ-T ¹	Relational victimization*	.87	(.69)	.72	.	.79	.62	(.63)	.61	.	.62
	Physical victimization*	.57	(.55)	.	.45	.50	.37	(.50)	.	.31	.29
	Prosocial behavior*	2.68	(.66)	.	2.78	2.78	2.76	(.66)	.	2.76	2.83
DDPC	Depressed mood	10.63	(2.99)	10.37	10.06	9.98	10.35	(3.00)	9.93	9.82	9.88
	Self blame	15.31	(2.74)	14.43	14.17	14.44	15.01	(3.09)	14.63	14.28	14.27
	Low energy*	11.08	(3.46)	10.68	10.45	10.53	10.25	(3.25)	10.40	9.78	10.19
	Low self-worth	10.23	(3.30)	9.71	9.71	9.79	9.70	(3.25)	9.44	9.43	9.50
	Total depression	47.25	(9.14)	45.24	44.39	44.73	45.31	(8.89)	44.39	43.31	43.90
PEER	Peer nominations	.11	(.27)	.15	.18	.15	.13	(.29)	.17	.18	.17
PKBS	Social cooperation	31.65	(4.35)	32.59	32.35	32.24	32.25	(4.00)	32.69	32.02	32.32
	Social interaction*	24.45	(5.02)	25.76	25.44	26.07	24.54	(5.22)	25.46	24.79	25.55
	Social independence*	27.95	(4.11)	29.07	28.58	28.57	28.35	(4.16)	28.91	28.15	28.87
HSCS	Social and emotional skills*	21.61	(7.39)	23.63	23.55	23.98	24.47	(7.06)	25.08	24.45	25.48
LEAS	Emotional awareness	25.04	(11.27)	31.47	34.03	35.90	23.84	(10.69)	28.10	30.15	30.98
DERS	Lack emotional awareness	21.31	(5.67)	22.50	23.60	24.51	20.93	(5.50)	22.49	22.20	22.43
	Non-acceptance of emotional responses	11.91	(4.46)	10.86	10.14	9.91	11.61	(4.63)	10.50	10.04	10.7
	Difficulty in goal-directed behavior	13.19	(5.14)	12.73	11.87	11.30	13.19	(5.33)	12.13	11.54	11.51
	Impulse control difficulties	13.16	(5.09)	12.69	12.49	11.94	13.60	(5.67)	12.17	11.57	11.90
	Limited access strategies	17.42	(5.59)	17.20	15.99	15.83	17.60	(5.97)	15.97	15.38	15.76
	Lack of emotional clarity	15.52	(4.32)	14.54	14.64	14.60	15.35	(4.16)	14.19	14.06	14.44
EMPATHY	Empathy	5.20	(2.36)	5.21	4.91	4.94	4.99	(2.32)	5.14	5.19	5.24

1 To limit the number of questions teachers had to answer each wave, this questionnaire was sent at two instead of three reassessments.

* Significant difference between intervention condition and comparison condition at baseline ($p < .05$).

First, intervention effects were examined by longitudinal multilevel mixed-effects linear regression analyses (Table 3). We combined all four cohorts and tested whether changes over time varied as a function of the intervention.

This first set of analyses resulted in a positive intervention effect found on the emotional awareness scale. Note that this is the only significant intervention effect out of twenty-seven tests.

Table 3: Intervention effect (β coefficient) over all four assessments for all four cohorts combined

Measure	Scale	β	SE	z	ρ	95 percent CI	
PBSI	ADH	-.022	.029	-0.74	0.458	-.079	.035
	ODD	-.012	.028	-0.42	0.675	-.067	.043
	Conduct problems	.008	.020	0.41	0.681	-.032	.048
	Relational aggression	.005	.030	0.18	0.857	-.054	.065
	Anxiety	.017	.032	0.53	0.599	-.045	.078
	Depression	.005	.031	0.17	0.868	-.055	.065
SEQ-T	Relational victimization	.025	.039	0.64	0.522	-.051	.101
	Physical victimization	.102	.071	1.45	0.148	-.036	.241
	Prosocial behavior	.074	.075	0.99	0.323	-.073	.222
DDPC	Depressed mood	.016	.168	0.09	0.925	-.314	.345
	Self blame	.114	.173	0.66	0.509	-.225	.454
	Low energy	-.032	.182	-0.18	0.860	-.388	.324
	Low self-worth	.090	.188	0.48	0.632	-.279	.460
	Total depression	.104	.473	0.22	0.825	-.823	1.032
PEER	Peer nominations	.004	.012	0.33	0.738	-.019	.027
PKBS	Social cooperation	-.065	.174	-0.37	0.710	-.280	.411
	Social interaction	.466	.295	1.58	0.115	-.113	1.046
	Social independence	.187	.276	0.68	0.498	-.353	.727
HSCS	Social and emotional skills	.111	.328	0.34	0.736	-.532	.753
LEAS	Emotional awareness	1.186	.393	3.02	0.003 *	.416	1.957
DERS	Lack emotional awareness	.694	.305	2.27	0.023	.095	1.292
	Non-acceptance of emotional responses	-.059	.245	-0.24	0.809	-.539	.421
	Difficulty in goal-directed behavior	-.139	.277	-0.50	0.616	-.683	.404
	Impulse control difficulties	.135	.297	0.46	0.648	-.446	.717
	Limited access strategies	.123	.332	0.37	0.712	-.528	.774
	Lack of emotional clarity	.141	.298	0.47	0.635	-.442	.725
EMPATHY	Empathy	-.200	.114	-1.76	0.079	-.422	.023

* Significant difference between intervention condition and comparison condition ($p < .01$).

Note: A minus sign (-) before the ρ of the PBSI, SEQ-T (except prosocial behavior scale), DDPC, and DERS represents a decrease in these problem behaviors/skills in the intervention condition over time, relative to the control condition.

Second, we checked whether there was a positive effect in the first year of implementation. We performed the same analyses as above, but limited them to the T0 – T1 adjusted change score. We found no intervention effects on any of the variables in the first year.

Third, we tested whether intervention effects were moderated by grade. We found three significant moderations on the prosocial scale. Effects were stronger for children in kindergarten in the intervention condition and weaker for children in grades 1 and 5 in the intervention condition. The analyses also showed that the effects for the LEAS were stronger for children in kindergarten. Otherwise, grade did not moderate outcomes.

Fourth, we checked whether there was a difference in effect for boys and girls by adding an interaction term (sex x condition). We did not find any gender-related difference.

Fifth, we tested whether intervention effects depended on the level of program completeness. Mean completeness was 50 percent in the first year (SD = 23 percent) and 49 percent in the second year (SD = 24 percent). We summed these two proportions to obtain one score (mean = .99, SD = .42, range .16 – 1.70), and tested whether this score was related to intervention effects. The analyses showed that program completeness did not moderate intervention effects.

Sixth, we tested whether intervention effects depended on implementation quality. The mean level of “conceptual use” was around 3.05 in the first year (SD = .27, range 2.30 – 3.75) and 3.07 in the second year (SD = .37, range 2.20 – 3.70). We calculated a mean of these two scores for each class and tested whether this score was related to intervention effects. Conceptual use did not moderate intervention effects.

4. Discussion

The purpose of this study was to examine the effectiveness of the efficacious school-based prevention program PATHS when implemented through Dutch municipal health services. In contrast to other studies, the intervention was implemented by health promotion professionals from a regular health service system rather than dedicated PATHS

trainers. As health promotion professionals implement school-based prevention programs on a regular basis, this linkage group could in the long run potentially support the national licensee in the dissemination of PATHS. However, virtually no intervention effects were found in this study.

It seems unlikely that the underlying theoretical model of the PATHS intervention can explain the lack of intervention effects. Previous research has demonstrated that PATHS is effective when delivered adequately. The implementation process therefore seems a more likely reason. In the present study, there appears to be a large gap between the intended intervention and the intervention participants actually received. The implementation strategy most probably affected the teacher support negatively. This resulted in low program completeness and probably affected some other implementation variables negatively as well. This finally resulted in poor intervention outcomes.

Although we cannot exactly identify the (implementation) variables that moderated the outcomes, we can conclude that the tested PATHS implementation strategy is not a recipe for effective prevention of problem behavior on a large scale.

4.1. Strengths and Limitations

This is the first study to test the effectiveness of PATHS when implemented through a regular health service, and sought to explore how problem behavior could be prevented on a large scale. Furthermore, this study had very high participation (97 percent) and low attrition (5 percent). We also had a fairly large sample that gave us enough power to detect small effects, even for the four grades/cohorts separately. It must further be recognized that the results are fairly comprehensive, and the measures included constructs of social and emotional skills and problem behavior of both teachers and children, on four occasions.

However, the study also suffers from important limitations. First, we did not randomize all schools. This may have caused baseline differences between the two conditions. However, we corrected for these differences with a thorough statistical technique specially developed for this kind

of problem (Twisk and De Vente 2008). A second limitation is the limited duration of implementation. As PATHS is an eight-year curriculum, two years of implementation may have been insufficient to achieve the desired change. However, previous studies were able to report positive outcomes within a similar timespan. Third, instead of using independent observations, self-reporting was used to assess teachers' quality and quantity of program implementation. Self-reports may have led to an over-estimation of the quantity and quality of implementation. Besides, our number of implementation measures was limited. Fourth, although we did not monitor this, the lack of effects could have been caused by the level of standard social and character development activities that were given in the control group, for example by other preventive programs. Nonetheless, the practical value of PATHS would have been reflected in additional effects on top of care as usual, and this is exactly what we studied.

4.2. Conclusions

It seems unfortunate that the combination of an efficacious prevention program and a health service system specifically designed to be a linking agent for the implementation of school-based health promotion interventions does not provide better results – especially as universal prevention programs can only have population impact on the prevention of problem behavior when they are implemented on a large scale.

One could question whether this implementation strategy was the best possible. Obviously, the implementation strategy could have been enhanced with greater levels of support for both health promotion professionals and teachers. However, our aim was not to study the effectiveness of an intervention using a theoretically optimal implementation strategy, but rather to study its effectiveness when employing an implementation strategy that has been used successfully before, but, even more importantly, is commonly used and achievable within the Dutch public health sector. Along with efficacy trials, this kind of study helps us to identify interventions and implementation strategies that could be helpful (or not) in preventing problem behavior.

In addition, we believe that the present findings are probably not just relevant for PATHS or limited to prevention in the realm of social behavior. Implementation seems likely to play a key role in establishing societal impact. Our study underlines the importance of studying the process of transferring outcomes of efficacy studies to the more naturalistic settings for program implementation, and of monitoring program application in different countries and settings. More, well-designed large-scale field trials are urgently needed to provide policymakers with realistic estimates of the investments required to obtain intervention effects that are replicable at population level.

References

- Achenbach, Thomas M. 1991. *Manual for the Teacher's Report Form and 1991 Profiles*. Burlington: University of Vermont, Department of Psychiatry.
- Bajgar, Jane, Joseph Ciarrochi, Richard Lane, and Frank P. Deane. 2005. Development of the Levels of Emotional Awareness Scale for Children (LEAS-C). *British Journal of Developmental Psychology* 23:569–86. doi:10.1348/026151005X35417
- Beelmann, Andreas, and Tobias Raabe. 2009. The Effects of Preventing Antisocial Behavior and Crime in Childhood and Adolescence: Results and Implications of Research Reviews and Meta-Analyses. *European Journal of Developmental Science* 3:260–81. doi:10.3233/DEV-2009-3305
- Bryant, Brenda K. 1982. An Index of Empathy for Children and Adolescents. *Child Development* 53 (2): 413–25. doi:10.2307/1128984
- Cillissen, Antonius H. N., and Lara Mayeux. 2004. Sociometric Status and Peer Group Behavior: Previous Findings and Current Directions. In *Children's Peer Relations: From Development to Intervention*, ed. Janis B. Kupersmidt and Kenneth A. Dodge, 3–20. Washington, DC: American Psychological Association.
- Coie, John D., Kenneth A. Dodge, and Heide Coppotelli. 1982. Dimensions and Types of Social Status: A Cross-Age Perspective. *Developmental Psychology* 18:557–70.
- CPPRG (Conduct Problems Prevention Research Group). 1999. Initial Impact of the Fast Track Prevention Trial for Conduct Problems: II. Classroom Effects. *Journal of Consulting and Clinical Psychology* 67:648–57. doi:10.1037/0022-006X.67.5.648
- Crick, Nicki R., and Jennifer K. Grotpeter. 1996. Children's Treatment by Peers: Victims of Relational and Overt Aggression. *Development and Psychopathology* 8:367–80. doi:10.1017/S0954579400007148
- De Wied, Minet, Cora Maas, Stephanie Van Goozen, Marjolijn Vermande, Rutger Engels, Wim Meeus, Walter Matthys, and Paul Goudena. 2007. Bryant's Empathy Index: A Closer Examination of its Internal Structure. *European Journal of Psychological Assessment* 23:99–104. doi:10.1027/1015-5759.23.2.99
- Dodge, Kenneth A. 2009. Community Intervention and Public Policy in the Prevention of Antisocial Behaviour. *Journal of Child Psychology and Psychiatry* 50:194–200. doi:10.1111/j.1469-7610.2008.01985.x
- Domitrovich, Celene E., Rebecca C. Cortes, and Mark T. Greenberg. 2001. Head Start Competence Scale Technical Report. Unpublished manuscript, Pennsylvania State University.
- Domitrovich, Celine E., Rebecca C. Cortes, and Mark T. Greenberg. 2007. Improving Young Children's Social and Emotional Competence: A Randomized Trial of the Preschool "PATHS" Curriculum. *Journal of Primary Prevention* 28 (2): 67–91. doi:10.1007/s10935-007-0081-0
- Domitrovich, Celine E., Mark T. Greenberg, Rebecca C. Cortes, and Carol Kusché. 1999. *Manual for the Preschool PATHS Curriculum*. South Deerfield, MA: Channing-Bete.
- Dunn, Lloyd M., and Leota M. Dunn. 1997. *Peabody Picture Vocabulary Test – Revised*. Circle Pines, MN: American Guidance Service.
- Durlak, Joseph A., Roger P. Weissberg, Allison B. Dymnicki, Rebecca D. Taylor, and Kriston B. Schellinger. 2011. The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions. *Child Development* 82:405–32. doi:10.1111/j.1467-8624.2010.01564.x
- Eisner, Manuel. 2009. No Effects in Independent Prevention Trials: Can We Reject The Cynical View? *Journal of Experimental Criminology* 5:163–83. doi:10.1007/s11292-009-9071-y
- Elliott, Delbert S., and Sharon Mihalic. 2004. Issues in Disseminating and Replicating Effective Prevention Programs. *Prevention Science* 5:47–53. doi:10.1023/B:PREV.0000013981.28071.52
- ErasmusMC. 2000. *Problem Behaviour at School Interview*. Rotterdam: Department of Child and Adolescent Psychiatry.
- Flay, Brian R., Anthony Biglan, Robert F. Boruch, Felipe González Castro, Denise Gottfredson, Sheppard Kellam, Eve K. Mo cicki, Steven Schinke, Jeffrey C. Valentine, and Peter Ji. 2005. Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science* 6:151–75. doi:10.1007/s11121-005-5553-y
- Gratz, Kim L., Lizabeth Roemer. 2004. Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure and Initial Validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioural Assessment* 26:41–54. doi:10.1023/B:JOBA.0000007455.08539.94
- Greenberg, Mark T., and Carol A. Kusché. 1993. *Promoting Social and Emotional Development in Deaf Children: The PATHS Project*. Seattle, WA: University of Washington Press.
- Greenberg, Mark T., Carol A. Kusché, Elizabeth T. Cook, and Julie P. Quamm. 1995. Promoting Emotional Competence in School-Aged Children: The Effects of the PATHS Curriculum. *Development and Psychopathology* 7:117–136. doi:10.1017/S0954579400006374
- Harter, Susan, and M. Nowakowski. 1987. *Manual for the Dimensions of Depression Profile for Children and Adolescents*. Denver, CO: University of Denver, Department of Psychology.
- Hox, Joop J. 2010. *Multilevel Analysis. Techniques and Applications*, 2d ed. New York: Routledge.
- Kam, Chi-Ming, Mark T. Greenberg, and Carla T. Walls. 2003. Examining the Role of Implementation Quality in School-Based Prevention using the PATHS Curriculum. *Prevention Science* 4:55–63. doi:10.1023/A:1021786811186
- Kusché, Carol A., and Mark T. Greenberg. 1994. *The PATHS Curriculum*. South Deerfield, MA: Channing-Bete.
- Louwe, Jos J., Cornelis W. van Overveld, Welmoet Merk, Bram Orobio de Castro, and Willem Koops. 2007a. De invloed van het Programma Alternatieve Denkstrategieën op reactieve en proactieve agressie bij jongens in het primair onderwijs: Effecten na één jaar. *Pedagogische Studiën* 84:277–92.
- Louwe, Jos J., Cornelis W. van Overveld, Welmoet Merk, Bram Orobio de Castro, and Willem Koops. 2007b. Het PAD van de school en de leraar. In *Jaarboek Lectoraten Faculteit Educatie 2007*. Utrecht: Hogeschool Utrecht.
- Malti, Tina, Denis Ribeaud, and Manuel Eisner. 2011. The Effectiveness of Two Universal Preventive Interventions in Reducing Children's Externalizing Behaviour: A Cluster Randomized Controlled Trial. *Journal of Clinical Child and Adolescent Psychology* 40 (5): 677–92. doi:10.1080/15374416.2011.597084
- Merrill, Kenneth W. 1996. Socio-Emotional Assessment in Early Childhood: The Preschool and Kindergarten Behaviour Scales. *Journal of Early Intervention* 20:132–45. doi:10.1177/105381519602000205
- Riggs, Nathaniel R., Mark T. Greenberg, Carol A. Kusché, and Mary Ann Pentz. 2006. The Mediation Role of Neurocognition in the Behavioural Outcomes of a Social-Emotional Prevention Program in Elementary School Students: Effects of the PATHS Curriculum. *Prevention Science* 7:91–102. doi:10.1007/s11121-005-0022-1
- Rubin, Kenneth H., Julie C. Wojslawowicz, Linda Rose-Krasnor, Cathryn Booth-LaForce, and Kim B Burgess. 2006. The Best Friendships of Shy/Withdrawn Children: Prevalence, Stability, and Relationship Quality. *Journal of Abnormal Child Psychology* 34 (2): 143–57. doi:10.1007/s10802-005-9017-4
- Salmivalli, Christina, Ari Kaukiainen, and Marinus Voeten. 2005. Anti-Bullying Intervention: Implementation and Outcome. *British Journal of Educational Psychology* 75:465–87. doi:10.1348/000709905X26011
- SCDR (Social and Character Development Research Consortium). 2010. *Efficacy of School Wide Programs to Promote Social and Character Development and Reduce Problem Behaviour in Elementary School Children (NCER 2011–2001)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.
- StataCorp. 2009. *Stata Statistical Software: Release 11*. College Station, TX: Stata-Corp LP.
- SvO (Seminarium voor Orthopedagogiek). 2005. *PAD, Programma Alternatieve Denkstrategieën, Lessen voor sociale en emotionele vaardigheden in de basisschool*. Utrecht: Trimbos Institute.
- Twisk, Jos W. R. 2006. *Applied Multilevel Analysis*. Cambridge: Cambridge University Press.

- Twisk, Jos W. R., and Wieke De Vente. 2008. The Analysis of Randomised Controlled Trial Data with More than One Follow-Up Measurement. A Comparison Between Different Approaches. *European Journal of Epidemiology* 23:655–60. doi:10.1007/s10654-008-9279-6
- Weisz, John R., Geri R. Donenberg, Susan S. Han, and Danika Kauneckis. 1995. Child and Adolescent Psychotherapy Outcomes in Experiments versus Clinics: Why the Disparity? *Journal of Abnormal Child Psychology* 23:83–106. doi:10.1007/BF01447046
- Wiefferink, Karin. H., Jos Poelman, M. Linthorst, Ine. Vanwesenbeeck, J. C. M. van Wijngaarden, and Theo G. W. Paulussen. 2005. Outcomes of a Systematically Designed Strategy for the Implementation of Sex Education in Dutch Secondary Schools. *Health Education Research* 20 (3): 323–33. doi:10.1093/her/cyg120
- Wilson, Sandra J., and Mark W. Lipsey. 2007. School-Based Interventions for Aggressive and Disruptive Behavior. Update of a Meta-Analysis. *American Journal of Preventive Medicine* 33 (2S): S130–43. doi:10.1016/j.amepre.2007.04.011
- Witvliet, Miranda, Pol A. C. van Lier, Pim Cuijpers, and Hans M. Koot. 2010. Change and Stability in Childhood Clique Membership, Isolation from Cliques, and Associated Child Characteristics. *Journal of Clinical Child and Adolescent Psychology* 39:12–24. doi:10.1080/15374410903401161

Ferry X. Goossens

fgoossens@trimbos.nl

Evelien M. J. C. Gooren

emjc.gooren@psy.vu.nl

Bram Orobio de Castro

b.orobiodecastro@uu.nl

Kees W. van Overveld

kees.vanoverveld@hu.nl

Goof J. Buijs

gbuijs@nigz.nl

Karin Monshouwer

kmonshouwer@trimbos.nl

Simone A. Onrust

sonrust@trimbos.nl

Theo G. W. M. Paulussen

theo.paulussen@tno.nl

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors

Tina Malti, Department of Psychology, University of Toronto, Canada

Denis Ribeaud, Institute of Sociology, Swiss Federal Institute of Technology Zurich (ETH), Switzerland

Manuel Eisner, Institute of Criminology, University of Cambridge, United Kingdom

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

► **Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors** Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Ilwilade (pp. 338 – 350)

Open Section



Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors

Tina Malti, Department of Psychology, University of Toronto, Canada

Denis Ribeaud, Institute of Sociology, Swiss Federal Institute of Technology Zurich (ETH), Switzerland

Manuel Eisner, Institute of Criminology, University of Cambridge, United Kingdom

An evaluation of the effectiveness of a school-based social competence curriculum PATHS (Promoting Alternative Thinking Strategies) on teacher-rated aggressive behavior, ADHD, and prosocial behavior in children. The one-year prevention program was administered to children in 28 of 56 Swiss elementary schools (N = 1,675). Outcomes were assessed at pretest and posttest with a follow-up 2 years later. Moderator interactions involving baseline child characteristics and economic factors were tested. There were significant treatment effects for ADHD/impulsivity and aggression at the follow-up. Baseline development variables predicted higher prosocial behavior as well as lower aggressive behavior and ADHD at the follow-up. Economic risk factors predicted poor behavior outcomes at the follow-up. Development variables moderated the impact of PATHS on ADHD and aggression at the follow-up. However, for most outcomes, no main effects or moderation of treatment effects were found.

Developmental scientists have noted the interaction between child characteristics and ecological conditions in the development of psychopathology and social competence (Bronfenbrenner and Evans 2000). The need to prevent children's mental health problems is broadly acknowledged as well, and a number of school-based curricula have been designed to prevent the development of problem behaviors and mental disorders such as aggression and ADHD. Meta-analyses of school-based aggression intervention programs indicate that programs that focus on the promotion of social skills reduce aggressive behavior and mental health problems (Hahn et al. 2007). What is missing are dissemination trials by researchers evaluating these programs independently of the program developers, using rigorous methodological designs and implementing the program "as marketed" (Eisner, Malti, and Ribeaud 2011). Additionally, relatively few large-scale studies have investigated the moderating role of child characteristics at baseline and economic factors in determining program outcomes, and most of the existing

studies were conducted in the United States. We therefore examined how certain child characteristics (baseline behavior, initial social-cognitive development, initial emotional development) and socioeconomic and demographic factors (socioeconomic status, financial problems, single-parent household) can moderate the impact of a universal intervention.

1. The Intervention

Our study employed PATHS (Promoting Alternative Thinking Strategies), which is a school-based universal intervention program that is widely reported to reduce mental health problems and improve social competence in primary-school children (Greenberg and Kusché 2002). It was chosen because it is evidence-based and was assessed positively in a feasibility study conducted in collaboration with the school authorities of the City of Zurich, Switzerland (Eisner et al. 2011).

PATHS relies on an integrative model of children's risk-and-resiliency development. The underlying assumption is that

the promotion of various aspects of social development helps to reduce a set of well-known risk factors for aggression (Greenberg et al. 2003). These risk factors include poor social-cognitive skills (Crick and Dodge 1996), poor emotional skills (Arsenio, Gold, and Adams 2006), and poor inhibition control (Riggs et al. 2006). To reduce these risks, the PATHS lessons promote social-cognitive development, positive social behavior, and understanding of emotions. PATHS is one of only eleven programs recommended as effective by Blueprints of Violence Prevention at the University of Colorado (Greenberg and Kusché 2002). Several rigorous trials of PATHS have been undertaken (Conduct Problems Prevention Research Group 2002; Riggs et al. 2006). Taken together, these trials show that PATHS has a strong evidential base. However, in contrast to the present study, most of these evaluations were supervised by the developer of the program and were conducted in the United States.

2. Moderators of Program Impact

There is emerging evidence that children at high risk for behavior problems are the most likely to benefit from school-based interventions. For example, two large-scale prevention projects for aggressive elementary grade school children achieved long-term reduction in antisocial behavior with their interventions, which included social competence training (Conduct Problems Prevention Research Group 2002, Conduct Problems Prevention Research Group and Dodge 2007). The MACS (2002) study, which included universal school-based prevention programs plus an intensive intervention for students with high levels of initial aggression, showed reduction of aggressive behavior. Because of this evidence that intervention outcomes are moderated by initial behavioral risk, the present study included the child's initial behavior as a moderator of intervention effects.

Contemporary research on the prevention of aggressive behavior also emphasizes the importance of social-cognitive and emotional development in reducing aggression. For example, studies have shown that social-cognitive problems and lack of emotional skills predict increased levels of aggression over time (Lansford et al. 2006). This research, combined with the PATHS approach to promoting improvement in social-cognitive development and emotional literacy, makes it necessary to explore the possible moder-

ation of social competence interventions by initial social-cognitive and emotional development. The present study therefore included two indicators of social-cognitive and emotional development: social problem-solving and moral emotions. These domains were chosen because children's social problem-solving is related to aggressive behavior (Orobio de Castro et al. 2002). Likewise, emotions about moral conflicts (i.e., emotions that children attribute to an actor as a consequence of aggression) have been shown to reduce aggression (Arsenio et al. 2006; Malti and Krettenauer 2012). To the best of our knowledge, no independent, large-scale randomized controlled trial of PATHS has investigated this possible moderation of behavioral, social-cognitive, and emotional development in children.

In addition, it is necessary to examine economic factors as possible moderator variables. Research suggests that up to 25 percent of children living in economic hardship have negative mental health outcomes (Costello et al. 1996; Keenan et al. 1997). It is thus important to understand whether the effectiveness of a given intervention varies as a function of the socio-economic background of the children.

3. The Present Study

Our study was designed to compare the effectiveness of the PATHS program to a control group using a factorial design with a post-test and two-year post-intervention follow-up. We tested moderators of program effects (child characteristics and economic factors). We predicted that children in the treatment (PATHS) condition would manifest greater reductions in aggressive behavior and ADHD/impulsivity, and greater increases in prosocial behavior, than children in the control condition. Based on the results of previous randomized controlled trials of PATHS in the United States, we hypothesized that the reduction in aggressive behavior and ADHD would be greatest among children showing the highest levels of aggressive behavior and ADHD at the beginning of the study. In addition, we expected that child characteristics (social-cognitive skills and moral emotions) as well as economic factors would play a moderating role in the effectiveness of the PATHS intervention. Socioeconomic status, occurrence of financial difficulties, and single parent households were used to represent economic risk. Based on related research in the

United States, we expected that children with higher economic risk would benefit more from the intervention (e.g., Raver et al. 2009). Given that longitudinal and intervention research has consistently identified the importance of gender, nationality, and special-needs education as predictors for the outcomes of interest (Raver et al. 2009), we controlled for these variables.

The present study aimed at analyzing moderators of treatment effects (baseline characteristics associated with variation in the achieved intervention effect) because understanding the active components of intervention trials is important for informing future intervention research. An analysis of moderators helps to understand whether some children benefit more or less from the intervention (Gardner et al. 2010; Hinshaw 2002). This is important, because it can spur further research on targeted interventions that match the needs of particular subgroups.

3.1. Method

3.1.1. Participants

The data were drawn from the Zurich Project on the Social Development of Children (z-proso), which is an ongoing prospective longitudinal study (for a detailed overview, see Eisner et al. 2011). For sampling, we used a cluster-randomized approach with the school as the unit of randomization (see Malti, Ribeaud, and Eisner 2011). Two universal prevention programs, one-school-based (PATHS) and the other family-based (Triple-P), were compared in a factorial design with schools randomly assigned to one of four treatment conditions (PATHS only, Triple-P only, PATHS and Triple-P, control). The sample at Time 1 (T1) consisted of 1,675 first graders (48 percent girls; mean age 7.0 years, $SD = 0.42$; see Malti et al. 2011, for response rates across data waves).

3.1.2. Intervention

The version of PATHS used in the present study was the same as that used in the Fast Track Project during the second school year (Bierman et al. 2010). It is a one-year program that includes forty-six primary lessons. The content, methods, and materials were culturally adapted to the Swiss school system, and the materials were tested in a pilot study (Eisner et al. 2011). PATHS lessons address problem-solving skills, social relationships, self-regulation, rule

understanding, emotion understanding, and positive self-esteem. The classes were taught for 67 minutes per week. The teachers who implemented PATHS attended a two-day training course prior to the start of the experimental sessions. The five coaches were trained and supervised by an experienced Dutch expert who also manages the PATHS teacher-training institute in the Netherlands. To increase implementation quality, the coaches visited each class four to six times during the implementation period, after which they discussed the lesson with the teacher and provided feedback. A refresher seminar was held midterm, and regular PATHS newsletters helped to create a sense of cohesion among the teachers. Zurich city council made the PATHS curriculum compulsory for teachers in the intervention group. The procedures used to monitor implementation closely followed methods described by Greenberg and Kusché (2002) and included teacher and child questionnaires in addition to observations by the coach. The overall implementation quality was evaluated as high (Eisner et al. 2011).

3.2. Variables

3.2.1. Dependent Variables

As dependent variables, we assessed teacher-reported aggressive behavior, ADHD, and prosocial behavior. In the z-proso study, behavior outcomes were assessed according to reports by teachers, parents, and children (Eisner et al. 2012; Malti et al. 2011). However, in the present analysis we focus on teacher-assessed outcomes only because treatment effects of a school-based intervention are most likely to be observed in the school context and we have shown elsewhere that the main effects were very limited, and mostly occurred in the teacher-rated data at follow-up (Malti et al. 2011). Furthermore, we have documented elsewhere that the PATHS plus Triple-P treatment does not have any stronger effects on externalizing behaviors than PATHS alone (Malti et al. 2011).

All outcomes were rated by the teachers at T1–T4 using the Social Behavior Questionnaire (SBQ) (Tremblay et al. 1991). The Zurich school system requires that children remain in the same class with the same teacher from the first to the third grade, but they enter new classes in the fourth grade; thus, all the teacher assessments at T4 were made by

new teachers who were blind as to treatment condition. The items in the teacher questionnaire were assessed on a five-point Likert scale. For aggressive behavior, eleven items were assessed. The average reliability (Cronbach's α) across the four waves was .93. For ADHD, eight items of the SBQ were used. The average α for the ADHD scale was .91. For prosocial behavior, the SBQ subscale contained seven items. Across the four waves, α was .92.

3.2.2. Moderator Variables: Child Characteristics

The following moderator variables representing child characteristics were assessed: social-cognitive skills, moral emotions, and baseline behavior (aggression, ADHD, prosocial behavior).

Social-cognitive skills. The children's social-cognitive skills were measured before the intervention by having them respond to four hypothetical vignettes: playing on a swing, participating in a game, laughing at someone, and stealing a ball. These four scenarios were adapted from previous research (Crick and Dodge 1996). After the child had been read the respective story text, he or she was asked for his or her problem solution strategies (e.g., "What could you say or do so that you could play on the swing?"). The responses were audiotaped and later coded in the following categories: (a) aggressive strategies (e.g., "I'd just push him off the swing"); (b) socially competent strategies (e.g., "I'll ask to take turns"); and (c) other strategies. Two independent coders rated the total content of all the transcripts. The inter-rater agreement (Krippendorff's α) across the categories was .80 (Krippendorff 1978). Proportional mean scores for aggressive and socially competent problem-solving strategies were then calculated.

Moral emotions. The children's moral emotions before the intervention were measured by a revised version of the original happy-victimizer task, which has been widely validated in the developmental literature (e.g., Malti et al. 2009). The children responded to four hypothetical rule violations and were asked to attribute emotion to self-as-victimizer ("How would you feel afterwards if you had done this? Why?"). Self-attributed emotions were coded on a four-point Likert scale from 1 (very good) to 4 (very bad). The four scores were aggregated across stories ($\alpha =$

.67), and the scale was labeled "moral emotions." Because the final score was skewed, it was log transformed.

3.2.3. Moderator Variables: Economic Factors

As economic moderator variables, we assessed household socioeconomic status (SES), financial difficulties, and single- versus two-parent household. SES was defined by coding the caregiver's current profession; the codes were then transformed into an International Socio-Economic Index (ISEI) occupational-status score (Ganzeboom et al. 1992). The final SES score was a derivative of the highest ISEI score of the two caregivers. Financial problems were assessed in the parental interview at T1. The parents were asked if they had experienced periods of financial difficulty resulting in arrears in payment of household bills during the last year. Household composition was assessed in the parental interview at T1 as well (see Table 2).

3.2.4. Control Variables

We controlled for gender, special-needs education, and nationality in all of the multilevel analyses. For nationality, a dummy variable was created that was coded 0 if at least one parent was Swiss and 1 if both parents were non-Swiss. The latter included more than eighty nationalities.

3.3. Procedure

The parents were asked to sign an informed consent form at the beginning of the first interview; informed consent was renewed at wave 4. The computer-assisted face-to-face interviews with parents lasted an average of one hour. In the first three waves, computer-assisted personal child assessments lasting 45 minutes were conducted at the school. In the fourth wave, 90-minute classroom-based paper-and-pencil surveys were utilized. The interviews were conducted by forty-four interviewers who had been thoroughly trained by the research team. The child's teacher completed a questionnaire on each child's social behavior.

3.4. Analysis Approach

Hierarchical linear modeling (HLM Version 6.08) was used to assess the intervention effects on child outcomes. The original design of the study combined PATHS and a family-based intervention (Triple-P) (for details, see Malti et al. 2011). We recoded treatment assignment as two

dummy variables to compare the PATHS and Triple-P conditions separately with the control condition. Thus, a standard approach to coding a 2 x 2 design (two levels of factor A crossed with two levels of factor B) was used to analyze program effects. This 2 x 2 design allowed us to specify the different timings of the interventions as well as the inclusion of the interaction involving the PATHS plus Triple-P condition. The cross-product of the PATHS plus Triple-P interaction answers the question of whether adding PATHS improves the effects of Triple-P and whether adding Triple-P improves the effects of PATHS. The models incorporated three levels: data-collection wave (level 1), child (level 2), and school (level 3). These levels were em-

ployed in conjunction with a two-way interaction between time and intervention to measure the treatment effects. Moderator effects were tested by three-way interactions between intervention, the respective moderating variable, and time point.

3.5. Results

3.5.1. Initial Equivalence and Attrition

Table 1 presents descriptive statistics for all the outcome variables, and Table 2 presents descriptive statistics for all the moderator variables. Here we present data for the PATHS only ($n = 360$) versus the control condition ($n = 356$) because of our analytic interest in the PATHS program.

Table 1: Outcome variables by treatment condition and time period

Outcome variable	Treatment condition							
	Control group				PATHS-only group			
	<i>M (SD)</i>				<i>M (SD)</i>			
	Pre 1	Pre 2 ^a	Post	Follow-up	Pre 1	Pre 2 ^a	Post	Follow-up
Teacher reports								
Aggressive behavior	0.51 (0.68)	0.41 (0.55)	0.45 (0.60)	0.53 (0.68)	0.56 (0.63)	0.56 (0.61)	0.62 (0.69)	0.48 (0.68)
ADHD	1.11 (0.97)	0.85 (0.94)	0.87 (0.94)	1.14 (1.04)	1.27 (1.02)	1.15 (1.00)	1.13 (1.01)	1.00 (0.89)
Prosocial behavior	2.16 (0.82)	2.14 (0.88)	2.42 (0.92)	2.26 (0.79)	2.34 (0.79)	2.43 (0.76)	2.54 (0.77)	2.24 (0.80)

^a Pre 2 is a second baseline score accounting for the time-lagged implementation of the PATHS and family interventions.

Table 2: Moderator variables by PATHS-only treatment condition versus control condition at T1

Moderator	Treatment condition	
	Control <i>M (SD)/ %</i>	PATHS-only <i>M (SD)/ %</i>
Child characteristics ^a		
Moral emotions	3.47 (0.59)	3.32 (0.68)
Aggressive problem-solving	0.15 (0.22)	0.16 (0.21)
Competent problem-solving	0.69 (0.30)	0.71 (0.27)
Economic characteristic		
Socioeconomic status (ISEI)	47.12 (18.10)	45.46 (17.08)
Financial difficulties: Yes	16	17
Single-parent household: Yes	13	15
Control variable		
Gender: Female	45	50
Nationality: Non-Swiss	39	28
Special education: Yes	8	6

Notes:

ISEI = International Socio-Economic Index.

^a For descriptive statistics of initial child behavior as moderator, see Time 1 variables in Table 1.

ANOVAs were conducted to determine the equivalence of the treatment and control groups across outcomes. The models took account of the nesting of students within schools, treating schools as a random effect. Although at baseline some of the outcome measures were higher in the PATHS group than in the control group, our preliminary analyses revealed that none of the baseline differences are significant. Table 3 displays the correlations between the

study variables across the four treatment conditions. Attrition was low across the waves of data collection and was comparable in the four treatment conditions. Specifically, children in the control condition completed an average of 3.68 waves, children in the PATHS condition an average of 3.76 waves, children in the Triple-P condition an average of 3.64 waves, and children in the PATHS plus Triple-P condition an average of 3.72 waves.

Table 3: Intercorrelations between study variables

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1. Aggressive behavior ^a	-											
2. ADHD	.66***	-										
3. Prosocial behavior	-.39***	-.36***	-									
4. Moral emotions	.03	.05	-.02	-								
5. Aggressive problem-solving	.07*	.10***	-.08**	.02	-							
6. Competent problem-solving	-.07*	-.10***	.12***	-.02	-.67***	-						
7. Socioeconomic status	-.16***	-.19***	-.03	-.11***	.04	-.02	-					
8. Financial difficulties	.13***	.12***	-.01	.04	-.03	.02	-.13***	-				
9. Single-parent household	.06	.07*	.02	.02	.01	.01	-.02	.15***	-			
10. Gender	-.24***	-.27***	.38***	-.03	-.13***	.14***	-.02	.01	.05	-		
11. Nationality	.11***	.08**	-.07*	.05	-.02	-.01	-.40***	.01	-.12***	.02	-	
12. Special education	.12***	.14***	-.07*	.02	.11***	-.13***	-.22***	.10***	-.04	-.10***	.17***	-

Notes:
 ADHD = Attention deficit/impulsivity
^a Because all behavior scales were related across time points, aggregate scores across time are reported.
 p* < .05. *p* < .01. ****p* < .001.

3.5.2. Treatment Effects at Follow-up (T4)

We report treatment effects only at follow-up because we have shown elsewhere that there were no treatment effects at post-test (Malti et al. 2011). Preliminary, unconditional models were run to ascertain the proportion of variance of each dependent variable that could be attributed to school level. Intraclass correlations were estimated in the control group using unconditional three-level hierarchical linear modeling. The intraclass correlation for school level was .25 across all teacher-reported outcome variables.

The multilevel models were then run to obtain intent-to-treat estimates of the intervention effects on aggres-

sive behavior, ADHD, and prosocial behavior as measured by the Social Behavior Questionnaire. We used continuous moderator variables in all multilevel analyses and computed interaction effects between the treatment variable and the respective continuous moderator variable. For clarity’s sake, we again describe only the findings for the PATHS only condition versus the control condition, because of our focus on PATHS effects. However, the full factorial design was used to analyze the data.

Table 4 presents the multilevel findings for the outcomes at follow-up. For each outcome variable, Model 1 included all the level 2 covariates, and Model 2 included

the respective interaction terms between treatment, the moderator variable, and time point. For the sake of brevity and clarity, treatment interaction terms invol-

ving the moderators are reported only if at least one of them is significant across outcomes.

Table 4: Parameter estimates of treatment effects on selected teacher-rated outcomes at follow-up

Parameter	Outcome		
	Aggressive behavior	ADHD	Prosocial behavior
Time x treatment	-.08 (.04)*	-.11 (.06)**	-.08 (.08)
	Child characteristic		
Aggressive problem-solving	-0.05 (0.08)	0.25 (0.15)*	0.03 (0.09)
Aggressive problem-solving x time x treatment	0.10 (0.06)	-0.04 (0.07)	-0.06 (0.07)
Aggressive problem-solving x time x treatment	0.10 (0.06)	-0.04 (0.07)	-0.06 (0.07)
Competent problem-solving	-0.08 (0.06)	-0.19 (0.10)*	0.14 (0.07)*
Competent problem-solving x time x treatment	-0.03 (0.03)	-0.08 (0.03)*	-0.04 (0.03)
Moral emotions	0.01 (0.02)	0.05 (0.04)	-0.01 (0.02)
Moral emotions x time x treatment	-0.03 (0.01)*	-0.02 (0.01)*	-0.05 (0.04)
Baseline behavior	1.09 (0.03)***	1.06 (0.02)***	1.10 (0.03)***
Baseline behavior x time x treatment	0.02 (0.03)	-0.03 (0.02)	-0.01 (0.01)
	Economic characteristics ^a		
Socioeconomic status	-0.01 (0.01)**	-0.01 (0.01)**	-0.01 (0.02)*
Financial problems	0.13 (0.05)*	0.22 (0.07)**	-0.01 (0.05)
Single-parent household	0.11 (0.04)*	0.17 (0.08)*	-0.05 (0.05)
	Control variable		
Girl	-0.18 (0.03)***	-0.38 (0.05)***	0.49 (0.04)***
Non-Swiss nationality	0.13 (0.03)***	0.03 (0.06)	-0.09 (0.04)*
Special-class education	0.02 (0.08)	0.16 (0.11)	-0.12 (0.06)*

Notes:

ADHD = Attention deficits/impulsivity

^a No moderation by economic factors was found.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Aggression. Children in the PATHS group were reported by their teachers as having a greater decrease in aggressive problem behaviors than children in the control group (effect size = 0.42; see Table 4). The treatment effect on teacher-rated aggression was moderated by level of moral emotions at baseline (effect size = 0.12). Children who exhibited higher levels of moral emotions and received the

intervention showed larger reductions in aggression at follow-up (by teacher report) than did children who started school with low levels of moral emotions. In addition, baseline aggression, financial problems, single-parent household, and non-Swiss nationality predicted higher aggression at the follow-up. Furthermore, SES and female gender predicted lower aggression at the follow-up.

ADHD. Children in the PATHS group were reported by their teachers as having a significantly greater decrease in ADHD related problems than children in the control group (effect size = 0.46; see Table 4). However, the treatment effect on teacher-rated ADHD was moderated by the level of moral emotions at baseline (effect size = 0.10). The main effect was also moderated by initial level of competent problem-solving strategies (effect size = 0.40). Children who exhibited higher levels of moral emotions and competent problem-solving strategies, and received the intervention, showed larger reductions in ADHD at follow-up (by teacher report) than did children who started school with low levels of moral emotions and competent problem-solving strategies. Furthermore, aggressive problem-solving strategies, baseline ADHD, financial problems and single-parent household predicted higher ADHD at follow-up. In contrast, competent problem-solving strategies, SES and female gender predicted lower ADHD at follow-up.

Prosocial behavior. The results suggest that overall, children in the PATHS condition did not differ from children in the control condition on prosocial behavior (Table 4). In addition, no significant interaction terms representing potential moderation by child characteristics or economic factors was found. Competent problem-solving strategies, baseline prosocial behavior and female gender predicted higher prosocial behavior at follow-up, whereas non-Swiss nationality and special needs education predicted lower prosocial behavior at follow-up.

4. Discussion

A large-scale randomized controlled trial of a cohort of children attending public elementary schools in Zurich, Switzerland investigated the effectiveness of the PATHS curriculum on teacher reports of children's aggressive problem behavior, ADHD, and prosocial behavior at follow-up (two years later), as well as the moderating roles of baseline child characteristics and economic factors (as measured before the intervention) on the effects of the interventions at follow-up.

As reported elsewhere, the analyses revealed a main effect of PATHS on ADHD and aggression as reported by teachers at follow-up (see Malti et al. 2011). The findings support pre-

vious research in the United States demonstrating the positive effects of social competence programs such as PATHS on ADHD/impulsivity problems (Riggs et al. 2006). However, as reported elsewhere, there were no significant treatment effects for most teacher-rated outcomes and almost all parent- and child-rated outcomes (Malti et al. 2011).

Overall, we found only very limited evidence that the intervention effects were moderated by baseline child characteristics. There was evidence that children in the PATHS condition who at baseline anticipated more negative emotions following transgressions were rated by their teachers as showing slightly larger reductions in ADHD at follow-up than children who started with low levels of moral emotions. Other research indicates that these moral emotions reduce externalizing symptoms (Arsenio et al. 2006; Malti and Krettenauer 2012).

Additionally, children in the PATHS condition who displayed high levels of competent problem-solving strategies before the intervention were rated by their teachers as showing slightly larger reductions in ADHD at follow-up than children who started with low levels of competent problem-solving skills. This finding is consistent with longitudinal research showing that socially competent problem-solving skills distinguish the trajectories of aggressive behavior (Lansford et al. 2006), as children who have strong social-cognitive skills before the intervention benefit most from the intervention in terms of a decrease in maladaptive behavior. Contrary to the results of previous randomized controlled trials of PATHS in the United States (Bierman et al. 2010), we did not find a moderating effect of baseline behavior on treatment outcomes. Further research is needed to disentangle these inconsistencies and investigate questions surrounding treatment dose for children at risk of behavioral problems.

Our findings confirm several of the expected direct effects of economic variables such as family socioeconomic status, financial problems, and single-parent household on the outcome variables. Overall, our findings support research on the role of economic risk in exacerbating negative behavioral outcomes (Keenan et al. 1997). The effects of PATHS were not moderated by these factors, contradicting

the results of other studies (Raver et al. 2009). This difference may be related to the fact that socioeconomic disparities in Switzerland are smaller than in the United States, where most comparable randomized controlled trials have been conducted.

This study was not without limitations. First, we only examined one informant (i.e., teacher report) because treatment effects of a school-based intervention are most likely to be observed in the school context and in teacher-assessed outcomes; we have documented elsewhere that there were only a few positive effects on teacher-rated child outcomes at follow-up, but not at post-test, (Malti et al. 2011). We also do not know whether any similar moderation effects hold for other informants (Malti et al. 2011). Second, the positive treatment effects were only found in the PATHS only group, but the PATHS plus Triple-P treat-

ment did not have any stronger effects on teacher-rated externalizing behavior than PATHS alone. It remains to be disentangled why no additional effects were found in the combined group. Third, as in many studies, the moderators were examined post-hoc rather than based on prior theoretical considerations. Post-hoc analyses are always at a risk of finding effects by pure chance and should hence be treated with caution. Fourth, at present we do not know whether any of the effects are maintained over a longer period of time. We did not find any positive treatment effects at post-test and most of the outcomes at follow-up were not significant either (Malti et al. 2011).

Despite these limitations, the present study contributes useful knowledge about the influence of moderators on the effects of a school-based intervention on children's externalizing behavior.

References

- Arsenio, William F., Jason Gold, and Erin Adams. 2006. Children's Conceptions and Displays of Moral Emotions. In *Handbook of Moral Development*, ed M. Killen and J. Smetana (581–610). Mahwah, NJ: Lawrence Erlbaum.
- Bierman, Karen L., John D. Coie, Kenneth A. Dodge, Mark T. Greenberg, John E. Lochman, Robert J. McMahon, Ellen E. Pinderhughes, and Conduct Problems Prevention Research Group. 2010. The Effects of a Multiyear Universal Social-Emotional Learning Program: The Role of Student and School Characteristics. *Journal of Consulting and Clinical Psychology* 78 (2): 156–68.
- Bronfenbrenner, Urie, and Gary W. Evans. 2000. Developmental Science in the 21st Century: Emerging Theoretical Models, Research Designs, and Empirical Findings. *Social Development* 9:115–25.
- Conduct Problems Prevention Research Group. 2002. Evaluation of the First Three Years of the Fast Track Prevention Trial with Children at High Risk for Adolescent Conduct Problems. *Journal of Abnormal Child Psychology* 30:19–35.
- Conduct Problems Prevention Research Group and Kenneth A. Dodge. 2007. Fast Track Randomized Controlled Trial to Prevent Externalizing Psychiatric Disorders: Findings from Grades 3 to 9. *Journal of the American Academy of Child and Adolescent Psychiatry* 46 (10): 1250–62.
- Costello, E. Jane, Adrian Angold, Barbara J. Burns, Dalene K. Stangl, Dan L. Tweed, Alaattin Erkanli, and Carol M. Worthman. 1996. The Great Smoky Mountains Study of Youth. *Archives of General Psychiatry* 53 (12): 1129–36.
- Crick, Nicki R., and Kenneth A. Dodge. 1996. Social Information-Processing Mechanisms in Reactive and Proactive Aggression. *Child Development* 67:993–1002.
- Eisner, Manuel P., Tina Malti, and Denis Ribeaud. 2011. Large-Scale Criminological Field Experiments. In *Sage Handbook of Criminological Research Methods*, ed S. Gadd, S. Karstedt, and S. Messner (410–24). London: Sage.
- Eisner, Manuel P., Daniel Nagin, Denis Ribeaud, and Tina Malti. 2012. Effects of a Universal Parenting Program for Highly Adherent Parents: A Propensity Score Matching Approach. *Prevention Science* 13 (3): 252–66.
- Ganzeboom, Harry B. G., Paul M. Degraaf, Donald J. Treiman, and Jan de Leeuw. 1992. A Standard International Socio-Economic Index of Occupational Status. *Social Science Research* 21:1–56.
- Gardner, Frances, Judy Hutchings, Tracey Bywater, and Chris Whitaker. 2010. Who Benefits and How Does It Work? Moderators and Mediators of Outcome in an Effectiveness Trial of a Parenting Intervention. *Journal of Clinical Child and Adolescent Psychology* 39 (4): 568–80.
- Greenberg, Mark T., and Carol A. Kusché. 2002. *Promoting Alternative Thinking Strategies: Blueprint for Violence Prevention* (vol. 10, 2d ed). Boulder: University of Colorado.
- Greenberg, Mark T., Roger P. Weissberg, Mary Utne O'Brien, Joseph E. Zins, Linda Fredewricks, Hank Resnik, and Maurice J. Elias. 2003. Enhancing School-Based Prevention and Youth Development Through Coordinated Social, Emotional and Academic Learning. *American Psychologist* 58:466–74.
- Hahn, Robert, Dawna Fuqua-Whitley, Holly Wethington, Jessica Lowy, Akiva Liberman, Alex Crosby, Mindy Fullilove, et al. 2007. Effectiveness of Universal School-Based Programs to Prevent Violent and Aggressive Behavior: A Systematic Review. *American Journal of Preventive Medicine* 33 (2, supplement 1): S114–S129.
- Hinshaw, Stephen P. 2002. Intervention Research, Theoretical Mechanisms, and Causal Processes Related to Externalizing Behavior Patterns. *Development and Psychopathology* 14:789–818.
- Keenan, Kate, Daniel S. Shaw, Brenda Walsh, Eric Delliquadri, and Joyce Giovannelli. 1997. DSM-III-R Disorders in Preschool Children from Low-Income Families. *Journal of American Academy of Child and Adolescent Psychiatry* 36:620–27.
- Krippendorff, Klaus. 1978. Reliability of Binary Attribute Data. *Biometrics* 34 (1): 142–44.
- Lansford, Jennifer E., Patrick S. Malone, Kenneth A. Dodge, Joseph C. Crozier, Gregory S. Pettit, and John E. Bates. 2006. A 12-Year Prospective Study of Patterns of Social Information Processing Problems and Externalizing Behaviors. *Journal of Abnormal Child Psychology* 34: 715–24.
- Malti, Tina, Michaela Gummerum, Monika Keller, and Marlis Buchmann. 2009. Children's Moral Motivation, Sympathy, and Prosocial Behavior. *Child Development* 80: 442–60.
- Malti, Tina, and Tobias Krettenauer. 2012. The Role of Moral Emotion Attributions in Children's and Adolescents' Prosocial and Antisocial Behavior: A Meta-Analysis. *Child Development*. Early online publication, 24 September. doi:10.1111/j.1467-8624.2012.01851.x
- Malti, Tina, Denis Ribeaud, and Manuel Eisner. 2011. The Effects of Two Universal Preventive Interventions to Reduce Children's Externalizing Behavior: A Cluster-Randomized Controlled Trial. *Journal of Clinical Child and Adolescent Psychology* 40 (5): 677–92.
- Metropolitan Area Child Study Research Group (MACS). 2002. A Cognitive-Evolutiological Approach to Preventing Aggression in Urban Settings: Initial Outcomes for High-Risk Children. *Journal of Consulting and Clinical Psychology* 70 (1): 179–94.
- Orobio de Castro, Bram, Jan W. Veerman, Willem Koops, Joop D. Bosch, and Heidi J. Monshouwer. 2002. Hostile Attribution of Intent and Aggressive Behavior: A Meta-Analysis. *Child Development* 73 (3): 916–934.
- Petrosino, Anthony, and Haluk Soydan. 2005. The Impact of Program Developers as Evaluators on Criminal Recidivism: Results from Meta-Analyses of Experimental and Quasi-Experimental Research. *Journal of Experimental Criminology* 1:435–50.
- Raver, C. Cybele, Stephanie M. Jones, Christine Li-Grining, Fuhua Zhai, Molly W. Metzger, and Bonnie Solomon. 2009. Targeting Children's Behavior Problems in Preschool Classrooms: A Cluster-Randomized Controlled Trial. *Journal of Consulting and Clinical Psychology* 77 (2): 302–16.
- Riggs, Nathaniel R., Mark T. Greenberg, Carol A. Kusche, and Mary Ann Pentz. 2006. The Medial Role of Neurocognition in the Behavioral Outcomes of a Social-Emotional Prevention Program in Elementary School Students: Effects of the PATHS Curriculum. *Prevention Science* 7:91–102.
- Tremblay, Richard E., Rolf Loeber, Claude Gagnon, Pierre Charlebois, Serge Larivee, and Marc LeBlanc. 1991. Disruptive Boys with Stable and Unstable High Fighting Behavior Patterns During Junior Elementary School. *Journal of Abnormal Child Psychology* 19:285–300.

Tina Malti
tina.malti@utoronto.ca

Denis Ribeaud
denis.ribeaud@soz.gess.ethz.ch

Manuel Eisner
mpe23@cam.ac.uk

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK

Michael Little, The Social Research Unit, Dartington, United Kingdom
Vashti Berry, The Social Research Unit, Dartington, United Kingdom
Louise Morpeth, The Social Research Unit, Dartington, United Kingdom
Sarah Blower, The Social Research Unit, Dartington, United Kingdom
Nick Axford, The Social Research Unit, Dartington, United Kingdom
Rod Taylor, University of Exeter Medical School, United Kingdom
Tracey Bywater, Institute for Effective Education, University of York, United Kingdom
Minna Lehtonen, The Social Research Unit, Dartington, United Kingdom
Kate Tobin, The Social Research Unit, Dartington, United Kingdom

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

► **The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK** Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK

Michael Little, The Social Research Unit, Dartington, United Kingdom
Vashti Berry, The Social Research Unit, Dartington, United Kingdom
Louise Morpeth, The Social Research Unit, Dartington, United Kingdom
Sarah Blower, The Social Research Unit, Dartington, United Kingdom
Nick Axford, The Social Research Unit, Dartington, United Kingdom
Rod Taylor, University of Exeter Medical School, United Kingdom
Tracey Bywater, Institute for Effective Education, University of York, United Kingdom
Minna Lehtonen, The Social Research Unit, Dartington, United Kingdom
Kate Tobin, The Social Research Unit, Dartington, United Kingdom

The Birmingham Brighter Futures strategy was informed by epidemiological data on child well-being and evidence on “what works,” and included the implementation and evaluation of three evidence-based programmes in regular children’s services systems, as well as an integrated prospective cost-effectiveness analysis (reported elsewhere). A randomised controlled trial (RCT) of the Incredible Years BASIC parenting programme involved 161 children aged three and four at risk of a social-emotional or behavioural disorder. An RCT of the universal PATHS social-emotional learning curriculum involved children aged four–six years in 56 primary schools. An RCT of the Level 4 Group Triple-P parenting programme involved parents of 146 children aged four–nine years with potential social-emotional or behavioural disorders. All three studies used validated standardised measures. Both parenting programme trials used parent-completed measures of child and parenting behaviour. The school-based trial used teacher reports of children’s behaviour, emotions, and social competence. Incredible Years yielded reductions in negative parenting behaviours among parents, reductions in child behaviour problems, and improvements in children’s relationships. In the PATHS trial, modest improvements in emotional health and behavioural development after one year disappeared by the end of year two. There were no effects for Triple-P. Much can be learned from the strengths and limitations of the Birmingham experience.

There is good reason to be concerned about the well-being of children in the United Kingdom. Research using repeated application of the same measures charted a downward trend in mental health outcomes over a three-decade period (Collishaw et al. 2004), and while conduct problems reached a plateau in the early 2000s (albeit at a level that is still concerning), this pattern is not so clear for emotional problems (Maughan et al. 2008). Indeed, the

proportion of young people reporting frequent feelings of depression or anxiety doubled between 1986 and 2006 (Collishaw et al. 2010). Cross-national comparisons using a range of indicators, including material well-being, education, behaviour, and family and peer relationships, find that children in the United Kingdom regularly perform poorly compared with other high-income nations (e.g. UNICEF 2007).

The authors would like to acknowledge the team of people at Birmingham City Council involved in this project, who have demonstrated that public services can be forward-thinking. Particular note should be given to Stephen Hughes and Cheryl Hopkins for their leadership. We also acknowledge the input from the Centre for Health Economics and Medi-

cines Evaluation at Bangor University, in particular Rhiannon Tudor Edwards and Pat Linck. All three trials were registered with the Current Controlled Trials Registry (ISRCTN): Incredible Years (ISRCTN 48762440); Triple-P (ISRCTN 10429692); and PATHS (ISRCTN32534848). The randomisation sequence for all three trials were cre-

ated by the North Wales Organisation for Randomised Trials in Health (& Social Care) (NWORTH). Interests declared: none. Ethical clearance: The Warren House Group Ethics Committee and the North Wales NREC (reference 10/WNo01/29).

A number of UK government initiatives from 1998 onwards sought to reverse this trend, including an assault on child poverty, the introduction of universal pre-school provision (Eisenstadt 2011), a drive to improve school performance, and efforts to better integrate social care, youth justice, education and mental health services. For example, in 2004, local authority services for children were integrated, creating single “children’s services” departments with one director and an elected politician who is accountable for children’s health and development.

There are 150 local authorities in England. The work described in this article took place in the largest, Birmingham, serving a child population of 260,000 with a budget of roughly £1.3 billion (€1.6 billion). Birmingham is one of a small number of English local authorities where the majority of children are from minority ethnic groups. The city excels in many areas, for example school performance, but there are significant risks to well-being, not least from poverty.

1. Bridging Science and Policy

The Social Research Unit at Dartington is an independent centre dedicated to improving child outcomes through research, service development, and dissemination. The methods and tools it uses are underpinned by an approach known as *Common Language*. This provides a conceptual framework or way of thinking that can allow people with different roles – say policy makers and scientists – or different disciplines – such as education, psychology, social work – to communicate effectively about how to improve the well-being of children. The approach includes various elements, including epidemiology, strategy development, service design, and training (see Axford and Morpeth 2012).

In Birmingham, *Common Language* was used to bring together leaders of children’s services to produce a single vision for children, clearly specifying the outcomes they wanted to achieve. The work required that the system leaders identify those activities most likely to deliver target outcomes, and that they be clear about the source of investments necessary to fund the activities. The *Common Language* methods ensured that this emerging strategy was backed up by strong logic; the best evidence on the well-

being of local children, obtained through an epidemiological study conducted in the city as part of the Common Language work (Axford and Hobbs 2011; Hobbs et al. 2011; Axford et al. 2012); and reliable information on what works, for whom, when, and why.

These methods have been developed over a decade and used to support both the integration of children’s services in Norfolk in the East of England and, later, a \$200 million (€156 million) philanthropic investment to improve outcomes for children in Ireland (Little and Abunimah 2007; Axford et al. 2008). The tools have been refined with each application.

2. Brighter Futures Strategy

The move to create single departments of children’s services led Birmingham to appoint a new strategic director, Tony Howell. He decided to take an inclusive approach to strategy development by involving all of the agencies working with children, including the third sector. The Social Research Unit at Dartington was commissioned to facilitate the strategy development using *Common Language*.

The result was a single strategy for the City known as “Brighter Futures” (BCC 2007). It prioritised six outcomes, including behaviour and emotional well-being. Brighter Futures supported increased use of evidence-based programmes with proven impact on child outcomes, and stressed the need to improve parenting.

Financial support for the strategy came from the “business transformation” applied by the City Council to public services across the city, which sought to encourage investments that would later generate savings in expenditure, for example building new low-maintenance, energy-efficient offices to replace older and more expensive ones. In the context of Brighter Futures, data from a range of international sources, strongly influenced by the work of the Washington State Institute for Public Policy (Aos et al. 2004), was used to calculate that a £42 million (€52 million) investment that included the evidence-based programmes as well as substantial expenditure on staff development and IT systems would produce an economic return of £101 million (€126 million) over a fifteen-year period.

The strategy development and the requirement to produce a return on investment created strong support for evidence-based programmes. The Social Research Unit used its Common Language approach to facilitate teams of systems leaders to select a portfolio of programmes that crossed several stages of child development and had the greatest impact on outcomes targeted by Brighter Futures. Four were eventually selected.

Family Nurse Partnership (Nurse Family Partnership in the US) was offered to high-risk and usually teenage mothers beginning pre-birth (Olds 1998/2006). Success is measured in terms of improved parenting, better child health and development, delay in the birth of a second child, and improved family income.

The Incredible Years BASIC parenting programme (Webster-Stratton 1994) was offered to the parents of three- and four-year-olds showing the symptoms of a conduct disorder (e.g. often having temper tantrums, often fighting with other children, being spiteful to other children, often argumentative with adults). Building on the work of Judy Hutchings in Wales (Hutchings et al. 2007; Hutchings, Bywater, and Daley 2007), the twelve-week version of this evidence-based programme was delivered in children's centres (the universal pre-school provision that became available in the United Kingdom from 2000 onwards). The success of Incredible Years is measured in terms of improved behaviour and better social relationships at home and with other children. The programme was delivered by a mixture of children's centre staff, family support workers, educational psychologists, and parenting practitioners. All implementation staff were trained by an accredited Incredible Years trainer (Judy Hutchings) and participated in weekly half-day supervision sessions with that trainer.

Promoting Alternative Thinking Strategies (PATHS) is a school-wide curriculum to improve regulation of emotions for children aged four–eleven (Greenberg and Kusché 2002), which in Birmingham took up one hour per week for three of the primary school years (reception, year one and year two). Success is measured in terms of better behaviour and improved emotional well-being. Systematic reviews, for example by Durlak and colleagues (2011), sug-

gest that social and emotional learning programmes like PATHS produce better-behaved, happier children who therefore learn more, generating increases in academic attainment. Teachers implementing PATHS received two days of training (initial and top-up) and technical assistance from trained PATHS coaches.

Triple-P, like Incredible Years, is a parenting programme. In Birmingham the Standard Level-4 Group Triple-P parenting programme (Bodenman et al. 2008) was offered to families with children aged four–nine years whose significantly poor behaviour was causing problems at home, in school and, for some, in the community. Unlike the other programmes, the implementation team inherited practitioners with varying degrees of training in Triple-P. These were a mixture of clinical psychologists and mental health practitioners as well as parenting coordinators. These practitioners received “top-up” support from the programme originator Matt Sanders. Outcomes for Triple-P are measured in terms of children's behaviour, aggression, and emotional well-being. Supervision is not mandatory for Triple-P facilitators, although online support is offered by Triple-P International Ltd. and Birmingham facilitators also had access to additional support from a nominated Triple-P trainer within the Parenting Support Service, if they felt they required it.

The Brighter Futures strategy was rooted in high-quality epidemiological data about the well-being of children living in the city (Hobbs et al. 2011). This data was gathered on over 5,000 children through a representative school-based survey of seven–eighteen year-olds and a representative household survey of parents of zero–six year-olds, and identified aspects of children's well-being in need of particular attention, for example because on standardised measures of health and development children were performing below national norms. The strategy also drew heavily on international data on “what works”, particularly as captured in online clearing-houses of effective programmes, such as Blueprints for Violence Prevention (<http://www.colorado.edu/cspv/blueprints/>). The business transformation method required proof of return on investment. These forces led Birmingham to radically change its approach to evaluation. It subjected all four of the evidence-based programmes to

Randomised Controlled Trials (RCTs), summarised later in this article, a major departure from traditional local authority approaches to research and development.

All of the evaluations applied the “intention to treat” principle, meaning that results include those children, parents, or schools that dropped out of the study. The findings therefore reflect what happens in real-world situations, with many intervention recipients either not starting or not completing an intervention paid for by the local authority. Each of the trials used a “waiting list” design, meaning that children or schools not receiving the intervention were given priority to receive it in future if the results of the evaluation were positive. Children in the control conditions received “services as usual”, which in some cases involved substantial support – for example, the SEAL (Social and Emotional Aspects of Learning) programme in the case of the PATHS trial. Participants in the programme groups could also continue to receive services as usual – that is, no services were withdrawn – although it is acknowledged that logistically this may have been difficult (for example, if PATHS lessons used curriculum time previously allocated to SEAL).

Typically, experimental evaluation is expensive. In order to reduce costs, the Social Research Unit sought only to replicate the findings established in other trials, thereby collecting considerably less data than is usually the case. The experimental approach was taken, randomly allocating units to control and intervention groups. Sample sizes reflect a calculation of the statistical power needed for any programme effect identified by the evaluations to be greater than chance. Robust measurement was also required. These elements are typical of a good RCT. The focus on replicating findings from other trials offers a different angle, however. Specifically, the data collection was restricted to the factors in the logic model underpinning the evidence-based programme, including the risks targeted, the fidelity of implementation of core elements of the intervention, and the outcomes sought. Other hypothesised moderators and other contextual information are

excluded. The net result is a high-quality evaluation with less data and therefore less cost.

3. Evaluation

Family Nurse Partnership (FNP) was introduced as part of a national evaluation. The largest RCT of FNP is being undertaken by the University of Cardiff and will report in 2013. Evaluations of Incredible Years, PATHS and Triple-P by RCT were undertaken by the Social Research Unit in collaboration with the Peninsula Medical School, which provided statistical advice, and the Centre for Health Economics and Medicines Evaluation at Bangor University, which examined the cost-effectiveness of Incredible Years, Triple-P and PATHS. Process and qualitative data were collected alongside the trials. This article presents a summary of outcome results from the three trials, with brief reference to fidelity of implementation where it may explain results (fidelity is covered in more depth in papers on each of the trials, to follow).

The Incredible Years evaluation was a parallel RCT with pre-post test design, which involved the parents of 161 children aged three and four, identified through referral from other agencies, self-referral, and screening of children served by children’s centres. In order to be eligible for the programme children needed to be at risk of a social-emotional or behavioural disorder, which meant reaching the “high need” threshold (17 or above out of 40) of the “total difficulties” score of the parent-completed Strengths and Difficulties Questionnaire (SDQ) (Goodman 1997). There were 101 males and 60 females, with a mean age of 44 months (SD = 6) at baseline. The sample comprised a high proportion of low-income families: 50 percent of families relied on benefits as their main source of income.

The 161 children who reached the threshold were randomised to the study on a 2:1 ratio – 110 to intervention, 51 to control. Baseline data (Wave 1) were collected on all 161 children before randomisation. Follow-up (Wave 2) took place six months after baseline.¹ It included 147 children,

¹ It was possible to conduct a long-term follow-up at one-year post-baseline (Wave 3) as well but this was for the intervention group only, not the control group. Results are reported elsewhere.

leaving 14 missing cases (10 intervention and four control): seven formal withdrawals from the study and seven who could not be contacted. As with all three trials, values were imputed for missing data, based on baseline scores. The SDQ and the Eyberg Child Behaviour Inventory (ECBI) (Eyberg and Ross 1978), both completed by parents, were used to measure child outcomes. Parenting behaviour was measured using the Arnold and O’Leary Parenting Scale (APS) (Arnold et al. 1993). As Table 1 shows, there were differences between intervention and control groups at baseline. The impact of Incredible Years is therefore calculated using an estimated mean difference between intervention and control groups. ANCOVA tests controlled for children’s start scores on the respective measure, the age and sex of the child, and the centre from which families were recruited.

The PATHS evaluation was a cluster-randomised design involving 56 schools, 29 of which were allocated to implement the PATHS curriculum and 27 of which were allocated to a waiting list. The Bangor Trials Unit (NORTH) generated the randomisation sequence and the sample was stratified by percentage of free school meals and size of school. As a universal intervention PATHS was offered to all children in reception and year one classes (aged four–six years). This cohort received the programme for two years. There were 5,397 pupils in the schools at baseline, attending 196 classes (102 intervention and 94 control). Data were collected at three points: baseline (Wave 1) in September 2009; first follow-up (Wave 2) in June 2010 (after one year of implementation); and second follow-up (Wave 3) in June 2011 (on completion of the intervention). Data were collected on 183 classes ($n = 5,074$) at Wave 1, on 176 classes ($n = 4,998$) at Wave 2, and on 178 classes ($n = 4,994$) at Wave 3. There were 4,006 complete cases with data at all three waves.

Outcomes were assessed using the SDQ teacher report and the PATHS Teacher Rating Survey (PTRS), a composite measure of seven scales (e.g. Child Behaviour Questionnaire). These two measures provide a picture of

children’s behavioural and emotional difficulties, social competence, and emotional regulation. To identify the unique contribution of PATHS to the outcomes measured, hierarchical linear models were run to take account of class/school clustering in outcomes and also to control for available covariates at all levels. Data were first analysed using only complete cases, ignoring missing data. A second set of analyses used multiple imputation methods to account for missing data. However, this did not fundamentally change the results.

The evaluation of Triple-P was a parallel randomised controlled trial, with pre-post test design. It involved 146 children aged four–nine years whose symptoms indicated a potential social-emotional or behavioural disorder, determined using the “high need” threshold on the SDQ “total difficulties” score (17 or above out of 40). The sample comprised 105 boys and 41 girls. The mean age was 82 months ($SD = 21$). The sample also comprised a high proportion of low-income families: 62 percent of children were entitled to free school meals compared to 33 percent for Birmingham as a whole.

The parent(s) of half (73) of these children were randomly assigned to attend Triple-P parenting groups, with the remaining half placed on a waiting list and receiving services as usual. Researchers performed the randomisation for each eligible child using an online programme, designed by NORTH. Children were randomised on a 1:1 ratio, using a dynamic allocation method, stratified by age and sex. Baseline (Wave 1) data was collected on all children. Follow-up (Wave 2) occurred six months after baseline and included 137 children.² The programme was delivered to intervention group parents at some point during those six months. The missing nine cases (three control, six intervention) were made up of two formal withdrawals from the study and seven that could not be contacted. The primary outcome instruments were the SDQ and ECBI. Parenting behaviour was measured using the Arnold and O’Leary Parenting Scale (APS). Estimated mean differences

² It was possible to conduct a long-term follow-up at one-year post-baseline (Wave 3) as well but this was for the intervention group only, not the control group. Results are reported elsewhere.

were used to calculate the impact of Triple-P. ANCOVA tests controlled for children's start scores on respective measures, the age and sex of the child, and the area from which families were recruited.

4. Results

The three evaluations revealed both success and failures for the Birmingham Brighter Futures strategy.

4.1. Incredible Years

The results summarised in Table 1 indicate significant benefits from the Incredible Years programme offered as part of children's centre provision (universal early years). There are reductions in reported negative parenting behaviours among the parents attending Incredible Years groups compared to controls. There are strong reductions in child behaviour problems and improvements in relationships for children receiving the intervention.

Table 1: Child outcomes for Incredible Years control and intervention groups

Child measure (cut-off)	Mean (SD) raw scores				Estimated mean difference (95% CI) ¹	Effect Size (d)
	Control (n=51)		Intervention (n=110)			
	Baseline	6 mth	Baseline	6 mth		
SDQ conduct ² problems (4)	6.53 (2.1)	4.43 (2.7)	6.29 (2.0)	3.62 (2.1)	0.78* (0.05 to 1.51)	0.39
SDQ emotion problems (5)	4.85 (2.5)	3.61 (2.6)	4.79 (2.4)	3.30 (2.3)	0.36 (-0.36 to 1.07)	
SDQ hyperactivity (7)	7.66 (1.9)	6.18 (2.4)	7.67 (1.9)	5.83 (2.5)	0.40 (-0.36 to 1.17)	
SDQ peer problems (4)	4.47 (1.9)	3.39 (2.1)	4.23 (1.8)	2.69 (1.8)	0.71* (0.85 to 1.34)	0.39
SDQ pro-social behaviour (<4) ³	5.19 (2.2)	6.35 (2.2)	5.72 (2.1)	6.77 (2.1)	-0.22 (-0.84 to 0.40)	
SDQ total difficulties (17)	23.50 (4.5)	17.60 (7.3)	22.98 (4.4)	15.44 (6.0)	2.23* (0.13 to 4.34)	0.50
SDQ impact (2)	0.92 (1.4)	0.58 (1.2)	0.59 (1.1)	0.14 (0.5)	0.37** (0.10 to 0.63)	0.31
ECBI-I (127) ⁴	143.86 (38.5)	134.35 (42.3)	142.70 (35.7)	123.10 (34.8)	13.48* (2.31 to 22.64)	0.37
ECBI-P (11) ⁵	17.31 (9.3)	14.33 (9.8)	16.71 (8.8)	11.24 (9.0)	2.62 (-0.07 to 5.32)	
APS total ⁶	3.58 (0.8)	3.32 (0.8)	3.49 (0.6)	3.01 (0.8)	0.29** (0.08 to 0.50)	0.43
APS laxness	3.79 (1.3)	3.43 (1.2)	3.58 (1.2)	3.04 (1.1)	0.30 (-0.01 to 0.61)	
APS verbosity	4.15 (0.9)	4.01 (1.0)	4.26 (0.9)	3.68 (1.0)	0.42** (0.12 to 0.72)	0.47
APS over-react	2.90 (1.0)	2.71 (1.1)	2.78 (0.8)	2.36 (0.8)	0.31* (0.06 to 0.57)	0.36

* significant at p < .05

** significant at p < .01

1 Difference in mean follow-up scores between intervention and waiting list control conditions, measured by analysis of covariance adjusted for baseline score, age of child, sex and area.

2 SDQ Strengths and Difficulties Questionnaire (on all scales higher scores equals greater problems, except for prosocial behaviour).

3 SDQ prosocial is measured positively: the higher the score, the better the behaviour of the children.

4 ECBI-I Eyberg Child Behaviour Inventory - Intensity Scale (higher scores equate to greater problems).

5 ECBI-P Eyberg Child Behaviour Inventory - Problem Scale (higher scores equate to greater problems).

6 APS Arnold and O'Leary Parenting Scale (and three sub-scales) (higher scores equate to greater problems).

The data suggest an effect size on SDQ conduct problems of 0.39, consistent with the 0.33 reported by Hutchings and colleagues (2007a) for their similar trial in Wales. In common with other evaluations of Incredible Years, benefits increased with dose. Training and supervision arrangements for Incredible Years facilitators, which boost fidelity of programme delivery, were also implicated in better outcomes. The effect size of 0.37 for the ECBI-I is smaller than in other UK community-based trials – 0.89 found by Hutchings et al. (2007) and 0.55 by Gardner et al. (2006) – but is nevertheless respectable, particularly given that implementation in Brighter Futures was arguably more “real world” than in these two earlier studies.

4.2. PATHS

The results for PATHS were more mixed. Table 2 presents the results from the hierarchical linear model analysis, examining the difference in outcomes between the PATHS group and control group, accounting for class- and school-level clustering in outcomes and available covariates at all levels. As indicated, at first follow-up there were modest improvements in pupils’ emotional health and behavioural development in the PATHS schools compared to those in control schools. However, at the two-year follow-up, these gains had all been lost.

Table 2: Hierarchical linear model results for the PATHS programme with N and between-group difference mean (95% CI)¹

	12-month follow-up adjusted for baseline	24-month follow-up adjusted for baseline
SDQ total difficulties	N=4255 -0.42 (-1.11 to 0.28)	N=3934 0.19 (-0.64 to 1.03)
SDQ impact	N=4123 -0.34 (-0.11 to 0.05)	N=4000 0.04 (-0.05 to 0.12)
SDQ conduct	N=4265 -0.15 (-0.31 to 0.01)	N=3953 0.16 (-0.04 to 0.35)
SDQ emotion	N=4265 -0.12 (-0.33 to 0.10)	N=3953 0.06 (-0.18 to 0.30)
SDQ pro-social	N=4265 0.18 (-0.16 to 0.52)	N=3953 0.16 (-0.27 to 0.59)
PTRS emotional regulation	N=4203 0.11 (-0.04 to 0.27)	N=4019 -0.18 (-0.35 to 0.00)
PTRS pro-social behaviour	N=4203 0.16 (-0.01 to 0.32)	N=4019 -0.06 (-0.25 to 0.13)
PTRS social competence	N=4203 0.14 (0.01 to 0.29)*	N=4019 -0.11 (-0.29 to 0.01)
PTRS aggressive	N=4203 -0.13 (-0.23 to -0.04)*	N=4019 0.01 (-0.09 to 0.13)
PTRS internalising	N=4203 -0.16 (-0.27 to -0.04)	N=4009 0.01 (-0.12 to 0.22)
PTRS hyperactive behaviour	N=4226 -0.07 (-0.13 to -0.001)*	N=4040 0.03 (-0.04 to 0.11)
PTRS peer problems	N=4217 -0.12 (-0.22 to -0.02)*	N=4003 0.08 (-0.04 to 0.21)
PTRS relational aggression	N=4217 -0.08 (-0.20 to 0.01)	N=3998 0.05 (-0.07 to 0.16)
PTRS learning behaviours	N=4180 0.05 (0.003 to 0.10)*	N=3974 -0.01 (-0.07 to 0.05)

* indicates a significant difference between PATHS and control ($p < .05$)

¹ A negative mean score indicates the follow-up score in the intervention group is lower than in the control group for that construct. Measures reporting on positive behaviours, where higher scores equate to better outcomes, are shaded grey on the tables. All other measures are reported negatively, higher score equates to greater problems. The ICC values at the level of the classroom on SDQ constructs ranged between 0.07 and 0.28.

There were some sub-group differences. The benefits of PATHS increased with age, although not significantly. There was a significant impact at two years on students who tested as depressed and/or anxious at baseline. White students benefited more than other ethnic groups, though not significantly so. Poverty did not emerge as a moderator of results.

4.3. Standard Level-4 Triple-P

As Table 3 illustrates, the results for this programme are not promising. Children of parents attending Triple-P

sessions improved their behaviour and were happier six months after the course concluded, but at roughly the same rate as children in the control group receiving services as normal. These results are not consistent with most other Triple-P trials around the world. However, as far as we are aware, only four randomised trials (including this one) have been undertaken independent of the programme originator (see also Gallart and Matthey 2005; Hahlweg et al. 2010; Malti, Ribeaud, and Eisner 2011). When these four studies are viewed together, the evidence of impact on child development is equivocal.

Table 3: Child outcomes for Triple-P control and intervention groups

Child measure (cut-off)	Mean (SD) raw scores				Estimated mean difference (95% CI)	p value
	Control (n=73)		Intervention (n=73)			
	Baseline	6 mth	Baseline	6 mth		
SDQ conduct problems (4)	5.28 (2.0)	4.13 (2.2)	5.42 (2.3)	4.33 (2.2)	-0.15 (-0.79 to 0.50)	0.65
SDQ emotion problems (5)	5.19 (2.9)	4.25 (2.8)	5.40 (2.4)	4.20 (2.8)	0.84 (-0.70 to 0.86)	0.83
SDQ hyperactivity (7)	8.13 (1.8)	7.08 (2.5)	7.66 (1.9)	6.52 (2.4)	0.21 (-0.48 to 0.90)	0.54
SDQ peer problems (4)	4.29 (2.1)	3.63 (2.1)	4.75 (1.8)	3.97 (2.3)	-0.12 (-0.77 to 0.53)	0.72
SDQ pro-social behaviour (<4)	5.70 (2.1)	6.56 (2.0)	5.98 (2.4)	6.81 (2.2)	-0.10 (-0.63 to 0.43)	0.71
SDQ total difficulties (17)	22.89 (4.3)	19.09 (7.0)	23.23 (4.4)	19.02 (7.6)	0.22 (-1.78 to 2.21)	0.83
SDQ impact (2)	3.99 (2.5)	2.73 (2.8)	4.67 (2.9)	3.01 (3.4)	-0.09 (-1.01 to 0.84)	0.85
ECBI-I (127)	156.07 (39.8)	141.51 (43.2)	155.00 (38.8)	143.64 (45.3)	-4.39 (-14.64 to 5.86)	0.40
ECBI-P (11)	20.47 (8.5)	15.89 (9.8)	19.96 (8.3)	15.96 (9.5)	-1.97 (-4.28 to 0.35)	0.10
APS total	3.62 (0.7)	3.34 (0.7)	3.62 (0.7)	3.29 (0.9)	0.05 (-0.16 to 0.26)	0.65
APS laxness	3.53 (1.2)	3.13 (1.1)	3.52 (1.3)	3.15 (1.3)	-0.45 (-0.37 to 0.28)	0.79
APS verbosity	4.40 (0.7)	4.23 (1.0)	4.41 (0.9)	4.19 (1.0)	0.04 (-0.26 to 0.33)	0.81
APS over-reactivity	3.17 (1.0)	2.92 (0.9)	3.12 (0.9)	2.77 (1.1)	0.15 (-0.14 to 0.43)	0.32

* significant at $p < .05$
 ** significant at $p < .01$

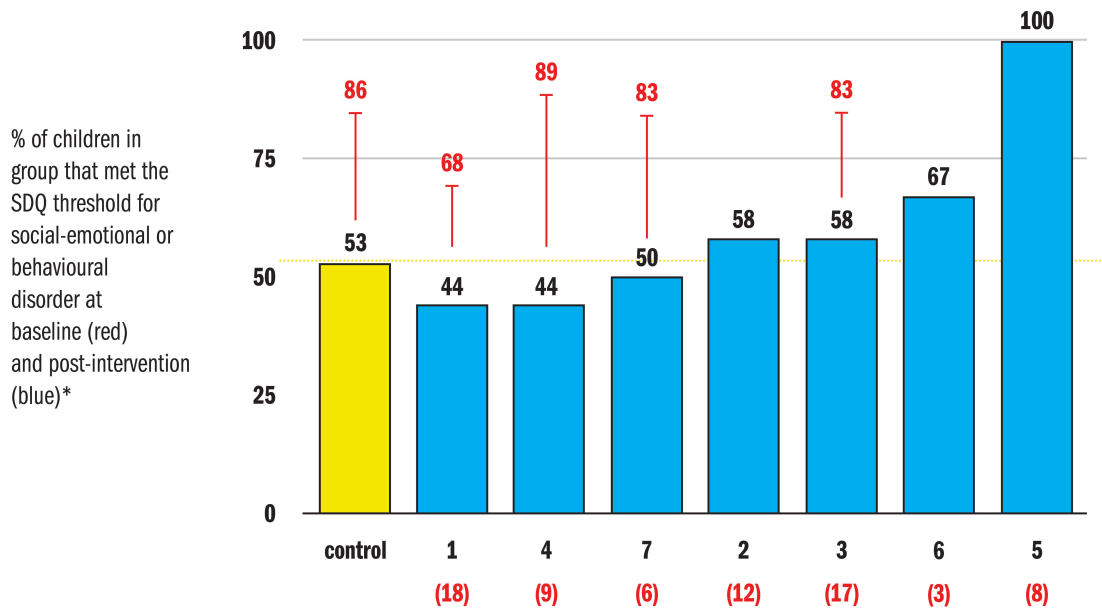
Poor fidelity of implementation may explain the failure to detect an impact on child development in the Birmingham Triple P trial (this is examined in more depth elsewhere). Brighter Futures inherited existing Triple-P practitioners who received “top-up” training from programme originator Matt Sanders. There was considerable variability in

the quality of provision and results achieved by these Triple-P practitioners. Not all parents received their full dose of Triple-P, with an average attendance across the groups of 40 percent. As Table 4 illustrates, one practitioner (group four) managed to reduce conduct disorders in the children of parents she or he was working with by 45 per

cent, while others failed to achieve any change. Emotional disorders for children whose parents participated in groups

led by three practitioners got worse, not better, than those in the control group receiving services as usual.

Table 4: Group-level outcomes for Triple-P



Control group (yellow) compared with Triple-P groups (one-seven)**

* Where a number only appears in blue (group two, five and six) the baseline and post-intervention % are identical

** Numbers in parentheses refer to number of parents

5. Discussion

The Brighter Futures work in Birmingham is a good example of how it is possible to create a “pull” for science to be used in social policy. Most attempts to link science and policy depend on “push”, injecting evidence into reluctant public systems. The *Common Language* work led Birmingham to demand and embed strong epidemiological data on the well-being of local children, good international evidence on “what works”, and reliable indications concerning the costs and benefits of competing investment options (Axford and Morpeth 2012). Perhaps most impressively, Birmingham committed to four experimental trials of new interventions. The local authority needed to know if these innovations worked. It would be surprising to find four

local authority-sponsored experimental evaluations in all of the other 149 English local authorities combined.

The resulting strategy had many positive features. It was proactive not reactive. It sought to forestall future problems rather than fire-fighting existing problems. It focused more on prevention and early intervention than on treatment, and it took a developmental approach to children. It tried to prevent children born into high-risk households falling behind. It sought to shift the distribution for all primary school children’s social and emotional regulation (see Rose 1998).

The strategy was unique at the time of its preparation in giving emphasis to evidence-based programmes proven by

the highest standards to work. And Brighter Futures introduced a new standard of financial accountability, seeking not only to invest in children's development but also to internalise the collection of high quality programme cost data at programme, school, group, or child level, as appropriate, with a view to producing estimates of short-term cost effectiveness and long-term return on investment.

The strategy had unanticipated benefits. The epidemiology had suggested that 15 percent of children aged three–four in the city would fall into the “high need” category of the SDQ “total difficulties” score. However, children's centres struggled to find these children in their catchment areas. Of an estimated pool of 437 children meeting that threshold in the relevant catchment areas, centres (collectively) needed to recruit 144 for the RCT but were only able to find 89 in the original recruitment period. This led to a concerted and ultimately successful effort to boost recruitment, including outreach and financial incentives, and taught important lessons about how children's centres can reach more disadvantaged families (Axford et al. 2012). This was valuable in the context of wider policy discussions about how children's centres need to refocus their activity.

In these and other respects Brighter Futures was radical and forward looking, and Birmingham should be commended for going first where many other local authorities are now hesitantly following. The leadership of the Chief Executive, elected politicians and the Strategic Director of Children's Services was a fundamental component in the success of Brighter Futures.

But there were also many limitations in the work. In the space available just two will be discussed. First, Brighter Futures started as and never quite progressed beyond the status of a “project” or “pilot”. It was a marginal not mainstream activity. It was a big project in a big pool. Over £40 million was set aside for Brighter Futures, but Birmingham was spending over £1.3 billion annually on its 260,000 children.

The marginal quality had significant effects when, almost inevitably in a local authority of its size, a preventable child death occurred (Radford 2010). In the political and media focus on this case, Brighter Futures was seen as a hindrance

– in the sense of significant resources being spent on prevention and early intervention rather than child protection in the traditional sense – when in other circumstances it might have been viewed as fundamental to righting the problem. Family Nurse Partnership (FNP), for example, is the most effective intervention known for preventing child maltreatment (e.g. MacMillan et al. 2009)

Second, the evidence-based programmes introduced by Brighter Futures had different impacts. Incredible Years was a success, as was PATHS initially (after one year of implementation). But after two years of implementation, PATHS only had an appreciable impact on children with emotional disorders. The Standard Level-4 Triple-P parenting programme not only had zero benefits overall but when poorly attended it generated potentially iatrogenic effects. The longer-term effects of these programmes – for example, six or twelve months after the programme ended – were not studied.

There is not space here to explore this variation fully. Although evidence-based programmes work, the size of effect is frequently small and can be diminished to nothing. In the case of Triple-P, fidelity of implementation was a problem. It is not sufficient simply to introduce an evidence-based programme; it has to be put into practice with great care and effort (Fixsen et al. 2005).

It is possible the local context may play a part in reducing impact of evidence-based programmes. PATHS, for example, has been proven to work in poorly funded schools in the United States serving high-risk communities (e.g. CPPRG 2010). European schools are invariably better funded, and so the existing provision – “services as usual” – against which programmes like PATHS are compared, might be stronger than in the US. In England, for instance, a project supported by national government, Social and Emotional Aspects of Learning (SEAL), had been addressing primary school pupils' social and emotional regulation for over five years. When Ross and colleagues (2012) evaluated PATHS in Northern Ireland they found only marginal benefits and an evaluation of the programme in Switzerland has produced similar results to those reported here (Malti et al. 2011).

It is possible that the different populations that the programmes served is part of the explanation for Incredible Years' success compared with Triple-P and PATHS. In Birmingham, Incredible Years was targeted at three–four years olds, while Triple-P was used for children aged four–nine years. Older children may have more entrenched difficulties. In addition, PATHS was delivered as a universal intervention to all classroom children; it is possible that greater effects may have been found with a targeted population.

It is known that the involvement of the programme originator in evaluations often skews the results in a positive direction (Eisner 2009). The four Birmingham trials were undertaken independent of the programme originators.

Another hypothesis is that some evidence-based programmes are more transportable than others. The consistent and positive findings for Incredible Years across contexts are striking. It works in the United States, Ireland, Wales, London, Birmingham and many other contexts worldwide. If transportability is a problem, what does Incredible Years have that other evidence-based programmes do not? This said, other studies show that Triple-P transports well in terms of impact, albeit – as indicated earlier in this article – with the involvement of the programme originator in the evaluation.

At this stage we can only speculate on reasons for the variability of results. What can be said for certain is that evidence-based programmes are a stepping-stone to future

improvement to children's health and development. They are not a panacea.

Although we were successful in applying a standard cost-effectiveness approach to each of the three trials, they could not be compared directly against one another due to differing ages and outcome measures. We calculated the cost per child for each programme and, in the case of Incredible Years, found an incremental cost-effectiveness ratio in line with previous economic evidence for this programme (Edwards et al. 2007). However, it has yet to be established whether there was any longer term return on investment from the Brighter Futures initiative in terms of benefits to the judicial system, education system or health care system.

In conclusion, Brighter Futures was a brave experiment. It demanded better science in policy formulation and involved testing evidence-based programmes in real world settings. Other studies had done this previously, including in relation to Incredible Years (Gardner et al. 2006; Hutchings et al. 2007a/b), but Brighter Futures went further along the efficacy-effectiveness spectrum by testing implementation city-wide in a range of regular services without a prior tradition of implementing these evidence-based programmes. It will leave an important legacy for the city, its children and other large-scale systems in Europe. Much can be learned from the Birmingham experience by those engaged in continuing attempts to improve children's health and development.

References

- Aos, Steve, Roxanne Lieb, Jim Mayfield, Marna Miller, and Annie Pennuci. 2004. *Benefits and Costs of Prevention and Early Intervention Programs for Youth*. Olympia: Washington State Institute for Public Policy.
- Arnold, David S., Susan G. O'Leary, Lisa S. Wolff, and Maureen M. Acker. 1993. The Parenting Scale: A Measure of Dysfunctional Parenting in Discipline Situations. *Psychological Assessment* 5: 137–44.
- Axford, Nick, Louise Morpeth, Michael Little, and Vashti Berry. 2008. Linking Prevention Science and Community Engagement: A Case Study of the Ireland Disadvantaged Children and Youth Programme. *Journal of Children's Services* 3 (2): 40–54.
- Axford, Nick, and Tim Hobbs. 2011. Getting the Measure of Child Health and Development Outcomes (1): A Method for Use in Children's Services Settings. *Child Indicators Research* 4 (1): 81–100.
- Axford, Nick, David Jodrell, and Tim Hobbs. 2012. Making Child Well-being Data Work Hard: Getting from Data to Policy and Practice. *Child Indicators Research*. doi:10.1007/s12187-012-9163-5
- Axford, Nick, Minna Lehtonen, Kate Tobin, Dwan Kaoukji, and Vashti Berry. 2012. Engaging Parents in Parenting Programs: Lessons from Research and Practice. *Children and Youth Services Review* 34 (10): 2061–71.
- Axford, Nick, and Louise Morpeth. 2012. The Common Language Prevention Operating System: From Strategy Development to Implementation of Evidence-based Practice. In *Handbook of Implementation Science for Educational Psychology*, ed. Barbara Kelly and Daniel F. Perkins, 443–58. Cambridge: Cambridge University Press.
- BCC (Birmingham City Council). 2007. *A Brighter Future for Children and Young People: The Birmingham Strategy*. http://ebriefing.bgfl.org/bcc_ebrief/content/resources/resource.cfm?id=4314&key=&zz=20111021114309327&zs=n
- Bodenman, Guy, Annette Cina, Thomas Ledenmann, and Matthew R. Sanders. 2008. The Efficacy of Positive Parenting Program (Triple P) in Improving Parenting and Child Behavior: A Comparison with Two Other Treatment Conditions. *Behaviour Research and Therapy* 46 (4): 411–42.
- Collishaw, Stephan, Barbara Maughan, Robert Goodman, and Andrew Pickles. 2004. Time Trends in Adolescent Mental Health. *Journal of Child Psychology and Psychiatry* 45 (8): 1350–62.

- Collishaw, Stephan, Barbara Maughan, Lucy Natarajan, and Andrew Pickles. 2010. Trends in Adolescent Emotional Problems in England: A Comparison of Two National Cohorts Twenty Years Apart. *Journal of Child Psychology and Psychiatry* 51 (8): 885–94.
- CPPRG (Conduct Problems Prevention Research Group) 2010. The Effects of a Multiyear Universal Social-emotional Learning Program: The Role of Student and School Characteristics. *Journal of Consulting and Clinical Psychology* 78 (2): 156–68.
- Durlak, Joseph A., Roger P. Weissberg, Allison B. Dymnicki, Rebecca D. Taylor, and Kriston B. Schellinger. 2011. The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis Of School-Based Universal Interventions. *Child Development* 82 (1): 405–32.
- Edwards, Rhiannon Tudor, Alan Ó Céilleachair, Tracey Bywater, Dyfrig A. Hughes, and Judy Hutchings. (2007) A Parenting Programme for Children at Risk of Developing Conduct Disorder: A Cost-Effectiveness Analysis. *British Medical Journal* 334: 682.
- Eisenstadt, Naomi. 2011. *Providing a Sure Start: How Government Discovered Early Childhood*. Bristol: Policy Press.
- Eisner, Manuel. 2009. No Effects in Independent Prevention Trials: Can We Reject the Cynics' Hypothesis? *Journal of Experimental Criminology* 5:163–83.
- Eyberg, Sheila M., and Arthur W. Ross. 1978. Assessment of Child Behavior Problems: The Validation of a New Inventory. *Journal of Clinical Child Psychology* 7:113–16.
- Fixsen, Dean L., Sandra F. Naoom, Karen A. Blase, Robert M. Friedman, and Frances Wallace. 2005. *Implementation Research: A Synthesis of the Literature*. Tampa: USE.
- Gardner, Frances, Jennifer Burton, and Ivana Klimes. 2006. Randomised Controlled Trial of a Parenting Intervention in the Voluntary Sector for Reducing Child Conduct Problems: Outcomes and Mechanisms of Change. *Journal of Child Psychology and Psychiatry* 47:1123–32.
- Gallart, Sonia C., and Stephen Matthey. 2005. The Effectiveness of Group Triple P and the Impact of the Four Telephone Contacts. *Behaviour Change* 22 (2): 71–80.
- Goodman, Robert. 1997. The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry* 38 (5): 581–86.
- Greenberg, Mark T., and Carol Kusché. 2002. *Promoting Alternative Thinking Strategies (PATHS)*, Blueprints for Violence Prevention 10. Boulder: Center for the Study and Prevention of Violence, University of Colorado.
- Hahlweg, Kurt, Nina Heinrichs, Annett Kuschel, Heike Bertram, and Sebastian Naumann. (2010). Long-term Outcome of a Randomized Controlled Universal Prevention Trial through a Positive Parenting Program: Is it Worth the Effort? *Child and Adolescent Psychiatry and Mental Health* 4:14–27.
- Hobbs, Tim, Nick Axford, and David Jodrell. 2011. Getting the Measure of Child Health and Development Outcomes (2): The Picture for a Local Authority in England. *Child Indicators Research* 4 (1): 81–100.
- Hutchings, Judy, Tracey Bywater, Dave Daley, Frances Gardner, Chris Whitaker, Karen Jones, Catrin Eames, and Rhiannon T. Edwards. 2007. Parenting Intervention in Sure Start Services for Children at Risk of Developing Conduct Disorder: Pragmatic Randomized Controlled Trial. *British Medical Journal* 334 (7595): 1–7.
- Hutchings, Judy, Tracey Bywater, and Dave Daley. 2007. Early Prevention of Conduct Disorder: How and Why did the North and Mid Wales Sure Start Study Work? *Journal of Children's Services* 2 (2): 4–14.
- Little, Michael, and Ali Abunimah. 2007. Improving Outcomes for Children in the Island of Ireland: The Role of Philanthropic Investment. *Journal of Children's Services* 2 (2): 60–67.
- MacMillan, Harriet L., C. Nadine Wathen, Jane Barlow, David M. Fergusson, John M. Leventhal, and Heather N. Taussig. 2009. Interventions to Prevent Child Maltreatment and Associated Impairment. *Lancet* 373 (9659): 250–66.
- Malti, Tina, Denis Ribeaud, and Manuel P. Eisner. 2011. The Effectiveness of Two Universal Preventive Interventions in Reducing Children's Externalizing Behavior: A Cluster Randomized Controlled Trial. *Journal of Clinical Child & Adolescent Psychology* 40 (5): 677–92.
- Maughan, Barbara, Stephan Collishaw, Howard Meltzer, and Robert Goodman. 2008. Recent Trends in UK Child and Adolescent Mental Health. *Social Psychiatry and Psychiatric Epidemiology* 43: 305–10.
- Olds, David L., Peggy G. Hill, Sharon F. Mihalic, and Ruth A. O'Brien. 1998, updated 2006. *Prenatal and Infancy Home Visitation by Nurses: Blueprints for Violence Prevention*. Boulder, CO: Center for the Study and Prevention of Violence, Institute of Behavioral Science, University of Colorado.
- Radford, John. 2010. *Serious Case Review Under Chapter VIII "Working Together to Safeguard Children" In respect of the Death of a Child, Case Number 14*. Birmingham: Birmingham Safeguarding Children Board.
- Rose, Geoffrey. 1998. *Rose's Strategy of Preventive Medicine*. Oxford: Oxford University Press.
- Ross, Steven M., Mary K. Sheard, Alan Cheung, Louise Elliott, and Robert Slavin. 2012. Promoting Primary Pupils' Social-emotional Learning and Pro-social Behaviour: Longitudinal Evaluation of the Together4All Programme in Northern Ireland. *Effective Education* 3 (2): 61–81.
- UNICEF. 2007. *Child Poverty in Perspective: An Overview of Child Wellbeing in Rich Countries*. Florence: UNICEF Innocenti Research Centre.
- Webster-Stratton, Carolyn. 1994. Advancing Videotape Parent Training: A Comparison Study. *Journal of Consulting and Clinical Psychology* 66: 715–30.

Michael Little

mlittle@dartington.org.uk

Vashti Berry

vberry@dartington.org.uk

Louise Morpeth

lmorpeth@dartington.org.uk

Sarah Blower

sblower@dartington.org.uk

Nick Axford

naxford@dartington.org.uk

Rod Taylor

rod.taylor@pms.ac.uk

Tracey Bywater

tracey.bywater@york.ac.uk

Minna Lehtonen

mlehtonen@dartington.org.uk

Kate Tobin

ktobin@dartington.org.uk

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components

Bryanna Hahn Fox, Department of Criminology, University of South Florida, Tampa, Florida

David P. Farrington, Institute of Criminology, Cambridge University, United Kingdom

Maria M. Ttofi, Institute of Criminology, Cambridge University, United Kingdom

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

► **Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components** Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components

Bryanna Hahn Fox, Department of Criminology, University of South Florida, Tampa, Florida
David P. Farrington, Institute of Criminology, Cambridge University, United Kingdom
Maria M. Ttofi, Institute of Criminology, Cambridge University, United Kingdom

Bullying prevention programs have been shown to be generally effective in reducing bullying and victimization. However, the effects are relatively small in randomized experiments and greater in quasi-experimental and age-cohort designs. Programs that are more intensive and of longer duration (for both children and teachers) are more effective, as are programs containing more components. Several program components are associated with large effect sizes, including parent training or meetings and teacher training. These results should inform the design and evaluation of anti-bullying programs in the future, and a system of accreditation of effective programs.

Bullying and victimization (being bullied) have been shown to have many serious and long-term effects on the physical and mental health of children (Ttofi and Farrington 2008). Specifically, results from a systematic review of bullying and its effect on later criminal behavior suggest that school bullying perpetration is a significant predictor of offending an average of nearly six years in the future, even after controlling for other major risk factors for criminality (adjusted Odds Ratio [OR] = 1.89; 95% confidence interval (CI) = 1.60–2.23) (Ttofi et al. 2011; Farrington et al. 2012). Being bullied is shown to be a significant predictor of depression an average of seven years later, even after controlling for other major childhood risk factors (Ttofi and Farrington 2011). Even more significant is the relationship between bullying perpetration and the commission of future violence (OR = 2.04; 95% CI = 1.69–2.45), with the effects again persisting many years later (Ttofi et al. 2012). As over half of all children are bullied, and half admit to bullying in school (Farrington 1993), these findings have significant and widespread implications.

Consequently, a great deal of resources have been invested in programs aimed at reducing school bullying and victimization (Ttofi and Farrington 2011), with several evaluations and systematic reviews conducted of the overall

effectiveness of the programs (for example, Smith et al. 2004; Vreeman and Carroll 2007; Farrington and Ttofi 2009). However, no firm conclusions have been drawn regarding the specific components of anti-bullying programs that yield the best, and most consistent, results. This paper addresses this issue by focusing on the fundamental components, implementation features, and methodological designs of successful anti-bullying programs, given that these core elements are the foundation upon which the success of any program is based. Drawing on the results of recent meta-analyses examining the effectiveness of anti-bullying programs from around the world (Farrington and Ttofi 2009; Ttofi and Farrington 2011), two aspects of program implementation – duration and intensity – were found to be highly significant in decreasing both bullying and victimization, while parent training and teacher training were among the most effective program components. Overall, anti-bullying programs were effective in reducing bullying by 20 to 23 percent, and victimization by 17 to 20 percent. Through the present research we aim to develop a better understanding of the research design, implementation features, and program components that are most effective in preventing school bullying and victimization, and lay the foundation for more successful future anti-bullying programs.

1. Research Design

Anti-bullying programs utilize four main research design types: 1) randomized experiments; 2) intervention/control comparisons with before-and-after measures of bullying and victimization; 3) other intervention/control comparisons; and 4) age-cohort designs (Ttofi and Farrington 2011). Of these, it has been widely recognized that randomized experiments are the “gold standard” in terms of demonstrating most convincingly whether a specific treatment has an effect on an outcome (Farrington and Welsh 2005). Provided that a sufficiently large number of units are randomly assigned during the experiment, and the participants in the control and treatment conditions are comparable on all measured and unmeasured extraneous variables (within the limits of natural fluctuation), the randomized experiment has the highest potential internal validity of all design types (Weisburd, Lum, and Petrosino 2001; Ttofi and Farrington 2011). While it would be expected that all prevention programs would opt to utilize randomized experiments because of the scientific advantages of this design, several difficulties and threats may prohibit the use of randomized experiments, or prevent the full benefits of the design from being achieved. Most notable of these limitations is the added time, cost, and cooperation necessary to enable a proper randomized experiment to occur. Some institutions refuse to participate in studies requiring so much effort and cooperation (Weisburd, Lum, and Petrosino 2001), and therefore those that do agree to participate may not be a representative sample of the whole population. This may limit the external validity of randomized experiments, while differential attrition from the treatment and control conditions may pose a threat to internal validity (Farrington 2003).

Finally, having an insufficient number of randomized units may threaten the validity of a randomized experiment. Unfortunately this is often the case for bullying prevention programs, which tend to randomize a small number of schools, rather than a large number of children within the schools (Ttofi and Farrington 2011, 30). Table 1 presents a detailed description of the units of randomization used in the bullying prevention programs included in the Campbell Collaboration meta-analyses conducted by Farrington and

Ttofi (2009). It can be seen that only Karna et al. (2011) randomized a reasonably large number of schools (78).

Table 1: Units of randomization used in bullying prevention experiments

Children:
De Rosier (2004) => 18 experimental students from each of eleven schools (N = 381)
Beran & Shapiro (2005) => 66 experimental students from two schools (N = 129)
Boulton & Flemington (1996) => 84 experimental students from one school (N = 164)
Meyer & Lesch (2000) => 18 experimental students from three schools (N = 36)
Classes:
Baldry & Farrington (2004) => 10 classes (N = 224)
Schools:
Cross et al. (2004) => 29 schools (N = 1,957)
Fekkes et al. (2006) => 50 schools (N = 2,221)
Fonagy et al. (2009) => 3 schools in experimental 1 condition; 3 schools in experimental 2 condition; 3 control schools (N = 1,345)
Frey et al. (2005) => 6 schools (N = 1126)
Hunt (2007) => 7 schools (N = 400)
Jenson & Dieterich (2007) => 28 schools (N = 668)
Karna et al. (2011) => 78 schools (N = 5,641)
Rosenbluth et al. (2004) => 12 schools (N = 1,763)
Sprober et al. (2006) => 3 schools (N = 144)

Note: N represents total sample size (number of students) in experimental and control conditions together.

Quasi-experimental evaluations with before-and-after measures of the outcome variable are widely considered to be the second-best option to randomized experiments, given that they avoid many of the most significant participant cooperation issues encountered by randomized experiments, although here too internal validity is threatened by differential attrition between control and treatment groups. The internal validity of the design is even more threatened when no measure of the outcome is taken prior to the study in both the control and treatment conditions, as is the case in other intervention/control studies. These studies have no way of establishing original comparability between the treatment and control groups, so if one group

is worse than the other to start with, regression to the mean may occur and threaten the internal validity of results.

The final design type, the age-cohort study, occurs when subjects of a given age after the intervention are compared with a different set of subjects of the same age in the same unit of examination (e.g. school) before the intervention. While this design is often considered methodologically inferior to the randomized and quasi-experimental (with before-and-after measures) designs, the age-cohort design has many advantages in eliminating selection, aging, regression, and differential attrition effects, resulting in high external validity (Olweus 2005; Ttofi and Farrington 2011).

In Farrington and Ttofi's meta-analysis of school bullying and victimization programs (2009), the design of

each of the forty-four evaluations was evaluated to determine which design type yielded the most significant effect size overall. Table 2 shows that the before-and-after quasi-experimental designs yielded the strongest effects on bullying (weighted mean OR = 1.60, $p < .0001$), while the other intervention/control studies were most successful for victimization (weighted mean OR = 1.43, $p < .006$). Very interestingly, age-cohort designs were found to be the next most effective for effects on both bullying and victimization (bullying weighted mean OR = 1.36, $p < .0001$; victimization weighted mean OR = 1.29, $p < .0001$). Randomized experiments yielded the lowest overall effect size of the four design types for victimization (OR = 1.17, $p < .050$), and no significant effects for bullying (Ttofi and Farrington 2011).

Table 2: Effect sizes for bullying and victimization programs with different designs

Program design	Bullying			Victimization		
	OR	CI	<i>p</i>	OR	CI	<i>p</i>
Randomized experiments Weighted mean (n=14)	1.10	0.97 - 1.26	n.s.	1.17	1.00 - 1.37	.050
Before/after intervention/control Weighted mean (n=17)	1.60	1.45 - 1.77	.0001	1.22	1.06 - 1.40	.007
Other intervention/control Weighted mean (n=4)	1.20	1.04 - 1.38	.010	1.43	1.11 - 1.85	.006
Age-cohort designs Weighted mean (n=9)	1.51	1.35 - 1.70	.0001	1.44	1.21 - 1.72	.0001
Total weighted mean (n=44)	1.36	1.26 - 1.47	.0001	1.29	1.18 - 1.42	.0001

Note: OR = odds ratio; CI = confidence interval.

2. Implementation Features

Program implementation features, such as the duration and intensity of the program for children and teachers, are related to a reduction in both bullying and victimization (Farrington and Ttofi 2009); see Table 3. The Farrington and Ttofi meta-analysis (2009) was among the first to successfully isolate program duration from inten-

sity, which is a highly important distinction (Carmody and Baer 2009, 636), with results suggesting that the longer-lasting and more intensive programs are more successful than shorter and less intensive programs, when controlling for other program elements (Ttofi and Farrington 2011).

Table 3: Effect sizes for implementation features and program components

	Cat (N) OR	Cat (N) OR	<i>p</i>
Bullying			
<i>Implementation Features</i>			
Intensity for children	19- (19) 1.25	20+ (13) 1.62	.0001
Duration for children	240- (20) 1.17	270+ (20) 1.49	.0001
Intensity for teachers	9- (16) 1.19	10+ (20) 1.52	.0001
Duration for teachers	3- (19) 1.22	4+ (19) 1.50	.0004
<i>Program Components</i>			
Parent training/meetings	No (24) 1.25	Yes (17) 1.57	.0001
Teacher training	No (13) 1.24	Yes (28) 1.46	.006
Total components	10- (23) 1.30	11+ (18) 1.48	.009
Victimization			
<i>Implementation Features</i>			
Intensity for children	19- (18) 1.21	20+ (14) 1.42	.002
Duration for children	240- (20) 1.15	270+ (20) 1.35	.001
Intensity for teachers	9- (15) 1.22	10+ (21) 1.37	.028
Duration for teachers	3- (18) 1.18	4+ (20) 1.41	.0003
<i>Program Components</i>			
Parent training/meetings	No (24) 1.20	Yes (17) 1.41	.0001
Teacher training	No (11) 1.24	Yes (30) 1.33	ns
Total components	10- (22) 1.33	11+ (19) 1.30	ns

Notes: Cat = dichotomized category of variable; OR =weighted mean odds ratio; duration in days; intensity in hours.

While the general consensus in the field of prevention is that longer intervention is better (Gottfredson and Wilson 2003, 29; Durlak 1995; Gottfredson 1997), not all meta-analyses of developmental intervention programs confirm this result. For example, Gottfredson and Wilson’s meta-analysis of school-based substance abuse prevention programs (2003), found the length of the intervention (a mixture of program duration and intensity) to have a positive but non-significant relationship with the outcome effect size. Closer analysis suggested that the positive relationship was driven by a single out-

lier, which was an “unusually intensive program” involving weekly contact between program staff and students over its two-year duration (Gottfredson and Wilson 2003, 33). This finding illustrates the importance of isolating program duration from intensity in any assessment of the impact of implementation features on a program’s effectiveness. Gottfredson and Wilson (2003) acknowledge this point too: “It may also be the case that program length is a poor proxy for program intensity. A more sensitive measure of program intensity may have produced different results” (36).

Farrington and Ttofi's meta-analysis (2009) created separate measures for each program's duration and intensity for both children and teachers. Program duration was defined as the length of the intervention from start to finish, while intensity of the program was defined as the amount of contact, in hours, between program staff and children across the duration of the program. Results indicate that the programs with higher intensity for children (20 hours or more of contact) were significantly more effective in reducing both bullying (OR = 1.62, $p < .0001$) and victimization (OR = 1.42, $p < .002$) than the lower-intensity programs (Ttofi and Farrington 2011). Intensity of training for teachers was also found to increase program effectiveness, as the more intensive programs had higher effect sizes for bullying (OR = 1.52, $p < .0001$) and victimization (OR = 1.37, $p < .028$).

Duration was also significantly related to effectiveness, with longer programs for children found to be more successful (bullying OR = 1.49, $p < .001$; victimization OR = 1.35, $p < .001$) as compared to shorter programs (Ttofi and Farrington 2011). The longer-duration teacher training programs were also significantly more effective in reducing both bullying (OR = 1.50, $p < .0004$) and victimization (OR = 1.41, $p < .0003$) than the shorter teacher training programs (Ttofi and Farrington 2011).

These results support the findings of other bullying prevention program evaluations (Olweus 2005; Smith 1997), that programs need to be long-lasting and intensive in order to create and maintain the necessary school ethos to effectively combat bullying (Ttofi and Farrington 2011). Still, it is feared that "longer time commitments may be a barrier to the ability and willingness of individuals to participate" (Carmody and Baer 2009, 627). This is of particular concern for the children at highest risk for bullying perpetration, as bullies tend to have negative attitudes towards school work and teachers, and tend to be unsuccessful in school (Farrington 1993). It is increasingly likely that these children miss long periods of school, and consequently avoid attending or participating in programs that are more intensive or long-lasting.

Similarly, victims may be at risk of not participating in programs of higher intensity and duration, although for

different reasons than the perpetrators. Given that victims typically experience a great deal of psychological and/or physical distress resulting from the bullying (Mellor 1991), they often find it difficult to concentrate on their school work, and may be afraid to go to school because of their fear of being victimized (Farrington 1993, 406). In fact, one study found that 15 percent of persistent school absentees reported being bullied as their primary reason for avoiding school, and 19 percent said that it was one of the major reasons for their continued absence (Reid 1989).

While rigorous analyses of forty-four international prevention programs indicate that program intensity and duration are two separate, but highly significant implementation features in reducing bullying and victimization in schools (Farrington and Ttofi 2009), it is important to recognize that even these critical elements have limitations that may inhibit their benefits from being fully reached, and that no program should be based on duration and intensity alone. Therefore, the program components with the most significant effects on both bullying and victimization must be considered as well.

3. Program Components

Ttofi and Farrington (2011) found that several components (notably firm disciplinary methods and improved playground supervision) were associated with large effect sizes, while work with peers was associated with small effect sizes (see also Ttofi and Farrington 2012).

As program design and implementation features alone may not impact those at highest risk of perpetration and victimization, it has been suggested that new anti-bullying initiatives must go beyond the scope of the school and target additional areas such as the family and teachers of the children (Ttofi and Farrington 2011, 46). Several prevention programs already include such components, with parent and teacher training among the most popular means of extending the program elsewhere in schools and families. As bullied children often do not share their victimization experiences with anyone, parents and teachers tend not to know of bad behavior or not to discuss it with the bullies (Fekkes, Pijpers, and Verloove-Vanhorick 2005), educating parents and teachers on what

to look out for, and how to handle bullies and victims, was thought to be a highly beneficial addition to bullying prevention programs to create awareness of the problem and knowledge about how to address it (Ttofi and Farrington 2011).

Relevant research on parent and teacher training suggests that positive outcomes occur when families and educators are included in school prevention programs (Flay 1999). For instance, trained teachers have been found to be more effective and have more favorable student outcomes, more likely to implement and support other components of the prevention program, and more likely to continue to use a program after its implementation than teachers without program training (Mihalic et al. 2004; Taggart et al. 1990; McCormick, Steckler, and McLeroy 1995; Gingiss 1992). This has led some to suggest that teachers are the “primary agents of school-based prevention efforts, and their support, motivation, and ‘buy-in’ is crucial to implementation success” (Fagan and Mihalic 2003, 238; also Hunter, Elias and Norris 2001).

Parent training programs have also been quite successful in leading to desired outcomes in a variety of prevention programs (Piquero et al 2009), as parent training was followed by significant improvements in children’s behaviors for at least two thirds of treated families in several studies (Webster-Stratton, Reid, and Hammond 2004; Brestan and Eyberg 1998; Taylor and Biglan 1998). Together, these findings are strongly related to the fact that parenting behaviors are known to be the most important risk factor for early-onset conduct problems in children (Webster-Stratton, Reid, and Hammond 2004, 105). Consequently, having meetings with parents and training them how to identify and prevent bullying in their children should predictably play a significant role in the success of anti-bullying programs.

Still, several exceptions have been found regarding the effectiveness of both parent and teacher training components of school-based prevention programs, with some studies finding no improvement, or even negative effects, when parent and teacher training is included (Griest and Forehand 1982; Ferber, Keeley, and Shemberg 1974; Taylor

and Biglan 1998; Webster-Stratton 1990; Wahler 1980). In one of the first studies to examine the added benefits of combining teacher training with parent training, child training, or both, to treat children with a conduct disorder, conditions including teacher training were found to significantly improve the children’s behavior at school. However, similar effects were also found when only child and parent training was utilized, indicating that no teacher intervention was needed for desired effects to occur (Webster-Stratton, Reid, and Hammond 2004, 121). On the other hand, some studies have found that programs utilizing parent training resulted in significant improvements in children’s behavior at home, but not in school or with peers (Webster-Stratton and Hammond 1997; Webster-Stratton, Reid, and Hammond 2004). Some families receiving parent training actually reported a significantly higher level of parenting stress and/or negative life events, leading to a negative impact on the child’s behavior (Kazdin 1995; Webster-Stratton 1985; Webster-Stratton and Hammond 1990). As many children who are at highest risk of bullying are disproportionately from lower socioeconomic status families with poor parenting techniques (Farrington 1993), or from single-parent families with high stress levels and family burdens (Strain, Young, and Horowitz 1981), those who stand to benefit most from parent training are those least likely to complete it due to life stress, work conflicts, or lack of motivation (Spoth et al. 1996).

In the meta-analysis conducted by Farrington and Ttofi (2009), both parent and teacher training had significant and positive effects on the reduction of bullying (parent training OR = 1.57, $p < .0001$; teacher training OR = 1.46, $p < .006$) compared to programs without these components (see Table 3). Parent training was also significantly related to reducing victimization (OR = 1.41, $p < .0001$) compared to programs without parent training, but teacher training was not found to have a significant effect on victimization (Ttofi and Farrington 2011) (see Table 3). Together, these findings indicate that parent training and teacher training are individually highly beneficial components of anti-bullying programs, though it is not possible to determine their combined, additive impact on bullying and victimization prevention in the original Farrington and Ttofi study (2009).

In addition to evaluating the effect of parent and teacher training on the success of anti-bullying programs, the impact of the total number of program components is also evaluated in order to account for the fact that when several different program components are included, there is a higher likelihood of influencing every child, family, and school. Similar results were found in a developmental intervention program evaluation, where the total number of program components successfully and significantly predicted future criminal convictions for the program attendees (Koegl et al. 2009, 429). This finding indicates that prevention programs comprised of more components will be more effective than programs with fewer components overall.

Supporting this prediction, a strong and significant effect size on bullying was found for programs containing eleven or more components ($OR = 1.48, p < .009$), though there was not a significant effect size for programs containing a high number of total components on victimization. Combined with the non-significant effect of teacher training in reducing victimization, it seems that the effects of prevention program components on victimization are weaker than the effects of similar components on bullying. It is possible that the programs have more effect on bullies than on victims because their main aim is to prevent bullying from occurring. Still, additional research could be conducted to investigate this issue further.

4. Conclusion

Taken together, these findings indicate that anti-bullying programs work, as the combined effect of the various program designs, implementations, and components is shown to decrease bullying and victimization by an average of 17 to 23 percent (Ttofi and Farrington 2011). This figure encompasses the full span of anti-bullying programs, including, for instance, programs of shorter duration, with lower intensity, without parent training, and with a small total number of components. Therefore it is possible that by refining future programs to comprise only elements, implementation features, and designs known to be most effective, the overall effectiveness of anti-bullying programs would be ever greater.

Still, certain program features turned out to be less successful than expected, including the use of randomized experiments, teacher training (effect on victimization), and total number of program components (effect on victimization). With respect to the randomized experiments, it is not contested that they are the most methodologically superior design in principle, but the manner in which they were utilized in the analyzed programs may have contributed to the lower than expected effectiveness. Specifically, few randomized experiments contained a sufficient number of randomized units (as schools or school classes were the most common units of randomization), leading to a decrease in internal validity and ultimately less significant results. Differential attrition also played a role in decreasing the effects in the randomized experiments, with one of the programs in the analysis suffering twice the attrition rate for the control condition, as compared to the experimental condition (Ttofi and Farrington 2011, 44). It is not surprising that intervention schools are more motivated than control schools to continue participation. If methodological issues such as these were overcome in future studies, it is possible that randomized experiments would yield better outcomes, though establishing this would require additional evaluations and meta-analyses.

The implementation of the programs is very important, with greater duration and intensity for children and teachers yielding better results for both bullying and victimization. Similarly, including parent and teacher training as program components was found to be highly effective for bullying, while parent training (but not teacher training) was found to be a significant predictor of effectiveness for victimization. The total number of program components is also shown to be important to a program's ability to reduce school bullying, while this effect did not apply to victimization.

It is possible to refine this analysis further, by including or prioritizing the program components that have been the most successful overall, and excluding those which are not significant, or perhaps detrimental. For instance, effective components such as improved playground supervision, firm disciplinary measures, good classroom management and clear rules, school conferences, information for par-

ents, cooperative group work, and a school-wide anti-bullying policy, as well as parent and teacher training, would be included in such a program. However, work with peers (peer mediation, peer mentoring, engagement of bystanders in bullying situations), which is actually found to have negative effects on bullying and victimization (Ttofi and Farrington 2011), would not be included in such a program, as peer-based components may “reinforce the aggressive behavior of school bullies and promote a cycle of violence” (Ttofi and Farrington 2012, line 209–210).

It should be noted that the figures quoted in this research show only correlations between program elements and effect sizes, and additional research is required to identify causal effects by randomly allocating elements to programs. This would allow researchers to compare children receiving a certain program with equivalent children receiving the same program but without component X.

References

- Brestan, Elizabeth V., and Sheila M. Eyberg. 1998. Effective Psychosocial Treatments of Conduct-Disordered Children and Adolescents: 29 Years, 82 Studies, and 5,272 Kids. *Journal of Clinical Child Psychology* 27:180–89.
- Carmody, James, and Ruth A. Baer. 2009. How Long Does a Mindfulness-Based Stress Reduction Program Need to Be? A Review of Class Contact Hours and Effect Sizes for Psychological Distress. *Journal of Clinical Psychology* 65:627–38.
- Durlak, James A. 1995. *School-Based Prevention Programs for Children and Adolescents*. Thousand Oaks, CA: Sage.
- Fagan, Abigail, and Sharon Mihalic. 2003. Strategies for Enhancing the Adoption of School-Based Prevention Programs: Lessons Learned from the Blueprints for Violence Prevention Replications of the Life Skills Training Program. *Journal of Community Psychology* 31:235–53.
- Farrington, David P. 1993. Understanding and Preventing Bullying. In *Crime and Justice*, vol. 17, ed. Michael Tonry, 381–458. Chicago: University of Chicago Press.
- Farrington, David P. 2003. Methodological Quality Standards for Evaluation Research. *Annals of the American Academy of Political and Social Sciences* 587:49–68.
- Farrington, David P., Friedrich Lösel, Maria M. Ttofi, and Nikos Theodorakis. 2012. *School Bullying, Depression and Offending Later in Life: An Updated Systematic Review of Longitudinal Studies*. Stockholm: Swedish National Council for Crime Prevention.
- Farrington, David P., and Maria M. Ttofi. 2009. School-Based Programs to Reduce Bullying and Victimization. *Campbell Systematic Reviews* 6.
- Farrington, David P., and Brandon C. Welsh. 2005. Randomized Experiments in Criminology: What Have We Learned in the Last Two Decades? *Journal of Experimental Criminology* 1:9–38.
- Fekkes, Minne, Frans I. M. Pijpers, and S. Pauline Verloove-Vanhorick. 2005. Bullying: Who Does What, When and Where? Involvement of Children, Teachers and Parents in Bullying Behavior. *Health Education Research* 20:81–91.
- Flay, Brian R. 1999. *Effects of Positive Action Program and Positive Action Training on Outcomes*. Twin Falls, ID: Positive Action.
- Ferber, Harold, Stuart M. Keeley, and Kenneth M. Shemberg. 1974. Training Parents in Behavior Modification: Outcomes of and Problems Encountered in a Program after Patterson's Work. *Behavior Therapy* 5:415–19.
- Gingiss, Phyllis L. 1992. Enhancing Program Implementation and Maintenance through a Multiphase Approach to Peer-Based Staff Development. *Journal of School Health* 62:161–66.
- Griest, Douglas L., and Rex Forehand. 1983. How Can I Get Any Parent Training Done with All these Other Problems Going On? *Child and Family Behavior Therapy* 4:73–80.
- Gottfredson, Denise C. 1997. School-Based Crime Prevention. In *Preventing Crime: What Works, What Doesn't, What's Promising: A Report to the United States Congress*, ed. Lawrence Sherman, Denise C. Gottfredson, Doris MacKenzie, John Eck, Peter Reuter, and Shawn Bushway. Washington, DC: U.S. Department of Justice, Office of Justice Programs.
- Gottfredson, Denise C., and David B. Wilson. 2003. Characteristics of Effective School-Based Substance Abuse Prevention. *Prevention Science* 4:27–38.
- Hunter, Lisa, Maurice J. Elias, and Jacqueline Norris. 2001. School-Based Violence Prevention: Challenges and Lessons Learned from an Action Research Project. *Journal of School Psychology* 39:161–75.
- Karna, Antti, Marinus Voeten, Todd D. Little, Elisa Poskiparta, Anne Kaljonen, and Christina Salmivalli. 2011. A Large-Scale Evaluation of the KiVa Anti-bullying Program: Grades 4–6. *Child Development* 82:311–30.
- Kazdin, Alan E. 1995. Child, Parent, and Family Dysfunction as Predictors of Outcome in Cognitive-Behavioral Treatment of Antisocial Children. *Behavior Research and Therapy* 33:271–81.
- Koegl, Christopher, J., David P. Farrington, Leena K. Augimeri, and David M. Day. 2008. Evaluation of a Targeted Cognitive-Behavioral Program for Children with Conduct Problems – The SNAP Under 12 Outreach Project: Service Intensity, Age and Gender Effects on Short- and Long-Term Outcomes. *Clinical Child Psychology and Psychiatry* 13:419–34.
- McCormick, Laura K., Allan B. Steckler, and Kenneth R. McLeroy. 1995. Diffusion of Innovations in Schools: A Study of Adoption and Implementation of School-Based Tobacco Prevention Curricula. *American Journal of Health Promotion* 9:210–19.

Using this method, it would be possible to obtain more accurate determinations of the most successful components and programs.

Through this study, and the future research recommended throughout this paper, we may move one step closer to developing a system of accreditation of anti-bullying programs, where only the most effective evidence-based programs are funded and utilized (McGuire 2001). This would ensure that programs or components that have undesirable effects or no effects on bullying and victimization would not be utilized, as it would encourage program funders and potential participants to avoid them. An accreditation system would not only increase our knowledge base concerning the success of bullying and victimization prevention programs, but would also enhance the ultimate goal of our efforts by reducing victimization and bullying in schools.

- McGuire, James. 2001. What Works in Correctional Intervention? Evidence and Practical Implications. In *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs*, ed. Gary A. Bernfeld, David P. Farrington, and Alan W. Leschied, 25–43. Chichester, UK: Wiley.
- Mellor, Andrew. 1991. Helping Victims. In *Bullying: A Practical Guide to Coping for Schools*, ed. Michele Elliott. Harlow: Longman.
- Mihalic, Sharon, Katherine Irwin, Abigail Fagan, Diane Ballard, and Delbert Elliott. 2004. *Successful Program Implementation: Lessons from Blueprints*. Washington, DC: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
- Olweus, Dan. 2005. A Useful Evaluation Design, and Effects of the Olweus Bullying Prevention Program. *Psychology, Crime and Law* 11:389–402.
- Piquero, Alex R., David P. Farrington, Brandon C. Welsh, Richard Tremblay, and Wesley G. Jennings. 2009. Effects of Early Family/Parent Training Programs on Antisocial Behavior and Delinquency. *Journal of Experimental Criminology* 5:83–120.
- Reid, Ken. 1989. Bullying and Persistent School Absenteeism. In *Bullying in Schools*, ed. Delwyn Tattum and David Lane. Stoke-on-Trent: Trentham.
- Smith, Peter K. 1997. Bullying in Schools: The UK Experience and the Sheffield Anti-Bullying Project. *Irish Journal of School Psychology* 18:191–201.
- Smith, Peter K., Debra Pepler, and Ken Rigby, eds. 2004. *Bullying in Schools: How Successful Can Interventions Be?* Cambridge: Cambridge University Press.
- Spoth, Richard, Cleve Redmond, Cathy Hockaday, and Chung Yeol Shin. 1996. Barriers to Participation in Family Skills Preventative Interventions and Their Evaluations: A Replication and Extension. *Family Relations* 45:247–54.
- Strain, Phillip S., Clifford C. Young, and Judy Horowitz. 1981. Generalized Behavior Change During Oppositional Child Training: An Examination of Child and Family Demographic Variables. *Behavior Modification* 5:15–26.
- Taggart, Virginia S., Patricia J. Bush, Alan E. Zuckerman, and Patricia K. Theiss. 1990. A Process Evaluation of the District of Columbia “Know Your Body” Project. *Journal of School Health* 60:60–66.
- Taylor, Ted K., and Anthony Biglan. 1998. Behavioral Family Interventions for Improving Child-Rearing: A Review for Clinicians and Policy Makers. *Clinical Child and Family Psychology Review* 1:41–60.
- Ttofi, Maria M., and David P. Farrington. 2008. Bullying: Short-Term and Long-Term Effects, the Importance of Defiance Theory in Explanation and Prevention. *Victims and Offenders* 3:289–312.
- Ttofi, Maria M., and David P. Farrington. 2011. Effectiveness of School-Based Programs to Reduce Bullying: A Systematic and Meta-Analytic Review. *Journal of Experimental Criminology* 7:27–56.
- Ttofi, Maria M., and David P. Farrington. 2012. Bullying Prevention Programs: The Importance of Peer Intervention, Disciplinary Methods and Age Variations. *Journal of Experimental Criminology*, online first August. doi:10.1007/s11292-01209161-0
- Ttofi, Maria M., David P. Farrington, Friedrich Lösel, and Rolf Loeber. 2011. The Predictive Efficiency of School Bullying versus Later Offending: A Systematic/Meta-Analytic Review of Longitudinal Studies. *Criminal Behavior and Mental Health* 21:80–89.
- Vreeman, Rachel C., and Aaron E. Carroll. 2007. A Systematic Review of School-Based Interventions to Prevent Bullying. *Archives of Pediatrics and Adolescent Medicine* 161:78–88.
- Wahler, Robert G. 1980. The Insular Mother: Her Problems in Parent-Child Treatment. *Journal of Applied Behavior Analysis* 13:207–219.
- Weisburd, David, Cynthia M. Lum, and Anthony Petrosino. 2001. Does Research Design Affect Study Outcomes in Criminal Justice? *Annals of the American Academy of Political and Social Science* 578:50–70.
- Webster-Stratton, Carolyn. 1985. Comparison of Abusive and Nonabusive Families with Conduct-Disordered Children. *American Journal of Orthopsychiatry* 55:59–69.
- Webster-Stratton, Carolyn. 1990. Long-Term Follow-Up of Families with Young Conduct Problem Children: From Preschool to Grade School. *Journal of Clinical Child Psychology* 19:144–49.
- Webster-Stratton, Carolyn, and Mary Hammond. 1990. Predictors of Treatment Outcome in Parent Training for Families with Conduct Problem Children. *Behavior Therapy* 21:319–37.
- Webster-Stratton, Carolyn, and Mary Hammond. 1997. Treating Children with Early-Onset Conduct Problems: A Comparison of Child and Parent Training Interventions. *Journal of Consulting and Clinical Psychology* 65:93–109.
- Webster-Stratton, Carolyn, M. Jamila Reid, and Mary Hammond. 2004. Treating Children with Early-Onset Conduct Problems: Intervention Outcomes for Parent, Child, and Teacher Training. *Journal of Clinical Child and Adolescent Psychology* 33:105–24.

Bryanna Hahn
bh349@cam.ac.uk

David P. Farrington
dpf1@cam.ac.uk

Maria M. Ttofi
mt394@cam.ac.uk

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools

Sonja Perren, Jacobs Center for Productive Youth Development, University of Zürich, Switzerland

Lucie Corcoran, School of Education, Trinity College Dublin, Ireland

Helen Cowie, Faculty of Health and Medical Sciences, University of Surrey, United Kingdom

Francine Dehue, Department of Psychology, Open University of the Netherlands, Heerlen, the Netherlands

D'Jamila Garcia, Instituto Universitário de Lisboa (ISCTE-IUL), Centro de Investigação e Intervenção Social – IUL, Portugal

Conor Mc Guckin, School of Education, Trinity College Dublin, Ireland

Anna Sevcikova, Faculty of Social Studies, Masaryk University, Brno, Czech Republic

Panayiota Tsatsou, Department of Political and Cultural Studies, Swansea University, United Kingdom

Trijntje Völlink, Department of Psychology, Open University of the Netherlands, Heerlen, the Netherlands

Vol. 6 (2) 2012

Focus:

Evidence-based Developmental Prevention of Bullying and Violence in Europe

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

► **Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools** Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



This work is licensed under the Creative Commons Attribution-NoDerivatives License.

ISSN: 1864-1385

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools

Sonja Perren, Jacobs Center for Productive Youth Development, University of Zürich, Switzerland

Lucie Corcoran, School of Education, Trinity College Dublin, Ireland

Helen Cowie, Faculty of Health and Medical Sciences, University of Surrey, United Kingdom

Francine Dehue, Department of Psychology, Open University of the Netherlands, Heerlen, the Netherlands

D'Jamila Garcia, Instituto Universitário de Lisboa (ISCTE-IUL), Centro de Investigação e Intervenção Social – IUL, Portugal

Conor Mc Guckin, School of Education, Trinity College Dublin, Ireland

Anna Sevcikova, Faculty of Social Studies, Masaryk University, Brno, Czech Republic

Panayiota Tsatsou, Department of Political and Cultural Studies, Swansea University, United Kingdom

Trijtje Völlink, Department of Psychology, Open University of the Netherlands, Heerlen, the Netherlands

A summary of current knowledge on successful responses to cyberbullying differentiating between three different response domains: reducing risks, combatting the problem, and buffering negative impact. A systematic literature search yielded thirty-six relevant studies, most of which report findings regarding general prevention strategies (e.g., anti-bullying policies or cybersafety strategies) and the use of coping strategies such as seeking support, responding (retaliation or confronting), technical solutions, and avoidant and emotion-focussed strategies. Whilst a few studies report perceived success, very few measure the success of the strategies in relation to risks and outcomes. There is a clear lack of evidence concerning successful responses.

Cyberbullying is generally considered to be bullying using technology such as the Internet and mobile phones (Menesini et al. 2012; Smith, Mahdavi, Carvalho, Fisher, Russel, and Tippett 2008). Cyberbullying takes a number of forms, such as sending insulting, rude or threatening messages, spreading rumours, revealing personal information, publishing embarrassing pictures, or exclusion from online communication. Recent studies have demonstrated that there is a significant conceptual and practical overlap between traditional bullying and cyberbullying, such that most young people who are cyberbullied also tend to be bullied by traditional face-to-face methods (Cross et al. 2009; Dooley, Pyzalski, and Cross 2009; Gradinger, Strohmeier, and Spiel 2009; Riebel, Jaeger, and Fischer 2009; Sourander et al. 2010).

Despite this overlap, cyberbullying differs from traditional bullying in several ways. First, a single upload of humiliating visual material to the internet is tantamount to repetition as the content can be permanent and available to a wide audience (Heirman and Walrave 2008). Second, power imbalance in cyberbullying can be expressed through (a) technological knowledge, (b) anonymity, (c) limited option of escape. Specifically, a perpetrator dominates a victim through greater knowledge of use of the internet and mobile phones and through the victim's limited possibilities of defence (not necessarily knowing the bully). Moreover, unlike traditional bullying, cyberbullying is not limited in time and space (Heirman and Walrave 2008; Smith et al. 2008; Vandebosch and Van Cleemput 2008).

Acknowledgement: The systematic literature review was conducted as part of COST ACTION IS0801 (Cyberbullying: Coping with Negative and Enhancing Positive Uses of New Technologies, in Relationships in Educational Settings). This study was also

supported by the Czech Ministry of Education, Youth, and Sports (LD11008). We greatly appreciate the support provided by the following graduate and doctoral students and COST-members in relation to the literature search,

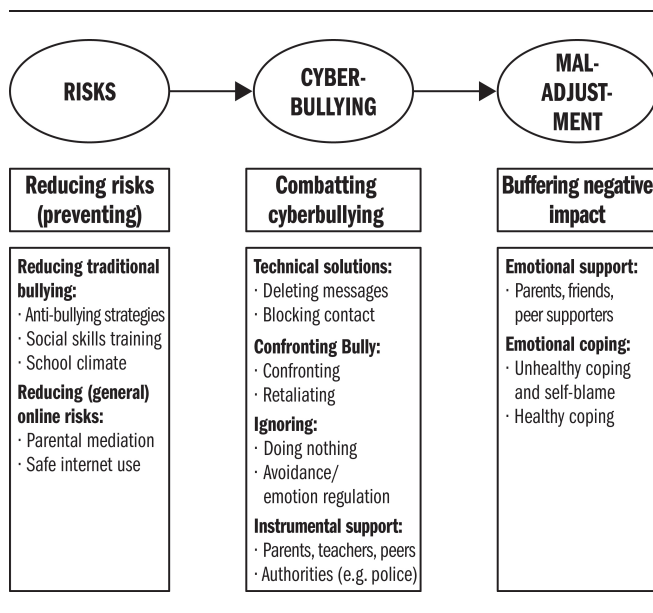
ratings, and/or preparation of the report: Monika Finsterwald, Alena Cerna, Hana Machackova, Katja Machmutow, Ursina Rumetsch.

Despite its overlap with traditional bullying, being a victim of cyberbullying has been identified as an additional risk factor for the development of depressive symptoms (Perren, Dooley, Shaw, and Cross 2010; Gradinger, Strohmeier, and Spiel 2009; Juvonen and Gross 2008) and of psychosomatic symptoms like headaches, abdominal pain and sleeplessness (Sourander et al. 2010). Moreover, adolescent victims of cyberbullying also engage in other types of problematic behaviour, such as increased alcohol consumption, a tendency to smoke and poor school grades (Mitchell, Ybarra, and Finkelhor 2007). Aggressors are at increased risk for school problems, conduct disorders, and substance use (Hinduja and Patchin 2008; Sourander et al. 2010). In sum, cyberbullying emerges as a significant concern for families, schools, and social and healthcare professionals. The present literature review summarizes current knowledge on responses to cyberbullying.

1. Responses to Cyberbullying

In the present review, responses to cyberbullying are conceptualized as reactions to this problem on the part of students, parents and schools. We differentiate between the following domains: reducing risks, combatting the problem, and buffering the negative impact (see Figure 1).

Figure 1: Conceptualization of responses to cyberbullying



First, from a preventive perspective, students, parents and schools may try to handle the emerging problem of cyberbullying by reducing known risks. As cyberbullying is strongly associated with traditional bullying (Cross et al. 2009; Dooley, Pyzalski, and Cross 2009; Gradinger, Strohmeier, and Spiel 2009; Riebel, Jaeger, and Fischer 2009), we may assume that taking action against traditional bullying and associated risk factors through such interventions as whole-school approaches and policies, social skills training, or improvement of the school climate could also reduce the risk of cyberbullying. As cyberbullying occurs via internet or mobile phone, it is also associated with general online risks such as risky online contacts or viewing inappropriate content (Livingstone, Haddon, Görzig, and Olafsson 2011). Therefore, parental mediation or internet safety measures might also be effective in reducing cyberbullying.

Secondly, when cyberbullying occurs, a different set of actions to combat these negative behaviours may be used by students, parents or schools. These responses include technical solutions (e.g., blocking contact), confronting the bully (e.g., constructive contacting or retaliation), ignoring (e.g., doing nothing, avoidant behaviour or emotion regulation) and instrumental support (e.g., asking someone else for help). As cyberbullying has negative consequences for victims such as depression or suicidal ideation (Gradinger, Strohmeier, and Spiel 2009; Juvonen and Gross 2008; Perren et al. 2010; Sourander et al. 2010), specific coping strategies might also be applied to enhance victims' well-being and buffer the negative impact: Victims themselves may try to cope emotionally with the problem; and parents, friends or peers may offer emotional and instrumental support.

The goal of the current review was to summarize the empirical database on successful responses to cyberbullying and identify what responses are successful. We conceptualized success in terms of (a) reducing cyberbullying risks (the *prevention* of cyberbullying), (b) *combatting* cyberbullying leading to stopping this problem, and (c) *buffering* its negative impact on victims.

2. Systematic Literature Search

A systematic literature search was conducted to identify relevant empirical studies. Relevant databases (PsychInfo, Pubmed, ERIC, SOCindex, Web of Science, etc) were systematically searched. Selected studies had to contain the keywords cyberbullying (or related terms), coping/re-sponses (or related terms), and youth/educational settings (or related terms).¹ Articles were rated for relevance in several steps and double-checked for inter-rater agreement. Publications up to September 2010 were included. Also included were findings from the EU Kids Online II study (initial findings published in October 2010, final publication in 2011).² The database search yielded 225 publications, which were rated regarding relevance and correspondence to inclusion criteria.

The following inclusion criteria were used (a) empirical studies on cyberbullying (new data and knowledge); (b) published papers only (scientific journals, book chapters, EU Kids Online report, dissertations, but excluding conference papers and posters); (c) parents, teachers (schools) or students/pupils responding to cyberbullying; (d) papers should include some measures of responses (listed in Figure 1); and (e) papers should address at least one of our predefined research questions (prevent, combat, buffer). Thirty-six articles were rated as being partly or highly relevant to our research question. All relevant papers were systematically analysed by seven different raters (mostly members of the current author team). The raters had to review methods (i.e. type of study, focus, sample, types of measures and their quality) and look for research evidence on the success of responses related to the domains, such as reducing cyberbullying risks, combatting cyberbullying, and buffering its negative impact. The raters were given a form with predefined responses to evaluate. Further, taking into consideration that the list of responses could not be exhaustive, the raters were asked to fill in the open-ended domain-related boxes with examined responses, including findings on the responses' success or otherwise (e.g.,

“Please provide the article’s results/conclusions/implications with regard to...”).

The current paper presents a selective narrative overview of the results of the systematic literature review, focussing on the question of measured *success* of responses.

3. Preventing Cyberbullying

This section first presents findings and suggestions for concrete measures to prevent cyberbullying.

3.1. Suggested Prevention Approaches

Based on general research findings on cyberbullying and the associated risks, several authors argue that we should draw upon experience from “face-to-face” bullying so as to prevent cyberbullying (Campbell 2005). In addition, the following preventative actions were suggested with emphasis on the whole school approach:

- Awareness-raising initiatives targeting teachers, parents and students in order to heighten awareness of cyberbullying and its risks and create a context for facilitating trust on the part of victims with regard to adult authorities (Campbell 2005; Juvonen and Gross 2008; Li 2007; Wright, Burnham, Inman, and Ogorchock 2009; Young, Young, and Fullwood 2007);
- School policies to respond to the challenge of cyberbullying and implement a range of preventive policies such as
 - the direct teaching of values education, empathy training and the use of stories and drama in the curriculum, along with direct teaching of “netiquette” (Campbell 2005; Dranoff 2008; Mason 2008; Stacey 2009), and last but not least to create an open line of communication between students and adults in school (Genz 2009);
 - the inclusion of social and curriculum programmes to motivate students towards taking action against cyberbullying (e.g., peer help programmes, buddy programmes, etc.) (Campbell 2005; Stacey 2009);

¹ A report describing details of the methodological approach (search terms, inclusion and exclusion criteria, etc.), the category system and complete tables with descriptive results can be obtained from the first author on request.

² EU Kids Online II is the largest (representative) study in Europe concerning this topic. It surveyed a sample of 25,142 children aged 9–16 years across twenty-five European countries (Livingstone et al. 2011).

- adult supervision, especially with regard to children's computer education and usage of technology (Campbell 2005; Rosen, Cheever and Carrier 2008) as well as education of parents concerning these matters (Stacey 2009).

The suggested prevention strategies emphasize the importance of both family and education/school (Smith et al. 2008) for preventing cyberbullying, while stressing the need to empower children and make them the key actors in deciding about, and implementing prevention strategies (Stacey 2009; Ybarra and Mitchell 2004; Young, Young, and Fullwood 2007). However, most of the studies described above drew their conclusions and suggested implications for prevention strategies from general empirical findings (at best).

3.2. Evidence Regarding Successful Coping Strategies to Prevent Cyberbullying

Although different strategies are recommended based on general research findings, only a few studies investigated the success of particular strategies in actually preventing cyberbullying. It has been suggested that peer support in the form of peer-intervention by student leaders in school may play a role in prevention of cyberbullying through creating bullying awareness in the school, developing leadership skills among students, developing bullying intervention practices and team-building initiatives in the student community, and encouraging students to intervene. DiBasilio (2008) showed that such peer intervention successfully led to a decline in cyberbullying, while students' understanding of bullying widened.

A second key category of prevention strategies reported in the literature focuses on parental supervision and parenting behaviour. As time spent online is considered as a risk factor for cyberbullying, parental restrictive mediation (which decreases the amount of time children spend online) was found to reduce cyberbullying risks (Livingstone et al. 2011; Rosen, Cheever, and Carrier 2008). Research has found that higher levels of parental warmth are negatively correlated with involvement in both traditional bullying and cyberbullying (Wang, Iannotti, and Nansel 2009). Conversely, a poor parent-child relationship, which may indicate insufficient parental supervision, has been found to be associated

with a higher risk of involvement in cyberbullying both as a perpetrator and as a victim (Ybarra and Mitchell 2004).

4. Combatting Cyberbullying and Buffering its Negative Impact

Besides the question concerning which strategies parents, schools and students can apply to prevent cyberbullying, research has also addressed the question about what victims of cyberbullying (or persons close to them) should do to cope with the problem. We will first outline what responses have been investigated and then present empirical evidence regarding their successfulness.

4.1. Responses to Ongoing Cyberbullying

In the reviewed studies research attention focussed predominantly on victims of cyberbullying and their responses to the problem. Victims report a range of coping strategies which can be classified as being problem-focussed or emotion-focussed (or mixed). According to coping theory (Lazarus and Folkman 1984), people tend to use problem-focussed coping when they believe that their own resources or critical aspects of the situation can potentially be changed, i.e. a person attempts to handle the stressful situation by tackling the problem that causes it. On the other hand, people use emotion-focussed coping when they believe that they can do little to change the stressful situation; here a person attempts to control their emotional response to the stressful situation by redefining or ignoring it or by focussing on the positive aspects of the situation.

Several types of coping strategies have been identified in relation to cyberbullying: reactions towards cyberbullies (retaliation, confronting), technical solutions (e.g., report abuse buttons, blocking the sender), supportive strategies (seeking support by adults, teachers, friends or external institutions), and avoidant and emotion-focussed strategies (e.g., ignoring). The next section gives an overview of the research on the use of coping strategies and their perceived successfulness in dealing with cyberbullying.

Reactions towards the bully. Confronting the bully is commonly reported by adolescents as an approach, where the victim knows the bully or is able to contact her or him (Arıcak et al. 2008; DiBasilio 2008; Stacey 2009). Students

consider retaliation a less constructive way of contacting the perpetrator. Hoff and Mitchell (2009), who asked students what they had done to stop cyberbullying, report that the answers mentioned active and physically retaliatory behaviour, especially among boys. Although the strategy of confronting the bully is often mentioned by those affected, this strategy has proven to be less helpful in retrospect (Price and Dalgleish 2010).

The assumption in some studies that online retaliation is more easily done, due to greater anonymity, and therefore occurs more often than “face-to-face contact”, was not confirmed. Juvonen and Gross (2008) found that, whereas 60 percent of the cybervictims defended themselves against the bully with traditional face-to-face methods, only 12 percent retaliated solely in cyberspace, and 28 percent used both traditional and online forms of retaliation.

Technical solutions. Specific cyberspace coping strategies, such as deleting or blocking threatening messages, are generally used and considered as being helpful (Aricak et al. 2008; Juvonen and Gross 2008; Kowalski, Limber, and Agatston 2008; Smith et al. 2008; Stacey 2009). Livingstone et al. (2011) report that the most popular technical coping strategies are blocking the person (46 percent), deleting nasty messages (41 percent), and stopping use of the internet (20 percent). Blocking was considered an effective strategy by study participants.

Using a mixed methodological approach, Price and Dalgleish (2010) found that blocking was the most widely used technical strategy; self-identified cybervictims considered this to be the most helpful online action taken. Technical solutions are often reported along with preventive strategies like banning websites and setting age-appropriate limits for using the computer and internet by parents (see also above; Kowalski, Limber, and Agatston 2008).

Support-seeking. Many students recommend asking parents for help in relation to a cyberbullying incident (Aricak et al. 2008; Smith et al. 2008; Stacey 2009; Topcu, Erdur-Baker, and Capa-Aydin 2008). However, some students recommend not consulting adults because they fear that they may lose privileges (e.g., having and using mobile phones

and internet access), and because they fear parents would simply advise them to ignore the situation or that they would not be able to help them as they are not accustomed to cyberspace (Hoff and Mitchell 2009; Kowalski, Limber, and Agatston 2008; Mishna, Saini, and Solomon 2009; Smith et al. 2008; Stacey 2009). In a web-based survey of 12–17-year olds, of whom most had experienced at least one cyberbullying incident in the last year, Juvonen and Gross (2008) found that 90 percent of the victims did not tell their parents about their experiences and 50 percent justified it with “I need to learn to deal with it myself”.

Students also have a rather negative and critical attitude to teachers’ support: many students consider telling a teacher or the school principal as rather ineffective (Aricak et al. 2008; DiBasilio 2008; Mishna, Saini, and Solomon 2009). Although 17 percent of students did report a cyberbullying incident to a teacher, in 70 percent of the cases the school did not react to it (Hoff and Mitchell 2009).

Asking for help from peers is a commonly used approach and is recommended (Aricak et al. 2008; DiBasilio 2008; Stacey 2009; Topcu, Erdur-Baker, and Capa-Aydin 2008), although prevalence rates vary widely. Price and Dalgleish (2010) report that Australian cybervictims consider “telling a friend” as the most helpful strategy. Livingstone et al. (2011) report that in terms of confiding in others, respondents were most likely to tell a friend (52 percent) or a parent (42 percent).

Avoidant and emotion-focussed strategies. In a study by Dehue, Bolman, and Vollink (2008) students reported that when they were victimised online they would usually “pretend to ignore it” (31 percent of victims) and/or “would ignore it” (30 percent). When asked how they coped with the problem, 36 percent of the respondents in the EU Kids Online II study reported that they tried to “fix the problem”, whereas 24 percent “hoped the problem would go away”, and 12 percent said that they “felt guilty” (Livingstone et al. 2011).

In sum, a range of coping strategies used by victims in relation to cyberbullying have been investigated. However, most of the studies investigated the use (and not the suc-

cess) of coping strategies among cybervictims, or in relation to hypothetical cyberbullying scenarios. For example the EU Kids Online II study showed that of those bullied online in the last 12 months (6 percent of participants), 85 percent reported being upset (Livingstone et al. 2011). However, the majority of victims (62 percent) “got over it straight away”. This finding led Livingstone et al. (2011) to conclude that children’s coping strategies were most likely effective, at least for those who do not continue to be upset. As this is a very general conclusion, we do not know what kind of coping strategies are “likely to be effective”.

4.2. Evidence Regarding Successful Responses

To investigate the success of responses, various methodological approaches have been applied, targeting different populations and using different study designs and assessment methods. From a purely methodological point of view, these approaches range from yielding no evidence (subjective evaluations) to a strong evidence base (experimental designs). In reviewing the selected studies, we have identified the following taxonomy of studies.

1. What do people, in general, think is effective?
2. Retrospective accounts of cybervictims regarding the success of chosen coping strategies.
3. Cross-sectional studies investigating associations between certain responses, cybervictimisation and victim’s well-being.
4. Longitudinal studies investigating whether certain responses or coping strategies are related to decreasing levels of cybervictimisation (combatting) or victim’s well-being (buffering).
5. Experimental studies investigating the impact of selected responses on changes in cybervictimisation and victim’s well-being.

Studies reporting on perceived success from a general perspective have already been described and are considered as yielding no real evidence.

Retrospective accounts of victims. Hensler-McGinnis (2008) examined the effect of coping on psychological trauma and impaired academic/career functioning following victimisation through cyberstalking. A sample of 452 college/university students aged between 18 and 43 years (female:

81.2 percent) participated in the research. Victimisation was found to be predictive of psychological trauma and impaired academic functioning. The following responses were rated by victims as being the most effective strategies decreasing the cyberstalking: “retaliating using electronic methods” (65.5 percent), “blocking my electronic accessibility” (63 percent), “limited disclosure of my personal information on the internet” (56.9 percent), and “decreased use of internet, cell phone etc.” (54 percent). Effective coping was characterised by limiting exposure and accessibility. Psychological trauma and academic/career functioning impairment were both found to be positively correlated to the number of coping strategies used by the victim, suggesting that these were victims who had tried many strategies but without success. Additionally, there was no evidence that resilient coping moderated the relationship between victimisation and trauma, or the relationship between victimisation and academic/career functioning.

Cross-sectional associations between coping strategies and cybervictimisation. Only one study reported on the relationship between different coping strategies and cyberbullying (Lodge and Frydenberg 2007). The results revealed that children with increased experience of cybervictimisation used more passive coping strategies, such as wishful thinking and mental distraction, compared to children with low levels of cybervictimisation. In general, children with an optimistic, relaxed and active mode of coping reported fewer cyberbullying experiences (Lodge and Frydenberg 2007). Results of this study yield first indications about what kind of coping strategies might be effective. However, as this was a cross-sectional study, we do not know whether any of the reported associations are causal.

Longitudinal associations between coping strategies and well-being. The study by Hay and Meldrum (2010) is one of the rare longitudinal studies on the topic; they measured the role of authoritative parenting and high self-control in buffering the negative impact of traditional bullying and cyberbullying. In a sample of 426 students aged between 10 and 21 years (female: 50 percent) they found that victimisation was associated with increased reporting of self-harm and suicidal ideation. Authoritative parenting and high levels of self-control moderated the negative impact of victi-

misation. The authors concluded that cognitive behavioural therapy could benefit vulnerable adolescents, by helping them to acknowledge their maladaptive coping and to change their behaviour. The longitudinal design advances our understanding of potential buffering effects. However, this study did not investigate specific coping strategies, but more general personal and parental characteristics that aimed to buffer the negative impact.

Experimental studies. The literature review yielded one intervention study investigating the impact of actions taken against cyberbullying. Chi and Frydenberg (2009) investigated the impact of two programmes (Best of Coping, BOC, and Cyber Savvy Teens, CST) on adolescents' psychological distress and ability to cope online. The BOC programme educates participants on general coping techniques such as decision-making whereas the CST programme was designed to increase adolescents' safety skills online, including coping strategies for cyberbullying. A sample of 50 adolescents (female: $n = 28$) aged 13 to 14 years was divided into three categories: a control group; a group with the CST programme; and a group with the BOC programme. Three coping styles (Productive Coping [P]; Non-productive Coping [N]; and Reference to Others [R]) and psychological distress were examined. Following the intervention, the CST group showed an increase in using the strategies "keep to self" (N) and "ignore the problem" (N), and a decrease in "focus on positive" (P) and "seek to belong" (P). However, a small increase in Productive Coping was identified. CST participants displayed increased willingness to report cyberharassment to teachers and parents post-intervention. The BOC group showed an increase in using "social action" (R), "physical recreation" (P), and "focus on solving the problem" (P) post-intervention, and a decrease in using "ignore the problem" (N), "wishful thinking" (N), and "worry" (N). The BOC group was also more likely to report cyberharassment to a trusted adult. An increase in Reference to Others (R) was identified after the intervention. Both groups showed a decrease in the use of Non-productive Coping (N). In terms of buffering negative effect, the authors concluded that both programmes reduced participants' anxiety and symptoms of depression.

5. Discussion

This systematic literature review identifies a number of studies which reported some empirical data on responses to cyberbullying. However, the conclusions which can be drawn are limited. We found very little empirical evidence concerning the success of responses. Despite this, the studies provided some insight into what students do to cope with cyberbullying. Most of the reported coping strategies are general problem-solving strategies such as looking for social support, trying to ignore or avoid the problem. Some are related to bullying (e.g., confronting the bully); others are specifically related to cyberbullying, such as the use of technological strategies. To reduce possible risk factors and to prevent cyberbullying, parental supervision, general anti-bullying or social skills development strategies, and education in cybersafety have been suggested.

In addition, many of the identified studies suffer from similar methodological limitations. Most of the studies used cross-sectional self-reports among adolescent samples. Responses including coping strategies were frequently only assessed with single items. Because of methodological shortcomings, the reactions of victims to cyberbullying did not reflect the context and ways they were victimised; more precisely there was a lack of studies on how victims responded to different forms of cyberbullying and to what extent the form of cyberbullying may relate to successful solutions. Moreover, there was a lack of theoretical background regarding selected coping strategies, their potential effects and underlying mechanisms. These issues are not new and are not specific to cyberbullying. For example Mc Guckin, Cummins, and Lewis (2010) emphasize similar issues as being of critical importance regarding research studies exploring traditional bully/victim problems.

Future studies utilising longitudinal approaches and methodologically sound intervention designs are required. Longitudinal studies should address the question of whether the use of specific coping strategies is more effective in combatting cyberbullying occurrences or in buffering the negative effects. In these studies, coping strategies should be investigated as potential mediators or moderators. For example, a recent study by Machmutow, Perren,

Sticca, and Alsaker (2012), which was published after the present literature review was conducted, shows that social support can buffer the negative impact of cybervictimization.

Intervention studies (preferably randomised controlled trials) should investigate the effectiveness of prevention strategies, either in relation to reducing risks or in relation to teaching specific response strategies for victims, bystanders, parents and teachers. Preliminary results of the Finish anti-bullying programme KiVa (Kiusaamista Vastaan, “against bullying”), whose findings, again, were published after compilation of studies for the literature review, suggest that prevention strategies targeting traditional bullying are able to reduce cyberbullying (Salmivalli, Kärnä, and Poski-

parta 2011). Conversely, another recently published longitudinal study about the effect of “netiquette” on the reduction of cyberbullying found no significant relationship between these two variables (Kumazaki, Suzuki, Katsura, Sakamoto, and Kashibuchi 2011). This indicates that more intervention studies are needed to understand which measures are successful in reducing cyberbullying risks.

Our review only included studies published up to September 2010. As there are a number of studies currently being carried out, and there may be relevant papers under review, we might soon expect more empirical evidence regarding the success of coping strategies. Only then can we seriously recommend guidelines and coping strategies to students, parents, and schools.

References

- Aricak, Tolga, Sinem Siahhan, Aysegul Uzunhasanoglu, Sevda Saribeyoglu, Songul Ciplak, Nesrin Yilmaz, and Cemil Memmedova. 2008. Cyberbullying among Turkish Adolescents. *CyberPsychology and Behavior* 11 (3): 253–61. doi:10.1089/cpb.2007.0016
- Campbell, Marilyn A. 2005. Cyber Bullying: An Old Problem in a New Guise? *Australian Journal of Guidance and Counselling* 15:68–76. doi:10.1375/ajgc.15.1.68
- Chi, Cecilia Wing, and Erica Frydenberg. 2009. Coping in the Cyberworld: Program Implementation and Evaluation – A Pilot Project. *Australian Journal of Guidance and Counselling* 19: 196–215. doi:dx.doi.org/10.1375/ajgc.19.2.196
- Cross, Donna, Therese Shaw, Lydia Hearn, Melanie Epstein, Helen Monks, Leanne Lester, Laura Thomas, and Data Analysis Australia 2009. *Australian Covert Bullying Prevalence Study (ACBPS)*. Perth, Australia: Child Health Promotion Research Centre, Edith Cowan University.
- Dehue, Francine, Catherine Bolman, and Trijntje Vollink. 2008. Cyberbullying: Youngsters’ experiences and parental perception. *CyberPsychology & Behavior* 11: 217–223. doi: 10.1089/cpb.2007.0008
- DiBasilio, Amy. 2008. *Reducing Bullying in Middle School Students through the Use of Student-Leaders*. M.A dissertation, Saint Xavier University.
- Dooley, Julian J., Jacek Pyzalski, and Donna Cross. 2009. Cyberbullying Versus Face-to-Face Bullying: A Theoretical and Conceptual Review. *Zeitschrift für Psychologie/Journal of Psychology* 217:182–88. doi:10.1027/0044-3409.217.4.182
- Dranoff, Robert. 2008. *Teacher and Student Perceptions of Bullying and Victimization in a Middle School*. Ph.D. Dissertation, Dowling College, New York.
- Genz, Heather. 2009. *Education, Prevention, and Intervention of Traditional and Electronic Bullying in Middle School*. Ph.D. dissertation, College of Education, Walden University, Minnesota.
- Gradinger, Petra, Dagmar Strohmaier, and Christiane Spiel. 2009. Traditional Bullying and Cyberbullying: Identification of Risk Groups for Adjustment Problems. *Zeitschrift für Psychologie/Journal of Psychology* 217:205–13. doi:10.1027/0044-3409.217.4.205
- Hay, Carter, and Ryan Meldrum. 2010. Bullying Victimization and Adolescent Self-Harm: Testing Hypotheses from General Strain Theory. *Journal of Youth and Adolescence* 39:446–59. doi:10.1007/s10964-009-9502-0
- Heirman, Wannes, and Michel Walrave. 2008. Assessing Concerns and Issues about the Mediation of Technology in Cyberbullying. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 2 (2). <http://cyberpsychology.eu/view.php?cisloclanku=2008111401&article=1>
- Hensler-McGinnis, Nancy Felicity. 2008. *Cyberstalking Victimization: Impact and Coping Responses in a National University Sample*. Ph.D. dissertation, University of Maryland.
- Hinduja, Sameer, and Justin W. Patchin. 2008. Cyberbullying: An Exploratory Analysis of Factors Related to Offending and Victimization. *Deviant Behavior* 29:129–56. doi:10.1080/01639620701457816
- Hoff, Dianne L., and Sidney N. Mitchell. 2009. Cyberbullying: Causes, Effects, and Remedies. *SO – Journal of Educational Administration* 47:652–65.
- Juvonen, Jaana, and Elisheva F. Gross. 2008. Extending the School Grounds? Bullying Experiences in Cyberspace. *Journal of School Health* 78:496–505. doi:10.1111/j.1746-1561.2008.00335.x
- Kowalski, Robin M, Susan P. Limber, and Patricia W. Agatston. 2008. Cyber Bullying: Bullying in the Digital Age. *American Journal of Psychiatry* 165:780–81. doi:10.1176/appi.ajp.2008.08020226.
- Kumazaki, Ajuchi, Kanae Suzuki, Rui Katsura, Akira Sakamoto, and Megumi Kashibuchi 2011. The Effects of Netiquette and ICT Skills on School-bullying and Cyber-bullying: The Two-wave Panel Study of Japanese Elementary, Secondary, and High School Students. *Procedia – Social and Behavioral Sciences* 29:735–41. doi:10.1016/j.sbspro.2011.11.299
- Lazarus, Richard. S., and Susan Folkman. 1984. *Stress, Appraisal, and Coping*. New York: Springer.
- Li, Qing. 2007. New Bottle but Old Wine: A Research of Cyberbullying in Schools. *Computers in Human Behavior* 23:1777–91.
- Livingstone, Sonia, Leslie Haddon, Anke Görzig, and Kjartan Olafsson 2011. *Risks and Safety on the Internet: The Perspective of European Children: Full Findings and Policy Implications from the EU Kids Online Survey of 9–16 Year Olds and Their Parents in 25 Countries*. London: LSE.
- Lodge, Jodie, and Erica Frydenberg. 2007. Cyber-bullying in Australian Schools: Profiles of Adolescent Coping and Insights for School Practitioners. *Australian Educational and Developmental Psychologist* 24:45–48.
- Machmutow, Katja, Sonja Perren, Fabio Sticca, and Françoise D. Alsaker 2012. Peer Victimization and Depressive Symptoms: Can Specific Coping Strategies Buffer the Negative Impact of Cybervictimization? *Emotional and Behavioral Difficulties*, 17:3-4, 403-420.
- Mason, Kimberly L. 2008. Cyberbullying: A Preliminary Assessment for School Personnel. *Psychology in the Schools* 45:323–48. doi:10.1002/pits.20301

- Mc Guckin, Conor, Pauline K. Cummins, and Christopher Alan Lewis. 2010. f2f and Cyberbullying among Children in Northern Ireland: Data from the Kids Life and Times Surveys. *Psychology, Society, and Education* 2:83–96.
- Menesini, Ersilia, A Nocentini, Anja Schultze-Krumbholz, Herbert Scheitauer, Ann Frisen, Sofia Berne, Juan Calmaestra, Catherine Blaya, and Piret Luik 2012. Cyberbullying Definition among Adolescents: A Comparison across Six European Countries. *Cyberpsychology, Behavior, and Social Networking*, in press.
- Mishna, Faye, Michael Saini, and Steven Solomon. 2009. Ongoing and Online: Children and Youth's Perceptions of Cyber Bullying. *Children and Youth Services Review* 31:1222–28. doi:10.1016/j.chilyouth.2009.05.004
- Mitchell, Kimberly J., Michel Ybarra, and David Finkelhor. 2007. The Relative Importance of Online Victimization in Understanding Depression, Delinquency, and Substance Use. *Child Maltreatment* 12:314–24. doi:10.1177/1077559507305996
- Perren, Sonja, Julian Dooley, Therese Shaw, and Donna Cross 2010. Bully/Victim Problems in Schools and in Cyberspace: Associations with Depressive Symptoms in Swiss and Australian Adolescents. *Child and Adolescent Psychiatry and Mental Health* 4:28.
- Price, Megan, and John Dalgleish. 2010. Cyberbullying: Experiences, Impacts and Coping Strategies as Described by Australian Young People. *Youth Studies Australia* 29:51–59.
- Riebel, Julia, Reinhold S. Jaeger, and Uwe C. Fischer. 2009. Cyberbullying in Germany – An Exploration of Prevalence, Overlapping with Real Life Bullying and Coping Strategies. *Psychology Science Quarterly* 51 (3): 298–314.
- Rosen, Larry D., Nancy A. Cheever, and L. Mark Carrier. 2008. The Association of Parenting Style and Child Age with Parental Limit Setting and Adolescent MySpace Behavior. *Journal of Applied Developmental Psychology* 29:459–71. doi:10.1016/j.appdev.2008.07.005
- Salmivalli, Christina, Antti Kärnä, and Elisa Poskiparta. 2011. Counteracting Bullying in Finland: The KiVa Program and Its Effects on Different Forms of Being Bullied. *International Journal of Behavioral Development* 35:405–11.
- Smith, Peter K., Jess Mahdavi, Manuel Carvalho, Sonja Fisher, Shanette Russel, and Neil Tippett 2008. Cyberbullying: Its Nature and Impact in Secondary School Pupils. *Journal of Child Psychology and Psychiatry* 49:376–85. doi:10.1111/j.1469-7610.2007.01846.x
- Sourander, Andre, Anat Brunstein Klomek, Maria Ikonen, Jarna Lindroos, Terhi Luntamo, Merja Koskelainen, Terja Ristkari, and Hans Helenius 2010. Psychosocial Risk Factors Associated with Cyberbullying among Adolescents. *Archives of General Psychiatry* 67:720–28.
- Stacey, Elisabeth. 2009. Research into Cyberbullying: Student Perspectives on Cybersafe Learning Environments. *Informatics in Education* 8:115–30.
- Topcu, Cigdem, Ozgur Erdur-Baker, and Yesim Capa-Aydin. 2008. Examination of Cyberbullying Experiences among Turkish Students from Different School Types. *CyberPsychology & Behavior* 11 (6):643–48.
- Vandebosch, Heidi, and Katrien van Cleemput. 2008. Defining Cyberbullying: A Qualitative Research into the Perceptions of Youngsters. *Cyberpsychology and Behavior* 11 (2): 499–503. doi:10.1089/cpb.2007.0042
- Wang, Jing, Ronald J. Iannotti, and Tonja R. Nansel. 2009. School Bullying among Adolescents in the United States: Physical, Verbal, Relational, and Cyber. *Journal of Adolescent Health* 45:368–75. doi:10.1016/j.jadohealth.2009.03.021
- Wright, Vivian H., Joy J. Burnham, Christopher T. Inman, and Heather N. Ogorchock 2009. Cyberbullying: Using Virtual Scenarios to Educate and Raise Awareness. *SO – Journal of Computing in Teacher Education* 26:35–42.
- Ybarra, Michel L., and Kimberley J. Mitchell. 2004. Youth Engaging in Online Harassment: Associations with Caregiver-Child Relationships, Internet Use, and Personal Characteristics. *Journal of Adolescence* 27:319–36. doi:10.1016/j.adolescence.2004.03.007
- Young, Adena, Atrisha Young, and Harry Fullwood. 2007. Adolescent Online Victimization. *Prevention Researcher* 14:8–9.

Sonja Perren
perren@jacobscenter.uzh.ch
Lucie Corcoran
corcorl@tcd.ie
Helen Cowie
h.cowie@surrey.ac.uk

Francine Dehue
francien.dehue@ou.nl
D'Jamila Garcia
djamila.garcia@gmail.com
Conor Mc Guckin
conormcguckin@gmail.com

Anna Sevcikova
asevciko@fss.muni.cz
Panayiota Tsatsou
p.tsatsou@swansea.ac.uk
Trijntje Völlink
trijntje.vollink@ou.nl

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland

Christina Salmivalli, Department of Psychology, University of Turku, Finland

Elisa Poskiparta, Centre for Learning Research, University of Turku, Finland

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

► **KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland** Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland

Christina Salmivalli, Department of Psychology, University of Turku, Finland
Elisa Poskiparta, Centre for Learning Research, University of Turku, Finland

The effects of a Finnish national school-based antibullying program (KiVa) were evaluated in a randomized controlled trial (2007–2009) and during nationwide implementation (since 2009). The KiVa program is been found to reduce bullying and victimization and increase empathy towards victimized peers and self-efficacy to support and defend them. KiVa increases school liking and motivation and contributes to significant reductions in anxiety, depression, and negative peer perceptions. Somewhat larger reductions in bullying and victimization were found in the randomized controlled trial than in the broad rollout, and the largest effects were obtained in primary school (grades 1–6). The uptake of the KiVa program is remarkable, with 90 percent of Finnish comprehensive schools currently registered as program users.

Bullying, defined as repeated aggressive behavior against a victim who cannot readily defend him- or herself (Olweus 1991) is recognized as a recurrent and serious problem among school-age children across the world (Craig and Harel 2004). The targets of such treatment suffer psychosocial problems such as depression, anxiety, and peer rejection (Card and Hodges 2008; Hawker and Boulton 2003). For a number of victimized students, these experiences continue to affect their lives later on in adulthood (Isaacs, Hodges, and Salmivalli 2008). Besides the targets, bullying constitutes a risk for the healthy development of the perpetrators (Sourander et al. 2007) as well as for bystanders merely witnessing victimization (Rivers et al. 2009). The need for evidence-based interventions against bullying is therefore indisputable and urgent.

Numerous initiatives to prevent and tackle bullying have emerged during past decades, many of these taking the form of school-based programs. Although bullying may take place both in and out of school, school is a context where bullied students cannot escape their tormentors (school attendance being compulsory) while bullies are often socially rewarded by peers who join their behavior or just reinforce it by verbal or nonverbal signals (Salmivalli

2010). As bullying is typically related to peer group dynamics, such as norms emerging in classrooms, targeting any individual child's behavior might not be enough: group problems need group solutions. The fact that large numbers of children and adolescents (i.e., whole cohorts) can be easily reached in the school context makes school-based programs potentially very cost-effective as well. On the other hand, schools often struggle with multiple expectations combined with limited resources; their readiness to implement prevention programs cannot be taken for granted. Program developers should provide them programs that are not only shown to be effective but that are also system-ready, possible to implement with fidelity in a school system.

The evidence on the effectiveness of school-based antibullying programs is overall much thinner than we would like it to be. The effects show great variability across programs and across studies, typically being small to modest in size (Ferguson et al. 2007; Merrell et al. 2008; Smith et al. 2004; Ttofi, Farrington, and Baldry 2008; Vreeman and Carroll 2007). Some of the antibullying programs even seem to bring about undesirable outcomes, i.e. increases in bullying problems. In the most recent meta-analysis on the

topic, however, Ttofi and Farrington (2011) conclude that school-based anti-bullying programs are effective. The programs selected for their analysis showed average reductions of 17–23 percent for bullying others and 17–20 percent for being bullied. Ttofi and Farrington point out that even relatively small effect sizes correspond to a substantial amount of bullying and victimization prevented and thus to huge amounts of suffering avoided.

Despite the conclusion of Ttofi and Farrington that school-based interventions work, it should be kept in mind that there are numerous programs “in the market” that 1) have not been evaluated at all; 2) have not been evaluated with methodological rigor; 3) have been evaluated but shown only minimal positive effects, no effects at all, or negative effects; or 4) have been shown to work in a small-scale trial but not during broad rollout, i.e., perhaps are not system-ready. Hopefully, with growing awareness of the importance of evidence among stakeholders and practitioners, the demand for rigorous evaluations of antibullying programs will increase – to the benefit of both research and good practice.

In Finland, the Ministry of Education and Culture has funded the development, evaluation, and national implementation of the KiVa antibullying program. Our exceptionally large data sets will enable us to continue studies on the effects of the program and its mechanisms for years to come. In the present article we summarize the studies evaluating KiVa, based on data sets from a large-scale randomized controlled trial, and from a nonrandomized trial during the first year of broad implementation.

1. KiVa Antibullying Program

In 2006, the Finnish Ministry of Education and Culture commissioned the University of Turku to develop and evaluate an anti-bullying program for Finnish comprehensive schools (the basic nine-year education from grades 1 to 9). From the very beginning, the vision was to develop a program that would be suitable for nationwide implementation.

The resulting KiVa antibullying program is predicated on the idea that how peer bystanders react when witnessing bullying is crucial for putting an end to (or sustaining) it

(e.g., Salmivalli et al. 1996; Salmivalli 2010). Influencing the peer context is therefore an essential part of the KiVa program. The program is designed to produce its effects, first of all, by encouraging students to support victimized peers instead of providing social rewards to the bullies. In addition, adults (teachers as well as parents) are provided with information about bullying and efficacy to intervene and prevent it.

KiVa includes both universal and indicated actions. The core of the universal actions consists of theme lessons (primary school) and theme days (secondary school). The topics cover a variety of issues related to group interaction and group pressure, the mechanisms and consequences of bullying, and most importantly, what students can do together to counter bullying and support their victimized peers. Virtual learning environments (such as anti-bullying computer games) are an integral part of the universal actions. Their content is closely connected to the topics of the student lessons and themes, enhancing the learning process and motivating students to apply the learnt skills in everyday interactions with peers (Poskiparta et al. 2012). In addition, the universal actions include a parents’ guide, web resources for teachers, and materials reminding both students and school personnel of KiVa (posters, highly visible, bright-colored vests with the KiVa logo for teachers supervising recess time).

The indicated actions involve discussions with victims and bullies, as well with selected prosocial classmates, who are challenged to support the victimized classmate. The discussions with the bullies and victims are conducted by KiVa teams within the schools, while the classroom teacher organizes separate meetings with potential supporters of the victim (for a more detailed description of program content, see Salmivalli, Kärnä, and Poskiparta 2010a, 2010b).

2. Evaluation Strategy and Methods Used

Besides testing the effectiveness of the KiVa antibullying program, the evaluation studies were designed to answer the most pressing questions in the area of bullying interventions: when (under which circumstances) the program works, for whom does it work, and how does it produce its

effects. At the same time, the studies were designed to meet the highest methodological standards and to include assessments that have been neglected in much of the previous research (e.g., implementation fidelity). Many of these studies are still ongoing. In the present article we provide an overview of the findings concerning the main effects of KiVa after one school year (nine months of implementation) under two different circumstances: a randomized controlled trial and large-scale implementation. In addition, we briefly summarize the first results concerning implementation fidelity.

We put the effectiveness of the KiVa program to a stringent test in a randomized controlled trial across all grade levels in comprehensive education (grades 1 to 9). Students in these grade levels are aged seven to fifteen years. All schools in mainland Finland were informed about the evaluation study and invited to participate either as intervention or control schools. From the volunteering schools, 234 were chosen and assigned to the two conditions. The control schools were told that they could start implementing the KiVa program after serving as controls for one year – which most of them did.

Most data were collected by means of online questionnaires where students logged in using the same personal user ID at each assessment point – thus, individual students were followed over time. Bullying and victimization were assessed by self-reports, peer-reports, and dyadic data (e.g., “who has been bullying you?”) enabling us to link bullied children to their tormentors and to identify bully-victim dyads. In addition, data were collected on numerous variables related to bullying (attitudes, bystander behaviors, perceptions of teachers’ efficacy to tackle bullying) and defending of victimized peers (empathy towards victims, self-efficacy, outcome expectations regarding defending). Students’ sociometric status and psychosocial (e.g. self-esteem, peer perceptions, depression, social anxiety, loneliness) and academic (e.g., school liking, academic motivation, academic performance) adjustment were also assessed. Finally, data were collected from teachers who delivered the program lessons as well as from school staff who were members of KiVa teams tackling the cases of bullying that came to attention. The choice of vari-

ables was guided by our preliminary hypotheses on the mediating mechanisms of program effects, possible moderators of the effects, as well as constructs that might be positively influenced by reductions in bullying or victimization.

After the randomized controlled trial, KiVa became available for all Finnish schools. There were several reasons to continue collecting data in schools that adopted the program. First, since our control schools became intervention schools after the first year, continuing the randomized controlled trial was impossible. We wanted to know, however, what happened to the effects of the KiVa program in the long run. We also wanted to find out what happens to implementation fidelity when the program is implemented on a large scale. Moreover, ongoing data collection enabled us to provide annual feedback to the schools implementing the program: The annual online survey served as a tool to monitor their success, as well as the quality of implementation.

The first large pre-test survey for schools starting the implementation in fall 2009 took place in May 2009, followed by an annual assessment every May. Already after the first year we were able to estimate program effects utilizing a cohort-longitudinal design (Olweus and Alsaker 1991). In this design, post-test data from students in each age cohort were compared with baseline data from same-aged students from the same schools (i.e., in the previous cohort), who had not yet been exposed to the intervention. For instance, grade two students who had been targeted by KiVa for one year (post-test data collected in May 2010) were compared with grade two students from the same schools who had not yet been involved (pre-test data from May 2009).

3. Results of Evaluation Studies

3.1. Randomized Controlled Trial

The findings from the randomized controlled trial were promising: the first phase of evaluation involving grades 4–6 (Kärnä et al. 2011b) showed that KiVa significantly reduced both victimization and bullying. After the first school year (i.e., nine months of implementation of KiVa) the odds of being bullied systematically (at least two to three times a week) were about 1.5 times higher for a con-

trol school student than for a student in an intervention school, whereas the odds of bullying others systematically were 1.3 times higher for a control school student than for an intervention school student. These effect sizes correspond to approximately 30 percent and 17 percent reductions in victimization and bullying. Importantly, peer-reported victimization and bullying were significantly reduced as well (effect sizes equal or larger than in the case of self-reporting). The program also resulted in reductions in negative bystander behaviors (reinforcing the bully), as well as increases in empathy towards victimized peers and self-efficacy to support and defend them.

In a further study focusing on different forms of victimization in grades 4–6 (Salmivalli, Kärnä, and Poskiparta 2011) KiVa was shown to reduce each of the examined nine forms (physical, verbal, social exclusion, social manipulation, threatening, racist, material, and sexual victimization, and cybervictimization), the reductions varying between -20 percent (threatening) and -63 percent (material victimization, such as taking or breaking the target peer's belongings).

The second phase of the evaluation involving also younger (grades 1–3) and older (grades 7–9) students indicated that the effectiveness of the program varied considerably across grade levels (Kärnä et al. forthcoming). Overall, the effects were larger in primary than secondary grade levels. The average weighted odds ratios across all grade levels 1–9 were 1.28 for victimization and 1.30 for bullying, corresponding to reductions of about 20 percent in each case.

During the randomized controlled trial, we also tested two different approaches to dealing with children who had been involved in bullying others, which we refer to as confronting and nonconfronting approaches (Garandeau, Poskiparta, and Salmivalli 2012). In the former approach, the bully was firmly told that his/her behavior was not tolerated and had to cease immediately whereas in the latter, the adults did not blame the bully but rather shared their concern about the victim with him or her. In half of the schools involved in the randomized controlled trial as intervention schools, KiVa team members were instructed to use the confronting approach, whereas the other half used

the nonconfronting approach. The outcome variable was the victim's report (in the follow-up discussion with KiVa team members) on whether the bullying had stopped, decreased, increased, or remained unchanged. In 79 percent of all cases, victims reported that the bullying had stopped completely, and overall the two methods were equally effective at making the bullying stop. There were some moderators of their effectiveness, however: The nonconfronting approach was more successful than the confronting approach in cases of long-term victimization and in primary school. The confronting approach, on the other hand, was more effective in cases involving more than one bully.

Few anti-bullying program evaluations have investigated potential positive “side-effects” of the programs, in other words, beneficial outcomes that go beyond the initially intended reductions in bullying and victimization. In a study based on the randomized controlled trial data from grades 4 to 6 (Williford et al. 2012), the KiVa program was found to be effective in reducing students' internalizing symptoms (anxiety and depression) and improving their peer-group perceptions. Furthermore, Salmivalli, Garandeau, and Veenstra (2012) found that KiVa increased school liking, academic motivation, and even academic performance among students in KiVa schools (compared to control schools).

3.2. Going to Scale: National Rollout

The national launch of the KiVa antibullying program started in the fall of 2009, when 1,450 schools started to implement it. In 2010 and in 2011 new schools joined, and at present there are about 2,500 schools implementing the program. They represent 90 percent of all schools providing comprehensive education in Finland. We continue collecting data in these schools, both from students and staff. Whereas most evidence concerning the effects of anti-bullying programs comes from short-term, small-scale efficacy studies, our aim is to study both implementation and effectiveness during national rollout over several years.

The evaluation of program effects during the national rollout is based on students' responses to web-based surveys that are completed every spring in schools implementing the program. The effects were generally weaker during the

broad rollout than in the randomized controlled trial (Kärnä et al. 2011a). They were statistically significant at the primary school level (Grades 1–6) with respect to bullying others, as well as being bullied. At the secondary, or junior high school level (Grades 7–9), the effects for bullying others were in the right direction but not significant, and the effects for being bullied were just at the border of being significant (except in grade eight, where the reduction in victimization was significant). On average, the prevalence of children bullying others, as well as of those being bullied systematically (at least two to three times a week) were both reduced by about 15 percent during the first year of national rollout, the odds ratios being 1.22 for victimization and 1.18 for bullying. Effects of this magnitude mean that had all schools in Finland been implementing KiVa, the reductions would amount to about 7,500 bullies and 12,500 victims during the first one-year period. This demonstrates how even rather small effect sizes can make a huge difference in the lives of numerous children and adolescents.

3.3. Implementation Fidelity

Implementation fidelity, referring to the extent to which an intervention program is delivered as planned (Dusenbury et al. 2003), is a critical precondition for success of any prevention/intervention program. However, in studies evaluating antibullying programs, implementation has often been assessed at a very general level – if at all. In the evaluation of KiVa we placed concerted emphasis on the assessment of various aspects of implementation (such as preparation, dose, coverage, and student responsiveness) of the different program components.

In the randomized controlled trial we attempted to capture, besides the *level* of implementation, also the *implementation process* by gathering monthly data from teachers delivering the student lessons involved in the program. Furthermore, KiVa teams tackling acute cases of bullying reported back to us each step taken when handling a case of bullying. During the broad rollout of KiVa (since fall 2009) we continued gathering data on implementation via annual online surveys for the school personnel. There is already evidence of a positive association between the level of implementation and reduction in victimization, both

from the randomized controlled trial (Haataja, Voeten, and Salmivalli 2011) and broad rollout (Kärnä et al. 2011a).

Although the level of implementing KiVa program was overall high (e.g., most teachers delivered the majority of program lessons), it tended to decrease already during the first academic year (from fall to spring; Haataja, Voeten, and Salmivalli 2011) and even more so in subsequent years. Overall, implementation fidelity was somewhat lower during the broad rollout than during the randomized controlled trial (Salmivalli, Haataja, and Poskiparta 2011). Whereas primary school teachers delivered on average 8.7 out of 10 lessons during the randomized controlled trial, the corresponding number was 7.8 lessons during the first and 7.2 lessons during the second year of broad implementation. An important future task is to identify individual and school-level factors enhancing the likelihood of high-quality implementation. For instance, support from the school principal for antibullying work seems to be a crucial precondition for the successful delivery of program lessons (Ahtola et al. 2012).

4. Discussion

Studies of our randomized controlled trial and evaluation of large-scale implementation of the KiVa antibullying program show that the program is effective in reducing bullying and victimization (Kärnä et al. 2011a, 2011b, forthcoming). During the randomized controlled trial, the average reductions in bullying and victimization amounted to -20 percent whereas the corresponding reduction during broad rollout of KiVa was -15 percent. The former number corresponds to the average effects found in the meta-analysis by Farrington and Ttofi (2009). It should be noted, however, that no previous study has included such a wide age range as ours. Even though the effects of KiVa were small (even minimal) at secondary school level (grades 7–9) they were clearly above average in primary school grades. Moreover, the effects of KiVa (even when calculated across all grade levels) were stronger than the effects obtained in previous randomized controlled trials, and they were confirmed by multiple informants (children themselves as well as their peers). It should also be remembered that all effects reported in the present paper concerned changes after only nine months of program implementation.

Besides reducing bullying and victimization, KiVa was shown to reduce internalizing problems (Williford et al. 2012). It also led to increases in well-being and school liking among a much wider group of children than just previous victims and bullies (Salmivalli et al. 2012). The indicated actions (i.e. discussions between schools' KiVa team members and children involved in bullying) were highly effective (Garandeau et al. 2012).

We are not aware of any country where a school-based anti-bullying program has spread at the pace KiVa did in Finland. How can the willingness of Finnish schools to adopt the program be explained? One explanation is the support provided by the government, which enabled the schools that introduced the program during 2009–2011 get all the materials as well as the two-day pre-implementation training free of charge. Second, we believe that the KiVa program was regarded as feasible by school staff and its reputation as an effective and user-friendly program spread fast – partly because of the relatively broad media attention devoted to the program. Third, two school massacres took place in Finland in 2007 and 2008, just before the broad rollout of KiVa started. They were widely linked to bullying problems and thus generated a lot of discussion about bullying and the need for effective prevention.

4.1. Strengths and Limitations

Throughout the KiVa evaluation project, we made concerted efforts to meet the highest methodological standards by including an appropriate control condition, random assignment, multilevel modeling of hierarchical data, multimethod and multi-informant outcome measures, systematic monitoring of implementation, adequate sample size, and accurate handling of missing data. We are not aware of any other evaluation of a bullying intervention with similar methodological rigor, an equally large and representative sample, and such numerous outcome variables (both self- and peer-reported).

Several limitations of our evaluation should be pointed out, however. It was unfortunate that our randomized controlled trial only lasted for one school year: after that, it was not possible to assess the effects of the KiVa pro-

gram by comparing changes occurring in KiVa schools with those taking place in control schools. We can only study the long-term effects of KiVa utilizing the cohort-longitudinal design. However, as this design is potentially sensitive to history effects it is not considered as strong as the RCT design.

All our data were collected by (mostly online) questionnaires. It would have been desirable to collect additional observational data or perhaps interview students in a randomly selected subsample. Whereas observational data is free from response bias, interviews might have provided insight into how the students perceived the program lessons and what they felt made them think and behave differently. Assessing implementation by observation, in addition to teacher and student reports, might have provided a good addition to questionnaire data.

4.2. Future Directions

Many planned studies, for example concerning the mediators and moderators of the effects of the KiVa anti-bullying program, are in progress. Besides the already existing data sets, it will be exciting to follow the implementation and effectiveness of the KiVa program during years to come. Going to scale is not unproblematic: one of the challenges is the quality of implementation in a situation where schools get limited external guidance. As much as program developers would like school staff to put effort into delivering the program as recommended, reality often restrains their possibilities for doing so. Multiple projects running alongside the usual curriculum, lack of resources, organizational changes, and work stress are among the factors that might decrease school staff's willingness to implement a program, even where they believe it has positive outcomes. The greater the standardization of a program, however, the easier it is for local agents to grasp the model and to identify deviations from it (Bradach 2004). As KiVa includes concrete guidelines and schedules for the different measures, it is relatively easy to implement with fidelity even with minimal external support. The motivation to implement the program may diminish over time, however, and support for schools is essential. It remains to be seen whether the effects obtained so far will be sustained and – hopefully – grow further.

References

- Ahtola, Annarilla, Anne Haataja, Antti Kärnä, Elisa Poskiparta, and Christina Salmivalli. 2012. Implementation of Antibullying Lessons in Elementary Classrooms: Principal Support Matters. Unpublished manuscript.
- Bradach, Jeffrey. 2004. *Going to Scale: The Challenge of Replicating Social Programs*, Stanford Social Innovation Review. Palo Alto, CA: Stanford University.
- Card, Noel, and Ernest Hodges. 2008. Peer Victimization among Schoolchildren: Causes, Consequences, and Considerations in Assessment and Intervention. *School Psychology Quarterly* 23:451–61.
- Craig, Wendy, and Yossi Harel. 2004. Bullying, Physical Fighting, and Victimization. In *Young People's Health in Context: International Report from the HBSC 2001/02 Survey*, ed. Candace Currie, Chris Roberts, Antony Morgan, Rebecca Smith, Wolfgang Settertobulte, Oddrun Samdal, and Vivian Barnekow Rasmussen, WHO Policy Series: Health Policy for Children and Adolescents 4. Copenhagen: WHO Regional Office for Europe.
- Dusenbury, Linda, Rosalind Brannigan, Mathea Falco, and William B. Hansen. 2003. A Review of Research on Fidelity of Implementation: Implications for Drug Abuse Prevention in School Settings. *Health Education Research* 18:237–56.
- Farrington, David P., and Maria Ttofi. 2009. School-Based Programs to Reduce Bullying and Victimization. *Campbell Systematic Reviews* 6.
- Fekkes, Minne, Frans I. M. Pijpers, and S. Pauline Verloove-Vanhorick. 2006. Effects of Antibullying School Program on Bullying and Health Complaints. *Archives of Pediatrics and Adolescent Medicine* 160:638–44.
- Ferguson, Christopher J., Claudia San Miguel, John C. Kilburn, Jr., and Patricia Sanchez. 2007. The Effectiveness of School-Based Anti-Bullying Programs: A Meta-Analytic Review. *Criminal Justice Review* 32:401–14.
- Garandeau, Claire, Elisa Poskiparta, and Christina Salmivalli. 2012. Tackling Acute Cases of Bullying: Comparison of Two Methods in the Context of the KiVa Antibullying Program. Unpublished manuscript.
- Haataja, Anne, Aaron Boulton, Elisa Poskiparta, Marinus Voeten, and Christina Salmivalli. 2012. KiVa Antibullying Program in Primary Schools: Implementation and Outcome. Paper presented at the conference of the Society of Research on Adolescence, Vancouver, Canada.
- Haataja, Anne, Marinus Voeten, and Christina Salmivalli. 2011. Variation in the Implementation Process of the KiVa Antibullying Program in Finland. Poster presented at the Conference on Evidence-Based Prevention of Bullying and Youth Violence, Cambridge, UK.
- Hawker, David, and Michael Boulton. 2000. Twenty Years' Research on Peer Victimization and Psychosocial Maladjustment: A Meta-Analytic Review of Cross-Sectional Studies. *Journal of Child Psychology and Psychiatry* 41:441–55.
- Isaacs, Jenny, Ernest V. E. Hodges, and Christina Salmivalli. 2008. Long-Term Consequences of Victimization: A Follow-Up from Adolescence to Young Adulthood. *European Journal of Developmental Science* 2:387–97.
- Kärnä, Antti, Marinus Voeten, Todd Little, Erkki Alanen, Elisa Poskiparta, and Christina Salmivalli. Forthcoming. Effectiveness of the KiVa Antibullying Program: Grades 1–3 and 7–9. *Journal of Educational Psychology*.
- Kärnä, Antti, Marinus Voeten, Todd Little, Elisa Poskiparta, Erkki Alanen, and Christina Salmivalli. 2011a. Going to Scale: A Nonrandomized Nationwide Trial of the KiVa Antibullying Program for Grades 1–9. *Journal of Consulting and Clinical Psychology* 79:796–805.
- Kärnä, Antti, Marinus Voeten, Todd Little, Elisa Poskiparta, Anne Kaljonen, and Christina Salmivalli. 2011b. A Large-Scale Evaluation of the KiVa Antibullying Program: Grades 4–6. *Child Development* 82:311–30.
- Merrell, Kenneth W., Brabara A. Gueldner, Scott W. Ross, and Duane M. Isava. 2008. How Effective Are School Bullying Intervention Programs? A Meta-Analysis of Intervention Research. *School Psychology Quarterly* 23:26–42.
- Olweus, Dan. 1991. Bully/Victim Problems among Schoolchildren: Basic Facts and Effects of a School-Based Intervention Program. In *The Development and Treatment of Childhood Aggression*, ed. Debra Pepler and Kenneth Rubin, 411–48. Hillsdale, NJ: Erlbaum.
- Olweus, Dan, and Francoise Alsaker. 1991. Assessing Change in a Cohort-Longitudinal Study with Hierarchical Data. In *Problems and Methods in Longitudinal Research: Stability and Change*, ed. David Magnusson, Lars Bergman, George Rudinger, and Bertil Törestad, 107–32. New York: Cambridge University Press.
- Poskiparta, Elisa, Ari Kaukiainen, Virpi Pöyhönen, and Christina Salmivalli. 2012. Anti-Bullying Computer Game as Part of the KiVa Program: Students' Perceptions of the Game. In *The Impact of Technology on Relationships in Educational Settings*, ed. Angela Costabile and Barbara Spears, 158–68. New York: Routledge.
- Rivers, Ian, V. Paul Poteat, Nathalie Noret, and Nigel Ashurst. 2009. Observing Bullying at School: The Mental Health Implications of Witness Status. *School Psychology Quarterly* 24: 211–23.
- Salmivalli, Christina. 2010. Bullying and the Peer Group: A Review. *Aggression and Violent Behavior* 15:112–20.
- Salmivalli, Christina, Claire Garandeau, and Rene Veenstra. 2012. KiVa Anti-Bullying Program: Implications for School Adjustment. In *Peer Relationships and Adjustment at School*, ed. Gary Ladd and Allison Ryan. Charlotte, NC: Information Age Publishing.
- Salmivalli, Christina, Anne Haataja, and Elisa Poskiparta. 2011. Implementation Fidelity of the KiVa Antibullying Program During Randomized Controlled Trial and Broad Dissemination. Presentation in symposium on large-scale implementation of school-based anti-bullying programs: strategies and some challenges, chaired by D. Olweus at the 15th European Conference of Developmental Psychology, Bergen, Norway.
- Salmivalli, Christina, Antti Kärnä, and Elisa Poskiparta. 2010a. From Peer Put-downs to Peer Support: A Theoretical Model and How It Translated into a National Anti-Bullying Program. In *Handbook of Bullying in Schools: An International Perspective*, ed. Shane R. Jimerson, Susan M. Swearer, and Dorothy L. Espelage, 441–54. New York: Routledge.
- Salmivalli, Christina, Antti Kärnä, and Elisa Poskiparta. 2010b. Development, Evaluation, and Diffusion of a National Anti-Bullying Program, KiVa. In *Handbook of Youth Prevention Science*, ed. Beth Doll, William Pfohl, and Jina S. Yoon, 240–54. New York: Routledge.
- Salmivalli, Christina, Antti Kärnä, and Elisa Poskiparta. 2011. Counteracting Bullying in Finland: The KiVa Program and Its Effects on Different Forms of Being Bullied. *International Journal of Behavioral Development* 35:405–11.
- Salmivalli, Christina, Ari Kaukiainen, and Marinus Voeten. 2005. Anti-Bullying Intervention: Implementation and Outcome. *British Journal of Educational Psychology* 75:465–87.
- Salmivalli, Christina, Kirsti M. J. Lagerspetz, Kaj Björkqvist, Karin Österman, and Ari Kaukiainen. 1996. Bullying as a Group Process: Participant Roles and Their Relations to Social Status within the Group. *Aggressive Behavior* 22:1–15.
- Smith, J. David, Barry H. Schneider, Peter K. Smith, and Katerina Ananiadou. 2004. The Effectiveness of Whole-School Antibullying Programs: A Synthesis of Evaluation Research. *School Psychology Review* 33:547–60.
- Sourander, Andre, Peter Jensen, John A. Ronning, Solja Niemelä, Hans Helenius, Lauri Sillanmäki, Kirsti Kumpulainen, Jorma Piha, Tuula Tamminen, Irma Moilanen, and Fredrik Almqvist. 2007. What is the Early Adulthood Outcome of Boys Who Bully or Are Bullied in Childhood? The Finnish "From a Boy to a Man" Study. *Pediatrics* 120:397–404.
- Ttofi, Maria, and David P. Farrington. 2011. Effectiveness of School-Based Programs to Reduce Bullying: A Systematic and Meta-Analytic Review. *Journal of Experimental Criminology* 7:27–56.

- Ttofi, Maria, David P. Farrington, and Anna Baldry. 2008. *Effectiveness of Programmes to Reduce School Bullying: A Systematic Review*, report for the Swedish National Council for Crime Prevention. Västerås, Sweden: Edita Norstedts.
- Williford, Anne, Aaron Boulton, Brian Noland, Antti Kärnä, Todd Little, and Christina Salmivalli. 2012. Effects of the KiVa Antibullying Program on Adolescents' Depression, Anxiety, and Perception of Peers. *Journal of Abnormal Child Psychology* 40:289–300.
- Vreeman, Rachel C., and Aaron E. Carroll. 2007. A Systematic Review of School-Based Interventions to Prevent Bullying. *Archives of Pediatrics and Adolescent Medicine* 161:78–88.

Christina Salmivalli
ejjasal@utu.fi

Elisa Poskiparta
elisa.poskiparta@utu.fi

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program

Rosario Ortega-Ruiz, Department of Psychology, University of Cordoba, Spain, and Visiting Professor at
Department of Psychology and Counselling, University of Greenwich, United Kingdom

Rosario Del Rey, Department of Developmental and Educational Psychology, University of Seville, Spain

José A. Casas, Department of Psychology, University of Cordoba, Spain

Vol. 6 (2) 2012

Focus: Evidence-based Developmental Prevention of Bullying and Violence in Europe

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

► **Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program** Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program

Rosario Ortega-Ruiz, Department of Psychology, University of Cordoba, Spain, and Visiting Professor at
Department of Psychology and Counselling, University of Greenwich, United Kingdom

Rosario Del Rey, Department of Developmental and Educational Psychology, University of Seville, Spain

José A. Casas, Department of Psychology, University of Cordoba, Spain

An evaluation of the success of the evidence-based ConRed program, which addresses cyberbullying and other emerging problems linked with the use of the internet and seeks to promote a positive use of this new environment. The main aims of the ConRed program are a) to improve perceived control over information on the internet, b) to reduce the time dedicated to digital device usage, and c) to prevent and reduce cyberbullying. The impact of the program was evaluated with a quasi-experimental design with a sample of 893 students (595 experimental and 298 control). The results of the mixed repeated measures ANOVAs demonstrate that ConRed contributes to reducing cyberbullying and cyber-dependence, to adjusting the perception of information control, and to increasing the perception of safety at school.

Research into bullying and action programs aimed at preventing bullying or alleviating its effects have a history stretching back over more than three decades (Rigby and Smith 2011). In recent years traditional forms of bullying – physical, verbal, and relational aggression – have been joined by cyberbullying, a new phenomenon which reflects the increasingly widespread use of digital devices in peer interaction among adolescents and young adults (Baldry and Farrington 2007; Ttofi and Farrington 2011). The use of information and communications technologies, henceforth referred to as ICTs, can be said to be altering many aspects of young people's social lives, with traditional bullying now being replaced by more specific forms of abuse, intimidation, and harassment perpetrated via the digital devices they use to contact and communicate with each other (Mitchell, Finkelhor, and Wolak 2004). Action programs therefore need to include scientifically proven strategies focusing not only on bullying but also on cyber-behavior, the prevention of cyber-aggression and support for victims of cyberbullying.

1. Cyberbullying

Many researchers consider cyberbullying as merely an extension of traditional bullying and therefore define it as a series of intentional, repeated acts of aggression based on the establishment of some kind of power imbalance and carried out using technological devices (Slonje and Smith 2008; Tokunaga 2010). The use of such devices partially alters the nature of the contact between victims and aggressors and introduces specific new factors and risks, such as the anonymity of the aggressor, the greater social dissemination of the abuse being perpetrated, and the practical difficulties involved in halting the aggression and, by extension, shortening the victims' suffering (Patchin and Hinduja 2006). Some authors believe that these factors aggravate cyberbullying's effect on its victims (Dooley, Pyżalski, and Cross 2009), while others argue that cyberbullying, which is less common than traditional bullying in schools (Olweus 2012), offers victims opportunities to respond and defend themselves that are not available in face-to-face bullying scenarios (Law et al. 2012).

This work was produced as part of the project Cyberbullying in Adolescence: Investigation and Intervention in Six European Countries (DAPHNE III,

JLS/2008/CFP/DAP/2008-1) and the VIEJUCO project (School and Juvenile Violence: Risks of Violent Dating, Sexual Aggression, and Cyberbullying; I+D

Ps-2010-17246–National R&D Plan, Spain). The authors would like to thank these agencies for their funding.

Scientific literature on bullying risk factors has established two basic categories: factors based on the personal characteristics of the people involved (basically aggressors and victims, although bystanders also play a significant role) and factors based on certain elements in the social context in which the bullying takes place. These contextual elements include empathy (or its absence among aggressors), social incompetence in victims, and school climate (Mera-viglia et al. 2003; Nickerson, Mele, and Princiotta 2008; Sherer and Nickerson 2010). School climate, which essentially encompasses interpersonal affection and relationships and commonly accepted rules for social interaction (both implicit and explicit), is the setting in which bullies and victims play out their roles. Because cyberbullying is an indirect form of bullying, it should be remembered that risk factors present in the traditional bullies' and victims' social system are also risk factors for cyberbullying, although cyberbullying also has its own more specific risk factors (Bear et al. 2011). In terms of school climate, perceived safety and the absence of problems at school impede the emergence and consolidation of cyberbullying in relationships between classmates (Brand et al. 2003). Of the personality-based factors, empathy is a particularly important factor which is typically absent or deficient among bullies (Jolliffe and Farrington 2004); it may also be lacking among cyber-aggressors (Gini et al. 2007). Card and Hodges (2008) found a lack of social skills/competence among the victims of violent bullying, and this may also be mirrored in cyberbullying (Gradinger et al. 2012).

Factors associated exclusively with cyberbullying include lack of control over personal information made available on the internet and the compulsive use of the internet, which may lead to addiction and personality disorders (Ybarra and Mitchell 2004) and increases the risk of exposure to abuse via the internet (Dinev, Hart, and Mullen 2008). High-risk actions such as sharing passwords, talking to strangers, and uploading intimate information on social networks make victims more vulnerable (Gradinger et al. 2012; Hinduja and Patchin 2009). The disordered, compulsive use of the internet or social networks also distances individuals from direct social relationships and productive work or leisure time, leading to personality disorders and increasing the possibilities of indulging in

or becoming exposed to aggressive behavior (Ybarra and Mitchell 2004).

2. Tackling Cyberbullying

Thirty years of psycho-educational research into bullying have provided us with a wide range of preventive and palliative resources for dealing with the phenomenon (Ttofi and Farrington 2011), and much of this knowledge has been found also to be valid when addressing cyberbullying (Pearce et al. 2011). However, programs are needed that are capable of combining bullying prevention procedures of proven efficiency (Olweus 2012) with initiatives geared towards the prevention of cyberbullying and its associated contextual risks. And that is the aim of ConRed (Programa Conocer, Construir y Convivir en la Red, or the Knowing, Building, and Living Together on the Internet Program).

ConRed is an evidence-based intervention. Implemented using the procedures described in successful anti-bullying programs (Baldry and Farrington 2007; Olweus 2012; Pearce et al. 2011; Ttofi and Farrington 2011), it focuses on the cyberbullying risk factors mentioned above. ConRed is based on the following previously successful strategies:

- a) *Proactive policies, procedures, and practices*: the implementation of clear policies with practical procedures for reducing bullying (Ttofi and Farrington 2011) and organizational support (Rigby and Slee 2008; Vreeman and Carroll 2007). ConRed implements a specific action plan to combat the risks involved in using the internet and social networks, improving technical and procedural skills with digital devices, and teaching how to use ICTs safely and healthily.
- b) *School community key understandings and competencies*: the implementation of mechanisms which help to develop skills for preventing, identifying, and reacting to the problem (Baldry and Farrington 2007; Ttofi and Farrington 2011). ConRed's basic function is to instruct schoolchildren, teachers, and parents and improve their skills, to facilitate the safe, healthy use of the internet and social networks. The program focuses mainly on raising individuals' awareness and procedural skills in digital communication, the aim being to improve students' online social competence.

- c) *Protective school environment*: the provision of safe spaces and facilities positively influences student behavior (Pearce et al. 2011; Ttofi and Farrington 2011). ConRed helps schools to create safe, healthy virtual communication environments for students, fomenting in them a culture of mutual support, empathy with the weakest, and better social relationships (including digital communication) between the three groups involved in the school: students, teachers, and families.
- d) *School–family–community partnerships* to promote cooperation between the school, the families, and the leading local organizations through greater participation, as a means of encouraging support and reducing intimidating behavior (Hemphill et al. 2009; Ttofi and Farrington 2011; Hong and Espelage 2008). The ConRed program encourages cooperation between the three groups – students, teachers, and families – through joint activities, offering a virtual environment where the school community can meet to discuss the problems of bullying and cyberbullying (www.uco.es/laecovi/conred).

3. The Theory of Cyber-Behavior Risk Analysis and the Educational Criteria of the ConRed Program

The ConRed program embraces the theory of normative social behavior (Rimal et al. 2005) that has been employed in various action programs to modify juvenile attitudes and behavioral patterns such as alcohol consumption (Borsari and Carey 2000). This theory argues that human behavior and attitudes are heavily influenced by perceived social conventions. Applying the theory of normative social behavior to cyberbullying, adolescents may possibly see much of their own online communication and exchanges of information as quite normal and inevitable, without being aware of the consequences of their conduct. According to Rimal and Real (2003), this influence of social conventions on individual behavior takes place via three mechanisms: a) injunctive norms; b) social expectations; and c) group identity processes. The ConRed program takes these three mechanisms into account, making them part of the key training content in instructional work sessions conducted with students. They thus become material for debate among the schoolchildren. Laws which regulate and sanction certain forms of conduct on the in-

ternet and in social networks are analyzed, along with the consequences of breaking them. With respect to expectations, defined as the perceived benefits and/or disadvantages of engaging in certain forms of conduct (Bandura 1986), ConRed stresses the importance of critical awareness regarding the compulsive use of the internet and social networks, the naivety and mistakenness of believing that one has total control over the personal information uploaded to cyber-environments, and the negative consequences of misusing language online. Finally, with regard to group identity, defined as the unquestioning, uncritical adoption of the peer group's attitudes and conduct by an individual (Tajfel 2010), ConRed engages adolescents in debate about the morally devastating effects that may ensue when an individual is attacked on the internet and/or in social networks.

The ConRed program was designed and developed to prevent cyberbullying by raising levels of technical, procedural, and communications expertise and improving social skills in virtual scenarios, especially the internet and social networks. Although the approach was “holistic,” taking into consideration all three social groups in the school community – students, teachers and families –, the most important element was the work carried out with the students, who received eight training sessions conducted by external experts (the researchers). The experts worked in collaboration with each school's school climate planning team for three months.

The work carried out with the students was aimed at:

- a) Improving the schoolchildren's ICT usage habits, especially those related to controlling personal information as a form of reducing vulnerability;
- b) raising their awareness of time spent using ICTs, especially excessive time devoted to internet activities, and the risk of addiction; and
- c) analyzing the morally unjust, unhealthy nature of cyberbullying and the risks facing victims of abuse perpetrated via digital devices.

The ConRed program concentrated on working directly with the schoolchildren. Over a period of three months, weekly contact was maintained with the participating schools and eight classroom sessions were conducted (see Table 1). These sessions were structured to form three

units: a) The internet and social networks unit focused on the importance of privacy and control over shared content and processes and highlighted the negative consequences of failing to control or establish safety measures in online communication processes; b) In the unit on the benefits of using the internet and social networks healthily and intelligently, students were taught to improve their technical skills, to prioritize prosocial spaces and practices, and to exercise moral awareness and fairness by avoiding and reporting cyberbullying; c) The unit on dealing with the problems that may arise if the internet and social networks are used in a naive or malicious manner provided students with strategies for addressing the problems associated with inappropriate, irresponsible usage, with special attention being paid to the prevention of cyberbullying and internet addiction (abuse). Table 1 details the three conceptual units covered in the eight sessions conducted with the students. The same content was addressed in two sessions with the teachers and in one session with the students' families.

Table 1: ConRed teaching sessions

Session 1: What do the ICTs mean to you? And to people generally?
Session 2: How do you use social networks?
Session 3: Our plan of action to become an expert
Session 4: How do I feel doing different activities on the internet?
Session 5: How can the internet help me and others? How can I help others?
Session 6: What do we do on the internet and why it may be damaging?
Session 7: The advantages and disadvantages of social networks
Session 8: Reflection: quiz game for consolidating knowledge

The instructional stage with the students began by exploring their preconceived ideas about the issues involved. A picture, video, news item, or case description was then used to generate a debate, chaired and guided by one of the researchers. The aim was to provoke cognitive conflict and sensitize the participants to conceptual errors and false beliefs. The

session ended with a personalized exercise on internet and social network use which drew together what the students had learned about internet practice. The results were published in a manual (Ortega-Ruiz, Del Rey, and Casas 2012).

Concurrently with this direct intervention involving students, teachers, and families, the ConRed program also implemented an awareness-raising campaign, using materials like leaflets, posters, stickers, bookmarks, etc., to support the continuity of the measures being taken in the schools. Simple, clear messages were presented, providing information about how to use the internet and social networks correctly and how to prevent the risks that may be encountered if such resources are used inappropriately (see Table 2).

Table 2: Advice for teachers and families in the ConRed awareness-raising campaign

Advice for teachers	Advice for families
1. Make knowledge and command of the potential of ICTs, internet and social networks one of your objectives.	1. Teach your children to move around on the internet in the same way that you taught them to move around in the street: to be careful not to bump into anyone or let anyone bump into them.
2. Creating spaces for dialog and engagement is crucial for bringing the school closer to students and avoiding alienating them.	2. Protect your children from harmful elements on the internet just as you taught them to protect themselves against the cold, the rain, and dangers in the street
3. Include the social climate in cyberspace part of your school climate project, because relationships between students are continued in social networks.	3. Teach your children to be wary of invitations and messages from strangers. On the internet not all friends are real friends.
4. Adapt detection and deterrence procedures to emerging problems such as cyberbullying.	4. Don't forget the keys. On social networks the keys are the passwords. Teach your children how to use them safely.
5. Ask for guidance if our intervention is not having the desired effect.	5. Help your son or daughter to make their own decisions when they are online, and not to be swept along by what others do or say

4. ConRed Evaluation: Evidence of the Program's Preventive Efficiency

4.1. Hypothesis

Our starting hypothesis was that after implementing ConRed with the three groups (students, teachers, and families) and carrying out the accompanying information campaign, improvements would be seen in all three proposed objectives.

4.2. Objectives

- Improve perceived control over information on the internet and promote safety and privacy.
- Promote healthy use of the internet and a reduction in time dedicated to digital device usage, in order to prevent possible overuse and addiction.
- Reduce involvement in cyberbullying, in all roles, by reducing risk factors in order to create a greater sense of safety at school.

5. Methodology

5.1. Participants

The sample comprised 893 students at secondary schools in the city of Cordoba, Spain: 595 (45 percent female) in the experimental group and 298 (47.6 percent female) in the control group. The students' ages ranged from 11 to 19 years ($M = 13.8$; $SD = 1.47$).

5.2. Instruments

- The Perceived Information Control Scale (Dinev and Hart 2004), comprising four Likert-type items with seven answer options ranging from "totally disagree" to "strongly agree" ($\alpha = .896$).
- Adaptation of the Internet-Related Experiences Questionnaire (CERI) (Beranuy et al. 2009), comprising ten Likert-type items with four answer options (1 to 4) reflecting behavior frequency ranging from "never" to "quite a lot." This questionnaire has two scales: intrapersonal conflicts, covering aspects of substance abuse and addictive/pathological gambling, and interpersonal conflicts, covering key elements in ICT-based interpersonal relationships. Its reliability levels are acceptable (α total = .78; α intrapersonal = .719; α interpersonal = .631).
- The European Cyberbullying Intervention Project Questionnaire (ECIPQ) (Brighi et al. 2012a), comprising twenty-two Likert scale items with five answer options for

frequency ranging from never to more than once a week. This questionnaire has two dimensions, cybervictimization and cyberaggression, with good reliability levels (α total = .87; α victimization = .80; α aggression = .88).

- The European Bullying Intervention Project Questionnaire (ECIPQ) (Brighi et al. 2012b), comprising fourteen Likert-type items with five answer options for frequency ranging from "never" to "yes, more than once a week." This questionnaire has two scales, victimization and aggression, with acceptable reliability levels (α total = .82; α victimization = .85; α aggression = .77).
- The Basic Empathy Scale (Jolliffe and Farrington 2006), comprising twenty Likert items with five answer options reflecting level of agreement. This questionnaire has two dimensions, cognitive empathy and affective empathy, with acceptable reliability levels (α total = .70; α cognitive = .79; α affective = .85).
- The School Climate Scale (Brand et al. 2003), in which we used the Safety Problems subscale Likert items with five answer options (6 items, $\alpha = .71$).

5.3. Procedure

As mentioned above, ConRed is a holistic program similar to those models which have proven successful in preventing traditional bullying (Olweus 2012) and which were already present in the schools in which it was implemented. Three schools were chosen for the program. Two of them were public schools (one with relatively high socioeconomic indicators, the other less so). The third was a private school. At each school a meeting was arranged with management and the person in charge of improving school climate (*convivencia* in Spanish) and the nature, objectives, and implementation conditions of the ConRed program were explained. The three schools accepted the conditions, the timetabling proposal, and the program agenda. The program researchers also agreed with the schools which classes would take part as the experimental and control groups. The experimental group was larger at the specific insistence of the managements of the participating schools. The program was evaluated with a quasi-experimental, ex post facto, longitudinal design, with pre-post measurement, covering two groups (one experimental and one quasi-control) (Montero and León 2007).

5.4. Analysis

To evaluate the impact of the program, repeated measures general linear models or mixed repeated measures ANOVAS were applied. Homogeneity of covariance matrices and covariance matrix sphericity, or multi-sample sphericity requirements (Keselman and Keselman 1988; Hyunh 1978), were tested using Box's M-test, which gave $p > 0.05$ in all cases except for the cyberbullying and bullying dimensions. As violence dimensions, these were corrected using Friedman's F-test. As an analysis strategy we chose

the repeated measures ANOVA because, when the univariate conditions for the matrix are satisfied, as in this case, this technique is stronger and more powerful than other analyses in longitudinal studies (Albert 1999; Rogan, Keselman, and Mendoza 1979).

6. Results

The mean scores obtained highlight the differences between the experimental group and the control group (see Table 3).

Table 3: Mean scores

	Experimental group (n=595)		Control group (n=296)	
	Pre-test M (SD)	Post-test M (SD)	Pre-test M (SD)	Post-test M (SD)
Information control	5.51 (1.630)	5.01 (1.961)	5.25 (1.860)	5.33 (1.793)
Addiction to internet	1.18 (0.641)	1.16 (0.687)	1.19 (0.596)	1.18 (0.625)
Intrapersonal addiction	0.90 (0.687)	0.94 (0.725)	0.92 (0.658)	0.93 (0.698)
Interpersonal addiction	1.45 (0.713)	1.39 (0.727)	1.46 (0.699)	1.45 (0.683)
Cyberbullying	0.09 (0.196)	0.07 (0.165)	0.11 (0.313)	0.11 (0.285)
Aggressor cyberbullying	0.06 (0.220)	0.05 (0.147)	0.09 (0.272)	0.09 (0.304)
Victim cyberbullying	0.12 (0.262)	0.08 (0.258)	0.14 (0.277)	0.14 (0.341)
Bullying	0.37 (0.430)	0.24 (0.317)	0.34 (0.401)	0.30 (0.397)
Aggressor bullying	0.25 (0.398)	0.18 (0.306)	0.22 (0.353)	0.19 (0.290)
Victim bullying	0.50 (0.641)	0.32 (0.463)	0.49 (0.645)	0.42 (0.657)
Empathy	1.84 (0.370)	1.85 (0.350)	1.88 (0.418)	1.89 (0.344)
Cognitive empathy	2.19 (0.467)	2.16 (0.439)	2.21 (0.464)	2.22 (0.397)
Affective empathy	1.53 (0.460)	1.57 (0.486)	1.63 (0.508)	1.61 (0.448)
Safety problem	0.31 (0.409)	0.28 (0.403)	0.31 (0.439)	0.30 (0.400)

The mixed model analysis of repeated measurement variables (see Table 4) shows that for both boys and girls perception of control over information was significantly lower in the experimental group than in the control group ($F=10.320$ $df=1$, $p<.01^*$ $d=0.278$). Among boys, there was a significant drop in internet addiction ($F=4.353$ $df=1$, $p<.05^*$ $d=0.1$), especially interpersonal internet addiction ($F=4.708$ $df=1$, $p<.05^*$ $d=0.126$). The level of cyberbullying also dropped significantly ($F=6.695$ $df=1$, $p<.01^*$ $d=0.2$), with regard both to aggression ($F=6.047$ $df=1$, $p<.05^*$ $d=0.5$) and victimization ($F=5.530$ $df=1$, $p<.05^*$ $d=0.154$), as did that of traditional bullying ($F=7.859$ $df=1$, $p<.01^*$ $d=$

0.348). Statistical testing showed that these changes occurred principally in the aggression scale corresponding to boys (versus girls) ($F=11.940$ $df=1$, $p<.01^*$ $d=0.243$) and in victimization among both boys and girls ($F=6.571$ $df=1$, $p<.05^*$ $d=0.326$), although there was a much more marked drop among boys ($F=8.131$ $df=1$, $p<.01^*$ $d=0.433$). With regard to empathy there was a significant increase in affective empathy ($F=3.953$ $df=1$, $p<.05^*$ $d=-0.085$), this change being more noticeable among girls ($F=17.822$ $df=1$, $p<.01^*$ $d=-0.2$). The level of perception of safety problems was significantly lower among boys than it was for girls ($F=8.545$ $df=1$, $p<.01^*$ $d=0.221$).

Table 4: Repeated measurement ANOVA

	Group	M		F (Anova index)	P	M \pm SD		F (Anova Index)	P
		Pre	Post			Pre	Post		
Information control	Experimental	5.51	5.01	10.320	0.001*	5.39	5.01	0.857	0.355
	Control	5.25	5.33			5.66	5.28		
Addiction to internet	Experimental	1.18	1.16	0.002	0.968	5.33	5.29	4.353	0.037*
	Control	1.19	1.18			5.16	5.38		
Intrapersonal addiction	Experimental	0.90	0.94	0.087	0.768	1.11	1.07	0.010	0.921
	Control	0.92	0.93			1.27	1.29		
Interpersonal addiction	Experimental	1.45	1.39	2.361	0.125	1.22	1.16	4.708	0.030*
	Control	1.46	1.45			1.15	1.19		
Cyberbullying	Experimental	0.09	0.07	6.695	0.010*	0.81	0.85	0.525	0.469
	Control	0.11	0.11			1.02	1.07		
Aggressor cyberbullying	Experimental	0.06	0.05	6.047	0.014*	0.94	0.95	0.345	0.557
	Control	0.09	0.09			0.89	0.90		
Victim cyberbullying	Experimental	0.12	0.08	5.530	0.019*	1.38	1.29	0.151	0.698
	Control	0.14	0.14			1.54	1.51		
Bullying	Experimental	0.37	0.24	7.859	0.005*	1.47	1.41	15.005	0.000*
	Control	0.34	0.30			1.44	1.50		
Aggressor bullying	Experimental	0.25	0.18	0.022	0.882	0.10	0.07	11.940	0.001*
	Control	0.22	0.19			0.08	0.06		
Victim bullying	Experimental	0.50	0.32	6.571	0.011*	0.11	0.08	8.131	0.004*
	Control	0.49	0.42			0.11	0.08		
Empathy	Experimental	1.84	1.85	1.287	0.257	0.29	0.23	0.030	0.673
	Control	1.88	1.89			0.33	0.36		
Cognitive empathy	Experimental	2.19	2.16	0.911	0.340	0.34	0.25	0.024	0.878
	Control	2.21	2.22			0.31	0.21		
Affective empathy	Experimental	1.53	1.57	3.953	0.047*	0.18	0.13	17.822	0.000*
	Control	1.63	1.61			0.24	0.23		
Safety problem	Experimental	0.31	0.28	0.081	0.775	0.20	0.15	8.545	0.004*
	Control	0.31	0.30			0.55	0.31		
						0.45	0.33		
						0.49	0.48		
						0.48	0.36		
						1.81	1.77		
						1.88	1.94		
						1.84	1.84		
						1.93	1.94		
						2.20	2.10		
						2.19	2.24		
						2.18	2.10		
						2.24	2.26		
						1.51	1.51		
						1.57	1.66		
						1.55	1.53		
						1.71	1.70		
						0.34	0.31		
						0.27	0.26		
						0.35	0.35		
						0.26	0.24		

7. Discussion and Conclusions

The ConRed program produced positive results with regard to the three main objectives: the experimental group showed a noticeable global improvement both in comparison with the control group and in the pre-post measurement. With regard to the first objective, there was a significant decrease in perception of control over personal information on the internet and in social networks. We interpret this as an increase in awareness of the risks that may affect personal information and of the need to enhance safety measures to protect private content made available on the internet. Given that all the information displayed on the internet has effects on the construction of adolescents' personal identity (Nosko, Wood, and Molema 2010) there is a need to control it. Therefore, the above-mentioned decrease can be identified as a better-adjusted adolescent perception of the real control they have over their personal information on the internet which, in turn, can be identified as a greater awareness of situations of potential insecurity.

With regard to the second objective, that of training for healthier online behavior, the decrease in addiction-related problems differed between boys and girls. Among boys, there was a significant decrease in the need to be online interacting with others on social networks (interpersonal addiction), whereas the girls maintained higher levels of frequency of online communication with others and their behavior was not significantly modified by ConRed. Similar results have been obtained in other studies. For example, Echeburua and Corral (2009) and Ruiz-Olivares, Lucena, Pino, and Herruzo (2010) describe how girls are more likely to develop a certain level of addiction to online activity in the fields of communication and social networking. More work therefore seems to be necessary to achieve this objective among the female population.

The drop in cyberbullying, in terms of both aggression and victimization, proves that ConRed successfully achieved objective three. Levels of involvement in both roles decreased, and we consider this a major success of the program: other equally ecological programs encompassing all the agents involved have not altered the frequency with which schoolchildren play the role of aggressor, although

they have brought about a decrease in the number of those involved as victims (Ttofi and Farrington 2011). The decrease in offensive behavior and, by extension, in the number of schoolchildren who describe themselves as online bullies, may be explained by the training the children have received and the awareness-raising concerning the moral implications of aggressive or offensive content in social network communication. Raising awareness of the harm that can be caused to others by content manipulation, offensive language, social exclusion, threats, etc. has proved to be one of ConRed's most interesting achievements. The program was designed specifically to prevent teenagers from perpetuating the old problem of traditional bullying, which actually decreases during the years of adolescence, in the new online environment. The evaluation shows that risk awareness and the training of teachers and parents to monitor and guide youth behavior reduce high risk conduct, induce the taking of precautionary measures, and encourage protective attitudes in online activity. This is important because it offers victims a way out of their isolation, helping them to feel supported by influential adults and better able to handle cases of gratuitous and sometimes cruel aggression (Hunter and Boyle 2004). This interpretation is reinforced by the changes observed in ConRed's empathy measurements, with increases in feelings of understanding, recognition, and affection towards cyberbullying victims.

ConRed produced gender-differentiated benefits: among boys, bullying dropped in terms of both aggression and victimization, but that was not the case with girls involved as aggressors. In contrast, affective empathy rose significantly among girls but not among boys. That is to say, after the ConRed program had been implemented perception of school *climate* was seen to have changed as intended, but there are a number of sex-related considerations which require further attention. The same thing happened with perception of safety: the boys now feel their school is safer, but the girls do not.

Although the overall evaluation of ConRed as an action program was positive with regard to its proposed objectives, attention must be drawn to some limitations: since the program was implemented by the researchers them-

selves working directly with schoolchildren, teachers, and families, ConRed still lacks a mechanism by which responsibility for training can be transferred to the members of the school community. This would give education agents the autonomy to directly implement the program themselves. The main limitation of the research design was that the experimental and control groups came from the same schools, with the consequent risk of contamination. We fully acknowledge this limitation and the risks it implies for achieving a greater level of homogeneity and comparability

between the groups (Trochim 1984). One important topic for future research and action will be an evaluation of the impact of the program once a certain period of time has elapsed. Accordingly, it would be necessary to verify if the positive effects of ConRed are long-lasting even when there is no ongoing intervention or if the benefits disappear gradually. It would be significant to confirm whether these positive effects remain when the program is implemented by the natural agents, the teachers, in which it could be considered a valid program for general use by schools.

References

- Albert, Paul S. 1999. Longitudinal Data Analysis (Repeated Measures) in Clinical Trials. *Statistics in Medicine* 18 (13): 1707–32. doi:10.1002/(SICI)1097-0258(19990715)18:13<1707::AID-SIM138>3.0.CO;2-H
- Baldry, Anna C., and David P. Farrington. 2007. Effectiveness of Programs to Prevent School Bullying. *Victims and Offenders* 2 (2): 183–204.
- Bandura, Albert. 1986. *Social Foundations of Thought and Action*. Englewood Cliffs, NJ: Prentice-Hall.
- Bear, George G., Clare Gaskins, Jessica Blank, and Fang Fang Chen. 2011. Delaware School Climate Survey-Student: Its Factor Structure, Concurrent Validity, and Reliability. *Journal of School Psychology* 49 (2): 157–74. doi:10.1016/j.jsp.2011.01.001
- Beranuy, Marta, Andrés Chamarro, Carla Graner, and Xavier Carbonell-Sánchez. 2009. Validation of Two Brief Scales for Internet Addiction and Mobile Phone Problem Use. *Psicothema* 21 (3): 480–85.
- Borsari, Brian, and Kate B. Carey. 2000. Effects of a Brief Motivational Intervention with College Student Drinkers. *Journal of Consulting and Clinical Psychology* 68 (4): 728–33. doi:10.1037/0022-006X.68.4.728
- Brand, Stephen, Robert Felner, Minsuk Shim, Anne Seitsinger, and Thaddeus Dumas. 2003. Middle School Improvement and Reform: Development and Validation of a School-level Assessment of Climate, Cultural Pluralism, and School Safety. *Journal of Educational Psychology* 95 (3): 570–88. doi:10.1037/0022-0663.95.3.570
- Brighi, Antonella, Rosario Ortega, Jacek Pyzalski, Herbert Scheithauer, Peter K. Smith, C. Tsormpatzoudis, V. Barkoukis, et al. 2012a. European Bullying Intervention Project Questionnaire (ECIPQ). University of Bologna. Unpublished manuscript.
- Brighi, Antonella, Rosario Ortega, Jacek Pyzalski, Herbert Scheithauer, Peter K. Smith, C. Tsormpatzoudis, V. Barkoukis, et al. 2012b. European Cyberbullying Intervention Project Questionnaire (ECIPQ). University of Bologna. Unpublished manuscript.
- Card, Noel A., and Ernest V. E. Hodges. 2008. Peer Victimization Among Schoolchildren: Correlations, Causes, Consequences, and Considerations in Assessment and Intervention. *School Psychology Quarterly* 23 (4): 451–61. doi:10.1037/a0012769
- Dinev, Tamara, Paul Hart, and Michael R. Mullen. 2008. Internet Privacy Concerns and Beliefs about Government Surveillance: An Empirical Investigation. *Journal of Strategic Information Systems* 17 (3): 214–33. doi:10.1016/j.jsis.2007.09.002
- Dinev, Tamara and Paul Hart. 2004. Internet privacy concerns and their antecedents—measurement validity and a regression model. *Behaviour and Information Technology* 23 (6): 413–22. doi:10.1080/01449290410001715723
- Dooley, Julian J., Jacek Pyzalski, and Donna Cross. 2009. Cyberbullying Versus Face-to-Face Bullying. *Zeitschrift Für Psychologie/Journal of Psychology* 217 (4): 182–88. doi:10.1027/0044-3409.217.4.182
- Echeburúa, Enrique, and Paz Corral. 2009. Las adicciones con o sin droga: una patología de la libertad. In *Adicción a las nuevas tecnologías en adolescentes y jóvenes*, ed. Enrique Echeburúa, Francisco J. Labrador, and Elisardo Becoña, 29–44. Madrid: Pirámide.
- Gini, Gianluca, Paolo Albiero, Beatrice Benelli, and Gianmarco Altoè. 2007. Does Empathy Predict Adolescents' Bullying and Defending Behavior? *Aggressive Behavior* 33 (5): 467–76. doi:10.1002/ab.20204
- Gradinger, Petra, Dagmar Strohmeier, Eva Maria Schiller, Elisabeth Stefanek, and Christiane Spiel. 2012. Cyber-victimization and Popularity in Early Adolescence: Stability and Predictive Associations. *European Journal of Developmental Psychology* 9 (2): 228–43. doi:10.1080/17405629.2011.643171
- Hemphill, Sheryl A., Rachel Smith, John W. Toumbourou, Todd I. Herrenkohl, Richard F. Catalano, Barbara J. McMorris, and Helena Romaniuk. 2009. Modifiable Determinants of Youth Violence in Australia and the United States: A Longitudinal Study. *Australian and New Zealand Journal of Criminology* 42 (3): 289–309. doi:10.1375/acri.42.3.289
- Hinduja, Sameer, and Justin W. Patchin. 2009. *Bullying Beyond the Schoolyard: Preventing and Responding to Cyberbullying*. Thousand Oaks, CA: Corwin.
- Hong, Jun Sung, and Dorothy L. Espelage. 2008. A Review of Research on Bullying and Peer Victimization in School: An Ecological System Analysis. *Aggression and Violent Behavior* 17 (4): 311–32. doi:10.1016/j.avb.2012.03.003
- Hunter, Simon C., and James M. E. Boyle. 2004. Appraisal and Coping Strategy Use in Victims of School Bullying. *British Journal of Educational Psychology* 74 (1): 83–107. doi:10.1348/000709904322848833
- Hyunh, Hyunh. 1978. Some Approximate Tests for Repeated Measurement Designs. *Psychometrika* 43: 383–86.
- Jolliffe, Darrick, and David P. Farrington. 2004. Empathy and Offending: A Systematic Review and Meta-Analysis. *Aggression and Violent Behaviour* 9 (5): 441–76. doi:10.1016/j.avb.2003.03.001
- Jolliffe, Darrick, and David P. Farrington. 2006. Development and Validation of the Basic Empathy Scale. *Journal of Adolescence* 29 (4): 589–611. doi:10.1016/j.adolescence.2005.08.010
- Keselman, Harvey J., and Joanne C. Keselman. 1988. Comparing Repeated Measures Means in Factorial Designs. *Psychophysiology* 25 (5): 612–18. doi:10.1111/j.1469-8986.1988.tb01898.x
- Law, Danielle M., Jennifer D. Shapka, José F. Domene, and Monique H. Gagné. 2012. Are Cyberbullies Really Bullies? An Investigation of Reactive and Proactive Online Aggression. *Computers in Human Behavior* 28 (2): 664–72. doi:10.1016/j.chb.2011.11.013
- Meraviglia, Martha G., Heather Becker, Barri Rosenbluth, Ellen Sanchez, and Trina Robertson. 2003. The Expect Respect Project: Creating a Positive Elementary School Climate. *Journal of Interpersonal Violence* 18 (11): 1347–60. doi:10.1177/0886260503257457

- Mitchell, Kimberly J., David Finkelhor, and Janis Wolak. 2004. Victimization of Youths on the Internet. *Journal of Aggression, Maltreatment and Trauma* 8 (1–2): 1–39. doi:10.1300/J146v08n01_01
- Montero, Ignacio, and Orfelio G. León. 2007. A Guide for Naming Research Studies in Psychology. *International Journal of Clinical and Health Psychology* 7 (3): 847–62.
- Nickerson, Amanda B., Danielle Mele, and Dana Princiotta. 2008. Attachment and Empathy as Predictors of Roles as Defenders or Outsiders in Bullying Interactions. *Journal of School Psychology* 46 (6): 687–703.
- Nosko, Amanda, Eileen Wood, and Seija Molema. 2010. All About Me: Disclosure in Online Social Networking Profiles: The Case of FACEBOOK. *Computers in Human Behavior* 26 (3): 406–18. doi:10.1016/j.chb.2009.11.012
- Olweus, Dan. 2012. Cyberbullying: An Overrated Phenomenon? *European Journal of Developmental Psychology* 9 (5): 520–38. doi:10.1080/17405629.2012.682358
- Ortega-Ruiz, Rosario, Rosario Del Rey, and José A. Casas. 2012. *El programa ConRed: Un programa Psicoeducativo para el uso seguro y saludable de Internet y las Redes Sociales*. Córdoba: Laecovi.
- Patchin, Justin W., and Sameer Hinduja. 2006. Bullies Move Beyond the Schoolyard: A Preliminary Look at Cyberbullying. *Youth Violence and Juvenile Justice* 4 (2): 148–69. doi:10.1177/1541204006286288
- Pearce, Natasha, Donna Cross, Helen Monks, Stacey Waters, and Sarah Falconer. 2011. Current Evidence of Best Practice in Whole-School Bullying Intervention and Its Potential to Inform Cyberbullying Interventions. *Australian Journal of Guidance and Counselling* 21 (1): 1–21.
- Rigby, Ken, and Phillip Slee. 2008. Interventions to Reduce Bullying. *International Journal of Adolescent Medicine and Health* 20 (2): 165–83.
- Rigby, Ken, and Peter K. Smith. 2011. Is School Bullying Really on the Rise? *Social Psychology of Education* 14 (4): 441–55.
- Rimal, Rajiv N., Maria K. Lapinski, Rachel J. Cook, and Kevin Real. 2005. Moving Toward a Theory of Normative Influences: How Perceived Benefits and Similarity Moderate the Impact of Descriptive Norms on Behaviors. *Journal of Health Communication* 10 (5): 433–50.
- Rimal, Rajiv N., and Kevin Real. 2003. Understanding the Influence of Perceived Norms on Behaviors. *Communication Theory* 13 (2): 184–203. doi:10.1111/j.1468-2885.2003.tb00288.x
- Rogan, Joanne C., H. Joanne Keselman, and Jorge L. Mendoza. 1979. Analysis of Repeated Measurements. *British Journal of Mathematical and Statistical Psychology* 32 (2): 269–86. doi:10.1111/j.2044-8317.1979.tb00598.x
- Ruiz-Olivares, Rosario, Valentina Lucena, Maria J. Pino, and Javier Herruzo. 2010. Analysis of Behavior Related to Use of the Internet, Mobile Telephones, Compulsive Shopping and Gambling Among University Students. *Adicciones* 22 (4): 301–10.
- Sherer, Yiping C., and Amanda B. Nickerson. 2010. Anti-Bullying Practices in American Schools: Perspectives of School Psychologists. *Psychology in the Schools* 47 (3): 217–29. doi:10.1002/pits.20466
- Slonje, Robert, and Peter K. Smith. 2008. Cyberbullying: Another Main Type of Bullying? *Scandinavian Journal of Psychology* 49 (2): 147–54. doi:10.1111/j.1467-9450.2007.00611.x
- Tajfel, Henri. 2010. *Social Identity and Intergroup Relations*. New York: Cambridge University Press.
- Tokunaga, Robert S. 2010. Following You Home from School: A Critical Review and Synthesis of Research on Cyberbullying Victimization. *Computers in Human Behavior* 26 (3): 277–87. doi:10.1016/j.chb.2009.11.014
- Trochim, William M. K. 1984. . Beverly Hills, CA: Sage.

Rosario Ortega
ed1orrur@uco.es

Rosario Del Rey
delrey@us.es

José A. Casas
m22caboj@uco.es

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model

Ersilia Menesini, Department of Psychology, University of Florence, Italy

Annalaura Nocentini, Department of Psychology, University of Florence, Italy

Benedetta Emanuela Palladino, Department of Psychology, University of Florence, Italy

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

► **Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model** Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model

Ersilia Menesini, Department of Psychology, University of Florence, Italy
Annalaura Nocentini, Department of Psychology, University of Florence, Italy
Benedetta Emanuela Palladino, Department of Psychology, University of Florence, Italy

An investigation of whether and to what extent a peer-led model is able to counteract mechanisms underlying bullying in peer groups, seeking clarification of divergence in reported results on the efficacy of peer-led models. Two studies were carried out in Italy within a project tackling bullying and cyberbullying in secondary schools. In the first study ($n= 386$), concerning the first phase of the project, a significant decrease was found only for cyberbullying, most of all for male peer educators. For the second study ($n= 375$) the model was improved and significant effects were found for several participating groups (peer educators and the experimental classes), who exhibited a decrease in bullying, victimization, and cybervictimization. Results suggest that peer educators can act as agents of change in the broader context.

School bullying has become a global problem in Western society, with potentially high social costs. Relevant percentages of primary and secondary school students are involved in peer-to-peer bullying as perpetrators or victims or both. Along with traditional types, a new form of bullying has appeared: cyberbullying, defined as bullying perpetrated by the use of electronic devices (Smith et al. 2008; Menesini, Nocentini, and Calussi 2011). In a recent Italian survey of a nation-wide sample of adolescents aged 12–18 years, 25.2 percent reported experiencing face-to-face bullying and 10 percent cyberbullying (Eurispes 2011).

Bullying is associated with externalizing behaviors, while being victimized causes psychological distress, low self-esteem, depression, anxiety, and psychosomatic symptoms (Arsenault, Bowes, and Shakoor 2010; Menesini, Modena, and Tani 2009; Ttofi et al. 2011; Veenstra et al. 2010). School bullying has also a negative impact on bystanders and on other children not involved in bullying problems (Gini et al. 2008).

Given these findings, the need for intervention to limit the harm caused by bullying is clear and urgent. The question is how to intervene: what are the psychological mechanisms underlying bullying attacks? What anti-bullying interventions are effective, and to what extent?

We know that bullies are usually motivated to gain dominance within the group (Pellegrini 2002; Salmivalli and Peets 2008), but we do not know why they are not stopped by the rest of the class. Several mechanisms explain bystanders' (non-)reactions. One is diffusion of responsibility: when an event occurs in front of a group of persons each individual feels less responsible (Salmivalli 2010). Other reasons include it being easier to be on the side of the bullies, as the dominant group in the class, and the attitudes of the majority of the class often influencing the bystanders' behavior (Gini et al. 2008).

In relation to victimization, Holt and Espelage (2007) find that moderate levels of peer support can reduce levels of

The authors wish to thank the Province of Lucca for its financial support for the project and the schools, teachers, and students for their collaboration.

anxiety and depression in victims. Flashpohler et al. (2009) find that perceived peer social support moderates the relationship between victimization and quality of life to a greater extent than teacher social support. Literature on victim support and on the bystanders' role underlines the value of involving the group and specifically uninvolved children (the so-called "silent majority") to change the dynamics of bullying and to stop negative behaviors (Menesini et al. 2003; Salmivalli 2010).

The "peer education" and "peer support actions" approaches focus on peer involvement (Cowie and Wallace 2000; Shiner 1999). These two models are both based on the assumption that peers learn from and have significant influence on each other, and that norms and behaviors are most likely to change when liked and trusted group members take the lead (Shiner 1999; Turner and Shepherd 1999). Peer-led models grow out of the spontaneous willingness of children and adolescents to help one another and create roles and structures where students, on the role, can be trained and helped to act in a responsible, sensitive, and empathic way towards other pupils. These programs can enhance active citizenship and prosocial behavior among pupils (Cowie and Wallace 2000; Naylor and Cowie 1999). Several studies support the effectiveness of peer education and peer support action in reducing bullying behavior and pro-bullying attitudes (Menesini et al. 2003), increasing support for victims (Houlston, Smith, and Jessel 2011), and generating possible benefits for peer supporters and schools in general (Cowie et al. 2002; Maticka-Tyndale and Barnett 2010; Naylor and Cowie 1999).

Recent meta-analyses show that on average bullying can be reduced by 20–23 percent and victimization by 20 percent in intervention schools compared with control schools (Ttofi and Farrington 2009). Findings on the effectiveness of peer support and peer mediation in schools are more controversial. A first meta-analysis found working with peers to be effective, particularly for reducing victimization (Ttofi and Farrington 2009), whereas a more recent paper (Ttofi and Farrington 2011) reports it as having a negative effect on bullying reduction.

The presents study aims to: 1) contribute to the literature in relation to the contradictory results on peer-led models;

2) understand whether and to what extent a peer-led model against bullying and cyberbullying applied in Italy is able to counteract some of the mechanisms underlying bullying in peer groups. We report results from two studies carried out in Italy within an ongoing project tackling bullying and cyberbullying in secondary schools. Although certain elements of rigorous program evaluation are lacking, the findings are nonetheless relevant to understanding the role of particular mechanisms and program components.

1. Study 1: Noncadiamointrappola Phase One

The web-based Noncadiamointrappola (Let's not fall into a trap) project was launched in 2008. It involved students from two schools in designing and developing a website to promote peer-to-peer content against bullying and cyberbullying. More schools became involved during the following school year (2009–2010). The present study examines the following stages between December 2009 and June 2010 (Menesini, Calussi, and Nocentini 2012):

- Initial evaluation (December 2009): questionnaires administered to the experimental and control groups (T1).
- Launch of the project and awareness-raising. Presentation of the project to the participating schools and classes to raise awareness and generate communication on issues related to cyberbullying and bullying.
- Selection of four on-line peer educators and four face-to-face peer educators in each participating class.
- Training day for peer-educators (eight hours), focused on communication skills, problem-solving, and social skills in real and virtual interactions.
- Intervention by online educators in the Noncadiamointrappola forum through a rotation schedule where each educator worked for a period of two weeks, each day controlling the forum posting new threads, answering questions posted by users, moderating discussions.
- Intervention by face-to-face peer educators. In particular, 1) conducting an awareness meeting on bullying and cyberbullying with a school class that had not participated in the previous steps; 2) participating in a meeting with local administrators, police, etc., to ask for

specific help making life safer in their city; 3) preparing a TV program about bullying and cyberbullying for a local network.

- Final evaluation (June 2010): the initial questionnaire was re-administered to evaluate the outcome (T2).

The present study evaluates the effectiveness of the intervention, comparing the peer educators, the awareness group, and the control group.

1.1. Methodology

The sample comprised 386 adolescents (62 percent females) enrolled in 9th to 13th grade at eight high schools in Tuscany, Italy. The age of participants ranged from 14 to 20 years (mean 16.29; SD=1.29). The schools were selected using a self-selection process and the classes were selected by the school staff. The consent procedure consisted of formal approval by the schools and consent by the parents. Overall 236 adolescents were included in the analyses on the basis of complete data at T1 and T2 (62 percent of the sample). Participants who dropped out of the study did not differ from those who remained with regard to the initial study variables. The sample was divided into three groups based on level of involvement in the intervention: 1) the control group (students who did not receive any kind of intervention: N=47); 2) experimental group 1 (awareness) (students who received only an intervention based on raising awareness of cyberbullying; N= 126); 3) experimental group 2 (peer educators) (students who were highly involved, took part in training, and worked actively in the real or virtual community; N= 63).

1.2. Measures

Bullying and Victimization

Bullying and victimization scales were used (Menesini, Calussi, and Nocentini 2012). Each scale consists of eleven items, asking how often respondents had experienced particular behaviors as perpetrator or victim during the past couple of months. Each item was evaluated on a 5-point scale from “never” to “several times a week.” Alpha coefficients at T1 and T2 were .80 and .80 for bullying and .59 and .69 for victimization. Although victimization showed low levels, the reliability of the victimization scale is confirmed by previous studies (see Menesini et al. 2012).

Cyberbullying and Cybervictimization

A revised version of the cyberbullying scale described by Menesini, Nocentini, and Calussi (2011) was used. It consists of two scales, one for perpetration and one for victimization. Each scale consists of eighteen items, asking how often respondents had experienced particular behaviors during the past couple of months. Each item was evaluated on a 5-point scale from “never” to “several times a week.” Alpha coefficients at T1 and T2 were .67 and .75 for cyberbullying and .72 and .84 for cybervictimization.

1.3. Results of the First Study

A series of mixed repeated measures ANOVAs analyses were carried out in order to evaluate the effect of time on bullying, victimization, cyberbullying, and cybervictimization across the three groups (peer educators, awareness, and control), controlling for gender. For cyberbullying, results showed a significant effect of time ($F_{(4, 228)} = 7.64$; $p < .001$; $\eta^2_p = .03$), and a significant interaction of time*group ($F_{(4, 228)} = 3.408$; $p < .05$; $\eta^2_p = .02$) and of time*group*gender ($F_{(4, 288)} = 3.039$; $p < .05$; $\eta^2_p = .02$). The main time effect is that the mean for cyberbullying decreases significantly from T1 to T2 (Figure 1). However, interaction effects reveal that this decrease varies across groups and gender. As Figure 2 shows, a significant decrease over time was found only for peer educators, and in particular male peer educators (respectively $F_{(2, 63)} = 4.277$; $p < .05$; $\eta^2_p = .07$ and $F_{(2, 21)} = 5.251$; $p < .05$; $\eta^2_p = .25$). No other significant effect was found for bullying, victimization, or cybervictimization.

Figure 1: Change in cyberbullying over time (total sample)

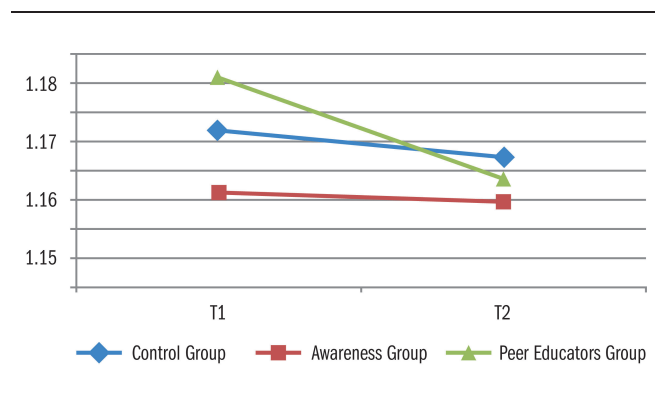
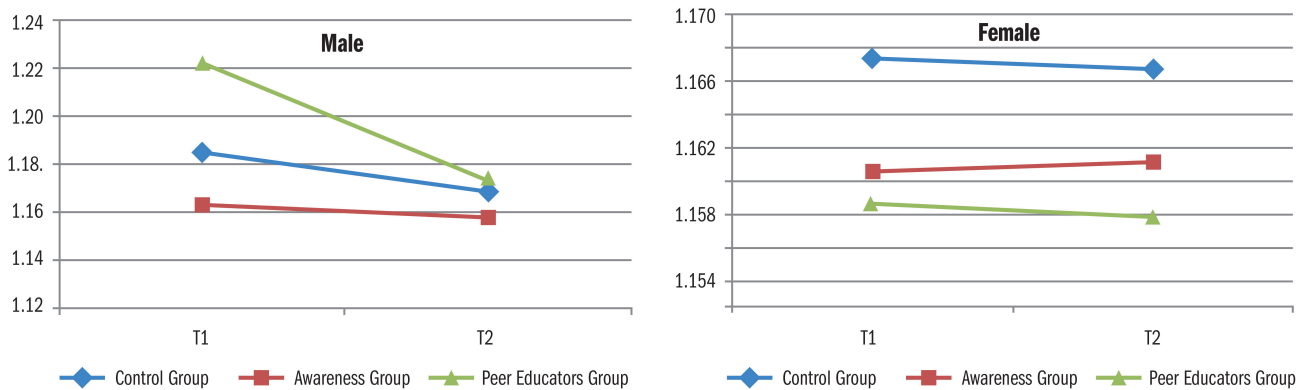


Figure 2: Change in cyberbullying over time by gender



1.4. Discussion

This first peer-led model produced more strongly positive effects for cyberbullying than for traditional bullying and stronger effects among male peer educators, who were the students involved in a very active and responsible role. The project was less effective for the other participants. This can be related to the type of intervention we carried out with these two groups: the peer educators' group worked more intensively, through the training and other tasks they were asked to fulfill. By comparison, the awareness group was less involved in the intervention and did not take part in an active process of empowerment. As an overall conclusion, the intervention showed some benefits but it was not so effective for the rest of the class, for victims, or for face-to-face bullying and victimization.

2. Study 2: Noncadiamointrappola Phase Two

Phase two of Noncadiamointrappola built on the initial results and sought to improve certain aspects of the model that were found to be underdeveloped in the first trial. Four elements were added:

- stronger attention to the victim's role and to support for the victims;
- more efforts to involve the bystanders;
- greater involvement of subject teachers in order to improve action on face-to-face bullying. Face-to-face educators were supported by class teachers and adapted their intervention to school needs. Specifically they tried

to involve the whole class and produced a short movie on cyberbullying, a guide for safer use of e-mail and social networks, and a poster against cyberbullying. In one school they ran a peer-to-peer counseling space.

- creation of a Facebook group to integrate the web forum: online peer educators posted photos, links, and video clips as facebook group administrators.

In order to evaluate the effects of Noncadiamointrappola phase two we analyzed data concerning bullying and cyberbullying in a before-after comparison of two groups: the control group (students who did not receive any kind of intervention) and the experimental group (all students attending classes participating in the project). In a second step, differences between peer educators and the other students of participant classes were analyzed: this last group comprised students who received the intervention provided by the trained peer educators within their class. This second study sought to evaluate the effectiveness of the intervention carried out by the peer educators by measuring the impact on the whole class. In particular, the question was: did the introduction of structured activities run by peer educators lead to change in the experimental classes as a whole? Were peer educators agents of change in these classes?

2.1. Methodology

The sample comprised 375 adolescents (males=20.3 percent), enrolled in 9th to 13th grade at four high schools in Tuscany. The experimental group composed 231 adoles-

cents (males = 15.4 percent; mean age = 16.80; SD = 1.92) attending ten classes at three high schools, and the control group comprised 144 adolescents (males = 20.8 percent; mean age = 15.15; SD = .90). Forty-two students from the experimental group were enrolled as peer educators (males = 23.8 percent). The schools were selected using a self-selection process and the classes were selected by the school staff. Self-report questionnaires were administered in class during school time by trained researchers (in December 2010 and May 2011). The consent procedure for research consisted of formal approval by the schools and consent by the parents. Participants who dropped out of the study represented 12 percent of the sample (N=55) and they did not differ from those who remained with regard to the initial study variables.

2.2. Measure

The same bullying and victimization scales and cyberbullying scales used in the first study were administered. Reliability coefficients at T1 and T2 were .75 and .82 for bullying, .74 and .71 for victimization, .79 and .82 for cyberbullying, and .80 and .87 for cybervictimization.

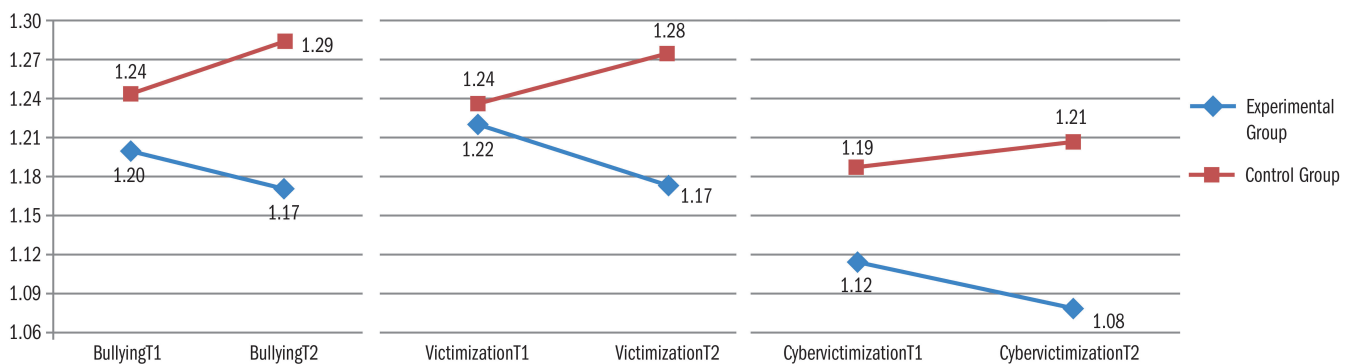
2.3. Results

2.3.1. Experimental vs Control group

Bullying and victimization: Repeated measures ANOVAs were conducted to evaluate the change in bullying and victimization over time in the two groups. Results showed no significant effect of time for both outcome measures but a significant interaction of time*group for bullying ($F_{(2, 375)} = 5.993$; $p < .05$; $\eta^2_p = .016$) and for victimization ($F_{(2, 375)} = 11.848$; $p < .01$; $\eta^2_p = .031$) (see Figure 3). For both dimensions, the experimental group showed a decrease across time as compared to the control group.

Cyberbullying and cybervictimization: Repeated measures ANOVAs were conducted on both variables. The results show a non-significant effect of time for both cyberbullying and cybervictimization, and a non-significant interaction of time*group for cyberbullying. For cybervictimization, a significant interaction of time*group was found ($F_{(2, 375)} = 5.706$; $p < .05$; $\eta^2_p = .015$), showing a decrease over time in the experimental group as compared to the control group (see Figure 3).

Figure 3: Change in bullying, victimization, and cybervictimization over time in control group and experimental group

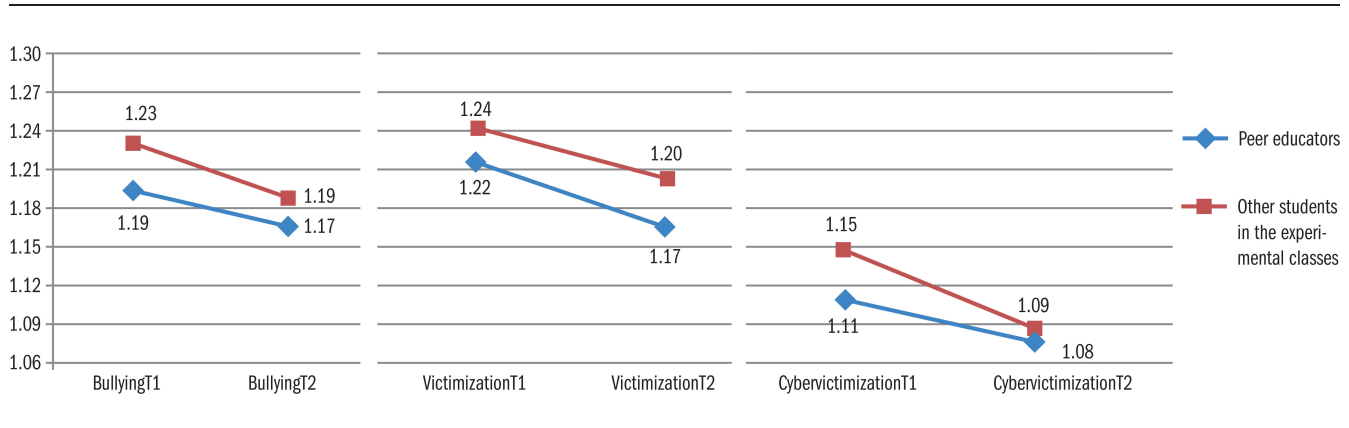


2.3.2 Peer Educators vs Other Students in Experimental Classes

Repeated measures ANOVAs were conducted to evaluate the effect of time on the dependent variable across the two groups (peer educators and other students in experimental classes). The results show: 1) for bullying an almost significant effect of time ($F_{(2, 231)} = 3.453$; $p = .06$) and a non-significant interaction of time*group; 2) for victimization a significant effect of time ($F_{(2, 231)} = 4.178$;

$p < .05$; $\eta^2_p = .018$) and a non-significant interaction of time*group; 3) for cybervictimization a significant effect of time ($F_{(2, 231)} = 8.919$; $p < .01$; $\eta^2_p = .037$) and a non significant interaction of time*group. These results show that the decrease across time in bullying, victimization, and cybervictimization was the same for both peer educators and the other students in experimental classes (see Figure 4).

Figure 4: Change in bullying, victimization, and cybervictimization over time for peer educators and other students in experimental classes



3. General Discussion

In the second study, the results show a significant pattern of decrease in bullying, victimization, and cybervictimization among peer educators and the other students in the experimental classes, as compared with the control group. Particularly, they highlight that Phase Two of Noncadiamointrappola is an effective approach to prevent and reduce bullying and cyberbullying among adolescents. It showed positive effects on the students involved (albeit the effect size is not very large), reducing bullying and cyberbullying in the whole class and not simply among peer educators. We can hypothesize that in this second study the peer educators had the capacity to act as agents of change, promoting a reduction of bullying and cyberbullying in the whole class.

The main effects applied to both victimization and cybervictimization, showing that greater attention to this side of the problem can help reduce the percentage of students victimized in the class. These approaches seem able

to work directly on peer educators and indirectly on the whole group, through awareness processes and group dynamics.

It appears that the underlying mechanisms behind these positive results are the new elements introduced in Phase Two, particularly deeper involvement of school teachers and of the whole class, and greater attention to victim support. Overall we obtained greater involvement by the majority of students by providing more intervention opportunities in class and online (forum and Facebook interactions). These results suggest that within a peer-led model the type of roles peer educators take on is highly relevant. If they start a process of personal change but are unable to involve the other students in this process, this approach can have limited effects (see Study 1). But if they are supported in their capacity to promote initiatives and active participation by other students, the process of change can involve the entire class. In this regard, a class approach and the involvement of class teachers as practiced

in the Italian model can be more promising than a school approach (Cowie and Wallace 2000).

Although these models highlight the importance of students' active involvement, it is crucial to promote adult in-

volvement and supervision in order to create space and time for student intervention. Finally, consideration should be devoted to cost-benefit evaluation of the peer-led model, given that this model usually has a low cost and can be highly profitable for schools and community.

References

- Arseneault, Louise, Lucy Bowes, and Sania Shakoor. 2010. Bullying Victimization in Youths and Mental Health Problems: "Much Ado About Nothing"? *Psychological Medicine* 40:717–29.
- Cowie, Helen, Paul Naylor, Lorenzo Talamelli, Peter K. Smith, and Preeti Chauhan. 2002. Knowledge, Use of and Attitudes towards Peer Support: A Two Year Follow-up to the Prince's Trust Survey. *Journal of Adolescence* 25:453–67.
- Cowie, Helen, and Patti Wallace. 2000. *Peer Support in Action: From Bystanding to Standing By*. London: Sage.
- Flashpohler, Paul D., Jennifer I. Elfstrom, Karin L. Vanderzee, Holli E. Sink, and Zachary Birchmeier. 2009. Stand by Me: The Effects of Peer and Teacher Support in Mitigating the Impact of Bullying on Quality of Life. *Psychology in the Schools* 46:636–49.
- Gini, Gianluca, Tiziana Pozzoli, Francesco Borghi, and Lara Franzoni. 2008. The Role of Bystanders in Students' Perception of Bullying and Sense of Safety. *Journal of School Psychology* 46:617–38.
- Holt, Melissa K., and Dorothy L. Espelage. 2007. Perceived Social Support among Bullies, Victims, and Bully-Victims. *Journal of Youth and Adolescence* 36:984–94.
- Houlston, Catherine, Peter K. Smith, and John Jessel. 2011. The Relationship between Use of School-Based Peer Support Initiatives and the Social and Emotional Well-Being of Bullied and Non-Bullied Students. *Children and Society* 25:293–305.
- Maticka-Tyndale, Eleanor, and Jessica Barnett Penwell. 2010. Peer-Led Interventions to Reduce HIV Risk of Youth: A Review. *Evaluation and Program Planning* 33:98–112.
- Menesini, Ersilia, Pamela Calussi, and Annalaura Nocentini. 2012. Cyberbullying and Traditional Bullying: Unique, Additive and Synergistic Effects on Psychological Health Symptoms. In *Cyberbullying in the Global Playground: Research from International Perspectives*, ed. Qing Li, Donna Cross, and Peter K. Smith, 245–262. Oxford: Wiley-Blackwell.
- Menesini, Ersilia, Elena Codecasa, Beatrice Benelli, and Helen Cowie. 2003. Enhancing Children's Responsibility to Take Action against Bullying: Evaluation of a Befriending Intervention in Italian Middle Schools. *Aggressive Behavior* 29:1–14.
- Menesini, Ersilia, Marco Modena, and Franca Tani. 2009. Bullying and Victimization in Adolescence: Concurrent and Stable Roles and Psychological Health Symptoms. *Journal of Genetic Psychology* 170:115–33.
- Menesini, Ersilia, Annalaura Nocentini, and Pamela Calussi. 2011. The Measurement of Cyberbullying: Dimensional Structure and Relative Item Severity and Discrimination. *Cyberpsychology, Behavior, and Social Networking* 14:267–74.
- Naylor, Paul, and Helen Cowie. 1999. The Effectiveness of Peer Support Systems in Challenging School Bullying: The Perspectives and Experiences of Teachers and Pupils. *Journal of Adolescence* 22:467–80.
- Pellegrini, Anthony D. 2002. Bullying, Victimization, and Sexual Harassment During the Transition to Middle School. *Educational Psychologist* 37:151–63.
- Salmivalli, Christina. 2010. Bullying and the Peer Group: A Review. *Aggression and Violent Behavior* 15:112–20.
- Salmivalli, Christina, and Kätlin Peets. 2008. Bullies, Victims, and Bully-Victim Relationships. In *Handbook of Peer Interactions, Relationships, and Groups*, ed. Ken Rubin, William Bukowski, and Brett Laursen, 322–40. New York: Guilford.
- Shiner, Michael. 1999. Defining Peer Education. *Journal Of Adolescence* 22:555–66.
- Ttofi, Maria M., and David P. Farrington. 2009. What Works in Preventing Bullying: Effective Elements of Anti-Bullying Programmes. *Journal of Conflict and Peace Research* 1:13–24.
- Ttofi, Maria M., and David P. Farrington. 2011. Effectiveness of School-Based Programs to Reduce Bullying: A Systematic and Meta-Analytic Review. *Journal Of Experimental Criminology* 7:27–56.
- Ttofi, Maria M., David P. Farrington, Friedrich Lösel, and Rolf Loeber. 2011. The Predictive Efficiency of School Bullying versus Later Offending: A Systematic/Meta-Analytic Review of Longitudinal Studies. *Criminal Behavior and Mental Health* 21:80–89.
- Turner, Glen, and Jonathan Shepherd. 1999. A Method in Search of a Theory: Peer Education and Health Promotion. *Health Education Research* 14:235–47.
- Veenstra, René, Siegwart Lindenberg, Anke Munniksma, and Jan K. Dijkstra. 2010. The Complex Relation between Bullying, Victimization, Acceptance, and Rejection: Giving Special Attention to Status, Affection, and Sex Differences. *Child Development* 81:480–86.

Ersilia Menesini

ersilia.menesini@unifi.it

Annalaura Nocentini

annalaura.nocentini@virgilio.it

Benedetta Emanuela Palladino

benedetta_palladino@yahoo.it

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing

Rezarta Bilali, Department of Conflict Resolution, Human Security, and Global Governance, University of Massachusetts Boston, United States

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Open Section

► **Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing** Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Ilwilade (pp. 338 – 350)



Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing

Rezarta Bilali, Department of Conflict Resolution, Human Security, and Global Governance, University of Massachusetts Boston, United States

Two U.S. studies report a differential effect of identity centrality and in-group superiority on reactions to in-group victimization and in-group harm-doing. Study 1 (N = 80) found that higher identity centrality predicted less justification for freely-recalled in-group victim events, whereas higher in-group superiority predicted more justification for freely-recalled in-group harm-doing events. Study 2 (N = 105) reexamined these findings in specific contexts of historical victimization (Pearl Harbor) and harm-doing (Hiroshima and Nagasaki), finding that in-group superiority was a predictor of reactions to historical in-group harm-doing (justification, emotional reactions, importance of events), whereas centrality was a predictor of reactions to historical in-group victimization.

Historical memory serves important functions for groups. History constitutes an important part of a group's self-image and serves to establish a sense of common fate in group members (Billig 1995). Group members rely on history to enhance group identity (Volpato and Licata 2010). History also informs a group's understanding of the present and shapes expectations for the future (Bilali and Ross 2012). For instance, historical memories play an important role in maintaining and exacerbating intergroup conflict (Devine-Wright 2003). They serve to justify outbreaks of violence and delegitimize the opponent (Bar-Tal 2003).

In recent years, a surge of social psychological research has investigated people's reactions to their groups' or nations' troubled histories (e.g., Branscombe and Doosje 2004; Doosje et al. 1998; also see the *International Journal of Conflict and Violence* focus section on collective memories, Volpato and Licata 2010). Generally this research has examined the consequences of historical memories, such as reactions to descriptions of historical events of an in-group's harm doing (e.g., Cehajic, Brown, and González 2009; Doosje et al. 1998; Zebel, Doosje, and Spears 2004; Iyer, Leach and Crosby, 2003), or the impact of reminders of past victimization on emotional reactions (Frijda 1997; Liu and László 2007) and on attitudes toward current conflicts (e.g., Wohl

and Branscombe 2008). The factors that shape individuals' historical memories are not as well understood, however.

People learn about the in-group's history through media, education, leaders, public images and symbols, and conversations with family and peers. Although historical memories are often shared within a group (Bar-Tal 2003), members exhibit clear differences in their endorsement of historical memories. Therefore, individual-level factors might also influence historical memories. Whereas some research has investigated factors that influence historical memories of the in-group's harm-doing (e.g., in-group identification, Bilali, Tropp, and Dasgupta 2012; Sahdra and Ross 2007; right wing authoritarianism, Sibley, Wilson, and Robertson, 2007), our knowledge of what shapes memories of past victimization is limited. The research described here investigates the factors that shape historical memories of past in-group victimization and harm-doing events. Building on previous research, I adopt social identity theory as the guiding framework in this research.

1. In-Group Identification and Historical Memories

According to social identity theory (Tajfel and Turner 1986), people derive their self-concepts, in part, from their membership in social groups. Because individuals are

motivated to view themselves positively, they are also motivated to view their groups favorably. History constitutes an important part of a group's image. The drive to maintain a positive self-image should encourage in-group-serving attributions in recollections of the group's past (Doosje and Branscombe 2003). For instance, Baumeister and Hastings (1997) observed that distortions of past events that portray the in-group positively are more frequent than distortions that portray the in-group negatively.

Similar patterns have emerged both in laboratory and field settings. For example, in a lab study, participants categorized into arbitrary groups expected in-group members to engage in more positive behaviors than out-group members and subsequently recalled more positive behaviors and fewer negative behaviors committed by members of their in-group than by out-group members (Howard and Rothbart 1980). In addition, because group members are not equally attached to their group, they are not equally motivated to protect the in-group's positive image. Individuals to whom group identity is important ought to be more motivated to maintain a positive image of their group, which in turn should lead to endorsing more in-group-favorable memories of the past (Sahdra and Ross 2007). For instance, Sahdra and Ross (2007) found that the more participants identified with their group (either Sikh or Hindu), the less they recalled events in which the in-group was a perpetrator. However, they did not find a relationship between in-group identification and recollections of in-group victim events. While the motivation to maintain a positive self-image helps to illuminate biases in historical memories of in-group's misdeeds, it does not explain individual differences in historical memories of in-group victimization. Is in-group identification then irrelevant to remembering an in-group's past victimization? I suggest that historical memories surrounding in-group victimization and in-group harm-doing are likely to be linked to distinct dimensions of in-group identification.

1.1. Dimensionality of Identification with Ethnic or National Group

There are two main approaches in the study of ethnic and national identification. The first draws from social identity theory and conceptualizes national identification as yet another form of in-group identification. Within this tradi-

tion, in-group identification has been usually treated as a unidimensional construct; however, a growing literature (e.g., Cameron and Lalonde 2001; Cameron 2004; Ellemers, Kortekaas, and Ouwerkerk 1999; Jackson 2002; Leach et al. 2008) shows that a multidimensional conceptualization is more appropriate. While there are disagreements about the number and the nature of the dimensions (see Leach et al. 2008), Tajfel's original conceptualization of social identity (1978) included evaluative, cognitive, and affective components.

The second approach conceptualizes national identity as a specific form of attachment to the group expressed either as nationalism or patriotism. Patriotism is perceived as a healthy national self-concept, and as positive love of one's own country (Bar-Tal 1993; Kosterman and Feshbach 1989) independent of out-group derogation (Brewer 1999). In contrast, nationalism is related to intergroup differentiation. The most important underlying dimension of nationalism is the view that one's own group is superior to other groups (Kosterman and Feshbach 1989).

In-group superiority. Drawing from both literatures, Roccas, Klar, and Livitan (2006) distinguish between two dimensions of identification: glorification and attachment. Glorification refers to beliefs in in-group superiority and deference to group norms and symbols (related to nationalism), whereas attachment refers to cognitive and emotional attachment to the in-group, such as self-definition as a group member or commitment to the group. The glorification dimension possesses a strong evaluative component. The evaluative component (i.e., evaluating the in-group as positive or negative) has been the most common way of thinking about in-group identification (Leach et al. 2008; for a review see Ashmore, Deaux, and McLaughlin-Volpe 2004), and has driven predictions regarding the effects of the strength of in-group identification. That is, the motivation to view one's group in a positive light may drive distortions and legitimizations of past events where the in-group was the perpetrator. Roccas and colleagues (2006; see also Leidner et al. 2010) used this rationale to suggest that the glorification dimension (i.e., positive evaluation dimension), rather than group attachment, drives denial of in-group responsibility for harm-

doing and legitimization of past in-group harms. Recently, Roccas, Sagiv, Schwartz, Halevy, and Eidelson (2008) further distinguish two dimensions of in-group glorification: deference and superiority. Of these two dimensions, only the latter constitutes an evaluative component. One way in which group members can maintain an in-group's favorable image is to view the in-group as better than other groups. This is particularly relevant in conflict contexts due to the comparative and competitive nature of intergroup conflict. Based on this discussion, it is in-group superiority that drives favorable in-group interpretations of an in-group's harm doing.

Identity centrality. Roccas and colleagues (2008) also distinguish between two components related to attachment: commitment and identity centrality (which they refer to as importance). I will specifically focus on identity centrality as an important component capturing the cognitive and affective aspects of in-group identification. Identity centrality is defined as the degree of importance and chronic salience of a group membership to a person's self-concept (Ashmore et al. 2004; Leach et al. 2008; Luhtanen and Crocker 1992). Identity centrality has been linked to increased perceptions of threat toward the in-group (Leach et al. 2008; Sellers and Shelton 2003). Indeed, various studies suggest a link between identity centrality and perceived in-group victimization (e.g., in-group discrimination). In one direction, discrimination experiences or prejudice toward the in-group might strengthen the importance of that group membership (e.g., Branscombe, Schmitt, and Harvey 1999; Jetten et al. 2001). In the other direction, high centrality of the group membership to a person's self-concept might intensify the sense of in-group discrimination (e.g., Major, Quinton, and Schmader 2003). In support of the latter directional link, a longitudinal study with Latino and White college students on a university campus in the United States showed that ethnic identification at Time 1 predicted heightened perceptions of ethnic victimization three years later, whereas the reverse link from a sense of victimization to ethnic identification was not significant (Thomsen et al. 2010). In that study, the ethnic identification measure was closely related to identity centrality: The three-item ethnic identification scale included two items tapping the identity centrality dimension, whereas one

item assessed in-group ties (i.e., how close respondents felt to other members of their ethnic group). Overall, these studies support the idea that the centrality dimension of in-group identification is closely related to perceptions of in-group victimization.

1.2. Identity Centrality, In-Group Superiority and Historical Memories

The conceptual distinctions between in-group superiority and identity centrality may lead to differences regarding how each dimension relates to memories and interpretations of past intergroup conflict. Recent literature suggests that harm-doing events pose a threat to a group's morality (e.g., Doosje et al. 1998; Wohl, Branscombe, and Klar 2006). Shnabel and Nadler (2008) argue that the conflictual past poses different concerns for victims and perpetrators. Whereas perpetrators are concerned with restoring the morality of their group, victims are motivated to restore lost power. These perspectives suggest that different types of events (i.e., in-group harm-doing vs. in-group victimization) raise different concerns for group members. For instance, terrorist attacks might threaten American identity and the well-being of American people, but not necessarily the evaluation of American identity as positive or negative. However, harm-doing events (e.g., the Abu Ghraib events) typically threaten the in-group's positive image. Furthermore, different ways of relating to the in-group might elicit sensitivity to different types of group threats. The desire to maintain a positive and moral self-image might prompt members of groups that have perpetrated harm to downplay in-group's negative events, minimize the negative consequences of these events, or shift the focus to mitigating conditions that serve to displace in-group's responsibility. Building on previous research (e.g., Roccas et al. 2006; Leidner et al. 2010), I suggest that in-group superiority should further strengthen these effects. In turn, identity centrality might elicit responses when an in-group's wellbeing or in-group identity (not its positive evaluation) is directly threatened. For instance, Baumeister and colleagues (e.g., Baumeister, Stillwell, and Wotman 1990; Baumeister and Catanese 2001) observe that victims' accounts of interpersonal transgressions emphasized the negative and lasting consequences of harm and perpetrators' responsibility for the acts. I predicted that in an intergroup context identity centrality

should strengthen these effects. The more central group identity is to self-concept, the more the members of victim groups will emphasize in-group victimization – the negative consequences of the harm and perpetrators' responsibility for the acts.

2. Study 1: U.S. as Victim or Perpetrator

The aim of Study 1 was to provide an initial test of these ideas. For this purpose I adopted the free recall task used by Sahdra and Ross (2007, study 1). I hypothesized that identity centrality and in-group superiority would differentially predict the number of freely-recalled victim and perpetrator events, as well as the degree to which these violent acts are perceived as justifiable. Based on this, I derived two hypotheses:

Hypothesis 1. In-group superiority should predict remembering of an in-group's past misdeeds, such as recollections of fewer events in which the in-group was a perpetrator (H1a), and increased justification for in-group perpetrator events (H1b).

Hypothesis 2. Identity centrality should predict remembering of an in-group's past victimization, such as recall of more events in which the in-group was a victim (H2a), and less justification for in-group victim events (H2b).

2.1. Methods

2.1.1. Participants and Procedures

Eighty participants (61 women, 19 men) were recruited at a university in the Northeastern United States. Participants were told that they were participating in a study examining their opinions on important events in U.S. history. First, participants were asked to complete measures of in-group (i.e., American) identification, then they were asked to freely recall and rate the justifiability of past events in which the U.S. was either a victim or a perpetrator.

2.1.2. Measures

Identity centrality and in-group superiority. Identity centrality was measured by the following three items: "Being an American is an important part of how I see myself," "I often think about the fact that I am American" (adapted from Cameron 2004; Leach et al. 2008), and "Being an American

is an important part of my self-image" (adapted from Cameron 2004). All items were measured using six-point scales ranging from 1 (strongly disagree) to 6 (strongly agree). The three items were aggregated to form a measure of identity centrality ($\alpha = .87$) ($M = 3.59$, $SD = 1.17$).

Two items that tapped the superiority dimension (rather than deference dimension) in Roccas et al.'s glorification scale (2006) were used to measure in-group superiority: "Relative to other nations, the U.S. is a very moral nation" and "The U.S. is better than all other nations in all respects." The two items were averaged to form a composite measure ($\alpha = .66$) of in-group superiority ($M = 2.85$, $SD = 1.09$).

The correlation between identity centrality and in-group superiority was moderate ($r = .40$) suggesting that the two identity dimensions are distinct and could be entered as simultaneous predictors in data analyses without raising multicollinearity concerns.

Free recall task. Participants were asked to think about the recent history of the United States (past one hundred years) and its role in international arena. Then, they were asked to recall up to six events in which the United States was a perpetrator of violent episodes committed toward another country or group, and up to six events in which the United States was a victim of violent episodes committed by other countries or groups.

Justification. For each recalled event, participants were asked to rate the extent to which they believed the events were justifiable in a scale ranging from 1 (not justified) to 6 (very justified). Justification ratings were averaged separately across perpetrator and victim events, forming one composite score for the degree of justification of victim events and one score for the degree of justification of perpetrator events.

2.1.3. Data Analysis

Repeated measures general linear models (GLM) were conducted with identity centrality and in-group superiority as continuous predictors. Type of event (perpetrator vs. victim) was the within-subject factor. The dependent variables were (1) the number of recalled events, and (2) justifi-

cation of the events. The simultaneous inclusion of identity centrality and in-group superiority as predictors made it possible to assess the unique effects of each identity dimension controlling for the other dimension. The interaction between in-group superiority and identity centrality was also tested in initial analyses. Because there were no significant interaction effects, the interaction term was excluded from the final analyses.

2.2. Results

2.2.1. Number of Recalled Events

The recalled events included the atomic bombings of Hiroshima and Nagasaki, slavery, abuse at Abu Ghraib and Guantanamo Bay, the recent wars in Iraq and Afghanistan as perpetrator events, and terrorist attacks on U.S. targets as victim events including the Pearl Harbor attack, the 9/11, the Iranian hostage crisis, and the USS Cole bombing. Some events, such as the Vietnam War, were included among perpetrator events by some participants, but among victim events by others.

The GLM yielded no difference between the number of perpetrator events and the number of victim events recalled, $F(1, 77) = 2.01, p = .16, \eta^2 = .025$. For each type of event, participants wrote down, on average, slightly more than two events (for means and standard deviations of all items see Table 1). The results revealed a superiority x type interaction, $F(1, 77) = 7.46, p < .01, \eta^2 = .09$. As expected

(H1a), higher in-group superiority marginally predicted recall of fewer events in which the in-group was a perpetrator, but in-group superiority did not predict the number of recalled events in which the in-group was a victim.

Contrary to the predictions (H2a), centrality was not associated with the number of recalled events, $F(1, 77) = 1.53, p = .22, \eta^2 = .02$, and its interaction with type of event was not significant, $F(1, 77) = 1.08, p = .30, \eta^2 = .01$. All standardized regression coefficients are shown in Table 1.

2.2.2. Justification

The GLM yielded a main effect of type of event, $F(1, 59) = 19.53, p < .001, \eta^2 = .25$, indicating that in-group victim events were perceived to be less justifiable than in-group harm-doing (perpetrator) events (see Table 1). As expected, the effect of type was qualified by two-way interactions with centrality, $F(1, 59) = 18.46, p < .001, \eta^2 = .24$, and Superiority, $F(1, 59) = 11.18, p = .001, \eta^2 = .16$. Supporting both hypotheses (H1b and H2b), higher identity centrality predicted less justification for events in which the in-group was a victim, whereas higher in-group superiority predicted more justification for events in which the in-group was a perpetrator (see Table 1). Identity centrality did not predict justification for perpetrator events, and superiority did not predict justification for victim events.

Table 1: Regressions predicting construals of historical events from identity centrality and in-group superiority (means, standard deviations, and standardized beta coefficients) (Study 1)

	M	SD	Centrality		Superiority	
			β	p	β	p
Victim events						
Number of recalled events	2.12	1.11	.14	.31	-.13	.36
Justification	1.89	.99	-.34	.01	-.16	.23
Perpetrator events						
Number of recalled events	2.25	1.73	.18	.29	-.24	.08
Justification	2.91	1.31	.14	.32	.36	.009

2.3. Discussion

This study lends support to the hypothesis that identity centrality and in-group superiority are differentially associated with justification of different types of events. As expected, the results revealed that higher identity centrality predicted less justification for events in which the in-group was a victim, whereas higher superiority predicted more justification for events in which the in-group was a perpetrator of attacks. In addition, the superiority dimension was a marginal predictor for the number of recalled events in which the in-group was a perpetrator. Interestingly, there was no difference between the number of perpetrator and victim events recalled, and contrary to predictions, identity centrality was not associated with the number of recalled victim events.

Unlike Sahdra and Ross's study (2007), in which participants were asked to recall violent events that occurred within a particular intergroup conflict context, participants in the present study were asked to recall a variety of perpetrator and victim events in U.S. history. As a consequence, participants in this study reported major historical events such as the Vietnam War, Pearl Harbor, atomic bombings of Hiroshima and Nagasaki, the 9/11 attacks, or the war in Iraq. Due to the prominence of these events in the history of the United States, it is possible that individual differences in dimensions of in-group identification (e.g., identity centrality) might influence construals of major events without impacting their recollection. This might explain the lack of relationship between the two in-group identification dimensions and the number of recalled events. Therefore, in Study 2, I extended the investigation by examining group members' construals of a specific in-group victim event (Pearl Harbor) and a specific in-group perpetrator event (Hiroshima and Nagasaki), which were mentioned by the majority of participants in Study 1.

3. Study 2: Pearl Harbor and Hiroshima

Study 2 aimed to replicate and extend the findings of Study 1. In Study 2, instead of asking participants to freely recall historical events, I assessed their construals of two important events in U.S. history: the atomic bombings of Hiroshima and Nagasaki (a harm-doing event) and the Pearl Harbor attack (a victimization event). There were two main

reasons for the choice of these two events. First, a majority of respondents in Study 1 freely recalled Pearl Harbor and Hiroshima/Nagasaki bombings, suggesting that American college students are familiar with these two events. Second, as Study 2 aimed to compare construals of a victim versus a perpetrator event, it was important to consider two historical events that differed mainly in one dimension (i.e., victim versus perpetrator), but were similar in other important dimensions that might drive differential construals. For instance, the events have similar temporal distance, and the out-group (i.e., the Japanese) is the same in both incidents. Therefore, any differences in participants' reactions to these events will be due to the type of event (i.e., victim versus perpetrator) rather than to different temporal distance or characteristics of the out-group.

One important weakness of Study 1 was the single-item justification measure. To better assess the construals of these events, and to complement the justification measure used in Study 1, in study 2 I also examined exonerating cognitions. In addition, I assessed emotional reactions (anger and sympathy) and the perceived importance of each historical event (personal importance and importance in U.S. history).

Similar to the predictions in Study 1, I expected that identity centrality would predict construals of in-group victim event, whereas in-group superiority would predict construals of in-group perpetrator event.

Hypothesis 1. In-group superiority should predict favorable in-group construals of the atomic bombings of Hiroshima and Nagasaki, but should not be associated with construals of the Pearl Harbor attack. Specifically, higher in-group superiority would predict more justification of the atomic bombings (H1a), less anger toward the United States, and less sympathy for the Japanese (H1b). Higher in-group superiority should also be associated with evaluating the atomic bombings as less important in U.S. history (H1c).

Hypothesis 2. In contrast, identity centrality should predict construals of the Pearl Harbor attack, but not of the atomic bombings. Those participants to whom identity is more central to their self-concept should view Pearl Harbor to be

less justifiable (H2a). They would also report more anger toward Japan, more sympathy for American victims (H2b), and would view Pearl Harbor as more important in U.S. history (H2c).

3.1. Methods

3.1.1. Participants

Participants were 105 undergraduates (86 women, 19 men) at a university in the Northeastern United States (mean age = 20.64, $SD = 3.17$). Participants were awarded research credits for their participation.

Sixteen participants reported coming from a working class or a lower-middle-class family, forty-eight from a middle-class family, and forty-one from an upper-middle-class family. Eighty-six participants identified themselves as White, five as Asian, three as Black, six as Hispanic, and the rest identified with an ethnic group not identified in the questionnaire.

One item asked participants to categorize themselves politically on a six-point scale from -3 (very liberal) to +3 (very conservative). The mean self-reported political orientation within the sample was slightly liberal ($M = -1.04$, $SD = 1.64$). Another item assessed participants' interest in history: "Generally speaking how much interest would you say you have in history?" (-3 = very uninterested; +3 = very interested). Participants were mildly interested in history ($M = .45$, $SD = 1.91$).

3.1.2. Procedures

Participants in the psychology department's subject pool were invited to participate in a survey research on "public opinions about contemporary and historical events in the United States." The first set of items in the questionnaire assessed identity centrality and in-group superiority (with regard to American identity). The second set assessed reactions toward two historical events: the Pearl Harbor attack and the atomic bombings of Hiroshima and Nagasaki. To control for order effects, half of the sample first completed the survey section on Pearl Harbor, whereas the other half first completed the Hiroshima and Nagasaki section. At the end, participants completed demographic items. Participants received research credits for their participation.

3.1.3. Measures

Identity centrality and in-group superiority. As in Study 1, three items were used to assess identity centrality. However, in order to increase the validity of the scale, the item "I often think about the fact that I am an American" (referring to identity salience, see Sellers et al. 1998) was replaced with another item specifically tapping the identity centrality construct: "The fact that I am American is an important part of my identity" (adapted from Leach et al. 2008).

The same items as in Study 1 were used to assess in-group superiority, with the addition of one more item to improve the scale's reliability: "My nation is superior to other nations in most respects." All items were measured on six-point scales (1 = strongly disagree; 6 = strongly agree). Each three-item scale revealed good reliability ($\alpha = .90$ for identity centrality; and $\alpha = .74$ for in-group superiority).

The mean identity centrality ($M = 3.98$, $SD = 1.23$) and in-group superiority ($M = 3.24$, $SD = 1.23$) were both slightly above the respective scale's mid-point. The correlation between identity centrality and in-group superiority was moderate ($r = .46$).

Justification. The same item as in Study 1 was used to assess the degree to which participants perceived the violent events to be justifiable.

Exonerating cognitions. Three items were constructed to assess the degree to which participants used exonerating cognitions to legitimize the attacks: "Considering the conditions of the World War, the Americans [Japanese] attacked Hiroshima and Nagasaki [Pearl Harbor] because they did not have any other choice of action," "The American [Japanese] attacks on Hiroshima and Nagasaki [Pearl Harbor] were intended to save American [Japanese] lives," and "The attacks on Hiroshima and Nagasaki [Pearl Harbor] can be considered to be a patriotic act of the Americans [Japanese] to save their country and their people." The three-item scales revealed good reliabilities ((= .75 for Pearl Harbor; (= .72 for atomic bombings).

Perceived importance of the event. Participants were asked to rate the importance of each event in U.S. history as

well as the personal importance of the events on a six-point scale (1 = not at all important; 6 = very important). The items read: “How important is Pearl Harbor attack [Hiroshima and Nagasaki bombings] in the United States history?” and “How important is Pearl Harbor attack [Hiroshima and Nagasaki bombings] to you personally?”

Emotional reactions. For each event, participants were asked to rate the degree to which they felt each of the following emotions when thinking about the historical event: (1) “I feel anger toward Japan [the U.S.],” (2) “I feel sympathy toward the victims of the attacks,” and (3) “I feel sympathy toward the Japanese [Americans] in general.” All items were measured on six-point scales ranging from 1 (strongly disagree) to 6 (strongly agree).

3.1.4. Data Analysis

As in Study 1, repeated measures General Linear Models (GLM) were used to analyze the data. Type of event (perpetrator vs. victim) was entered as the within-subject factor, whereas identity centrality and in-group superiority

were entered as continuous predictors. Target group (in-group vs. out-group) was included as an additional within-subject factor in the analyses of emotional reactions.

Similar to the procedure in Study 1, the interaction between identity centrality and in-group superiority was entered in initial models, but was later dropped as there were no significant interaction effects. The order in which the events were presented might also influence the construals of these historical events, particularly because Pearl Harbor and the atomic bombings of Hiroshima and Nagasaki are perceived to be causally linked (i.e., Pearl Harbor might provide justification for the atomic bombings). Therefore, order was entered as a factor in the initial GLM analyses; however it was excluded from the final reports as there were no significant effects.

Following the GLM analyses, as in Study 1, regression analyses for each type of event were conducted to clarify interaction effects. The standardized beta coefficients from these analyses, as well as the means and standard deviations of all dependent variables are presented in Table 2.

Table 2: Regressions predicting construals of historical events from identity centrality and in-group superiority (means, standard deviations, and standardized beta coefficients) (Study 2)

			Centrality		Superiority				Centrality		Superiority	
	M	SD	β	p	β	p	M	SD	β	p	β	p
	Pearl Harbor (victim event)						Atomic bombings (perpetrator event)					
Justification	2.6	1.28	-.22	.05	.13	.26	3.11	1.33	.07	.50	.26	.02
Exonerating cognitions	3.01	1.13	-.13	.27	.11	.33	2.91	1.17	-.05	.67	.46	.000
Anger toward:												
the United States	2.05	1.28	.01	.96	-.15	.19	2.75	1.43	.03	.80	-.30	.007
Japan	3.57	1.53	.44	.000	-.02	.87	2.62	1.31	.18	.11	.06	.59
Sympathy for:												
Victims	5.29	1.06	.34	.002	-.19	.08	4.94	1.20	.14	.17	-.46	.000
Americans	3.93	1.26	.34	.002	.01	.96	3.07	1.42	.17	.12	.02	.84
Japanese	2.79	1.44	-.14	.19	-.20	.06	3.54	1.33	-.02	.82	-.23	.04
Historical importance	5.40	.77	.24	.03	-.12	.28	5.07	1.07	.11	.32	-.28	.01
Personal importance	3.19	1.29	.34	.002	-.07	.53	3.13	1.29	.08	.46	-.15	.17

3.2. Results and Discussion

3.2.1. Justification

The GLM analysis revealed a marginal main effect of type of event, $F(1, 101) = 3.35, p = .07, \eta^2 = .03$, such that the perpetrator event (the atomic bombings) was rated as more justifiable than the victim event (Pearl Harbor) (see Table 2). There was also a marginal interaction between type of event and centrality, $F(1, 101) = 3.39, p = .07, \eta^2 = .03$. Supporting H2a, higher identity centrality predicted less justification for the Pearl Harbor attack, but

did not predict justification for the atomic bombings. Although type x superiority interaction was not significant, $F(1, 101) = .88, p = .35$, higher in-group superiority predicted higher justification for atomic bombings of Hiroshima and Nagasaki, but did not predict justification for Pearl Harbor (supporting H1a; see Table 2). Figures 1 and 2 respectively show the effects of identity centrality and in-group superiority on justification of Pearl Harbor and the atomic bombings. Overall, these results replicated the findings of Study 1.

Figure 1: The effect of identity centrality and in-group superiority on justification of the Pearl Harbor attack

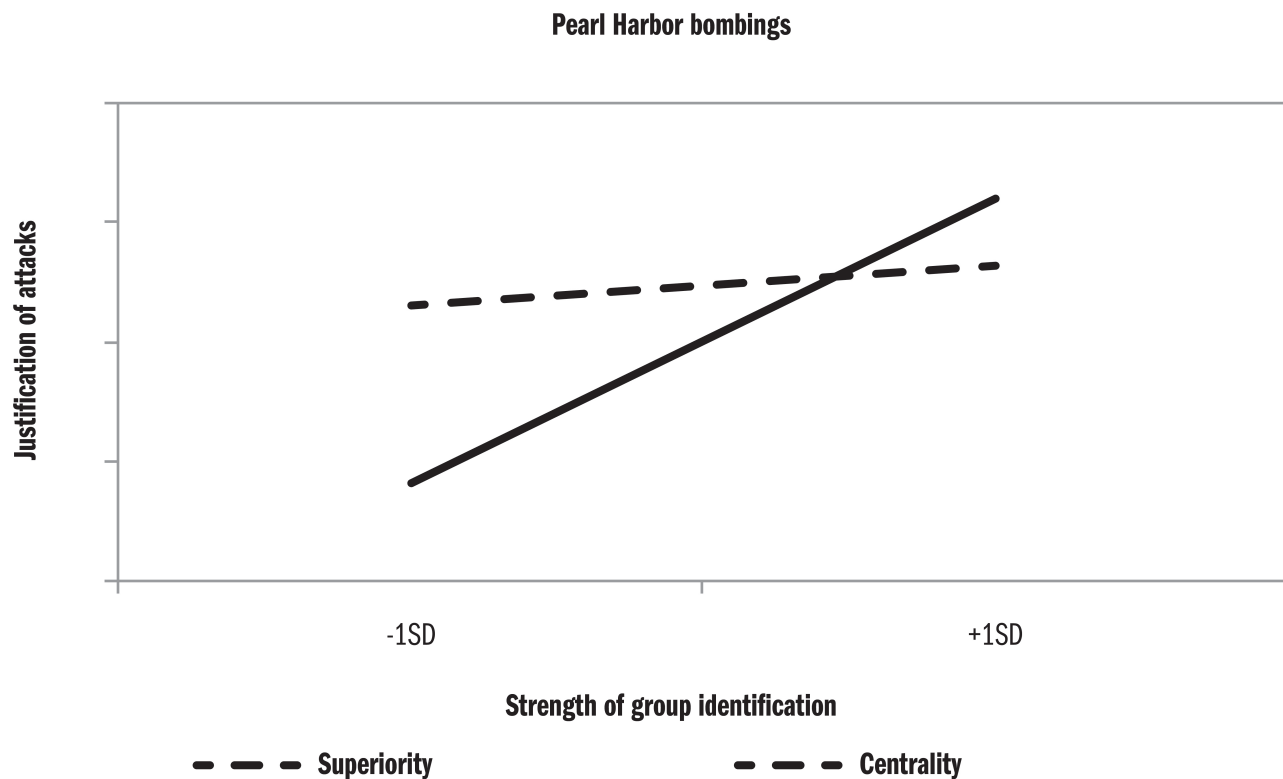
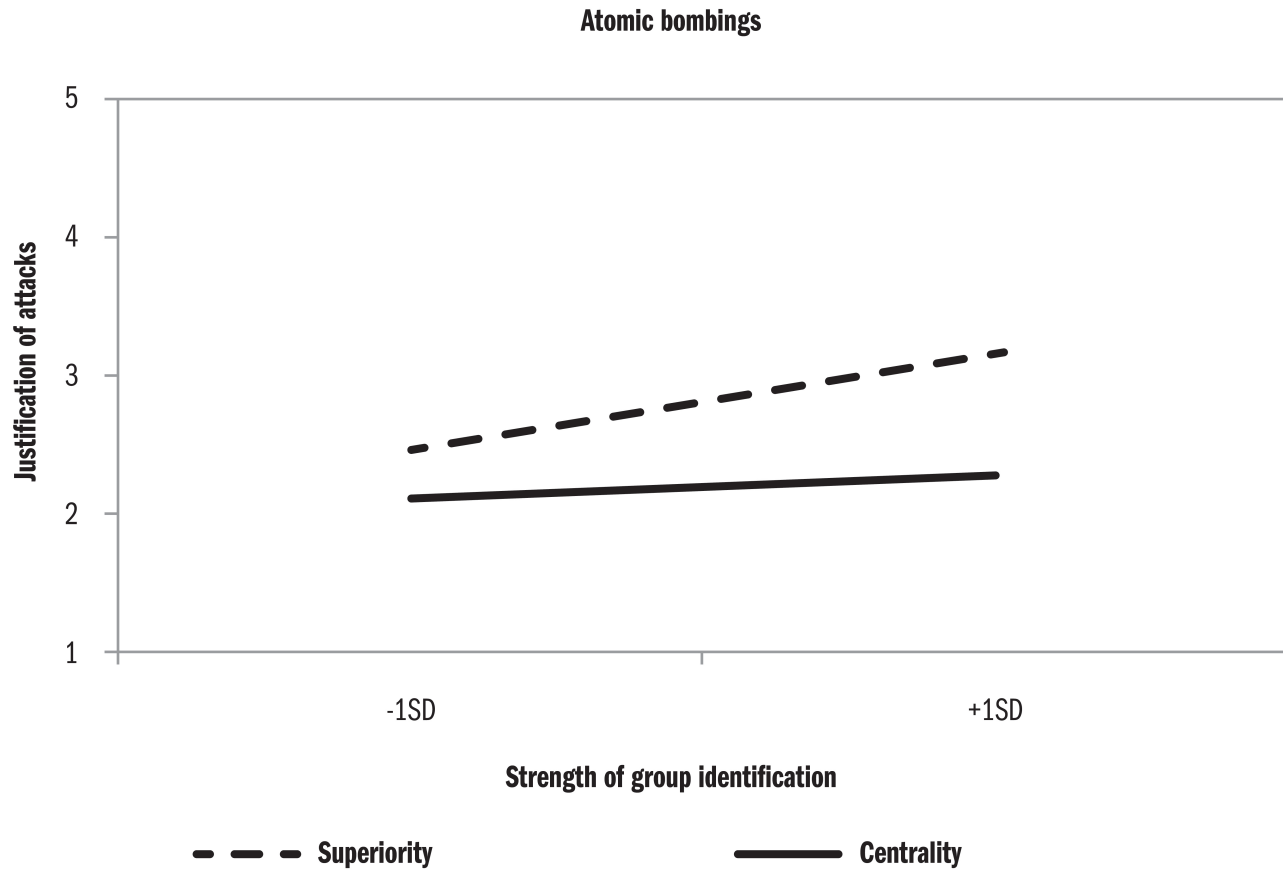


Figure 2: The effect of identity centrality and in-group superiority on justification of the Hiroshima and Nagasaki atomic bombings



3.2.2. Exonerating Cognitions

The GLM analysis yielded a main effect for type of event, $F(1, 90) = 9.17, p < .01$, partial $\eta^2 = .09$, which was further qualified by a type x superiority interaction, $F(1, 90) = 8.9, p < .01, \eta^2 = .09$. As shown in Table 2, the interaction effect revealed that higher in-group superiority predicted higher legitimization of the perpetrator event (atomic bombing), but was not related to legitimization of the victim event (Pearl Harbor).

The interaction between centrality and type was not significant, $F(1, 90) = .24, p = .62$, indicating that centrality was not a significant predictor of legitimization of either victim or perpetrator event. One explanation for the lack of relationship between identity centrality and (lower)

legitimization of the Pearl Harbor attack might be that exonerating cognitions are less relevant to in-group victimization. Exonerating cognitions constitute legitimization mechanisms that are activated when the in-group has committed misdeeds. Thus, there is no reason for group members to use exonerating cognitions in instances of in-group victimization.

3.2.3. Emotional Reactions

Anger. The repeated measures GLM revealed a marginal effect of type, $F(1, 102) = 3.53, p = .06, \eta^2 = .03$, such that less anger was evoked by Pearl Harbor ($M = 1.81, SE = .11$) than by Hiroshima and Nagasaki ($M = 2.67, SE = .10$). The results also yielded a main effect of target, $F(1, 102) = 11.93, p = .001, \eta^2 = .11$, such that participants re-

ported feeling less angry toward the in-group ($M = 2.41$, $SE = .11$) than toward the out-group ($M = 3.10$, $SE = .12$).

More importantly, the results yielded a target x centrality interaction, $F(1, 102) = 9.11$, $p < .01$, $\eta^2 = .08$, which was further qualified by a three-way interaction with type, $F(1, 102) = 4.41$, $p = .04$, $\eta^2 = .04$. Lending support to H2b, centrality predicted more out-group anger in response to Pearl Harbor, but did not predict out-group anger in response to atomic bombings, or anger toward the in-group in either event.

Target also interacted with superiority, $F(1, 102) = 5.52$, $p < .05$, $\eta^2 = .05$, such that higher in-group superiority predicted less anger toward the in-group (i.e., the United States), but did not predict anger toward the out-group (i.e., Japan). Although the three way interaction between target, type and superiority did not reach significance, $F(1, 102) = 2.31$, $p = .13$, $\eta^2 = .02$, superiority was a significant predictor of in-group anger (i.e., less anger toward the United States) for the atomic bombings, but not for Pearl Harbor (supporting H1b; see Table 2).

Not surprisingly, these results suggest that the two dimensions, identity centrality and in-group superiority, predict anger directed toward the harm-doer, but not toward the victim. Furthermore, identity centrality predicts experiencing more anger toward the harm-doer (i.e., the out-group) when the in-group is the victim of violence, whereas in-group superiority predicts less anger toward the in-group when the in-group is the harm-doer.

Sympathy. A 2 x 3 repeated measures GLM was conducted with in-group superiority and identity centrality as continuous predictors. Type of event (victim vs. perpetrator) and target (victims of attacks vs. in-group members vs. out-group members) were the within-subject factors. The dependent variable was the amount of sympathy felt toward each target.¹

The results yielded a main effect of target, $F(2, 204) = 13.19$, $p < .001$, $\eta^2 = .11$, such that participants reported feeling most sympathy toward the victims of attack ($M = 5.11$, $SE = .09$), then toward the in-group ($M = 3.74$, $SE = .11$), and least toward the out-group ($M = 2.93$, $SE = .11$).

In addition, there was a significant type x target interaction, $F(2, 204) = 16.78$, $p < .001$, $\eta^2 = .14$, which was further qualified by three-way interactions with centrality, $F(2, 204) = 8.52$, $p < .001$, $\eta^2 = .08$, and superiority, $F(2, 204) = 5.56$, $p < .01$, $\eta^2 = .05$. A decomposition of the three-way interaction effects lent support to H1b and H2b (see standardized regression coefficients in Table 2). In the context of the Pearl Harbor attack, the more American identity was central to participants' self-concept, the more sympathy they felt for the American victims and for Americans in general. By contrast, in the context of Hiroshima and Nagasaki, the more participants viewed the in-group as superior, the less sympathy they felt for the victims and for the Japanese in general.

Overall, the two in-group identification dimensions, identity centrality and in-group superiority, predicted sympathy for the victims of the attacks and for members of the attacked group. As expected, identity centrality predicted more sympathy for the victims and the victim group in general when the in-group was the victim of attacks, whereas in-group superiority predicted less sympathy for the victims and the group members of the victimized groups when the in-group was the harm-doer.

3.2.4. Importance of Historical Events

Importance of historical events for U.S. history. The GLM analysis did not yield any significant results ($F(1, 102) = 1.17$, $p = .28$ for type; $F(1, 102) = 2.69$, $p = .10$ for type x superiority interaction; $F(1, 102) = .30$, $p = .58$ for type x centrality interaction). However, in support of H1c and H2c, univariate analyses conducted for each type of event separately indicated that identity centrality predicted higher ratings of the importance of Pearl Harbor, whereas

1 Sympathy toward the victims of each event (Pearl Harbor and atomic bombings), as well as the importance of these events in U.S. history, were highly skewed

(see means in Table 1). Transformations of these variables did not restore normality. Thus, the variables were dichotomized at the mean, and logistic regres-

sions were conducted with centrality and superiority as predictors. The results of these analyses replicate the results of linear regressions presented in Table 1.

in-group superiority predicted lower ratings of the importance of the atomic bombings (see Table 2). Superiority was not related to the importance of Pearl Harbor, whereas centrality was not related to the importance of the atomic bombings.

Personal importance of historical events. The GLM analysis yielded a significant effect of type, $F(1, 102) = 8.27, p < .01, \eta^2 = .075$, such that Pearl Harbor was rated as more important ($M = 3.19, SD = 1.29$) than the atomic bombings ($M = 3.13, SD = 1.29$). There was also a significant type x centrality interaction, $F(1, 102) = 5.49, p = .02, \eta^2 = .05$, indicating that, as expected, higher identity centrality predicted heightened personal importance of Pearl Harbor, but not of the atomic bombings. The type x superiority interaction was not significant, however, $F(1, 102) = .59, p = .44$.

While in-group superiority did not influence the degree to which these historical events were personally important, the results confirm the expectation regarding identity centrality: That is, participants higher in identity centrality perceived the in-group's historical victimization event as more important to them.

4. General Discussion

The studies reveal initial evidence that identity centrality and in-group superiority differentially predict historical memories of in-group victimization and in-group harming. Specifically, in-group superiority predicted reactions to historical memories of in-group harm doing (e.g., justification, emotional reactions, and perceived importance of events), whereas identity centrality predicted reactions to historical memories of in-group victimization. The hypotheses were consistently supported in both studies, and across the dependent variables in Study 2. This research contributes to two areas of study in the context of intergroup violence: social identity and historical memories.

Consistent with previous studies (e.g., Leach et al. 2008; Roccas et al. 2008) the current findings emphasize the importance of using a multidimensional approach to social identity. Recent research (e.g., Leidner et al. 2010; Roccas et al. 2006) demonstrates that in-group glorification,

rather than attachment, drives the adverse effects of in-group identification in intergroup conflict and violence. However, by focusing only on the in-group's misdeeds, these previous findings capture only phenomena related to one aspect of intergroup conflict. The current research suggests that identity centrality, rather than in-group superiority, might drive responses to historical victimization.

The scales previously used to assess in-group glorification and attachment consist of items tapping deference and superiority dimensions (for in-group glorification), and identity centrality and commitment dimensions (for in-group attachment). The conceptual differences between these dimensions might, however, produce mixed research findings. For instance, Sellers et al. (2008) note the difference between identity centrality and other affective and evaluative dimensions of in-group attachment in the context of racial identity. Overall, the current findings call for further investigation of the role of identity centrality and other in-group identification dimensions in intergroup conflict.

Research on the magnitude gap in interpersonal transgressions suggests that both victims and perpetrators systematically, though differentially, distort memories of the past (e.g., Baumeister et al. 1990; Baumeister and Catanese 2001). Victims' accounts of the transgressions emphasize the negative and lasting consequences of the harm and the perpetrators' responsibility, whereas perpetrators focus on the mitigating circumstances that led them to carry the acts and minimize the consequences of their actions on the victims (Baumeister and Catanese 2001). At the level of intergroup conflict, Kraft (2009) observes similar discrepancies in accounts given by victims and perpetrators to the Truth and Reconciliation Commission in South Africa. However, previous research has not assessed how in-group identification might differentially influence these discrepancies. The results of the present studies indicate that different ways of relating to the in-group (i.e., different identity dimensions) might strengthen these systematic differences in historical memory. For instance, in-group superiority was associated with more exonerating cognitions (i.e., the use of mitigating circumstances to legitimize the events). However, when the in-group was

the victim, identity centrality was associated with increased anger toward the perpetrator of in-group harm, more sympathy toward in-group victims, less justification of in-group victimization, and heightened perceived importance of the events.

The present research has several limitations. First, both studies used correlational methods. Participants completed measures of in-group identification before they were reminded of historical events, and in-group identification dimensions were considered predictors of construals of historical events. However, causal direction might also be reversed, such that the in-group's history might play an important role in the construction of group identity (see the extended discussion below). A second limitation of the present research is the use of student samples. Special characteristics of the student samples (e.g., education, ideology, age) might have influenced the observed relations. In particular, prior research on collective memories reveals strong generational and cohort effects on remembering of collective events (e.g., Schuman and Corning 2012).

Although group histories always include both harm-doing and victimization episodes, these types of events have typically been addressed in separate areas of research. The interpretations of historical victimization and harm-doing do not occur in a vacuum, but within the in-group's broader historical narrative. Often historical events of victimization and perpetration are causally linked (whether real or perceived). For instance, the Pearl Harbor attack is often perceived to have led to the atomic bombings of Hiroshima and Nagasaki. Although order effects in the current analyses were not significant, the perceived causal link between Pearl Harbor and Hiroshima and Nagasaki in the United States represents a weakness of the repeated measures design of Study 2.

Nevertheless, the current research provides initial evidence that different dimensions of in-group identification are linked to different construals of the in-group's past. Future research should delve deeper into the nature of these relations, and determine potential moderating factors that might further explain the complex relation between group identity and historical memory.

5. The Nature of the Relationship between Group Identity and Historical Memories

Building on previous literature, I predicted that different cognitive and motivational factors underpinning each identity dimension would lead to biases in historical memories of in-group victimization and in-group harm-doing. According to this view, the adverse effects of in-group superiority (or in-group identification in previous research) are a result of a motivated defense to image threats posed by the in-group's misdeeds. However, there might be other explanations for the observed results. For instance, in-group superiority implies that group members have an inflated (positive) image of their group. Flattering national images are part of the national narrative of most nations. National narrative might be used as an in-group stereotype, which serves to perpetuate the glorified images of the in-group through selection and distortion of events in ways that confirm the stereotype (Hirshberg 1993). Historical events of in-group harm-doing are learned and interpreted through the existing knowledge frameworks about the in-group (i.e., in-group's master narrative, see Hammack 2009). Because those individuals who view the in-group as superior to other groups are likely to endorse flattering national images, they are also more likely to reinterpret negative historical events in ways that fit the existing images (Bilali, forthcoming). Therefore, a schema consistency effect is also a plausible interpretation of the association between in-group superiority and historical memories of in-group harm-doing.

Although the present research considered in-group identification dimensions as antecedents of construals of historical events, the relation between identity and historical memories is dynamic (e.g., Kurtis, Adams, and Yellow-Bird 2010). At a collective level, historical memories form the content of group identity (Billig 1995). Social representations of the in-group's history might in turn influence the degree and the way in which individuals identify with their group (Liu and Hilton 2005). For instance, to enhance their identity, groups often distort the past by silencing or reinterpreting the negative events in their history, and by embellishing and glorifying history to portray the in-group favorably (Baumeister and Hastings 1997). Such glorified portrayals of the in-group's history might lead group members to view their

in-group as superior to other groups (i.e., leading to higher in-group superiority). However, when the in-group's history is portrayed negatively, group members might disidentify with their group to avoid negative psychological consequences on the self (Liu and Hilton 2005). In contrast, historical memories of past victimization increase group solidarity and strengthen in-group identity (Devine-Wright 2003; Roe 2003), though they might also damage group esteem (Pratto and Glasford 2008). Additionally, the in-group's past victimization can also be used to provide moral legitimacy to current and future aggressive ventures of the in-group (Wohl and Branscombe 2008).

Overall, the above discussion suggests that although the degree (e.g., Sahdra and Ross 2007) and nature of in-group identification might lead to biases in historical memories (as shown by the present research), the characteristics and uses of historical memories might also influence how individuals relate to their groups (i.e., identity dimensions). Future research should further investigate this dynamic relationship. Longitudinal studies would be best suited to assessing how identification with the in-group influences construals of historical events, and in turn, how changes in the collective/social representations of the nation's history influence in-group identification.

6. Conclusion

The present research has important implications for the study of conflict and violence. Theoretically, the findings shed light on the psychological underpinnings of reactions toward historical victimization and harm-doing. Construals of historical events of perpetration and victimization are extremely important as they might either exacerbate conflicts or facilitate reconciliation. Thus, a better understanding of the psychological factors that contribute to construals of victim and perpetrator events is important in informing strategies to address the underlying motivations and needs arising from these events. Furthermore, the current research shifts the focus from the study of victim and perpetrator groups, to studying events in which a group has either perpetrated or experienced harm. This is important considering the cyclical nature of most violent conflicts. Often, establishing one group as a victim or a perpetrator is contentious, as even groups that have perpetrated mass violence and genocide might perceive themselves as victims (Bilali et al. 2012). The current research points to the benefits of integrating the two areas of research (i.e., on victim groups and perpetrator groups) to reach a better understanding of the dynamics underlying intergroup violence and conflict.

References

- Ashmore, Richard D., Kay Deaux, and Tracy McLaughlin-Volpe. 2004. An Organizing Framework for Collective Identity: Articulation and Significance of Multidimensionality. *Psychological Bulletin* 130 (1): 80–114. doi:10.1037/0033-2909.130.1.80
- Bar-Tal, Daniel. 2003. Collective Memory of Physical Violence: Its Contribution to the Culture of Violence. In *The Role of Memory in Ethnic Conflict*, ed. Ed Cairns and Michael D. Roe, 75–93. London: Palgrave-MacMillan.
- Baumeister, Roy F., and Kathleen Catanese. 2001. Victims and Perpetrators Provide Discrepant Accounts: Motivated Cognitive Distortions about Interpersonal Transgressions. In *The Social Mind*, ed. Joseph P. Forgas, Kipling D. Williams, and Ladd Wheeler, 274–93. New York: Cambridge University Press.
- Baumeister, Roy F., Arlene Stillwell, and Sara R. Votman. 1990. Victim and Perpetrator Accounts of Interpersonal Conflict: Autobiographical Narratives about Anger. *Journal of Personality and Social Psychology* 59 (5): 994–1005. doi:10.1037/0022-3514.59.5.994
- Baumeister, Roy F., and Stephen Hastings. 1997. Distortions of Collective Memory: How Groups Flatter and Deceive Themselves. In *Collective Memory of Political Events: Social Psychological Perspectives*, ed. James W. Pennebaker, Dario Paez, and Bernard Rimé 277–93. Hillsdale, NJ: Lawrence Erlbaum.
- Bilali, Rezarta. Forthcoming. National narrative and social psychological influences in Turks' denial of the mass killings of Armenians as genocide. *Journal of Social Issues*.
- Bilali, Rezarta, and Michael A. Ross. 2012. Remembering Intergroup Conflict. In *The Oxford Handbook of Intergroup Conflict*, ed. Linda R. Tropp, 123–35. New York: Oxford University Press.
- Bilali, Rezarta, Linda R. Tropp, and Nilanjana Dasgupta. 2012. Attributions of Responsibility and Perceived Harm in the Aftermath of Mass Violence. *Peace and Conflict: Journal of Peace Psychology* 18 (1): 21–39. doi:10.1037/a0026671
- Billig, Michael. 1995. *Banal Nationalism*. London: Sage.
- Branscombe, Nyla, and Bertjan Doosje. 2004. *Collective Guilt: International Perspectives*. New York: Cambridge University Press.
- Branscombe, Nyla R., Michael T. Schmitt, and Richard D. Harvey. 1999. Perceiving Pervasive Discrimination among African Americans: Implications for Group Identification and Well-being. *Journal of Personality and Social Psychology* 77 (1): 135–49. doi:10.1037/0022-3514.77.1.135
- Brewer, Marilynn B. 1999. The Psychology of Prejudice: Ingroup Love Or Outgroup Hate? *Journal of Social Issues* 55 (3): 429–44. doi:10.1111/0022-4537.00126
- Cameron, James E. 2004. A Three-Factor Model of Social Identity. *Self and Identity* 3 (3): 239–62.
- Cameron, James E., and Richard N. Lalonde. 2001. Social Identification and Gender-Related Ideology in Women and Men. *British Journal of Social Psychology* 40 (1): 59–77. doi:10.1348/014466601164696
- Čehajić, Sabina, Rupert Brown, and Roberto González. 2009. What Do I Care? Perceived Ingroup Responsibility and Dehumanization as Predictors of Empathy Felt for the Victim Group. *Group Processes and Intergroup Relations* 12 (6): 715–29. doi:10.1177/1368430209347727
- Devine-Wright, Patrick. 2003. A Theoretical Overview of Memory and Conflict. In *The Role of Memory in Ethnic Conflict*, ed. Ed Cairns and Michal D. Roe, 9–33. New York: Palgrave MacMillan.
- Doosje, Bertjan, and Nyla R. Branscombe. 2003. Attributions for the Negative Historical Actions of a Group. *European Journal of Social Psychology* 33 (2): 235–48. doi:10.1002/ejsp.142
- Doosje, Bertjan, Nyla R. Branscombe, Russell Spears, and Antony S. R. Manstead. 1998. Guilty by Association: When One's Group has a Negative History. *Journal of Personality and Social Psychology* 75 (4): 872–86. doi:10.1037/0022-3514.75.4.872
- Ellemers, Naomi, Paulien Kortekaas, and Jaap W. Ouwerkerk. 1999. Self-Categorization, Commitment to the Group and Group Self-Esteem as Related but Distinct Aspects of Social Identity. *European Journal of Social Psychology* 29 (2–3): 371–89. doi:10.1002/(SICI)1099-0992(199903/05)29:2/3<371::AID-EJSP932>3.0.CO;2-U
- Frijda, Nico H. 1997. Commemorating. In *Collective Memory of Political Events: Social Psychological Perspectives*, ed. James W. Pennebaker, Dario Paez, and Bernard Rimé, 103–27. Hillsdale, NJ: Lawrence Erlbaum.
- Hammack, Phillip L. 2009. Exploring the Reproduction of Conflict Through Narrative: Israeli Youth Motivated to Participate in a Coexistence Program. *Peace And Conflict: Journal Of Peace Psychology* 15 (1): 49–74.
- Hirshberg, Matthew S. 1993. The Self-Perpetuating National Self-Image: Cognitive Biases in Perceptions of International Interventions. *Political Psychology* 14 (1): 77–98. doi:10.2307/3791394
- Howard, John W., and Myron Rothbart. 1980. Social Categorization and Memory for In-Group and Out-Group Behavior. *Journal of Personality and Social Psychology* 38 (2): 301–10. doi:10.1037/0022-3514.38.2.301
- Iyer, Aarti, Colin Wayne Leach, and Faye J. Crosby. 2003. White Guilt and Racial Compensation: The Benefits and Limits of Self-Focus. *Personality and Social Psychology Bulletin* 29 (1): 117–29. doi:10.1177/0146167202238377
- Jackson, Jay W. 2002. Intergroup Attitudes as a Function of Different Dimensions of Group Identification and Perceived Intergroup Conflict. *Self and Identity* 1 (1): 11–33. doi:10.1080/152988602317232777
- Jetten, Jolanda, Nyla R. Branscombe, Michael T. Schmitt, and Russell Spears. 2001. Rebels with a Cause: Group Identification as a Response to Perceived Discrimination from the Mainstream. *Personality and Social Psychology Bulletin* 27 (9): 1204–13. doi:10.1177/0146167201279012
- Kosterman, Rick, and Seymour Feshbach. 1989. Toward a Measure of Patriotic and Nationalistic Attitudes. *Political Psychology* 10 (2): 257–74. doi:10.2307/3791647
- Kraft, Robert N. 2009. The Magnitude Gap: Revealing Differences in Recall between Victims and Perpetrators. In *Applied Memory*, ed. Matthew R. Kelley, 147–66. Hauppauge, NY: Nova Science.
- Kurti , Tu çe, Glenn Adams, and Michael Yellow Bird. 2010. Generosity Or Genocide? Identity Implications of Silence in American Thanksgiving Commemorations. *Memory* 18 (2): 208–24. doi:10.1080/09658210903176478
- Leach, Colin Wayne, Martijn van Zomeren, Sven Zebel, Michael L. W. Vliek, Sjoerd F. Pennekamp, Bertjan Doosje, Jaap W. Ouwerkerk, and Russell Spears. 2008. Group-Level Self-Definition and Self-Investment: A Hierarchical (Multicomponent) Model of In-Group Identification. *Journal of Personality and Social Psychology* 95 (1): 144–65. doi:10.1037/0022-3514.95.1.144
- Leidner, Bernhard, Emanuele Castano, Erica Zaiser, and Roger Giner-Sorolla. 2010. Ingroup Glorification, Moral Disengagement, and Justice in the Context of Collective Violence. *Personality and Social Psychology Bulletin* 36 (8): 1115–29. doi:10.1177/0146167210376391
- Liu, James H., and Denis J. Hilton. 2005. How the Past Weighs on the Present: Social Representations of History and their Role in Identity Politics. *British Journal of Social Psychology* 44 (4): 537–56. doi:10.1348/014466605X27162
- Liu, James H., and Janos László. 2007. A Narrative Theory of History and Identity: Social Identity, Social Representations, Society and the Individual. In *Social Representations and History: Content, Process, and Power*, ed. Gail Moloney and Iain Walker, 85–107. New York: Palgrave-Macmillan.
- Major, Brenda, Wendy J. Quinton, and Toni Schmader. 2003. Attributions to Discrimination and Self-Esteem: Impact of Group Identification and Situational Ambiguity. *Journal of Experimental Social Psychology* 39 (3): 220–31. doi:10.1016/S0022-1031(02)00547-4
- Pratto, Felicia, and Demis E. Glasford. 2008. How Needs Can Motivate Intergroup Reconciliation in the Face of Intergroup Conflict. In *The Social Psychology of Intergroup Reconciliation*, ed. Arie Nadler, Thomas E. Malloy, and Jeffrey D. Fisher, 117–44. New York: Oxford University Press. doi:10.1093/acprof:oso/9780195300314.003.0007
- Quinn, Diane M., and Stephenie R. Chaudoir. 2009. Living with a Concealable Stigmatized Identity: The Impact of Anticipated Stigma, Centrality, Salience, and Cultural Stigma on Psychological Distress and Health. *Journal of Personality and Social Psychology* 97 (4): 634–51. doi:10.1037/a0015815
- Roccas, Sonia, Yechiel Klar, and Ido Liviatan. 2006. The Paradox of Group-Based Guilt: Modes of National Identification, Conflict Vehemence, and Reactions

- to the In-Group's Moral Violations. *Journal of Personality and Social Psychology* 91 (4): 698–711. doi:10.1037/0022-3514.91.4.698
- Roccas, Sonia, Lilach Sagiv, Shalom Schwartz, Nir Halevy, and Roy Eidelson. 2008. Toward a Unifying Model of Identification with Groups: Integrating Theoretical Perspectives. *Personality and Social Psychology Review* 12 (3): 280–306. doi:10.1177/1088868308319225
- Roe, Michael. 2003. Cowlitz Indian Ethnic Identity, Social Memories and 150 Years of Conflict with the United States Government. In *The Role of Memory in Ethnic Conflict*, ed. Ed Cairns and Michael D. Roe, 55–74. New York: Palgrave Macmillan.
- Sahdra, Baljinder, and Michael Ross. 2007. Group Identification and Historical Memory. *Personality and Social Psychology Bulletin* 33 (3): 384–95. doi:10.1177/0146167206296103
- Schuman, Howard, and Amy Corning. 2012. Generational Memory and the Critical Period. *Public Opinion Quarterly* 76 (1): 1–31.
- Sellers, Robert M., Cleopatra H. Caldwell, Karen Schmeelk-Cone, and Marc A. Zimmerman. 2003. Racial Identity, Racial Discrimination, Perceived Stress, and Psychological Distress among African American Young Adults. *Journal of Health and Social Behavior* 44 (3): 302–17. doi:10.2307/1519781
- Sellers, Robert M., and Nicole J. Shelton. 2003. The Role of Racial Identity in Perceived Racial Discrimination. *Journal of Personality and Social Psychology* 84 (5): 1079–92. doi:10.1037/0022-3514.84.5.1079
- Sellers, Robert M., Mia A. Smith, Nicole J. Shelton, Stephanie A. J. Rowley, and Tabbye M. Chavous. 1998. Multidimensional Model of Racial Identity: A Reconceptualization of African American Racial Identity. *Personality and Social Psychology Review* 2 (1): 18–39. doi:10.1207/s15327957pspr0201_2
- Shnabel, Nurit, and Arie Nadler. 2008. A Needs-Based Model of Reconciliation: Satisfying the Differential Emotional Needs of Victim and Perpetrator as a Key to Promoting Reconciliation. *Journal of Personality and Social Psychology* 94 (1): 116–32. doi:10.1037/0022-3514.94.1.116
- Sibley, Chris G., Marc S. Wilson, and Andrew Robertson. 2007. Differentiating the Motivations and Justifications Underlying Individual Differences in Pakeha Opposition to Bicultural Policy: Replication and Extension of a Predictive Model. *New Zealand Journal of Psychology* 36 (1): 25–33.
- Tajfel, Henri, ed. 1978. *Differentiation between Social Groups: Studies in the Social Psychology of Intergroup Relations*. Oxford: Academic Press.
- Tajfel, Henri, and John C. Turner. 1986. The Social Identity Theory of Intergroup Behavior. In *Psychology of Intergroup Relations*, ed. Stephen Worchel and William G. Austin, 7–24. Chicago: Nelson-Hall.
- Thomsen, Lotte, Eva G. T. Green, Arnold K. Ho, Shana Levin, Colette van Laar, Stacey Sinclair, and Jim Sidanius. 2010. Wolves in Sheep's Clothing: SDO Asymmetrically Predicts Perceived Ethnic Victimization among White and Latino Students Across Three Years. *Personality and Social Psychology Bulletin* 36 (2): 225–38. doi:10.1177/0146167209348617
- Volpato, Chiara, and Laurent Licata, eds. 2010. Collective Memories of Colonial Violence," focus section, *International Journal of Conflict and Violence* 4 (1): 4–133.
- Volpato, Chiara, and Laurent Licata. 2010. Introduction: Collective Memories of Colonial Violence. *International Journal of Conflict and Violence* 4 (1): 4–10.
- Wohl, Michael J. A., and Nyla R. Branscombe. 2008. Remembering Historical Victimization: Collective Guilt for Current Ingroup Transgressions. *Journal of Personality and Social Psychology* 94 (6): 988–1006. doi:10.1037/0022-3514.94.6.988
- Wohl, Michael J. A., Nyla R. Branscombe, and Yechiel Klar. 2006. Collective Guilt: Emotional Reactions when One's Group Has Done Wrong or Been Wronged. *European Review of Social Psychology* 17:1–37.
- Zebel, Sven, Bertjan Doosje, and Russell Spears. 2009. How Perspective-Taking Helps and Hinders Group-Based Guilt as a Function of Group Identification. *Group Processes and Intergroup Relations* 12 (1): 61–78. doi:10.1177/1368430208098777

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century

Charles Ugochukwu Ukeje, Department of International Relations, Obafemi Awolowo University, Ile-Ife, Nigeria
Akin Iwilade, Department of International Relations, Obafemi Awolowo University, Ile-Ife, Nigeria

Vol. 6 (2) 2012

**Focus:
Evidence-based Developmental
Prevention of Bullying and
Violence in Europe**

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

► **A Farewell to Innocence? African Youth and Violence in the Twenty-First Century** Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



A Farewell to Innocence? African Youth and Violence in the Twenty-First Century

Charles Ugochukwu Ukeje, Department of International Relations, Obafemi Awolowo University, Ile-Ife, Nigeria
Akin Iwilade, Department of International Relations, Obafemi Awolowo University, Ile-Ife, Nigeria

This is a broad examination of the issue of youth violence in twenty-first-century Africa, looking at the context within which a youth culture of violence has evolved and attempting to understand the underlining discourses of hegemony and power that drive it. The article focuses specifically on youth violence as a political response to the dynamics of (dis)empowerment, exclusion, and economic crisis and uses (post)conflict states like Liberia, Sierra Leone, and Nigeria to explain not just the overall challenge of youth violence but also the nature of responses that it has elicited from established structures of authority. Youth violence is in many ways an expression of youth agency in the context of a social and economic system that provides little opportunity.

When a bird perches on a long rope, neither the bird nor the rope can expect stability.

Yoruba proverb

1. Introduction: Youth and Agency in the African Public Space

In most parts of modern Africa, the relationship between youth and society can be described by the metaphor of the bird and the rope. Like the rope, African society is faced with a growing youth population (the bird) whose very expression of “being” (the simple, and perhaps even necessary, act of perching) is potentially destabilizing for all concerned. Many societies in contemporary Africa are only now coming to terms with concerns that the youth question, and its potentially destabilizing impact on social relations, may be replacing ethnicity and religion as a more powerful framework for explaining dynamics of social change on the continent. This concern is neither unfounded nor misplaced, not just because more than two-thirds of the continent’s population are under the age of 35 years – making it the most “youthful” continent (Richards 2002), but more importantly because the variety of youth engagements is having significant impacts – for good or bad – in many communities. Therefore, rather than merely focusing on how society shapes youth and experiences of youth, understanding the dynamics of social change in

Africa also requires interrogating what De Boeck and Honwana (2005, 1) refer to as “the way young people in Africa reconfigure geographies of exclusion and inclusion.” That is, how they navigate the complex tapestry of exclusion in a post-colonial state that offers not only little opportunity for graduation into adulthood but also makes surviving the constraints and incentives of youth extremely difficult.

Using “African youth” as an analytical category, as we do in this paper, is not intended to suggest that youth can mean the same thing right across a continent as diverse as Africa. We acknowledge the diverse historical and material experiences of different African societies and how these may frame the way youth and their social roles are configured. In spite of this diversity however, there are congruencies that suggest that some limited generalization can be generated for analytical purposes. These congruencies are chiefly located in the generally accepted modern history of Africa, particularly as it relates to social pressures generated by violent conflict, economic crisis, and environmental challenges. These pressures appear to shape social responses in similar ways all over the continent, and youth is no exception.

To properly situate youth in the context of their engagement with the public space in Africa it is important to

find analytical congruencies, and also to broaden the conception of youth to more accurately reflect the realities of twenty-first-century Africa. A more accurate notion of youth will accommodate four basic perspectives. First, it will view youth as a social category within an intergenerational discourse. Second, it will place this intergenerational discourse within a broader discourse of power, authority, and control. Third, it will take due note of the shifting nature of the inherent relations of power and control and how this affects the social notions of youth. Finally, it will view youth as a lived experience rather than an imagined one.

With regard to these four perspectives, it becomes clear, as Bayart (1993) demonstrates, that generational categories such as childhood, youth, and adulthood are not neutral or even natural, but rather a “part of the struggle for influence and authority within almost every society” (Christiantine, Utas, and Vigh 2006, 11). Related to this struggle is a growing youth appropriation of the public sphere made possible largely by the empowering characteristics of emerging new media technologies. As a consequence, youth has become a very critical emergent category through which we may understand the dynamics of power, influence, and control in the emerging African public sphere. As important as youth participation in this sphere is, the implications for social mobilization and thus socio-political reconstruction, as demonstrated by recent youth-led social movements in North Africa, is perhaps much more significant. Thus, youth agency must be recognized not only as capable of freeing young people from dominance by a paternalistic and patrimonial system but also as capable of reconfiguring the nature of power within the broader society. It should be noted that the apprehensions of states, corporations, and traditional institutions about the agency of youth are most deeply embedded in this political implication. Any accurate reading of youth in Africa must thus take note of this political ramification and perhaps utilize it as a principal framework for analysis.

This paper takes due note of this political ramification and pays attention to how the four elements of our conceptualization of youth can help us understand the multidimensional challenges facing African youth, as well

as how these are altering and redefining the scope and content of their engagements within and with societies on the continent. This occurs within the particular context of youth-specific experiences of violence and (dis)empowerment, particularly in conflict societies like Nigeria, Liberia, and Sierra Leone. This is important as it demonstrates the intersection that can be discerned between such experiences and broader questions of identity, economic crisis, and human rights in Africa.

The paper is divided into five interconnected sections. After the introduction, we take a broad look at the context within which “youth” is constructed in Africa. We then go on, in the section that follows, to answer the questions of “why” and “how,” with respect to youth and violence in (post)conflict societies in Africa. We also look at the gendered peculiarities that may sometimes shape the way youth engage with and in violence, and the overall implications of all these for the ability of youth to effectively participate in the public sphere. In section four, we look at the various ways in which African states have responded to youth violence and examine what this means for processes of social change and politics in Africa. The final section summarizes the key arguments and draws attention to important lessons that may be learnt. These lessons include the need to reconfigure social conceptions of youth engagement with violence in ways that take due note of the broad context of crisis within which they develop rather than focusing on youth violence as an expression of criminality.

2. Youth and the African Crisis

While being young can be fraught with danger in many societies around the world, the peculiar African context of social deprivation within which many young people in Africa grow up makes the youth experience there particularly problematic. Therefore, taking a look at the structural conditions that shape youth experience and provide incentives for violent choices in the way they express “self” is critical to having a holistic conversation about the “youth problem.” In other words, beyond youths entering popular discussions as troublesome citizens – for instance, *township youths* in the heyday of apartheid in South Africa, *rarray boys* in the ghettos of Freetown, *egbesu boys* in Nigeria’s oil

delta, *area boys* in Lagos, *bakassi boys* in southeastern Nigeria – the circumstances pushing them towards the margins of society must also be privileged in social discourse (Rashid 1997; Ukiwo 2002). It is important to understand that the discourse on youth in Africa cannot and should not be dominated by narratives of violence which oftentimes tend to be too narrowly focused on youths as threats while ignoring the underlying social meanings of violence, for instance with regard to legitimate claims against an authoritarian and incapable state.

For Africa, the social circumstances within which state tyranny and ineptitude develop and which ultimately generate youth responses are generally well known. There appears to be a consensus in the literature that Africa faces a widespread and deepening crisis of development. Colonialism and its continued salience (Amin 2001), Africa's marginal place in the international system (Amin 2001; Bigsten and Dureval 2008), and its severe governance deficits (World Bank 1981, 2000; Fukuyama 2004) are frequently cited explanations. However, even though the economic numbers still remain comparatively low, there is growing optimism that Africa is at last showing signs of emerging from its underdevelopment. In fact, triumphalist literature has appeared, heralding the "institutionalization of political power in Africa" (Posner and Young 2007), the strengthening of civil society and democracy (Halperin, Siegle, and Weintein 2010), and the growth of its economies (Soludo 2005). What these suggest is that while there appears to be an enduring climate of socio-economic crisis, some opportunities for advancement are evident. The youth question evolves within this context and is therefore shaped by the intersection of crisis and opportunity.

To give clarity to our notion of youth developing within the context of crisis and opportunity, one may look to the controversial work of Patrick Chabal and Jean-Pascal Daloz (1999). In *Africa Works: Disorder as Political Instrument* they develop the thesis that even though Africa appears to be a place of chaos, its social formations have learnt to not only manipulate the crisis, but also to appropriate its discourses and patterns for self-advancement. It is in this context that youth in Africa have been able to define themselves by finding alternative social spaces for self-ex-

pression within a constraining polity. In this sense, crisis (with its informalization of politics, rule of law deficits, and progressively weakening value systems) creates opportunities (for instance for fraud and violence) which when taken by youth often generate further crisis. This is illustrated by the situation in the Niger Delta region of Nigeria, where governance deficits in the area of petroleum refining has allowed the emergence of a shadow economy of illegal refineries (crudely constructed and with only rudimentary technology) controlled by armed youth gangs who steal crude oil from pipelines, refine it, and export the products along the West African coast. In this regard, youth seize the opportunity provided by governance deficits to accumulate capital, a significant proportion of which funds militancy and criminality in the region. Durham (2000, 113) makes a similar point about what she describes as "occult economies" where the "potency and potential of youth is extracted to sustain the power of those in authority while young people themselves feel increasingly unable to attain the promises of the new economy and society." In this regard, young people sense the powers they possess for shaping society, albeit in shapes dictated by the elite, and yet feel powerless to do anything about their own lives. They thus increasingly define themselves by working both within and at the same time around the corrupt system.

The failures of governance and statehood in Africa have created societies in perpetual crisis within which legal opportunities for social mobility are, at best, few. This is the case for entire societies, but its expression is surely graver for marginalized social categories like children, women, and youth.

The implications of social marginalization of youth are easily discerned, among other things, in violent conduct (Cruise O'Brien 1996; Utas 2005). This exacerbates generational tensions that often end in further youth exclusion, thus creating a vicious cycle of exclusion and resistance.

Anecdotal evidence can shed light on popular apprehensions about youth in Africa, in the traditional and/or modern contexts, and on how these often lead to social exclusion. In December 2003, a Sierra Leonean academic who had lived and worked continuously in Nigeria since 1978,

was conferred with a chieftaincy title of *Baaluwe* (the leader of the learned people) by the king of a renowned and ancient city in southwest Nigeria. When he was handed the oath of office to publicly pledge his allegiance to the traditional ruler, the first item on the long list of dos and don'ts was that under no circumstance must he conspire with or support the youths of the town to undermine the traditional ruler. By making this pledge, he committed himself to an unwavering loyalty to the traditional monarch, significantly at a time when that institution is contemptuously derided as anachronistic in contemporary Nigerian politics. He also unwittingly committed himself to the defense of a governance structure that not only sees the youth as trouble, but also deems it necessary to exclude them from governance.

In conditions as described above, where economic accumulation is extremely difficult, political exclusion and social decay are rife, and the structures of the state are either too weak or too uninterested to resolve conflict, it is not difficult to see the social crucible that forges violent resistance in youth. As the example we gave above suggests, even informal avenues for seeking redress and expressing opinion appear to be becoming closed to youth, thereby driving more and more into violent conduct. Unfortunately, very few societies have subjected themselves to the kinds of critical introspection that could lead to sincere and open acknowledgements, if not acceptance, of responsibility for the pitiable ways youths are increasingly falling into in contemporary Africa. Rather than critical introspection, most societies in Africa are content with merely constructing conflicting public images of youth. The first of such touts them patronizingly as *the hope for the future*, the other castigating them as nuisances to public order (Waal 2002, 13). Perhaps the only consolation is that no matter the idiom society uses to qualify its youth, there is little controversy that at the start of the present millennium, youth themselves have become central and strategic to the making and unmaking of social order on the continent as the experiences of Liberia, Sierra Leone, Cote D'Ivoire, Tunisia, Libya, the DRC, Uganda, Rwanda, Burundi, Mozambique, and Sudan – to mention but a few of the recent flashpoints – have clearly demonstrated (Abdullah 1999; Richards 1994, 1997; Ukeje 2001; Zegeye 2003; Maxted

2003; Utas 2003). Indeed, there is concern that African societies ought to start guarding against a situation where their worst apprehensions and fears concerning youth may soon metamorphose into a self-fulfilling prophecy.

3. Youth and Conflict in Africa: Violence, Survival and Victimhood

Having provided a broad view of what it means to be young in Africa, we can now go on to examine the specific question of youth violence. Two broad categories of youth violence can be discerned: violence aimed at political goals and criminal violence. Except where expressly stated, we focus more on violence with political meaning. It should be noted however that the line between these two categories sometimes blurs, as the Niger Delta, Sierra Leone, and Liberia conflicts show (Kandeh 1999; Ukeje 2001).

Since the mid-1990s at least, one of the defining features of the youth discourse in Africa has been an attempt to understand the critical interface between youth and violent conflicts. This is not out of place, given the unprecedented involvement of young people in the civil wars and low intensity conflicts that have erupted in states like Liberia, Sierra Leone, and Nigeria. Of course young people have always been involved in violent conflict, with youth forming the bulk of the armed forces in most societies; in recent times however, they appear to increasingly take on the roles of instigators and leaders of violence rather than mere followers (youth-led social movements like the Movement for the Emancipation of the Niger Delta, MEND, for example, are notorious for violent engagement with the state and multinational oil companies). It is possible to establish a causal link between this emerging youth role in violence and broader questions about social decomposition, economic crisis, and the critical intersection of the local and the global (El-Kenk 1996; Macdonald 1997; Comaroff and Comaroff 2005; Aluiagba 2009). Deconstructing youth participation in violence in Africa is therefore incomplete without an engagement with this important phenomenon: not only does it demonstrate the deep-seated crisis of (dis)empowerment facing many societies, it also provides crucial insights into the way youth navigate this complex terrain and the weapons or tools they use to do so. The very nature of many of these conflicts, particularly their disproportionate linkage to resource management and

wealth distribution issues, may be a reflection of the attraction of accumulation to a youthful population that has long been excluded from productive engagement within the formal economies. One should however acknowledge, in the first instance, that there is a gendered perception of violence that takes note of gender-specific experiences of youth violence and methods of navigating the dangerous terrains created by conflict. In the first place, it should be noted that gender is all too often conflated with women and girls. As Amani El Jack (2003, 6) notes, however, gender simply “refers to perceptions of appropriate behavior, appearance and attitude for women and men arising from social and cultural expectations.” As a consequence, gendered perceptions of youth in violence must take cognizance of the local context and understandings of gender. Generally perceived notions of youth in relation to violence are, however, almost exclusively male. This is not unconnected to the gendered delineation of roles in conflict societies, which regards female identity in violent conflict in the context of victimhood. This blanket assumption of victimhood however often ignored the crucial role that women play in the outbreak, management, and resolution of violent conflict. As Iwilade (2011, 27) notes, “an ethnography of social tactics in conflict situations easily counters the reductionist portrayals of women as merely passive victims of conflicts.” There are good examples of young women in Liberia and Sierra Leone who acted as combatants in civil wars and some, like Colonel Black Diamond of the Women’s Auxiliary Corps in Liberia, even commanded elite units (Utas 2005, 404).

Notwithstanding the disproportionate emphasis on young men in the discourse on youth violence, there is still some analytical value to examining youth violence in socially separate but mutually reinforcing gender crucibles. Even though disempowerment is a shared misfortune of many African youths irrespective of gender, one may still discern differentiated experiences across gender divisions. For one, young women have unique pre-conflict experiences of disempowerment that provide important insights into how they respond to the dynamics of violence. As Brett and Sprecht (2004, 87) note, young women often participate in organized armed violence primarily to escape domestic violence, abuse, and poverty rather than in defense of relig-

ious or ethnic interests as can often be the case for young men. This indicates that female disempowerment and marginalization by a patriarchal system is a major reason young women participate in violent public conduct. This does not however provide adequate explanation for the methods by which they navigate the geography of violence.

In deconstructing youth participation in violence it is helpful to answer the questions of “why” and “how.” “Why” helps us to understand the specific factors that draw youths into violent conduct while “how” explains the tactics and tools with which they navigate the dangerous geography of violent conflict. Both questions collectively provide important insights into the dynamic engagement of youth with violence in Africa and the implications for social change.

To address the question of “why,” we can apply Murphy’s four models of youth participation in violence (2003, 64–66) and find appropriate parallels within contemporary Africa. The first is the “coerced youth model” which views youth as being brutally coerced into a (violent) military role and thus as passive victims of social upheaval. This model has been useful in providing some explanation for the “child soldier” phenomenon in Sierra Leone and Liberia (Richards 1994, 1997). The second is the “revolutionary youth model” which views youth as rebelling against political and economic marginalization. This model has been used to rationalize the engagement of youths within social movements involved in violent confrontation with the state/multinational oil coalition in the Niger Delta region of Nigeria (Ifeka 2006). The two models differ in what they choose to emphasize about youth participation in violence. In the first model, youth is denied agency as they are framed as unwilling or choiceless victims of a brutal and coercive apparatus of violence. The second fully acknowledges the agency of youth, making sure to point out their deliberate and rational rejection of marginalizing social systems and their creative responses to the opportunities created by social conflict.

The third is the “delinquent youth model” which views youth participants in violent conflicts not as revolutionary idealists but as “alienated and economically dispossessed opportunists exploiting the economic spoils of social tur-

moil” (Murphy 2003, 64). In this case, young people engage in violence in defense of no higher ideal, but rather for the heady adventure of violence itself (the *West Side Boys* in Liberia for instance; Abdullah 1998) or for the criminal benefits that can be derived from conflict (some criminal elements of insurgency movements in the Niger Delta for instance). This model is reflective of traditional notions of youth as a period that is carefree, rebellious, contemptuous of authority, and generally mischievous. It follows a path slightly different from the earlier two, straddling the realms of agency and agenthood. Agenthood here refers to the state of being an agent: lacking independent capacity to take decisions without direction from others. While it acknowledges the choice of youth to participate in delinquent violence, it frames that choice as natural and thus demeans it.

The fourth is the “youth clientelism model” which emphasizes how youth manage their dependency and agency within “an institutional structure of repressive patrimonialism in which their subordination to adults is based on a cruel mixture of brutality, personal benevolence and reciprocity” (Murphy 2003, 65). This model is markedly different from the three described earlier because it focuses on an extraneous factor to explain youth agency in violence: institutions built through client-patron relations. This model is particularly appropriate for analyzing relations between young combatants and the commanders who recruit, mentor, and discipline them within the ranks of rebel movements. The civil wars in Liberia and Sierra Leone are, of course, poster children for this phenomenon.

The other question of “how” relates primarily to methods and tools for navigating the complex geographies of violent conflict in Africa. It is important to understand that the tactics with which youth engage in or navigate violent situations cannot be explained with a monocultural or fossilized lens. It often involves a series of constantly adjusted tactics, developed in response to the constraints and incentives created, on the one hand by a hostile socio-economic context, and on the other by the immediate consequences of conflict.

One of the main features of youth violence in Africa was insightfully analyzed by Caroline Ifeka (2006) in a study of

the intersection between religion and violence. In that study, she developed a framework for analyzing youth cultures of resistance and violence in the context of customary and world religions in which old and new gods are important sources of ideological resistance (Ifeka 2006, 721). In this context, Ifeka (2006, 725) invariably builds on the “revolutionary youth model” to describe the way religious identities are appropriated as a tool to navigate the complex terrains of violence and resistance. By plugging into religious rituals and rhetoric, youth take advantage of violence and thus gain legitimacy for an essentially aberrant social form. By purporting to engage in violence on behalf of, in the name of, or through the agency of religion and its rituals, youth gain public support or at least acquiescence for what is often essentially violent resistance to a hostile patrimonial system. This method of gaining legitimacy for youth violence is easily discernible in Islamic revivalist movements in northern Nigeria and among the *egbesu* warriors of the oil communities in the Niger Delta. Of course, it must be noted that this method is often intricately connected to broader questions of culture, ecology, and economics, especially in the case of violent resistance in resource-rich communities like the Niger Delta.

In the context of major armed conflicts like civil wars, violence is in itself often a method to navigate the violent terrain created by war. In this regard, many young people simply join armed groups as a way of gaining some protection from brutal and unforgiving armies (that are sometimes the very same ones that vulnerable youth join). It is thus often a case of “if you can’t beat them, join them.” This tactic for navigating violence has been thoroughly addressed within the literature. The works of Utas (2003, 2005), MacMullin and Loughry (2004), and Murphy (2003) are particularly rich in this with regard to Liberia and Sierra Leone, while Ukeje (2001) and Obi (2006) make similar points about the Niger Delta. As MacMullin and Loughry (2004) note in the case of young women, escaping the heightened vulnerability of women and girls to violent abuse during armed conflict is one of the key motivations that drive many of them to enlist. As McKay and Mazurana (2004) also note, during the 1976–1992 civil war in Mozambique, many young girls joined FRELIMO to get away from the rural areas, to improve their education or career

opportunities, and to expand gender roles for women (issues that had become more difficult as a direct consequence of violent conflict). What these cases show is that violence is in itself sometimes a tactic to avoid violence or its consequences.

While the two navigation tactics discussed above are by no means exhaustive, they provide an adequate description of the broad scope and ingenuity of youth encounters with violent conflict. The implications of growing organized and unorganized violence on the psyche of youth and invariably on the society itself are dire. For one, youth violence deepens the intergenerational debate as well as the conflict therein, as clearly manifested in the *Mungiki* movement in Kenya.¹ It also deepens apprehension about the future of traditional notions of youth and adulthood, as well as raising critical questions about how to find the right balance between these notions and the new forms of social relations imposed by a rapidly changing global capitalist system. Some of these new forms include changes in the political economy of accumulation as a consequence of a growing emphasis on skills now more typically possessed by youth (for example software engineering, social networking, and so on). With violent conflict in Africa now increasingly erupting as a consequence of youth dissatisfaction with their chances for advancement, rather than disagreements within the traditional elite structures (Libya, Egypt, and to a lesser extent Nigeria; Boko Haram in the North and MEND in the Niger Delta are important examples), creating a constructive outlet for youth agency has become a national security concern across Africa.

While the recent youth-inspired violent eruptions in North Africa no doubt happened within specific social contexts, this could, one way or the other, be replicated in many other African countries. It is common therefore to see the discourses around new protest movements in diverse places across sub-Saharan Africa, from Nigeria to Mozambique,

Kenya and Ghana, taking a cue from and compared with the Arab Spring protests in Tunisia, Algeria, Libya and Egypt. This is an indication that these violent youth-led movements have entered into youth imagination across Africa. Such imagination is, however, not as important as the growing youth perception, no doubt also fuelled by the successes in North Africa, that such movements can actually bring about monumental changes in society and politics in many countries. This is the most ominous for established patterns of authority and control on the continent. Rather than simply working to expand opportunities for youth, it appears that the fear of youth losing their “innocence” is now becoming the discourse around which paternalistic social systems across Africa constructs responses to the “youth problem.”

4. Farewell to Innocence: Responses to Youth Violence in Africa

Because the more popular idioms relating to contemporary African youth present them as “troublemakers,” public policy interventions have focused on youths as a set of “bads,” mostly with disastrous consequences. At the broader level, a major consequence is that the dominant public idiom continues to cloud societal judgments on youth and youth-related matters. More specifically, such dominant idioms have totally diminished the remarkable enterprise and resilience of young people at critical periods of society-building and nation-building processes; for instance, the contributions of youths towards decolonization in the 1960s and their popular roles in the re-democratization projects in different countries from the early 1990s. Unfortunately, public discourse tends to suppress, ignore, or devalue such contributions of youths to socio-political development in different African countries. Quite often, the eventual beneficiaries of such developments end up distancing themselves from or working against youth aspirations and needs. This was the situation soon after young people played a frontline role in the decolonization struggles, after which the post-independence

¹ The Mungiki Movement in Kenya was a youth organisation that was notorious for its violent manipulation of generational conflicts and discourses during and after the 2002 general elections. A comprehensive analysis is provided by Kagwanja (2005).

socio-political project was hijacked by the ruling elites in different countries for personal and group aggrandizement. It still is the situation after youths spearheaded the socio-political “revolution” that led to the collapse of one-party civilian or military regimes and the process of re-democratization and democratic consolidation.

Some of the ex-combatants interviewed during earlier fieldwork in Monrovia, Liberia, in 2001 captured the exasperation and anger associated with this “use and dump” politics perfected by Charles Taylor’s regime. As one of them put it, “you cannot use a truck to build a house and not allow it even pass by after completion,” meaning that after all the sacrifices they made during years of fighting in the bush, they would not accept their neglect by the Taylor regime. This narrative of angst certainly played a major part in the emergence of a new rebel group, Liberians United for Reconstruction and Development (LURD) in 2000, and in the manner in which it successfully mobilized disgruntled ex-combatants to launch successful attacks that contributed to the downfall and exile of Charles Taylor. Presently, a similar undercurrent of neglect and disillusionment is driving the recent resurgence of gangsterism and violence in Nigeria’s oil region, especially in Port Harcourt. The main gangs terrorizing the city started their careers as political thugs to help different factions achieve electoral victory during the 2003 gubernatorial elections in the state. With little or no use for them after that election, political patronage for the gangs dried up – but not before they had reformed as organized armed groups engaged in oil bunkering and threatening public order and stability.

What the above suggests is that youth is only considered relevant by society to the extent that it serves the narrow interests of the political and social elite. This exploitative attitude to youth feeds into the discourses of generation and hegemony that shape state responses to youth violence.

These responses can be grouped into three broad and interconnected areas: co-option, exclusion and repression. African states have mostly responded to youth pressures by employing all of these strategies to varying degrees and within specific contexts. Depending on the overall democratic character of governance in the country concerned,

states have shown willingness to clamp down on youth protests and protest cultures and also to shape the youth discourse in ways that frame them as at best troublesome.

The 2009 amnesty deal for militants in the Niger Delta is an interesting policy that shows the deliberate use of all three strategies in response to incidences of youth violence and resistance. As background, it should be noted that resistance in the oil-producing region of the Niger Delta has a long-drawn history rooted in the character and activities of European foreign capital and colonial rule over several centuries. This is the case if one considers the common thread that connects different phases of social disorder in the region since at least the sixteenth century (Okwechime, 2011; Ukeje, 2011a, 2011b). Since the discovery of crude oil in commercial quantities in Oloibiri in 1958, the Niger Delta has grown to become the heart of the Nigerian economy, contributing about 80 percent of all federal revenue. This centrality of oil to the Nigerian economy and, perhaps more importantly, to the revenue accruing to the state, has led social convulsions in the region to be construed as direct threats to national security and even to the very survival of the country. It has also bred social agitations and violent resistance by oil communities who consider their meager receipts from the proceeds of oil exploitation grossly unfair and unacceptable. At the heart of resistance to the state in the Niger Delta are the youth who challenge the state/multinational oil company coalitions and demand a greater share of the resources accruing from oil exploration and exploitation. The generational dimension of the crisis further complicates the situation. In this case, traditional patterns of power and authority are increasingly being violently renegotiated, to the point where young people appear to have lost confidence in the ability of adults to effectively represent and defend their interests.

In response to widespread youth-led violence that undermined the vital oil industry, the government instituted an amnesty program that was meant to demobilize, disarm, and rehabilitate militants in the region. A critical assessment of the actual working of the policy reveals key indicators of the three response strategies mentioned earlier: repression, co-option and exclusion.

In the first place, the discourses that culminated in the amnesty policy were largely shaped by an elite structure that invariably excluded the youth. Obi and Rustad (2011, 204) report that the amnesty program was not the outcome of open negotiations or a formal peace agreement between the government and militants; instead, they were done at the highest levels of government, and involved members of the Niger Delta elite/elders and top government officials of Niger Delta origin negotiating with militia commanders. This implies that from the start, the amnesty was a tool of exclusion rather than inclusion.

We can easily discern a second policy of co-option in the way militant commanders were pulled into the embrace of the state elite, much to the chagrin of many of their erstwhile fighters who were inevitably left in the cold. As Davidheisser and Nyiyiana (2010) note, “while ex-militia ‘commanders’ enjoyed state patronage and largesse, their erstwhile foot soldiers in designated camps complained of the poor living conditions, lack of training facilities and programmes and delayed payment of allowances.” This sort of co-option has left many former militia commanders turning into mouth-pieces and defenders of the state (and the crop of political elites who now control it) against whom they had fought so hard. In fact, there are indications that many of these ex-militia now provide security services to oil companies, defending them against protesters.

The third element, repression, is also evident in the amnesty program and is perhaps best illustrated by the continued militarization of the region. The presence of the Joint Task Force, the armed security force charged with maintaining “order” in the region, is a constant reminder not only to the now co-opted militants, but also to their excluded counterparts, of the might of the state and its willingness to use force to achieve its goals. It is no surprise that since the amnesty program was launched in 2009, there have been reports of indiscriminate extrajudicial killings and harassment of youth all over the region. Similar patterns can be found in DDR projects in Liberia and Sierra Leone (MacMullin and Loughry 2004; Utas 2003) and in attempts at resolving the Boko Haram insurgency in Northern Nigeria (Adesoji 2009).

What these portend for youths is undoubtedly far-reaching. As opportunities for advancement and success fade away, Argenti (2002, 145) reminds us: “as local traditional orders become increasingly intermeshed with national political orders that have lost all legitimacy, young people are challenged to find alternative forms of representation.” Two of these alternative forms of youth representation for example, the growing socialization of youths into the informal sector and youth attraction to radical religion (Spinks: 2002), are not only producing distinctive socialization outcomes but are themselves complicating the matrix of order and stability in contemporary Africa.

Regarding the socialization of youths into the informal sector, for instance, the process seemed to have been aided, if not triggered, by declining family (household) incomes consequent upon the economic crises of the 1980 and after. Such informalization processes are creating street-level socializations, invariably also weakening the tenuous relationship among youth and between youth and family, youth and society, and youth and government. With shrinking access to subsidized socio-economic opportunities, the resilience of the family as the most basic unit for value creation, moral affection, and individual protection becomes threatened and too functionally weak to perform well. This adulteration is compounded by the harsh consequences of the present neo-liberal economic regime, particularly those associated with globalization (Meagher 2003). Another factor is what Ly described as the “eclipse of those traditional forms of solidarity that large kinship groups had generated and sustained” over decades in the continent (1988). As more young people achieve social puberty, therefore, they are confronted with the “fact of life” that they may not have the socio-economic wherewithal to live through this difficult phase of life, at least not independently, and are thus encouraged to evolve into a culture of violence and impunity.

5. Conclusion

What we have done above is to take a broad look at the context within which youth violence has developed in Africa, understanding it as both a failure of governance and a strategy for survival. Four models of youth participation in violence can be easily discerned in Africa. They include “coerced youth” (which denies youth agency by focusing

on factors that force youth into violence), “revolutionary youth” (which acknowledges youth agency and situates violence in state decay), “delinquent youth” (which views youth as economically dispossessed opportunists exploiting social turmoil), and “youth clientelism,” which focuses on how institutionalized client-patron structures shape youth engagement with violence. We also show that youth participation in violence is often a series of constantly adjusted tactics developed in response to constraints and incentives created by a hostile socio-economic context and the immediate consequences of conflict. We identify religious rhetoric as one key normative framework that has been exploited by youth to rationalize violence for and on behalf of “faith” and show that violence is often, in itself, a tactic to escape the implications of violent conflict. We also demonstrate the broad use of cooption, exclusion, and repression by the state in response to youth resistance and violence.

The “youth problem” thus flows from broader social crises faced by the state in Africa. The challenge is therefore chiefly about how to reconstruct African society (and the state) in ways that address the youth crisis as a developmental problem within a holistic framework. In Liberia, there is a Creole word of wisdom that “bad bush no dey for throwaway bad pikin,” literally meaning that there is no bad bush to throw away a bad child. This simply means that one cannot solve a problem by simply wishing it away. In the context of our discussion, this implies that Africa cannot afford to bid farewell to the innocence of youth, but rather needs to acknowledge, accommodate, and come to terms with the challenge they represent for the youth, in the first instance, and the society at large. A major point in this paper is therefore that the “problem” of youth is symptomatic of deeper and festering challenges facing African societies, and must be addressed from this holistic premise. It is important however that the symptom be kept separate from the cause, as it is the pervasive tendency to lump them together that has fueled public concern about the

“diabolical” exploits of youth and driven denial of their resilience. There is an even bigger concern in the twenty-first century, especially in the context of challenges posed by globalization (Hedley 2001). On the continent today, there is now ample evidence that the globalisation process is going to bypass, marginalize or completely neglect millions of people for several decades to come. For those people, a large percentage of whom fall within the social category of youth, according to Gus Speth, as cited by Agarwal (1998), “poverty is a denial of the most basic of all human rights: the Right to Life.” A new preoccupation in national, continental, and global policy debate and action should therefore focus on improving the quality and dissemination of human security, especially for the most marginalized and vulnerable social categories in the world.

There is a sound conviction that young people could reach their maximum potential without migrating to cities or engaging in dangerous social activities if they can secure subsidized access to educational, medical, economic, political, social, and cultural resources wherever they are located, whether in the rural or urban areas. This approach to youth was promoted by the United Nations Food and Agricultural Organization (FAO) between 2000–2005 through its program on “rural youth and food security.” According to the FAO, this must involve giving them the right training and education, supporting programs that ensure gender balance, training youths in leadership, communication, and group activities, and packaging programs of activities that raise the self-esteem of young people, to mention but a few. Youth empowerment means expanding the opportunities available to young people, taking cognizance of their ideas, vision, and skills, and channeling them towards development. It is clear that substantial ground still needs to be covered before African countries can adequately make sense of the concerns, yearnings, and aspirations of their youths, and turn these into the energy and drive necessary to claim the twenty-first century.

References

- Abdullah, Ibrahim, and Patrick Muana. 1998. The Revolutionary United Front of Sierra Leone: A Revolt of the Lumpen Proletariat. In *African Guerrillas*, ed. Christopher Clapham, 173–93. Oxford: James Currey.
- Abdullah, Ibrahim. 1998. Bush Path to Destruction: The Origin of the Revolutionary United Front/Sierra Leone. *Journal of Modern African Studies* 36:203–35.
- Abdullah, Ibrahim. 1999. The Role of Youths in Conflicts. In *The Role of Youth in Conflict Prevention in Southern Africa*, ed. Anne Palme, 3–29. Kauhava, Finland: KATU.
- Adesoji, Abimbola. 2010. The Boko Haram Uprising and Islamic Revivalism in Nigeria. *Africa Spectrum* 45 (2): 95–108.
- Agarwal, Anil. 1998. *Globalisation, Civil Society, and Governance: The Challenges for the 21st Century*. Lecture delivered at NORAD's Environment Day held in Oslo on December 15.
- Aina, Tade. 1996. *Globalization and Social Policy in Africa: Issues and Research Directions*, CODESRIA Working Paper 6/96. Dakar: CODESRIA.
- Aluaigba, Moses. 2009. Circumventing or Superimposing Poverty on the African Child? The Almajiri Syndrome in Northern Nigeria. *Childhood in Africa* 1 (1): 19–24.
- Amin, Samir. 2001. Africa: Living on the Fringe? *Africa Insight* 31 (2): 3–7.
- Argenti, Nicholas. 2002. Youth in Africa: A Major Resource for Change. In *Young Africa: Realising the Rights of Children and Youth*, ed. Alex de Waal and Nicholas Argenti, 123–53. Trenton, NJ: Africa World Press.
- Bigsten, Arne, and Dick Dureval. 2008. *The African Economy and Its Role in the World Economy*, Current African Issues 40. Uppsala: Nordic Africa Institute.
- Chabal, Patrick, and Jean-Pascal Daloz. 1999. *Africa Works: Disorder as Political Instrument*. Oxford: James Currey.
- Christiansen, Catrine, Mats Utas, and Henrik Vigh. 2006. Navigating Youth, Generating Adulthood. In *Navigating Youth, Generating Adulthood: Social Becoming in an African Context*, ed. Catrine Christiansen, Mats Utas, and Henrik Vigh, 9–28. Uppsala: Nordic Africa Institute.
- Comaroff, Jean, and John Comaroff. 2006. Reflections on Youth, from the Past to the Postcolonial. In *Frontiers of Capital: Ethnographic Reflections on the New Economy*, ed. Melissa Fisher and Greg Downey, 267–81. Durham: Duke University Press.
- Coulter, Chris, Mariam Persson, and Mats Utas. 2008. *Young Female Fighters in African Wars: Conflict and its Consequences*, NAI Policy Dialogue 3. Uppsala: Nordic Africa Institute.
- Cruise O'Brien, Donald. 1996. A Lost Generation? Youth Identity and State Decay in West Africa. In *Postcolonial Identities in Africa*, ed. Richard Werbner and Terrence Ranger, 55–74. London: Zed.
- Davidheisser, Mark, and Kialele Nyiayaana. 2010. Demobilization or Remobilization? The Amnesty Programme and the Search for Peace in the Niger Delta. Paper presented at the Second International Conference on Natural Resource, Security and Development in the Niger Delta, held at the Niger Delta Wetlands Center, Yenagoa, Bayelsa State, Nigeria, 8–11 March.
- de Boeck, Filip, and Alcinda Honwana. 2005. Children and Youth in Africa: Agency, Identity and Place. In *Makers and Breakers, Made and Broken: Children and Youth as Emerging Categories in Postcolonial Africa*, ed. Phillip de Boeck and Alcinda Honwana, 1–14. Oxford: James Currey.
- de Waal, Alex. 2002. Realising Child Rights in Africa: Children, Young People and Leadership. In *Young Africa: Realising the Rights of Children and Youth*, ed. Alex de Waal and Nicholas Argenti, 1–28. Trenton, NJ: Africa World Press.
- Durham, Deborah. 2000. Youth and the Social Imagination in Africa: Introductions to Parts 1 and 2. *Anthropological Quarterly* 73 (3): 113–20.
- El Jack Amani. 2003. *Gender and Armed Conflict: Overview Report*. <http://www.bridge.ids.ac.uk/reports/CEP-Conflict-Report.pdf>.
- El-Kenz, Ali. 1996. Youth and Violence. In *Africa Now: People, Policies and Institutions*, ed. Sylvester Ellis, 42–57. The Hague, London, and Portsmouth: DGIS, Currey, and Heinemann.
- Everatt, David. 1995. School Reject or Eject? Contextualizing “Out-of-School Youth” in the New South Africa. *Prospects* 25 (3): 451–68.
- Fukuyama, Francis. 2004. *State-Building: Governance and World Order in the Twenty-First Century*. London: Profile.
- Halperin, Morton, Joseph Siegle, and Michael Weinstein. 2010. *The Democracy Advantage: How Democracies Promote Prosperity and Peace*. New York: Routledge.
- Hedley, Jeremy. 2001. Globalization and Its Impact on Youth. *Journal of Future Studies* 6 (1): 89–106.
- Ifeka, Caroline. 2006. Youth Cultures and the Fetishization of Violence in Nigeria. *Review of African Political Economy* 110:721–36.
- Iwilade, Akin. 2011. Women and Peace Talks in Africa. *Journal of International Women's Studies* 12 (1): 22–37.
- Kagwanja, Peter. 2005. Power to Uhuru: Youth Identity and Generational Politics in Kenya's 2002 Elections. *African Affairs* 105 (418): 51–75.
- Kandeh, Jimmy. 1999. Ransoming the State: Elite Origins of Subaltern Terror in Sierra Leone. *Review of African Political Economy* 81:362–64.
- Kaplan, Louise. 1984. *Adolescence, the Farewell to Childhood*. New York: Simon and Schuster.
- Kenny, Charles. 2003. The Internet and Economic Growth in Less-developed Countries: A Case of Managing Expectations. *Oxford Development Studies* 31 (1): 99–113.
- Ly, Boubacar. 1988. The Present Situation of Youth in Africa. In *Perspectives on Contemporary Youth*, ed. Janusz Kuczynski, Shmuel Eistendadt, Boubacar Ly, and Lotikar Sarkar, 3–19. Tokyo: United Nations University.
- Macdonald, Robert. 1997. Youth, Social Exclusion and the Millennium. In *Youth, the “Underclass” and Social Exclusion*, ed. Robert Macdonald, 167–97. London: Routledge.
- MacMullin, Collin, and Maryanne Loughry. 2004. Investigating Psychosocial Adjustment of Former Child Soldiers in Sierra Leone and Uganda. *Journal of Refugee Studies* 17 (4): 460–72.
- Mamdani, Mahmood. 1996. *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton: Princeton University Press.
- Maxted, Julia. 2003. Children and Armed Conflict in Africa. *Social Identities* 9 (1): 51–72.
- McKay, Susan, and Dyan Mazurana. 2004. *Where are the Girls? Girls in Fighting Forces in Northern Uganda, Sierra Leone and Mozambique: Their Lives During and After War*. Quebec: Rights and Democracy.
- Meagher, Kate. 2003. A Back Door to Globalisation? Structural Adjustment, Globalisation and Transborder Trade in West Africa. *Review of African Political Economy* 30 (95): 57–76.
- Mische, Ann. 1995. Citizenship Across Youth Networks in Brazil. *International Review of Social History* 40 (3): 131–58.
- Murphy, William. 1980. Secret Knowledge as Property and Power in Kpelle Society: Elders Versus Youth. *Africa* 50:193–207.
- Murphy, William. 2003. Military Patrimonialism and Child Soldier Clientalism in the Liberian and Sierra Leonean Civil Wars. *African Studies Review* 46 (2): 61–87.
- Obi, Cyril, and Siri Rustad. 2011. Conclusion: Amnesty and Post Amnesty Peace, Is the Window of Opportunity Closing for the Niger Delta? In *Oil and Insurgency in the Niger Delta: Managing the Complex Politics of Petro-Violence*, ed. Cyril Obi and Siri Rustad—200–210. London: Zed.
- Obi, Cyril. 2006. *Youth and Generational Dimensions to the Struggle for Resource Control in the Niger Delta: Prospects for the Nation-State Project in Nigeria*, CODESRIA Monograph Series. Dakar: CODESRIA.
- Okwechime, Iwebunor. 2011. The International Dimensions of the Crises in the Niger Delta, 1956–2006. Ph.D. thesis, Department of International Relations, Obafemi Awolowo University, Ile-Ife, Nigeria.
- Posner, Daniel, and Daniel Young. 2007. The Institutionalization of Political Power in Africa. *Journal of Democracy* 18 (3): 126–40.
- Rashid, Ishmail. 1997. Subaltern Reactions: Lumpens, Students and the Left. *African Development* 23 (3–4): 13–28.
- Richards, Paul. 1994. Rebellion in Liberia and Sierra Leone: A Crisis of Youth? In *Conflict in Africa*, ed. Oliver Furley, 134–70. London: Taurus Academic.

- Richards, Paul. 1997. *Fighting for the Rain Forest: War, Youth and Resources in Sierra Leone*. Oxford and Portsmouth: James Currey and Heinemann.
- Richards, Paul. 2002. Youth, Food and Peace: A Reflection on Some African Security Issues at the Millennium. In *Africa in Crisis: New Challenges and Possibilities*, ed. Tunde Zack-William, Diane Frost, and Alex Thomson, 179–95. London: Pluto.
- Soludo, Charles. 2005. *The Political Economy of Sustainable Democracy in Nigeria*, paper presented at the 2005 Democracy Day Lecture, Abuja, Nigeria.
- Spinks, Charlotte. 2002. Pentecostal Christianity and Young Africans. In *Young Africa: Realising the Rights of Children and Youth*, ed. Alex de Waal and Nicholas Argenti, 193–206. Trenton, NJ: Africa World Press.
- Ukeje, Charles. 2001. Youth, Violence and the Collapse of Public Order in the Niger Delta of Nigeria. *Africa Development* 26 (1–2): 337–66.
- Ukeje, Charles. 2011a. From Oil Rivers to the Niger Delta: The Paradoxes of Domination and Resistance in Colonial Nigeria. In *Environment and Economics in Nigeria*, ed. Toyin Falola and Adam Paddock, 183–205. New York and London: Routledge.
- Ukeje, Charles. 2011b. Changing the Paradigm of Pacification? Oil and Militarization in Nigeria's Delta Region. In *Oil and Insurgency in the Niger Delta: Managing the Complex Politics of Petroviolence*, ed. Cyril Obi and Siri Aas Rustad, 83–98. London: Zed.
- Utas, Mats. 2003. *Sweet Battlefields: Youth and the Liberian Civil War*. Uppsala: Uppsala University Dissertations in Cultural Anthropology.
- Utas, Mats. 2005. Victimcy, Girlfriending, Soldiering: Tactic Agency in a Young Woman's Navigation of the Liberian War Zone. *Anthropological Quarterly* 78 (2): 403–30.
- Villalon, Leonardo. 1995. *Islamic Society and State Power in Senegal: Disciples and Citizens in Fatick*. Cambridge: Cambridge University Press.
- Weber, Max. 1978. *Economy and Society*. Berkeley: University of California Press.
- World Bank. 1981. *World Development Report 1981*. Washington, D.C.
- World Bank. 2000. *World Development Report 2000*. Washington, D.C.
- Wyn, Johanna, and Rob White. 1997. *Rethinking Youth*. London: Sage.
- Zegeye, Abebe. 2003. Amulets and Dreams: War Youth and Change in Africa. *Social Identities* 9 (1): 37–50.