

EQUIPping High School Students
Effects of a universal prevention program
on antisocial behavior

Floor van der Velden

ISBN 978-90-393-5244-1

Cover design and photography: Katelijne Arts

Printed by Ipskamp Drukkers, Enschede, the Netherlands

© 2009 Floor van der Velden

All rights reserved. No part of this thesis may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without the prior permission of the author.

EQUIPping High School Students
Effects of a universal prevention program
on antisocial behavior

EQUIP in het Voortgezet Onderwijs
Effecten van een universeel preventie programma op antisociaal gedrag
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen
op maandag 18 januari 2010 des ochtends te 10.30 uur

door

Floortje van der Velden

geboren op 8 november 1978, te Eindhoven

Promotoren: Prof. dr. D. Brugman
Prof. dr. W. Koops

Co-promotor: Dr. J. Boom

This dissertation was supported by a grant from the Netherlands Organization for Scientific Research (NWO PROO 411 04 604)

Contents

Chapter 1	General Introduction	7
Chapter 2	Moral cognitive processes explaining antisocial behavior in young adolescents	19
Chapter 3	The How I Think Questionnaire as a tool for evaluation of prevention programs and diagnostic purposes in young adolescents	37
Chapter 4	Effects of EQUIP for Educators on students' self-serving cognitive distortions, moral judgment, and antisocial behavior	53
Chapter 5	EQUIP for Educators: the longitudinal effects of a school based universal prevention program on antisocial behavior and moral cognitions in young adolescents	73
Chapter 6	General Discussion	101
	References	113
	Summary	125
	Samenvatting (Summary in Dutch)	129
	Dankwoord (Acknowledgements)	133
	Curriculum Vitae	135

Chapter 1

General Introduction

Antisocial behavior

Antisocial behavior can be conceptualized as outward behavior that either directly or indirectly harms others through the violation of important moral or social norms, and includes aggressive and delinquent acts (Barriga, Morrison, Liao, & Gibbs, 2001; Liu, 2004). The high prevalence of antisocial behavior among youth is of great societal concern (Redding, Goldstein, & Heilbrun, 2005), since it forms a substantial risk factor for severe problem behavior, psychopathological disorders, and crime later in life (Moffitt, 2003; Overbeek, Vollebergh, Meeus, Engels, & Luijpers, 2001). Therefore, the amount of research concerning the development and persistence of antisocial behavior increased substantially during the past decades (Verhulst, 2008).

Many theories of antisocial behavior adopted a developmental perspective, which means that origins of antisocial behavior in adulthood are supposed to be found in childhood and adolescence. However, developmental models differ with respect to the number of developmental trajectories in antisocial behavior that individuals are assumed to follow (Lacourse et al., 2002). Gottfredson and Hirschi (1990) proposed a single pathway, in which low levels of self-control in childhood are assumed to lead to delinquency later in life. Moffitt (1993; 2003) suggested the distinction in two pathways of antisocial behavior: the life course persistent pathway, starting in childhood and remaining stable throughout the life course, and the adolescence-limited pathway, starting during adolescence and (mostly) ending in early adulthood. Loeber (1991) made a division into three pathways, based on type of antisocial behavior (overt, covert, and conflict with authority). However, these theories were largely based on longitudinal research using techniques that allowed investigation for the presence of a priori defined groups (Lacourse et al., 2002). More recently, prospective techniques were used to examine developmental taxonomies. Results from these longitudinal studies display important heterogeneity in the development of antisocial behavior, and indicate that there are more developmental trajectories for antisocial behavior during adolescence, for both males and females (Lacourse et al., 2002; Odgers et al., 2008).

An underlying assumption of the developmental models is that diverse causal factors lead to different trajectories, sometimes for different types of antisocial behavior. Although there is a large body of research indicating diverse factors associated with the development and persistence of antisocial behavior, there is no absolute consensus regarding the most salient causes of antisocial behavior in adolescence (Odgers, Vincent, & Corrado, 2002). Since the multidimensional nature of antisocial behavior is better appreciated recently, the multi factorial integrative approach is increasingly used to study antisocial behavior, recognizing the role of both biological or individual and social or environmental factors (Liu, 2004), as well as the interaction between those two (Rutter, Moffitt, & Caspi, 2006). It is important to

gain insight into these age and developmental phase related factors and processes underlying antisocial behavior, in order to contribute to effective preventive strategies and ultimately alter the negative spiral of antisocial behavior (Boendermaker, 2008; Boxer, Goldstein, Musher-Eizenman, Dubow, & Heretick, 2005). In the current thesis, the emphasis will be on the evaluation of a prevention program that focuses on dynamic individual factors in a social setting – a lack of skills and competencies in at risk youth that participate in a negative peer culture – which are assumed to facilitate antisocial behavior. The aim of this thesis is to investigate the effectiveness of a program to alter the negative developmental pathway of antisocial behavior by reducing these delays or limitations of at risk youth.

Prevention of antisocial behavior

The prevalence and persistence of antisocial behavior among youth as described above (Flannery et al., 2003; Wilson & Lipsey, 2007) stresses the need for evidence-based prevention programs aimed at stabilizing or reducing antisocial behavior (Boxer et al., 2005). Although many preventive interventions concerning antisocial behavior have been developed and applied in recent years (Gottfredson & Gottfredson, 2002), relatively few have been thoroughly evaluated. Van Yperen and Van Bommel (2009) provide criteria to assess effectiveness and distinguish (preventive) intervention programs which are *promising* or *evidence-based*. The first criterion concerns the evidential value of the research design, based on diverse research characteristics (see Table 1). An intervention can be qualified as *promising* when the evidential value of the research design is moderate or quite strong and as *evidence-based* when the evidential value is strong. The second criterion is the number of available studies providing evidence for effectiveness.

Table 1. *Quality requirements for research (Van Yperen & Van Bommel, 2009)*

Evidential value	Research characteristics
Weak	Possibility of replication results
	Relation effects and aim/ population
	Use of reliable and valid instruments
	Inclusion of pre and post assessment(s)
	Performance of adequate statistical analysis
Moderate	Quasi-experimental design including control group
	High quality design
Quite strong	Implemented in practice or representative for practice
Strong	Follow up (min. 6 months) or other design
Very strong	Experimental design (random selection)

Although there are no golden rules, Van Yperen and Van Bommel (2009) used the following guideline: for the qualification of *promising* there should be at least three studies available with moderate to quite strong evidential value or two studies with strong to very strong evidential value. To be qualified as *evidence-based*, there should be at least two studies available with strong to very strong evidential value which provide evidence for effectiveness. The third criterion is the presence of a clear link between the effects and the aim and population of the study. Finally, the fourth criterion concerns the effect size, which should be in accordance with program format, and relevant in practice. Furthermore, in these evaluations, preventive intervention programs can be divided by program format, treatment modality, age of participants, and prevention setting.

Program format

Preventive interventions can be distinguished in *universal* interventions, *selective* or *targeted* interventions, and *clinical* or *indicated* interventions (Institute of Medicine, 1994). Universal interventions focus on entire populations such as complete schools or grade levels. Selective interventions focus on at risk subgroups already displaying precarious levels of problem behavior, which are in need for intervention. Clinical interventions focus on youth with detectable symptoms underlying antisocial behavior. The advantage of both selective and clinical preventive interventions is their high efficiency because of the focus on at risk youth. However, their success depends on correct identification of at risk adolescents, while the literature shows substantial rates of false positives (unnecessarily exposed to the intervention and the risks associated with labeling) or false negatives (not exposed to the intervention they could benefit from) (Offord, Kraemer, Kazdin, Jensen, & Harrington, 1998; Van Lier, Verhulst, & Crijnen, 2003).

Reviews concerning the most effective prevention program format are still inconclusive. A recent review by Limbos and colleagues (2007) shows that higher effectiveness concerning youth violence was reported as the level of intervention increased from universal to clinical. Wilson and Lipsey (2007) found in their meta-analysis that the most common and most effective program formats are universal interventions delivered to all the students in a classroom or school, and selected interventions for at risk students who participate in programs outside their classroom. However, since the universal programs included in the review by Wilson and Lipsey (2007) mainly used a cognitive approach and the selective interventions also used other approaches, it is not entirely clear whether effects can be attributed to the program format or to the approach.

Treatment modality

Furthermore, diverse approaches or treatment modalities (e.g. behavioral, cognitive and emotional oriented programs) can be distinguished in the literature concerning prevention of antisocial behavior. Recent review studies concerning the prevention of antisocial behavior in children and adolescents show that cognitive-behavioral approaches are most successful, showing positive effects (Bennett & Gibbons, 2000; Wilson, Gottfredson, & Najaka, 2001). Cognitive-behavioral therapy (CBT) is based on the assumption that dysfunctional patterns of thinking contribute to the development and persistence of antisocial behavior. By altering this biased thinking, it would be possible to modify antisocial aspects of the personality and consequent behaviors. Cognitive changes and behavioral changes are assumed to reinforce each other. In order to strengthen thoughts that lead to positive behaviors and to strengthen behavior due to the positive consequence of that behavior, CBT teaches new skills in areas where at risk youths show deficits, such as perspective-taking (Milkman & Wanberg, 2007). CBT is mostly taught in classroom settings, and sessions often include group exercises involving role-play, rehearsal, feedback, and generally follow a structured curriculum with detailed lesson plans (Milkman & Wanberg, 2007; Sukhodolsky, Kassinove, & Gorman, 2003). According to a meta-analysis concerning prevention of antisocial behavior by Wilson, Lipsey and Derzon (2003), multimodal and peer mediation programs were least successful, displaying smaller effect sizes compared to cognitive-behavioral programs. In addition, Wilson, Gottfredson and Najaka (2001) showed that non-cognitive-behavioral counseling, social work, and other therapeutic interventions were even consistently associated with negative effects.

Age of participants

Although the need for early prevention is illustrated by studies showing that antisocial behavior occurs along a developmental continuum of behavioral severity, meaning that antisocial behavior increases over time (Schaeffer, Petras, Ialongo, Poduska, & Kellam, 2003; Tremblay, 2006), and research indicating that early onset predicts antisocial behavior later in life, continuation of support in preventing problem behaviors across the transition from childhood to adolescence is of great importance as well. Adolescence is characterized by changes of substantial impact on the lives of children, e.g. adolescents' relationships to key adults are reconfigured (Eccles et al., 1993), and peers become of greater influence (Berndt, 1996). Adolescents' conformity to negative peer norms appears to be a major risk factor linked to negative outcomes, e.g. delinquency (Allen, Porter, & McFarland, 2006; Gest, Farmer, Cairns, & Xie, 2003). In addition, several studies show that antisocial behavior is most prevalent in adolescence (Krueger, 1999; Nagin & Tremblay, 1999), and the level of delinquency appears to increase during early and middle adolescence (Coie & Dodge, 1998; Moffitt, 2003). Although a substantial part of the adolescents adapt to these changes within a

reasonable time frame and only act out during this specific period in life – which is then being labeled as normative behavior by many researchers – prevention is needed for those who remain antisocial or are at risk. In addition, Bennett and Gibbons (2000) found a trend for child age to correlate positively with post-treatment effect size, suggesting that current child-based cognitive behavioral intervention programs targeting antisocial behavior are more effective for adolescents and older elementary-school aged children than for younger elementary-school aged children.

Prevention domain

Depending on the specific sample to be reached, prevention can take place in diverse domains. Several studies concerning the prevention of antisocial behavior emphasized the importance of family-based approaches (e.g. Côté, Vaillancourt, Barker, Nagin, & Tremblay, 2007; Farrington, 2005), and positive effects have been confirmed for parent management training in children and adolescents (Kazdin, 2005; Patterson, 2005). However, the preventive efforts in this setting mostly have a targeted character and are aimed at a selected group. Parent management training often takes place at individuals' homes or in small groups in community centers, which is costly and therefore limits the number of children or adolescents that can be reached. In addition, parent-focused interventions are associated with high attrition rates (DeMatteo & Marczyk, 2005).

Alternatively, schools provide an excellent domain to implement prevention programs targeting antisocial behavior (Cunningham & Henggeler, 2001), since they have access to a large number of youth for extended periods of time (Wilson & Lipsey, 2007). Schools play a significant role in shaping the lives of adolescents, and form a setting in which a lot of problem behavior (e.g. relational and physical aggression, vandalism) occurs (Cho, Hallfors, & Sánchez, 2005; Wilson et al., 2001). School based programs can be defined as 'interventions operated in a school building, by school staff or under school or school system auspices' (Wilson et al., 2001, p.251). Since the peer group forms an important context of development during adolescence, and it seems likely that peers have an important influence on adolescent behaviors (Ryan, 2001), interventions in school settings might be particularly beneficial. Regular school populations are heterogeneous and comprise substantial percentages of students showing prosocial skills and behavior, which can positively influence at risk students (Dishion, McCord, & Poulin, 1999).

EQUIP for Educators

EQUIP is one of the multi-component (including anger management, social skills, and social perspective taking, see below) interventions for adolescents with antisocial behavior problems, which was found to be effective for delinquent youth. Results of American research

using a preliminary version of the EQUIP program showed a substantial reduction of recidivism and significant improvements in institutional conduct after intervention (Leeman, Gibbs, & Fuller, 1993). The researchers concluded that EQUIP promoted therapeutic change in juvenile delinquents. In addition, findings of a Dutch study including incarcerated boys, showed a significant reduction of cognitive distortions and a more negative attitude towards antisocial behavior in the experimental group as compared to the control group (Nas, Brugman, & Koops, 2005). However, a follow up study could not provide evidence for a reduction in recidivism: there were no differences found in speed or seriousness of reoffending between adolescents in the experimental group and those in the control group (Brugman & Bink, 2009). Nonetheless, Brugman and Bink (2009) did establish a mediating effect of self-serving cognitive distortions on recidivism, since adolescents in the experimental group scoring low on egocentric bias reoffended later compared to those whose levels of self-serving cognitive distortions were average or high.

Based on these promising results, a preventive version of the original treatment program EQUIP for juvenile offenders (Gibbs, Potter, & Goldstein, 1995) was developed. This prevention program targeting antisocial behavior in students, 'EQUIP for Educators: Teaching youth (grades 5-8) to think and act responsibly' (DiBiase, Gibbs, Potter, & Spring, 2005), will be evaluated for the first time in the current thesis. EQUIP for Educators (EFE) is guided by a theoretical framework based on psycho education, referring to the teaching and training of skills, knowledge, and mature awareness required for competent daily living (DiBiase et al., 2005). By integrating a peer-helping approach (PPC, Vorrath & Brendtro, 1985) with the helping skills curriculum based on Aggression Replacement Training (Goldstein & Glick, 2001), EFE aims to equip youth to think and act more prosocially. Although EFE does not include mutual help meetings – which are part of the EQUIP program for juvenile delinquents (Gibbs et al., 1995) – , the PPC component is manifested in the curriculum by the interaction between youths, in which prosocial students function as an example. Youths who are at risk for the development of antisocial behavior often show limitations which can be defined in terms of the three D's: sociomoral developmental Delay – that is persistence into adolescence of immaturity in moral judgment and egocentric bias –, social cognitive Distortions, and social skill Deficiencies. By remediating these delays and deficiencies, a reduction in antisocial behavior is expected (Gibbs, Potter, Barriga, & Liau, 1996).

Sociomoral developmental Delay

Moral judgment can be defined in terms of justice and the capability of using moral criteria in the evaluation of behavior (Colby & Kohlberg, 1987). Delay in moral judgment can be seen as 'the persistence beyond early childhood of a pronounced degree of me-centeredness or

egocentric bias' (Gibbs, 1991, p.96) and implies immature moral judgment. Kohlberg (1984) originally proposed six invariant stages for the development of moral judgment, while Gibbs (2003) revised the stages of this theory and reduced the number of stages to four. In his revision, stage 5 and 6 are not regarded as truly developmental stages, but represent an existential phase. According to Gibbs (2003), stage 1 and 2 represent immature or superficial moral judgment, while with the transition to stage 3 moral judgment advances beyond superficiality to a mature understanding of moral norms and values. Mature moral thinking may function as a buffer against antisocial behavior (Stams et al., 2006).

Social cognitive Distortions

According to social-cognitive theories, people act upon their interpretation of social events and antisocial behavior is based on deficiencies in interpreting these events. Barriga, Gibbs, Potter, and Liao (2001, p.1) define cognitive distortions as 'inaccurate or biased ways of attending to or conferring meaning upon experiences'. Barriga, Gibbs and colleagues (2001) focus on those cognitive distortions that, at a high level of prevalence, facilitate aggression and other types of antisocial behavior. They distinguish four types of those self-serving cognitive distortions: self-centeredness (according status to one's own views and needs to such a degree that views of others are scarcely considered), blaming others (misattributing blame to external sources), minimizing/ mislabeling (depicting antisocial behavior as causing no real harm or being acceptable) and assuming the worst (gratuitously attributing hostile intentions to others). Self-centeredness is considered as a primary cognitive distortion that precedes antisocial behavior. Blaming others, minimizing/mislabeling and assuming the worst are secondary cognitive distortions that neutralize feelings of guilt and blame that accompany antisocial behavior (Cromwell & Thurman, 2003; Gibbs, 2003).

Social skill deficiencies

Social skills are crucial for social action and can be defined as 'balanced and constructive behaviors in difficult interpersonal situations' (Gibbs et al., 1995, p.165). Since at risk youth typically do not know what specific steps are involved in constructive problem solving and have problems with social perspective taking (Dodge & Rabiner, 2004), social skills training is intended to displace out-of-control destructive responses with incompatible constructive ones.

Evaluation of EQUIP for Educators

Based on the literature concerning effectiveness of prevention programs as described above, EQUIP for Educators is in the current thesis implemented as a school based universal prevention program in highschools, or more specifically: prevocational secondary education in the Netherlands. The Dutch educational system is characterized by tracking after

elementary school. Various levels or schooltypes are distinguished, based on students' cognitive abilities. Prevocational secondary education represents the lowest educational level in the Netherlands, which is associated with a higher risk for antisocial behavior (Boxer et al., 2005; Farrington, 2005). This population is of interest for the evaluation of preventive efforts, since higher rates of problem behavior may indicate a stronger need and create better opportunities for improvement (Flannery et al., 2003; Stoolmiller, Eddy, & Reid, 2000). Implementation of EFE started in eighth grade – which equals second grade in secondary school according to the Dutch educational system – and included complete classes, facilitating the preconditions of confidentiality and feeling comfortable in the group.

Design and participants

Recruitment of prevocational secondary schools took place in the Western, mostly urbanized, part of the Netherlands. Participation was granted by the ability of the school to meet research requirements, e.g. participation with at least three eighth grade classes, the ability to have the teachers well-trained, and the ability to implement the program in classroom settings for two school hours a week over a period of 16 weeks (32 EFE meetings in total). Data collection within the research project occurred over a period of twenty months, starting with two data waves separate from and preceding the evaluation studies, and followed by five data waves representing the pre and post assessments of the evaluation studies. A total of eleven schools were included in the research project. Seven of these schools agreed to implement EFE according to the research requirements and participated in the intervention condition. In addition, these schools allowed for data collection in their former eighth grades in the year preceding the implementation and evaluation of EFE. Furthermore, two schools participated in the non-intervention condition, following the data waves of the intervention condition. Finally, two schools which intended to implement EFE but eventually could not meet the research requirements only participated in the data waves during the college year preceding the evaluation.

The first two data waves were used to outline the theoretical background (Chapter 2), and to assess the psychometric quality and the suitability of the How I Think Questionnaire, an instrument to determine the level of self-serving cognitive distortions, as a tool for evaluation and diagnostic purposes (Chapter 3). The pre and post assessments were spread over three college years to gain more insight in the development of adolescents and effects of EFE over time (Chapter 4 and 5). Depending on the cohort schools were assigned to, students were assessed once at the end of seventh grade, followed by three assessments in eighth grade, or three times in eighth grade and once at the beginning of ninth grade.

Aim and outline of the thesis

The aim of the present thesis is to determine the effects of EQUIP for Educators as a prevention program for prevocational secondary school students. In the following chapters, four empirical studies are presented.

In order to explain the development and persistence of antisocial behavior, *Chapter 2* presents a two wave longitudinal model including moral judgment, self-serving cognitive distortions, and perception of community. A reduction of self-serving cognitive distortions and improvement of moral judgment, perception of community, and social skills are assumed to lead to a reduction of antisocial behavior. It is examined how these moral cognitions and antisocial behavior are related, both cross-sectionally and over time. It is examined to what extent moral cognitions and antisocial behavior are stable over time, and whether moral cognitions predict antisocial behavior or vice versa. Sex and ethnic differences are taken into account in these models.

Chapter 3 focuses on the psychometric quality of the HIT-Q – a questionnaire used to assess one of the main variables in the current thesis: self-serving cognitive distortions – and the suitability of this instrument as a tool for diagnostic purposes and the evaluation of prevention programs. Validity and reliability are investigated in a large mixed-gender sample, comprising a substantial percentage of at risk students. In addition, the relevance of the anomalous response scale – designed to detect ‘suspect responding’ – is investigated. Furthermore, profile norms classifying students into normal, borderline – clinical, or clinical groups are reexamined.

Next, two studies are carried out in order to find out whether EQUIP for Educators can be labeled as an evidence-based program, preventing antisocial behavior in prevocational secondary school students.

In *Chapter 4*, the effects of EQUIP for Educators (EFE) on prevalence of antisocial behavior, attitude towards antisocial behavior, self-serving cognitive distortions, and moral judgment are examined using a quasi-experimental pre-test/ post-test with control group design. In addition, an overview of the prevalence of different forms of self reported antisocial behavior and a categorisation of the seriousness of self-serving cognitive distortions and teacher reported antisocial behavior is provided to determine the suitability of the present sample for evaluation research.

Chapter 5 focuses on both the initial and long term (up till 9 months after the intervention) effects of EFE – implemented as a universal prevention program – on students’ self-serving cognitive distortions, moral judgment, social skills, attitude towards antisocial behavior, and prevalence of antisocial behavior. A sophisticated longitudinal design with four measurement occasions is used, which is more sensitive to trends in individual trajectories and the way these are affected by the intervention than a pretest-posttest design. Complex latent growth curve models for both the baseline trajectories and the effects of the intervention on these trajectories are combined to assess developmentally relevant effects. Models for each variable separately, models with sex and ethnic differences accounted for, and multivariate models (to capture the relevant relationships between the trajectories for the variables) were investigated. In addition, effects of EFE on teacher reported behavior, also controlling for sex and ethnic differences, are investigated using repeated measures.

Finally, in *Chapter 6*, a general discussion is presented in which findings from the previous studies are summarized and integrated. Implications for practice as well as directions for future research are discussed.

Chapter 2

Moral cognitive processes explaining antisocial behavior in young adolescents

Abstract

This study addresses the longitudinal relationships between three kinds of moral cognitions – self-serving cognitive distortions, moral judgment, perception of community – and antisocial behavior in young adolescents. Aims were to gain insight in direct and indirect relationships, stability, and causality. The sample included 724 students ($M\ age = 14.52$, $SD = .67$) from prevocational secondary schools in the Netherlands. Both self reports and teacher reports were filled out twice, with a time-interval of four months. Students exhibited high rates of aggression, vandalism, and stealing, indicating that they form an at risk group for antisocial behavior. Positive associations were found between self-serving cognitive distortions and antisocial behavior and between moral judgment and perception of community, while negative associations were found between perception of community and both self-serving cognitive distortions and antisocial behavior. Longitudinal structural equation models established a moderate to high stability of the moral cognitions and antisocial behavior, and indicated that self reported antisocial behavior primarily preceded self-serving cognitive distortions. Although moderation by sex and ethnicity was established, differences between the groups appeared to be modest. The consequences of these findings for prevention and treatment of antisocial behavior are discussed.

Introduction

Numerous studies in the past decades investigated the link between moral cognitions and moral behavior, mostly focusing rather narrowly on a single aspect of moral cognition: moral judgment. Moral judgment can be defined in terms of justice and the capability of using moral criteria in the evaluation of behavior (Colby & Kohlberg, 1987). Barriga, Morrison, Liao and Gibbs (2001) moved beyond this traditional approach and introduced a model using various moral cognitive variables – moral judgment and self-serving cognitive distortions – explaining antisocial behavior. Antisocial behavior can be conceptualized as outward behavior that either directly or indirectly harms others through the violation of important moral or social norms, and includes aggressive and delinquent acts (Barriga, Morrison et al., 2001; Liu, 2004). The present study builds on the findings of Barriga and colleagues (2001) and adds another aspect of morality – perception of community in school (a characteristic of the moral atmosphere) – to the explanatory model of antisocial behavior, using a longitudinal design with a sample of young adolescents.

Moral judgment and perception of community

The most influential psychological theory on moral development was formulated by Lawrence Kohlberg (1984). Moral judgment can be seen in the context of a person explaining and justifying a certain course of action. One of the most important tasks adolescents have to fulfill, in order to internally regulate their behavior, is to progress from a morality focused on outcomes to a morality based on ideal reciprocity. According to Kohlberg (1984), moral judgment develops through an invariant sequence of six hierarchically ordered stages. Each stage is considered to be more advanced than the preceding stage in providing more prescriptive and universally acceptable solutions to moral issues. When children follow a normal developmental pathway, they reach stage 3 (characterized by ideal reciprocity: “treat others as you would like them to treat you”) between age twelve to sixteen (Colby, Kohlberg, Gibbs, & Lieberman, 1983).

Gibbs (2003) presents a revision of Kohlberg’s stage sequence in which moral development is limited to the first four stages. Stage 5 and 6 are not regarded as truly developmental stages, but represent an existential phase that broadens the standard stage 3 and stage 4 reasoning. The immature level consists of stages 1 and 2, while the mature level comprises stages 3 and 4. Stage mixture is accepted and stages are described from an information processing viewpoint. Gibbs emphasizes the superficial and pronounced egocentric bias of the immature stages (1 and 2) as risk factors for antisocial behavior in adolescents. Breaking the law would more readily seem morally acceptable at the self-centered lower stages. In contrast, the deeper social understanding characterizing the higher stages may buffer against delinquent behavior, as it is imperative that the well-being of others will be taken into account (Stams et al., 2006).

Many studies have established a negative relationship between moral judgment and antisocial behavior, showing that delinquent adolescents reason at a lower moral stage than their non-delinquent peers. Nevertheless, there is a broad range in the magnitude of this relationship, depending on the seriousness of antisocial behavior. In general only low correlations are found between moral judgment and antisocial behavior in non-delinquent samples (Barriga, Morrison et al., 2001; Kuther & Higgins - D'Alessandro, 2000), while moral judgment is more strongly related to antisocial behavior in samples including incarcerated delinquents (Stams et al., 2006).

Although Kohlberg originally claimed that moral judgment was the only moral variable predicting moral behavior, he later introduced the moral atmosphere as an important environmental characteristic that influences (lack of) moral behavior (Higgins, Power, & Kohlberg, 1994). The perception of schools' moral atmosphere is characterized by judgments concerning school routines, which reflect an individual student's perception on the social relations for the majority of the students within school (Power, Higgins, & Kohlberg, 1989). An important aspect of the individual perception of moral atmosphere in school that will be investigated in this study is the 'perception of community', the extent to which a student experiences his/ her school as a social community. A community is characterized as '... a social organization whose members know, care about and support one another, have common goals and a sense of shared purpose, and to which they actively contribute and feel personally committed' (Solomon, Watson, Battistich, Schaps, & Delucchi, 1996).

It should be noted that, like moral judgment, perception of the moral atmosphere can be studied as an individual characteristic (Beem, Brugman, Høst, & Tavecchio, 2004). Beem and colleagues (2004) concluded that students' perception of school moral atmosphere as measured by a questionnaire should not be regarded primarily as a shared perspective among students within a school like a moral school culture, but rather as an instance of the social competence of the individual student.

Self-serving cognitive distortions

Several theories that aim to explain the origin, development and maintenance of antisocial behavior have stressed the importance of self-serving distortions in social cognitions (Bandura, 2002; Crick & Dodge, 1996). According to social-cognitive theories, people act upon their interpretation of social events and antisocial behavior is based on deficiencies in interpreting these events, i.e. cognitive distortions. Barriga, Gibbs, Potter, and Liao (2001, p.1) define cognitive distortions as 'inaccurate or biased ways of attending to or conferring meaning upon experiences'. Barriga, Gibbs and colleagues (2001) focus on those self-serving cognitive distortions that, at a high level of prevalence, facilitate aggression and other types of antisocial behavior. According to Gibbs (2003), a low moral judgment stage does not necessarily lead to antisocial behavior, unless it is combined with a high degree of cognitive

distortions. Cognitive distortions are assumed to block moral judgment development because one does not consider oneself to be responsible for one's delinquent behavior. In both delinquent and non-delinquent groups of adolescents, self-serving cognitive distortions relate positively to self reported antisocial behavior and account for a substantial part of the explained variance (Barriga, Morrison et al., 2001; Nas, Brugman, & Koops, 2008).

Present study

The present study uses a sample comprising students attending prevocational secondary education in the Netherlands. The Dutch educational system distinguishes various educational levels after elementary school, based on students' cognitive functioning. Prevocational secondary education represents the lowest educational level in the Netherlands, comprising mostly students with low socioeconomic background and high rates of students from ethnic minority groups. Although several studies (Boxer, Goldstein, Musher-Eizenman, Dubow, & Heretick, 2005; Farrington, 2005) show that a low educational level is associated with a higher risk for antisocial behavior, Barriga et al. (2001) investigated the influence of moral cognitions on antisocial behavior using highly educated and older respondents, e.g. college or university students. We propose that knowledge on the strength of effects of moral cognitive variables in at risk groups is of even more relevance for the development and implementation of preventive intervention programs.

To compensate for possible biases in self reports, we also include teacher reports in assessing antisocial behavior in students. Most studies investigating the relationship between self-serving cognitive distortions and antisocial behavior have been based on self reports only (Lardén, Melin, Holst, & Långström, 2006; Liau, Barriga, & Gibbs, 1998). It is well known that self reported antisocial behavior correlates only weakly with teacher reported antisocial behavior (Achenbach, Dumenci, & Rescorla, 2002).

The first aim of this study is to gain a better insight into the relationships between self-serving cognitive distortions, moral judgment, perception of community, and antisocial behavior. We hypothesize that a higher stage of moral judgment will be related to a more positive perception of community (cf. Høst, Brugman, Tavecchio, & Beem, 1998). Furthermore, a lower stage of moral judgment and a more negative perception of community will be related to both a higher level of self-serving cognitive distortions and to higher levels of antisocial behavior. It is further hypothesized that a high level of self-serving cognitive distortions will be related to high levels of antisocial behavior.

The second aim is to elaborate the understanding of the relationships between moral cognitions and antisocial behavior over time. So far, these relationships have been investigated cross-sectionally rather than longitudinally. In this study, follow-up data are used to investigate 1) to what extent moral cognitions and antisocial behavior are stable over time, and 2) whether moral cognitions predict antisocial behavior or vice versa. Firstly, it is

hypothesized that both moral cognitions and antisocial behavior are rather stable over a time period of four months. Secondly, it is hypothesized that moral cognitions will mainly precede antisocial behavior (e.g., Gibbs, 2003).

Sex and ethnic differences are taken into account in the analyses as moderators, since numerous studies established higher rates of antisocial behavior for males than females (e.g., Côté, Vaillancourt, Barker, Nagin, & Tremblay, 2007) and ethnicity has been found to be an important characteristic in studies of antisocial behavior in the Netherlands (Van der Laan & Blom, 2006). Ethnic status can be defined as an easily identifiable characteristic that implies a common cultural history with others possessing the same characteristic. The most common ethnic ‘identifiers’ are race, religion, country of origin, language and/ or cultural background (Okazaki & Sue, 1998). In the present study, a student is characterized as being part of an ethnic minority group when at least one of his/ her parents is born in another country.

Method

Participants

Participants were 724 students from eight preparatory vocational secondary schools located in both highly urbanized and rural parts of the Netherlands. All students attended second grade, which equals eight grade in the American school system. Since it is possible to repeat a class in the Netherlands or start later in a grade based on cognitive functioning, the age range is rather large, representing students from 13.3 to 17.2 years (ages 13-14, 28%; ages 14-16, 70%; ages 16 -17, 2%). It should be noted that the mean age was 14.52 ($SD = .67$) and only 6.8% of the students was older than 15.5. At the first wave, 667 students (375 boys) filled out the self report questionnaires, at the second wave 638 students (361 boys) did. Of these, 581 students (335 boys) participated in both waves of data collection. Additionally, teacher reports were completed for 624 students at the first wave and for 508 students at the second wave and included 18 students who did not complete the self reports at either wave. The gap between the numbers of respondents on teacher report at the two waves is largely due to the fact that all second grade teachers from one of the participating schools omitted to fill out the questionnaires at the second wave. Since there were no differences between this school and the other schools on rates of self reported antisocial behavior ($F(7, 664)_{\text{wave 1}} = 1.90, p > .05$), it is expected that this will not influence the results.

Class-size ranged from 9 to 28 students. A majority (62%) of the participants belonged to a minority group, meaning that at least one parent had a different country of origin than the Netherlands (e.g. Morocco, Surinam). By far, most of these parents had a non-western background (98.9%). Socioeconomic status was ascertained by asking about their father’s and mother’s highest completed education and appeared to be largely low, i.e. elementary school or secondary education (55.4% mother, 52.9% father) or moderate, i.e. non-university higher education (38% mother, 36.7% father). Note that the reported percentages are valid

percentages, meaning that missing values were not taken into account. The percentages of students who did not know the educational level of their parents were high (57.4% for mother data, 58.9 % for father data).

Procedure

Students were asked to fill out the same questionnaires at two time points: in January (Wave 1) and in June (Wave 2) of the same year. After a brief oral and written instruction, they completed the self report questionnaires within classroom settings. Students were encouraged to ask the supervisor for help when needed. Parents received written information about the project, participation was granted by means of passive consent. Participants were assured of the confidentiality of their responses. Small rewards (pencils) were provided.

Measures

The Antisocial Behavior Questionnaire. The Antisocial Behavior Questionnaire consisted of 8 items selected from self report scales used in previous research (Leenders & Brugman, 2005). Items (e.g. 'Beating or kicking someone, did you conduct this behavior?') were scored on a 4-point scale (1 = no, 2 = once, 3 = sometimes, 4 = often). The mean scale score was used in the analysis, with scores above 1 indicating that the respondent committed, at least once, one or more forms of antisocial behavior. Reliability and stability over time were both high (see Table 1).

Teacher's Report Form (Achenbach, 1991). The Dutch translation of the externalizing behavior scale from the Teacher's Report Form (Verhulst, Ende, & Koot, 1997) was used to assess antisocial behavior as perceived by teachers. The questionnaire comprised 34 items on delinquency and aggression such as 'Fights a lot' or 'Threatens other people', all of which are rated as 0 = not true (as far as I know); 1 = somewhat true; and 2 = very true or often true. Teachers rated problems based on students' functioning over the preceding two months. Total scores were computed and based on national norms for peers of the same age range and sex. Based on guidelines provided in the Teacher's Report Form Manual (Verhulst et al., 1997), students were assigned to normal (scores until the 67th percentile), borderline (scores between the 67th and 70th percentile) and clinical range (scores above the 70th percentile) groups. Both international and Dutch studies established high validity and reliability of the instrument (Rescorla et al., 2007; Verhulst et al., 1997). These results were replicated in the present study, with Cronbach's alphas of .96 at both waves.

How I Think Questionnaire. To measure self-serving cognitive distortions the How I Think Questionnaire (Barriga, Gibbs et al., 2001) was used. The HIT comprises 39 items addressing four types of cognitive distortions (i.e. 9 self-centered, 10 blaming others, 9 minimizing, 11 assuming the worst items), 8 anomalous response items (AR) designed to screen for suspect responding (for example: "In the past, I have lied to get myself out of

trouble"), and 7 positive filler items to encourage full use of the scale. We replaced the 7 positive fillers with 11 social desirability items (for example: "I am always honest") based on the Marlowe-Crowne questionnaire (Crowne & Marlowe, 1960). While the positive fillers were not meant to result in a meaningful score, the social desirability items do and fulfill the same function, i.e. to compensate for the rather negative formulated cognitive distortion items. Internal consistency of the social desirability subscale was found to be high (see Table 1).

Each of the 39 cognitive distortion items refers to one or another of four categories of conduct disorder and oppositional defiant disorder derived from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (American Psychiatric Association, 2000): disrespect for rules, laws or authorities (opposition defiance); physical aggression; lying; stealing. For example, the item "I can't help losing my temper a lot" represents a blaming others cognitive distortion item applied to an opposition defiance behavioral referent. Participants responded along a 6-point Likert scale (from 'disagree strongly' to 'agree strongly'), with responses to individual items of 4 (slightly agree) or higher indicating cognitive distortion and responses of 3 (slightly disagree) or lower indicating absence of distortion. Mean scale and subscale scores were used in the analysis. Based on the American normative sample (comprising non referred youth), students were assigned to normal, borderline - clinical or clinical range groups (Barriga, Gibbs et al., 2001).

Nas et al. (2008) found that the Dutch translation of the instrument had good validity and reliability. The reliability of the HIT in the present study was high, with Cronbach's alphas ranging from .72 to .81 on subscale level and a Cronbach's alpha of .92 (Wave 1) and .93 (Wave 2) for the overall score (see Table 1).

Moral Judgment Rating Task. This task assesses stage of moral judgment development as defined by Kohlberg (Colby & Kohlberg, 1987). The Moral Judgment Rating Task (MJRT) consists of two times short dilemmas containing two conflicting moral values, followed by a series of reactions favoring the first value and a series of reactions favoring the second value. The dilemmas and the quotes are derived from the Moral Judgment Interview (MJI) manual (Colby & Kohlberg, 1987) and were used earlier in the Moral Judgment Sorting Task (Boom, Brugman, & Van der Heijden, 2001). Quotes are stage typed representing stages 1 to 4 and in-between stages as defined in the MJI manual. This results in 4 sets of 7 items, which equals 28 items. The task is comparable to (and inspired by) the SROM-sf (Basinger & Gibbs, 1987), however, the response format is different in order to allow a Rating Scale analysis (Wright & Masters, 1982). Students categorized the quotes as 'for wiser people' (1), 'worthy of respect' (2), or 'childish' (3).

Rating scale analysis was used to assess how well the task was performed, because classical reliability (Cronbach's alpha) seems less appropriate since it assumes all items measure the same construct whereas our items are designed to reflect different stages.

Validity was supported in that, for full stages 1 to 3, the ordering in terms of difficulty estimates corresponded with the ordering proposed by Kohlberg (although separation between stages was weak). The higher the stage the quote represents, the more likely it is that it is judged to be wiser, keeping ability constant; and, the higher ability of the respondent is, the more likely an assertion is judged to be childish, keeping stage of quote constant. Since we found that stages higher than stage 3 are out of reach for most participants in our relatively young sample we decided to omit items reflecting these stages from further use in the analysis. In addition, the analysis revealed that the first set of quotes gave mixed results, not contributing positively to the expected ordering; therefore we omitted this set also from further use in the analysis. For further modeling we did not use the rather complex Rasch estimates and instead used the mean score over the items per person, since there is a strong correlation between raw mean scores and individual Rasch scores according to the general idea.

School Moral Atmosphere Questionnaire: Perception of Community. The scale ‘Perception of community’ of the School Moral Atmosphere Questionnaire (Høst et al., 1998) was used to assess students’ perception of the shared understanding of the school as a community. Previous research showed an item total correlation between the perception of community score and the total score for perception of the moral atmosphere of .69 (Høst et al., 1998). The scale perception of community contains 24 items with a 5-point Likert scale (totally disagree to totally agree). For example: ‘You can trust other students at this school’. Negative items were recoded and means were computed. Reliability and stability of the measure over time was moderate, see Table 1.

Table 1. *Psychometric characteristics of the measures and changes over time*

	Reliability			Stability		Changes over time		
	<i>n</i>	<i>a 1</i>	<i>a 2</i>	<i>N</i>	<i>r</i>	<i>M 1</i>	<i>M 2</i>	<i>t</i>
Antisocial Behavior Questionnaire								
Self reported antisocial behavior	8	.78	.73	581	.75**	1.43	1.33	6.53**
Teacher’s Report Form								
Teacher reported antisocial behavior	34	.96	.96	409	.64**	8.57	8.42	.30
How I Think Questionnaire								
Self-serving cognitive distortions	39	.92	.93	568	.75**	2.68	2.51	8.15**
Anomalous Response	8	.73	.74	575	.64**	3.57	3.59	-.81
Social Desirability	11	.65	.74	572	.56**	1.09	1.09	.31
Moral Judgment Rating Task								
Moral Judgment ¹	12	-	-	524	.23**	2.18	2.05	8.52**
School Moral Atmosphere Questionnaire								
Perception of community	24	.61	.69	506	.48**	3.17	3.47	-16.43**

¹ = Included only stage 1-3 and excluded first set of 7 assertions

** $p < .001$

Plan of Analysis

The descriptive statistics present attrition analysis and an overview of the prevalence of antisocial behavior and self-serving cognitive distortions. Longitudinal structural equation models were used to investigate relationships between moral cognitions and antisocial behavior, stability over time, and cross-lagged paths between the first and second wave (to find out whether moral cognitions predicted antisocial behavior or vice versa).

Results

Descriptive statistics

Missing values

Due to time constraints at assessment, the number of missing values fluctuated on the different measures used in this study. Moreover, for all measures individual scale scores were only computed if at least 80 percent of the items was scored. Attrition of scale scores ranged from 0.5% to 12% at the first wave and from 0.15% to 4.5% at the second wave.

For generalization purposes, univariate analyses of variance were conducted to find out whether students who were present at the data waves and filled out the questionnaires differed on the teacher reports of antisocial behavior from students who were absent at both data waves. There were no significant differences between these groups at either wave (respectively $F(1, 620)_{\text{wave 1}} = .66, p > .05$; $F(1, 504)_{\text{wave 2}} = 2.35, p > .05$).

Prevalence of antisocial behavior and self-serving cognitive distortions

In order to gain insight in the prevalence of the different forms of self reported antisocial behavior, some descriptive results are presented. Physical aggression such as beating and/ or kicking someone was reported by more than half of the respondents (56%). Vandalism was also common within this sample, with 43% of the students reporting to have deliberately destructed something and 35% reporting they started a fire. Stealing minor things from a store was reported by 23% of the students, while serious forms of stealing such as robbery and stealing expensive things from a store were the least reported crimes (5% and 7%, respectively).

Most respondents scored within the normal range on teacher reported antisocial behavior (66.2%) as well as on self-serving cognitive distortions (57.3%) at the first wave. Although few were assigned to the borderline group, a substantial percentage of the students showed high rates of antisocial behavior and/ or self-serving cognitive distortions and was marked as 'clinical'. In total 11% of the students could be assigned to the clinical group of both teacher reported antisocial behavior and self-serving cognitive distortions. Seventeen percent of the respondents could be divided into the normal range of teacher reported antisocial behavior and into the clinical range of self-serving cognitive distortions, while 11%

scored within the normal range of self-serving cognitive distortions and within the clinical range of teacher reported behavior.

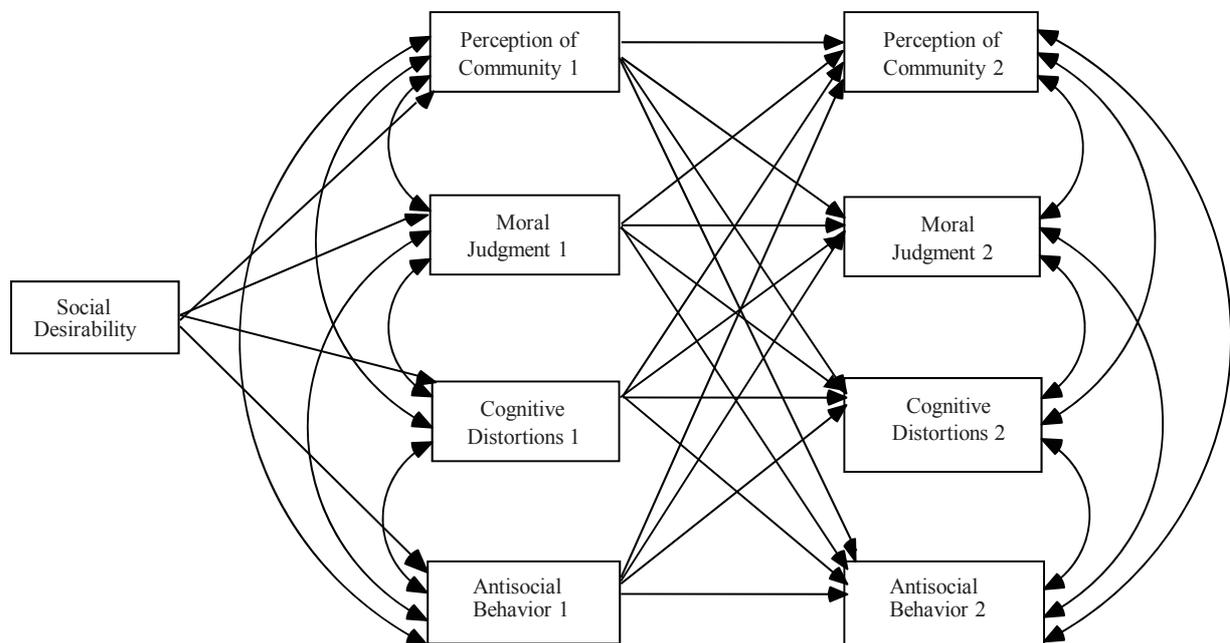


Figure 1. *Conceptual model longitudinal relationships moral cognitions and antisocial behavior*

Longitudinal relationships

To investigate the relationships between moral cognitions and antisocial behavior over time, structural equation models including all the main variables assessed at both waves, and social desirability as a control variable, were constructed (see Figure 1). After establishing an acceptable fit (Hu & Bentler, 1999) for the total self report model ($\chi^2(4) = 6.13, p = .19$; CFI = .99; RMSEA = .03) and the total teacher report model ($\chi^2(5) = 12.11, p = .03$; CFI = .99; RMSEA = .04), moderation by sex and ethnicity was investigated using multigroup analyses (Kline, 1998). Significant differences between the chi square values favoring the unconstrained models were found, confirming moderation by sex and ethnicity (see Table 2). Parameter estimates for the unconstrained models, revealing differences between boys and girls, and between students from Dutch ethnic background and students from ethnic minority groups, will be reported.

Table 2. Multigroup analysis of model fit

Model	Fit statistics - Longitudinal model											
	Self reported antisocial behavior					Teacher reported antisocial behavior						
	χ^2	$\Delta \chi^2$	df	p	CFI	RMSEA	χ^2	$\Delta \chi^2$	df	p	CFI	RMSEA
<i>Sex differences</i>												
Constrained	112.93	-	40	.00	.96	.05	89.82	-	41	.00	.96	.04
Unconstrained	14.30	98.63	8	.07	.99	.03	28.09	61.73	10	.00	.99	.05
<i>Ethnic differences</i>												
Constrained	62.01	-	40	.01	.98	.03	102.51	-	41	.00	.95	.05
Unconstrained	10.33	51.68	8	.24	.99	.02	19.37	83.14	10	.04	.99	.04

Note. Unconstrained = all paths left free, Constrained = stability paths, correlations and cross-lagged paths constrained

Contemporaneous relationships

Firstly, the relationship between moral judgment and perception of community showed a positive trend at the first wave, with significant associations for boys and students from ethnic minority groups. Secondly, self-serving cognitive distortion was negatively associated with moral judgment for boys at the first wave and for students from Dutch ethnic groups at both waves, and with perception of community at both waves for both sexes and ethnic groups.

Thirdly, although we could not establish systematic significant associations between antisocial behavior and moral judgment, results showed a trend of negative associations at the first wave for both sexes and ethnic groups. These findings could not be replicated at the second wave, where positive relations for boys and students from ethnic minority groups were found. Furthermore, negative associations were found between perception of community and self reported antisocial behavior (at the first wave stronger for girls and students with a Dutch ethnic background and at the second wave stronger for boys).

Finally, results showed strong positive relationships between self-serving cognitive distortions and self reported antisocial behavior at the first wave. This relationship was considerably lower for teacher reported behavior, and was in this respect even non-significant for students from ethnic minority groups. At the second wave, paths between self-serving cognitive distortions and both self and teacher reported antisocial behavior dropped substantially, due to strong stability of the measures. Although the relationship in the self report model remained moderate to high (especially for boys), the path between self-serving cognitive distortion and teacher reported behavior became non-significant, except for students with a Dutch ethnic background (Table 3).

Stability

The longitudinal models established a strong stability for self-serving cognitive distortions and antisocial behavior over time. Multigroup analyses showed stronger stability paths for girls on self-serving cognitive distortions and self reported antisocial behavior and stronger stability for students with a Dutch ethnic background on both self and teacher reported antisocial behavior (Table 3). Stability of moral judgment and perception of community was respectively low to moderate, with stronger stability paths for girls and students with a Dutch ethnic background.

Cross-lagged paths

Cross-lagged paths were included in the longitudinal models to investigate the direction of the associations between moral cognitions and antisocial behavior. Results for the self report model showed a positive relationship between antisocial behavior measured at the first wave and self-serving cognitive distortions measured at the second wave (Table 3). This association was found to be stronger for boys and appeared to be reciprocal for girls. However, these findings were not replicated for teacher reported antisocial behavior. Teacher reported

Table 3. *Multigroup analysis relationships between moral cognitions and antisocial behavior*

Paths	<i>Self reported model</i>				<i>Teacher reported model</i>			
	Boys	Girls	Dutch	Minority	Boys	Girls	Dutch	Minority
<i>Correlations</i>								
MJ – PC	.15*	.10	.11	.14*	.15*	.10	.10	.15*
MJ – CD	-.17*	-.07	-.18*	-.09	-.18*	-.06	-.18*	-.10
MJ – AB	-.06	-.06	-.10	-.05	.00	-.06	-.10	.02
PC – CD	-.35*	-.46*	-.53*	-.34*	-.34*	-.46*	-.53*	-.34*
PC – AB	-.16*	-.24*	-.29*	-.17*	.07	-.05	-.20*	.09
CD – AB	.56*	.42*	.54*	.54*	.11*	.17*	.21*	.10
SD – MJ	-.05	-.06	-.04	-.03	-.05	-.07	-.05	-.03
SD – PC	.20*	.20*	.18*	.21*	.21*	.19*	.15*	.21*
SD – CD	-.11*	-.27*	-.20*	-.18*	-.10	-.27*	-.17*	-.17*
SD – AB	-.26*	-.37*	-.32*	-.29*	-	-	-	-
MJ2 – PC2	-.08	.01	.06	-.08	-.09	.02	.05	-.08
MJ2 – CD2	.04	-.07	-.15*	.06	.05	-.05	-.14*	.07
MJ2 – AB2	.19*	-.08	.01	.13*	-.10	-.06	-.08	-.10
PC2 – CD2	-.45*	-.42*	-.40*	-.42*	-.45*	-.41*	-.39*	-.42*
PC2 – AB2	-.28*	-.13*	-.22*	-.23*	-.12	.02	-.03	-.06
CD2 – AB2	.40*	.32*	.33*	.42*	.03	.00	.34*	-.11
<i>Stability paths</i>								
MJ1 – MJ2	.17*	.24*	.24*	.15*	.17*	.24*	.24*	.15*
PC1 – PC2	.29*	.48*	.52*	.29*	.29*	.48*	.51*	.29*
CD1 – CD2	.59*	.75*	.61*	.65*	.64*	.79*	.70*	.69*
AB1 – AB2	.66*	.79*	.79*	.67*	.67*	.67*	.82*	.55*
<i>Cross-lagged paths</i>								
PC1 – MJ2	.08	.13	.05	.17*	.09	.12	.05	.17*
PC1 – CD2	-.16*	.01	-.17*	-.08	-.16*	.00	-.15*	-.09*
PC1 – AB2	.02	.06	-.04	.07	.09	-.11	-.01	.07
MJ1 – PC2	.01	.01	.05	-.01	.01	.02	.04	-.01
MJ1 – CD2	.00	.04	.02	.02	.00	.05	.03	.02
MJ1 – AB2	-.01	.03	.03	-.01	-.05	-.04	-.02	-.07
CD1 – MJ2	-.09	-.17*	-.11	-.16*	-.02	-.12	-.04	-.09
CD1 – C2	-.33*	-.20*	-.18*	-.33*	-.33*	-.18*	-.18*	-.33*
CD1 – AB2	.03	.11*	.03	.07	.06	-.03	.04	.04
AB1 – MJ2	.08	.10	.07	.12	-.10	-.02	-.10	-.01
AB1 – PC2	-.02	.05	-.01	-.01	-.07*	.02	-.09	-.03
AB1 – CD2	.10*	.13*	.15*	.09*	.06	.07	.04	.06

Note. AB = Antisocial behavior, CD = Self-serving cognitive distortions, MJ = Moral Judgment, PC = Perception of Community, SD = Social Desirability

** $p < .05$

antisocial behavior did not precede self-serving cognitive distortions; cross-lagged paths in both directions appeared to be non-significant.

Self-serving cognitive distortions were found to precede moral judgment in the self report model for girls and for students from ethnic minority groups, and perception of community (especially for boys and students from ethnic minority groups). For boys and students from ethnic minority groups, this relationship was found to be reciprocal, although the association between perception of community at the first wave and self-serving cognitive distortions at the second wave was lower.

Finally, results showed a trend in which perception of community appeared to precede moral judgment, although this relationship was only significant for students from ethnic minority groups.

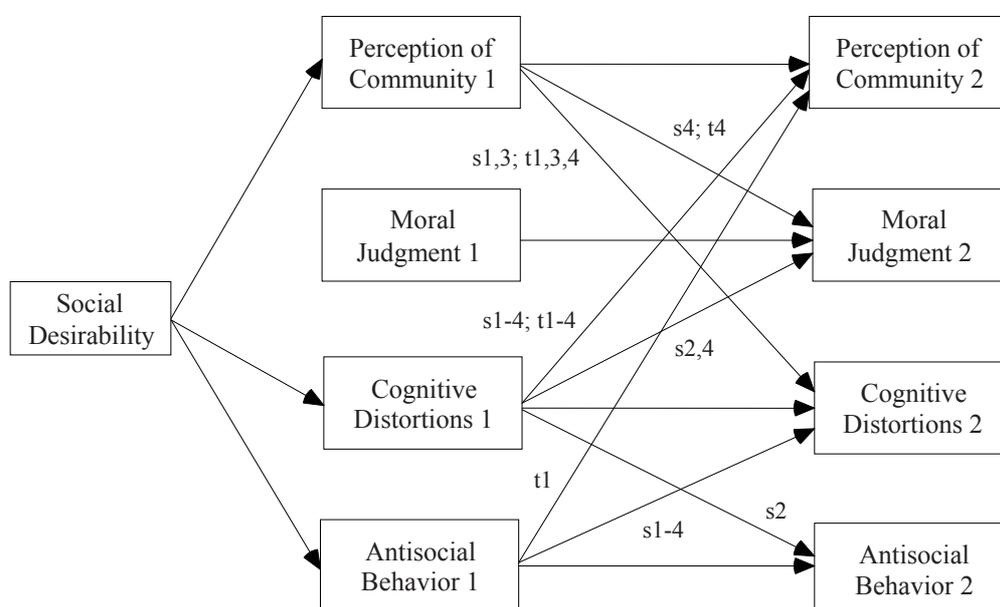


Figure 2. *Cross-lagged relationships moral cognitions and antisocial behavior (significant paths only)*

Note. S = Self reported model, T = Teacher reported model, 1 = Boys, 2 = Girls, 3 = Dutch ethnicity, 4 = Ethnic minority. Both stability paths and depicted correlations with social desirability were significant for both models and all groups (note that relationships between social desirability and teacher reported behavior were not computed)

Discussion

This study examined the longitudinal relationships between moral judgment, perception of community, self-serving cognitive distortions and antisocial behavior in a large sample of students from prevocational secondary education. In line with various other studies investigating prevalence of antisocial behavior among adolescents in the Netherlands (e.g., Van der Laan & Blom, 2006), results of the present study showed high rates of self reported aggression, vandalism, and stealing in these students. Teacher reports of aggressive and

delinquent behavior showed the same trend; a quarter of the students could be classified as clinical – scoring above the 70th percentile – according to the norms of the Teachers' Report Form (TRF, Achenbach, 1991). These results correspond to another recent Dutch study concerning antisocial behavior among young adolescents, reporting that approximately 20% of the boys and 23% of the girls scored within the problem range (Vollebergh et al., 2006). Self reports of self-serving cognitive distortions showed even higher percentages of students scoring within the clinical range – above the 83th percentile – according to the How I Think Manual (HIT, Barriga, Gibbs et al., 2001). The percentage of students scoring within the clinical range of both teacher reports of antisocial behavior and self-serving cognitive distortions was considerably lower (11%). This difference might be due to the fact that teachers are only able to observe and rate behaviors that take place within the classroom and on school grounds, which is a relatively limited part of students' social environment. This result is in line with previous research, showing higher rates of antisocial behavior on self reports compared to teacher reports (e.g., Achenbach et al., 2002), and our results showing that the percentage of students scoring within the normal range of the TRF and the clinical range of the HIT was higher than reversely.

The high rates of cognitive biases and various forms of antisocial behavior in this sample may indicate that these young students are threatened in their development and form an at risk group for externalizing behavior. Specifying at risk groups is very important for the design and implementation of prevention programs targeting antisocial behavior, since antisocial behavior in adolescence causes a high amount of (societal) damage (Flannery et al., 2003) and is a strong predictor for antisocial behavior in adulthood (Tremblay, 2006). In line with these assumptions, sex and ethnic differences were taken into account. Longitudinal structural equation models, using multigroup analysis, were moderated by sex and ethnicity showing mainly (mostly small) differences in strengths of relationships. Both sexes and the various ethnic groups were represented in the borderline and clinical groups of the TRF and HIT. These findings implicate that delay in moral cognitions represent risk factors for both sexes and different ethnic groups.

Confirming our hypotheses on the concurrent relationships between moral cognitions and antisocial behavior, analyses showed that students with a higher stage of moral judgment had a more positive perception of community (especially girls and students from Dutch ethnic groups) and reported a lower level of self-serving cognitive distortions (especially boys and students with a Dutch ethnic background). Furthermore, the higher students scored on self-serving cognitive distortions, the more negative their perception of community was and the more antisocial they were (especially boys) according to both self and teacher reports.

Although most studies report low associations between moral judgment and antisocial behavior in samples of regular youths (Barriga, Morrison et al., 2001), this study could not establish significant negative relationships between moral judgment and antisocial behavior. It is difficult to measure the development of moral judgment over just a short period of time,

and the time interval of only four months appeared to be not representative for the underlying developmental theories of moral reasoning. Another possible explanation might be the homogeneity (in terms of level of cognitive functioning) of the sample used in this study. Our desire to target the at risk group for antisocial behavior led to a restriction of range. The range might have become too narrow to detect significant moral judgment development in these samples comprising students of the same grade and educational level.

In line with our expectations, students with a more positive perception of community showed lower rates of antisocial behavior. Finally, self-serving cognitive distortions appeared to be most strongly associated with antisocial behavior, stressing the importance of a cognitive oriented approach in prevention or treatment of antisocial behavior.

Although the time span of the present study was limited, the follow up data could be used to gain insight in the relationships over time and to establish stability of the used measures. Most studies investigating the relationships between moral cognitions and antisocial behavior among adolescents were conducted cross-sectionally (Barriga, Morrison et al., 2001; Liao et al., 1998). Our longitudinal analysis indicated that the relationships between the moral cognitions and antisocial behavior were weaker in the second wave of these extended models. This finding can be explained by the stability of the measures, which was found to be moderate to high. Furthermore, in line with our expectations, results showed that the self report models predicting antisocial behavior were much stronger than the teacher report models, reflecting the limited observation possibilities by teachers which were mentioned before. Lastly, it should be noted that although there were some differences based on sex and ethnicity, most differences appeared to be small.

In contrast to our expectations, the findings of the present study could not provide evidence for an explanatory model of antisocial behavior by the proposed moral cognitions. Cross-lagged analyses even revealed the opposite for self reported antisocial behavior: self reported antisocial behavior preceded self-serving cognitive distortions, indicating that the higher the level of antisocial behavior at the first wave, the more cognitive distortions students made at the second wave. An explanation might be that self-serving cognitive distortions were primarily used as neutralizers in an attempt to reduce feelings of guilt or shame (Bandura, 2002; Festinger, Carlsmith, & Bem, 2007). In the present study, we only used the total score of self-serving cognitive distortions, while future research could investigate these relationships using primary cognitive distortions (which are theorized to precede antisocial behavior) on the one hand and secondary cognitive distortions (which are theorized to follow on antisocial behavior) on the other hand. Furthermore, for girls, this cross-lagged relationship was reciprocal, indicating that self-serving cognitive distortions and antisocial behavior influence each other. For girls, antisocial behavior seems to be a more individual characteristic, which is more cognitively driven than for boys. This assumption finds support in research such as that by Cillissen and Mayeux (2004), who stated that boys affiliate with a larger social network than girls do. To gain more insight in the impact of

antisocial behavior on cognitive distortions and sex differences in this perspective, different forms of antisocial behavior (e.g. physical aggression, relational aggression) should be further explored. The findings could not be replicated for teacher reported behavior, which can be explained by the limited ability of teachers to observe specific behaviors.

In addition, several causal relationships between moral cognitions were established. First of all, self-serving cognitive distortions appear to precede perception of community, indicating that a higher level of cognitive distortions leads to a more negative perception of community. This relationship was reciprocal for boys and students with a Dutch ethnic background, meaning that a more positive perception of community in these groups also leads to a lower level of cognitive distortions. The differences between the groups may be explained by the assumption that the social network or peer group is more important for boys and plays a more substantial role for students with a Dutch ethnic background. Overall, self-serving cognitive distortions are more likely to precede perception of community, since these relationships were stronger and applied for all groups (in both the self report and teacher report model). When perception of community is seen as an instance of the social competence (Beem et al., 2004), these findings are in line with the cognitive dissonance theory (Festinger et al., 2007).

Secondly, results show a trend in which perception of community appears to precede moral judgment, indicating that a more positive perception of community leads to a higher stage of moral judgment. Although this relationship is small and only significant for students from ethnic minority groups, they stress the importance of attention for a more contextually related aspect, such as perception of community, in the development or implementation of prevention programs. In this respect, a positive atmosphere in the group might be an import precondition for effectiveness of these programs.

Lastly, the self report model shows that a higher level of cognitive distortions leads to a lower stage of moral judgment for girls and students from ethnic minority groups. The stronger relationships for girls and students from ethnic minority groups may be explained by the assumption that cognitive distortions and moral judgment are both individual characteristics, which play a more substantial role for these groups of students. However, it should be noted that the differences between the groups based on sex and ethnicity are quite small. The lack of significant relationships in the teacher reported model in this respect may be explained by the fact that both cognitions and morality are individual characteristics, which are not accessible for teachers.

Certain characteristics of the present sample should be taken into account. Firstly, one may suggest that because of the behavioral referents of the cognitive distortions items, there is a potential confound with the HIT Questionnaire when using self-report measures. The validity of the HIT was investigated by Nas, Orobio de Castro and Koops (2005) using vignettes in which the actual intention of a peer was ambiguous. Significant relations were found between the HIT scores and the social information processing scores tapped by the

vignettes, e.g., hostile attribution of intent, attribution of other's emotions, response generation. Furthermore, the DSM-IV related behaviors addressed by the HIT are differently phrased from the behaviors of the self-report measure, while the behavioral referents are relatively hidden in the cognitive distortion items. Finally, cognitive distortions was not only related to self-reported behaviors but also to teacher reported behavior.

Secondly, although this sample represents the population that is of primary interest with respect to antisocial behavior, generalizability is limited. Schools participated on a voluntary basis in the present study. All schools were located in the western part of the Netherlands and contained a sample of students from different ethnic groups (Van der Laan & Blom, 2006). This probably has lead to a student population with a high level of antisocial behavior problems. To further specify at risk groups, future research should extent generalizability of the results of this study by including different grades. Furthermore, future studies should include more assessments over a longer period of time to provide a better representation of some of the underlying developmental processes concerning (the relationships between) moral cognitions and behavior.

In sum, this longitudinal study, using a large sample of young at risk adolescents, stresses the complexity of the processes underlying antisocial behavior. Although this study could not provide evidence for the prediction of antisocial behavior by moral cognitions, it did reveal that self reported antisocial behavior precedes self-serving cognitive distortions. Future research should continue on these findings, which can contribute to the understanding of the processes related to the development and maintenance of antisocial behaviors that are critical to both prevention and intervention for at risk youth.

Chapter 3

The How I Think Questionnaire as a tool for evaluation of prevention programs and diagnostic purposes in young adolescents

Abstract

This study addresses the psychometric quality of the Dutch translation of the How I Think Questionnaire (HIT-Q) measuring self-serving cognitive distortions, and investigates whether the HIT-Q is suitable as a tool for the evaluation of prevention programs and diagnostic purposes. A total of 1534 prevocational secondary students (56% male; $M_{\text{age}} = 14.2$, $SD = .72$) with diverse ethnic backgrounds participated in the study. Firstly, the present study confirmed that the reliability and validity of the HIT-Q was adequate, in line with previous findings by Barriga, Gibbs, Potter & Liau (2001) and Nas, Brugman & Koops (2008). Secondly, previous studies found evidence that deleting respondents scoring high on AR in a Dutch sample did not improve the validity of the results, in contrast to the original USA study. CFA's showed that AR and SD appeared to be separate constructs. Furthermore, cut off points for suspect responding could not be replicated in the present sample: the validity of the instrument did not improve by removing students with high AR scores. Thirdly, mean scores on all HIT-Q scales appeared to be higher in Dutch studies compared to the North American study, indicating that profile norms need to be revised for different groups and cultures.

Introduction

Cognitive distortions have long been posited to facilitate and maintain antisocial behavior (Barriga, Hawkins, & Camelia, 2008; Sykes & Matza, 1957). Antisocial behavior can be conceptualized as outward behavior that either directly or indirectly harms others through the violation of important moral or social norms, and includes aggressive and delinquent acts. Cognitive distortions can be defined as ‘biased ways of attending to or conferring meaning upon experiences’ (Barriga, Gibbs et al., 2001, p.1), and are presumed to contribute to problematic responses to those experiences. Individuals with a high level of cognitive distortions are likely to exhibit a higher prevalence of antisocial behavior, since cognitive distortions can be described as mechanisms of moral disengagement, separating antisocial behavior from self-evaluation (Bandura, 2001), or as mechanisms reducing empathy or guilt (Gibbs, 2003).

Barriga, Gibbs and colleagues (2001) developed a four-category typology of self-serving cognitive distortions: self-centeredness (according status to one’s own views and needs to such a degree that views of others are scarcely considered), blaming others (misattributing blame to external sources), minimizing/ mislabeling (depicting antisocial behavior as causing no real harm or being acceptable) and assuming the worst (gratuitously attributing hostile intentions to others). Self-centeredness is considered a primary cognitive distortion that precedes antisocial behavior. Blaming others, minimizing/mislabeling and assuming the worst are secondary cognitive distortions that neutralize feelings of guilt and blame that accompany antisocial behavior (Cromwell & Thurman, 2003; Gibbs, 2003).

The How I Think Questionnaire (HIT-Q, Barriga, Gibbs et al., 2001) is a paper-and-pencil instrument measuring self-serving cognitive distortions, which was developed since other questionnaires attempting to assess cognitive distortions exhibited a variety of psychometric limitations, did not include a response set, and did not include a norm group to determine the target group of respondents standing within the general population. The instrument was validated by Barriga et al. (2001) for an American sample of both referred and non-referred youth, and recently validated by Nas, Brugman and Koops (2008) for a Dutch sample of delinquent and non-delinquent male adolescents. Both the American and Dutch version of the HIT-Q showed acceptable reliability and validity. The Dutch study showed that delinquent adolescents exhibited more cognitive distortions than non-delinquent adolescents. However, in contrast to the American findings, the relationship with moral judgment was inconclusive and a negative relation with intelligence was found.

In addition, Nas and colleagues (2008) found that both delinquent and non-delinquent adolescents with lower levels of education could be assigned to the borderline-clinical range on several types of cognitive distortions according to the norms presented by Barriga et al. (2001), which would imply that cognitive distortions are much more prevalent in non-delinquent adolescents in the Netherlands than expected. These findings were replicated in a recent study on longitudinal relationships between moral cognitions and antisocial behavior

among prevocational secondary students (Van der Velden, Brugman, Boom, & Koops, in press). In addition, Van der Velden (in press) found that a substantial percentage (34%) could be assigned to the borderline or clinical range of the Teacher Report Form (an instrument which was validated and normed for the Dutch population). These findings are of great importance for the implementation and evaluation of cognitive-behavioral prevention programs targeting non-delinquent adolescents, like EQUIP for Educators (DiBiase, Gibbs, Potter, & Spring, 2005). Further investigation into cognitive distortions and psychometric quality of the HIT-Q in this population is required.

In order to find out whether the HIT-Q can be used as a tool for evaluation of prevention programs and diagnostic purposes, the present study aims at investigating the psychometric quality of this instrument, the meaning of the AR scale, and the relevance of the USA profile norms in a sample of young adolescents in the Netherlands.

The first research question concerns replication of the findings concerning reliability and validity of the HIT-Q by Barriga, Gibbs, et al. (2001) and Nas et al. (2008) in a mixed gender, non-delinquent sample representing students with a low educational level and diverse ethnic backgrounds. First of all, since numerous studies established higher rates of antisocial behavior for males than females (e.g., Côté, Vaillancourt, Barker, Nagin, & Tremblay, 2007), the present study uses a mixed-gender sample, in contrast to the previous Dutch validation study by Nas and colleagues (2008). In line with the findings of Côté and colleagues (2007), it is hypothesized that male students report higher levels of self-serving cognitive distortions than female students. Secondly, the Dutch educational system distinguishes various educational levels after elementary school, based on students' cognitive functioning. In the present study, all participants attended prevocational secondary education which represents the lowest educational level in the Netherlands. Several studies (Boxer, Goldstein, Musher-Eizenman, Dubow, & Heretick, 2005; Farrington, 2005) show that a low educational level is associated with a higher risk for antisocial behavior. Thirdly, ethnicity was found to be an important characteristic in studies of antisocial behavior in the Netherlands, in which it was concluded that the prevalence rates were higher among adolescents from ethnic minority groups compared to students with a Dutch ethnic background (e.g. Harland, Van der Laan, Smeenk, & Weerman, 2005; Junger-Tas, Cruijff, Van de Looij-Jansen, & Reelick, 2003). In line with these assumptions, it is hypothesized that students from ethnic minority groups will report higher rates of self-serving cognitive distortions as well. Since a low educational level comprises mostly students with low socioeconomic background and high rates of students from ethnic minority groups, we propose that knowledge on the psychometric quality of the HIT-Q in the present sample is of great relevance with regard to prevention of antisocial behavior.

The second question concerns the relevance of the Anomalous Response scale (AR) in the Dutch population and its relation to the Social Desirability scale (SD). Previous instruments attempting to measure cognitive distortions did not include questions designed to

detect disingenuous, incompetent, or otherwise suspect responding. To correct for this important omission – given the anticipated underreporting of self-serving cognitive distortions by antisocial adolescents due to impression management or social desirability concerns – the AR scale was added to the HIT-Q. According to the American HIT-Q manual (Barriga, Gibbs et al., 2001), AR scores greater than 4.00 should be considered suspect and interpreted cautiously, while respondents with AR scores greater than 4.25 should be excluded from data analyses in research contexts. In the validation samples, values between 4.00 and 4.25 were found to be appropriate cutoff points between the continuous distribution of AR scores and outliers (Barriga, Gibbs et al., 2001). In contrast, Nas, Brugman, and Koops (2005) did not find evidence that deleting respondents scoring high on the AR scale in a Dutch sample improved the validity of the results. Therefore, the present study aims to gain insight in the relevance of the AR scale in a Dutch sample of young adolescents.

Lastly, the HIT-Q manual provides American profile norms based on a sample comprising non-referred youth. Percentile equivalents of scale scores were derived from the sample and cutoff scores were calculated. Based on the profile plot, adolescents can be categorized in a ‘non-clinical’ range, a ‘borderline-clinical’ range, and a ‘clinical’ range. The profiles can be a helpful tool for evaluation of intervention or prevention programs. In the present study the relevance of the USA profile norms for a non-delinquent Dutch population of young adolescents will be investigated.

Method

Participants

Participants were 1534 students (56% boys, 44% girls) aged 12 to 18 ($M = 14.20$, $SD = .72$) from 11 prevocational secondary schools located in the western, mostly highly urbanized, part of the Netherlands. Schools mainly served socially disadvantaged students. Socioeconomic status (SES) was based on father’s highest completed education (‘low’ = elementary school or secondary education, ‘high’ = non-university higher education or university). It should be taken into account that the reported SES percentages are valid percentages, meaning that missing values were not included. The percentage of students who did not know the educational level of their father was high (57%). Furthermore, the majority (60%) of the students belonged to an ethnic minority group, meaning that at least one parent was born in another country than the Netherlands. A further distinction in ethnic groups was made based on father’s country of origin, with 43% Dutch, 11% Moroccan, 16% Turkish, 12% Surinamese, 4% Dutch Antillean, and 14% other.

Participants completed the self report questionnaires within classroom settings during school hours, after a brief oral and written instruction. During the assessment, a supervisor of the research team and one of their own teachers was present and students were encouraged to ask for help when needed. Participation was granted by means of active consent of the schools and passive consent of the parents (since students were older than 12 years). Participants were

assured of the confidentiality of their responses. Small rewards were provided after completion of the questionnaires. In addition, teachers completed a questionnaire about externalizing behavior for each student.

Instruments

How I Think Questionnaire. To measure self-serving cognitive distortions the How I Think Questionnaire (Barriga, Gibbs et al., 2001), was used. The HIT-Q comprises 39 items addressing four types of cognitive distortions (i.e. 9 items for self-centered; 10 items for blaming others; 9 items for minimizing; 11 items for assuming the worst), 8 anomalous response items (AR) designed to screen for suspect responding (for example: "In the past, I have lied to get myself out of trouble"), and 7 positive filler items (PF) to encourage full use of the scale. The 7 positive fillers were in the present study replaced by 11 social desirability items based on the Marlowe-Crown questionnaire (Crown & Marlowe, 1960). While the positive fillers were not meant to result in a meaningful score, the social desirability items do. An example of a social desirability item is 'I'm always friendly to others'.

Each of the 39 cognitive distortion items refers to one of four categories of antisocial behavior derived from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM IV; American Psychiatric Association, 2000): disrespect for rules, laws or authorities (opposition defiance); physical aggression; lying; stealing. For example, the item "I can't help losing my temper a lot" represents a blaming others cognitive distortion item applied to an opposition defiance behavioral referent. Participants responded along a 6-point Likert scale (from 'disagree strongly' to 'agree strongly'), with responses to individual items of 4 (slightly agree) or higher indicating cognitive distortion and responses of 3 (slightly disagree) or lower indicating absence of distortion. The AR scale was recoded, resulting in the True AR score, which was used in further analyses. Mean scale and subscale scores were used in the analysis. Based on the American normative sample (comprising non-referred high school and college students), participants could be assigned to normal (scores up to the 74th percentile), borderline (scores between the 74th and 84th percentile) and clinical range (scores above the 84th percentile) groups (Barriga, Gibbs et al., 2001).

The Antisocial Behavior Questionnaire. The Antisocial Behavior Questionnaire (ABQ) consisted of 12 items selected from self report scales used in previous research (Høst, Brugman, Tavecchio, & Beem, 1998; Leenders & Brugman, 2005). Mean scores for the subscales of 'Prevalence of antisocial behavior' (with scores above 1 indicating that the respondent committed, at least once, one or more forms of antisocial behavior) and the scale 'Attitude towards antisocial behavior' (with low scores indicating a more negative attitude towards antisocial behavior) were used in the analysis. Reliability of these scales were adequate, with Cronbach's alpha's of .76 and .87 respectively.

Teacher's Report Form (Achenbach, 1991). The Dutch translation of the externalizing behavior scale from the Teacher's Report Form (Verhulst, Ende, & Koot, 1997) was used to assess antisocial behavior as perceived by teachers. Total scores for the subscales 'Aggressive

behavior' and 'Delinquency' and for the total scale "Externalizing behavior' were computed. In addition, based on guidelines provided in the Teacher's Report Form Manual (Achenbach, 1991; Verhulst et al., 1997), students were assigned to normal (scores up to the 67th percentile), borderline (scores between the 67th and 70th percentile) and clinical range (scores above the 70th percentile) groups. Both international and Dutch studies established high validity and reliability of the instrument (Achenbach, 1991; Rescorla et al., 2007; Verhulst et al., 1997). These results were replicated in the present study, with Cronbach's alpha of .96.

Moral Judgment Rating Task. This task aims to assess stage of moral judgment development as defined by Kohlberg (Colby & Kohlberg, 1987). The Moral Judgment Rating Task (MJRT) consists of two short dilemmas containing two conflicting moral values, followed by stage-typed quotes favoring the first value or the second value. The task is comparable to (and inspired by) the SROM-sf (Basinger & Gibbs, 1987), however, the response format is different in order to allow a Rating Scale analysis (Wright & Masters, 1982). Students categorized the quotes as 'for wiser people' (1), 'worthy of respect' (2), or 'childish' (3). Validity was supported in that, for full stages 1 to 3, the ordering in terms of difficulty estimates corresponded with the ordering proposed by Kohlberg (although separation between stages was weak). Mean scores over the 12 items were calculated. Reliability was low, with Cronbach's alpha of .57.

School Moral Atmosphere Questionnaire: Perception of Community. The scale 'Perception of community' of the School Moral Atmosphere Questionnaire (Høst et al., 1998) was used to assess students' perception of the shared understanding of the school as a community. The scale perception of community contains 24 items with a 5-point Likert scale (totally disagree to totally agree). For example: 'You can trust other students at this school'. Negative items were recoded and means were computed. Reliability was adequate, with Cronbach's alpha of .69.

Academic Grades

Intelligence was assessed using academic grades. A mean score was calculated, based on the individual grades (overall scores at the end of the college year) of the three most important and obligatory courses for all students in the Dutch educational system: Dutch, English, and Mathematics.

Plan of analysis

Firstly, factor analytic validity will be assessed by means of confirmatory factor analysis (CFA). In determining model fit, diverse fit indexes can be used. According to Schermelleh-Engel, et al. (2003), there should not be too much emphasis placed on the significance of the χ^2 statistic, because of its dependence on sample size. Instead, it is recommended to use the χ^2/df ratio, with a value of 3 being indicative of an acceptable data-model fit. In addition, both absolute close-fit indexes (e.g. RMSEA) and incremental close-fit indexes (e.g. CFI) were used in the present study. Secondly, to determine the relevance of the Anomalous Response

scale, CFA's will be used to find out whether AR and SD are separate constructs. Further, CFA's will be used to find out whether internal validation improves when students with high AR scores are deleted. Finally, means and standard deviations for the American normative sample will be compared to these statistics derived from the samples used in the study by Nas et al. (2008) and the present study, and classification based on the HIT-Q Profile form will be compared to classification based on the Dutch TRF Profile Form.

Results

Factor analytic validity

The structure of the HIT-Q was assessed by means of confirmatory factor analysis (CFA). Like Nas et al. (2008) and Barriga, Gibbs et al. (2001), it was tested whether a six-factor solution (four cognitive distortion factors, one AR factor, and one SD factor) was statistically preferred to a three-factor solution (one cognitive distortions factor, one AR factor and one SD factor) and a four-factor solution (one primary and one secondary cognitive distortion factor, one AR factor, and one SD factor). CFA's showed that the six-factor model (χ^2 (1588) = 6501.00, $p < .001$, CFI = .76, RMSEA = .045) was supported against the three- and four-factor model, based on significant χ^2 difference tests (respectively χ^2 (1592) = 6530.45, $p < .001$, CFI = .76, RMSEA = .045, and χ^2 (1590) = 6515.50, $p < .001$, CFI = .76, RMSEA = .045).

To improve model fit of the preferred six-factor model, items were allowed to be correlated with each other within factors and items were allowed to be loaded on other factors. This resulted in a reasonable fit of the six-factor model to the data, χ^2 (1445) = 4408.05, $p < .001$, CFI = .86, RMSEA = .037, χ^2/df ratio = 3.0). The χ^2/df ratio with a value of 3 was indicative of an acceptable data-model fit (Schermelel - Engel et al., 2003). Furthermore, the absolute close-fit index, the RMSEA, was lower than .05, indicating a good fit. However, the incremental close-fit index, the CFI approached the adequate fit of .90, but was not higher than .86. It should be noted that a negative association has been established between the number of indicators in a model and the level of incremental close-fit associated with that model (Gignac, 2007). When deleting items with lower factor loadings (<.40; Table 2), the CFI indeed increased to an acceptable fit of .90 (χ^2 = 2298.43, df = 1108, N = 1534, $p < .001$, RMSEA = .039, CFI = .90, χ^2/df ratio = 3.2). However, to compare results with previous studies, it was not desirable to omit items. In addition, all factor loadings of the improved six-factor model were significant, although a few factor loadings were low (Table 1).

Internal consistency

The Cronbach's alphas of the four types of cognitive distortions ranged from .70 to .73, the Cronbach's alpha for Anomalous Responding was .73, and the Cronbach's alpha for Social Desirability was .71. In addition, the Cronbach's alpha for the HIT-Q Overall was found to be .91.

Table 1. *Confirmatory Factor Analysis by cognitive distortions – Factor loadings*

Item	Present study	Nas et al., 2008	Barriga et al., 2001
3 (3) < SC	.22	.53	.51
7 (7) < SC	.37	.44	.52
10 (10) < SC	.44	.54	.61
22 (23) < SC	.59	.72	.76
28 (29) < SC	.59	.61	.72
37 (40) < SC	.60	.30	.66
42 (45) < SC	.54	.72	.66
52 (55) < SC	.27	.54	.51
54 (57) < SC	.55	.70	.69
6 (6) < BO	.25	.13	.33
11 (11) < BO	.36	.51	.53
21 (22) < BO	.65	.63	.76
25 (26) < BO	.53	.55	.60
26 (27) < BO	.50	.50	.69
36 (38) < BO	.35	.44	.46
39 (42) < BO	.44	.55	.73
44 (47) < BO	.49	.49	.70
46 (49) < BO	.43	.64	.46
50 (53) < BO	.54	.69	.69
5 (5) < MM	.38	.52	.50
12 (13) < MM	.45	.63	.72
14 (15) < MM	.47	.47	.60
17 (18) < MM	.41	.57	.68
19 (20) < MM	.60	.54	.61
30 (32) < MM	.53	.59	.72
33 (35) < MM	.52	.58	.69
40 (43) < MM	.62	.59	.67
47 (50) < MM	.31	.52	.69
2 (2) < AW	.31	.29	.38
8 (9) < AW	.33	.45	.49
15 (16) < AW	.38	.45	.64
18 (19) < AW	.42	.49	.54
23 (24) < AW	.36	.54	.57
29 (30) < AW	.26	.44	.39
32 (34) < AW	.57	.65	.66
35 (37) < AW	.64	.67	.73
43 (46) < AW	.59	.68	.77
49 (52) < AW	.50	.65	.71
53 (56) < AW	.61	.70	.79
4 (4) < AR	.21	-	.37
13 (14) < AR	.41	.31	.30
20 (21) < AR	.43	.27	.42
27 (28) < AR	.62	.60	.64
31 (33) < AR	.52	.63	.67
38 (41) < AR	.62	.66	.69
45 (48) < AR	.56	.54	.68
51 (54) < AR	.63	.55	.67
(1) < SD	.45	-	-
(8) < SD	.51	-	-
(12) < SD	.19	-	-
(17) < SD	.43	-	-
(25) < SD	.67	-	-
(31) < SD	.25	-	-
(36) < SD	.49	-	-
(39) < SD	.36	-	-
(44) < SD	.63	-	-
(51) < SD	.36	-	-
(58) < SD	.47	-	-

Note. Item numbers between brackets are according to the numbering of the version of the HIT-Q as used in the present study

Convergent validity

Pearson correlations for the four types of cognitive distortions and the TRF scales ranged from .15 to .19 ($p < .01$). Furthermore, positive relationships between the TRF scales and the HIT-Q Overall scale (r 's ranging from .18 to .20, $p < .01$) were found. In addition, Pearson correlations for the four types of cognitive distortions and the Prevalence scale of the ABQ ranged from .42 to .51, ($p < .01$), while the correlation between the HIT-Q Overall scale and the Prevalence of antisocial behavior scale was $r = .53$ ($p < .01$).

Divergent validity

Pearson correlations for the four types of cognitive distortions and age ranged from .07 to .11 ($p < .01$). There was also a small but positive relationship between the HIT-Q Overall scale and age ($r = .09$, $p < .01$). Negative relationships with social desirability could be established for the HIT-Q scales, see Table 2, and HIT-Q Overall ($r = -.17$, $p < .01$). Correlations between the HIT-Q scales and academic grades were calculated. Note that academic grades were only available for a small part of the samples ($n = 271$). Pearson correlations for the four types of cognitive distortions and academic grades ranged from $-.12$ ($p < .05$) to $-.16$ ($p < .01$). The HIT-Q Overall scale also correlated negatively with academic grade ($r = -.16$, $p < .01$).

Table 2. Pearson's correlations among the HIT-Q scales, AR and social desirability

Scale	1	2	3	4	5
1 Self-Centered	-				
2 Blaming Others	.68	-			
3 Minimizing/ Mislabeled	.74	.71	-		
4 Assuming the Worst	.69	.71	.70	-	
5 Anomalous Responding	-.52	-.49	-.49	-.51	-
6 Social Desirability	-.15	-.14	-.18	-.12	.41

Note. All correlations were significant at the .01 level

Construct validity

Consistent with previous findings (Barriga, Gibbs et al., 2001), moral judgment correlated negatively with all cognitive distortions HIT-Q scales, with r values ranging from $-.07$ ($p < .05$) to $-.10$ ($p < .01$). The HIT-Q Overall scale also correlated negatively with moral judgment ($r = -.10$, $p < .01$).

As expected, the cognitive distortion HIT-Q scales correlated negatively with perception of community, with r values ranging from $-.41$ to $-.43$ ($p < .01$). The HIT-Q Overall scale also correlated negatively with perception of community ($r = -.48$, $p < .01$).

Further, in line with the expectations, positive correlations were established between the cognitive distortion HIT-Q scales and attitude towards antisocial behavior, with $r = .34$ for

Assuming the worst, $r = .38$ for Blaming others, $r = .47$ for Self-centeredness, and $r = .49$ ($p < .01$) for Minimizing/ Mislabeled. The HIT-Q Overall scale correlated positively with attitude towards antisocial behavior ($r = .48$, $p < .01$).

Discriminant validity

Group differences with respect to sex, ethnicity, and normal versus at risk youth (based on classification according to the TRF, see Instruments), were tested with ANOVA's. Differences between boys and girls on all HIT-Q scales were established (F -values ranged from 17.78 to 49.69, $p < .01$, $df = 1$), with boys reporting higher scores than girls. Effect sizes were small (η^2 ranging from .01 to .03).

Furthermore, significant differences between students with a Dutch ethnic background and students from ethnic minority groups were found for the scales Blaming Others, Assuming the Worst, and HIT-Q Overall (F -values ranged from 5.40 to 8.70, $p < .05$, $df = 1$). On these scales, students from ethnic minority groups scored higher than students with a Dutch ethnic background. Effect sizes were small (all $\eta^2 .01$).

Lastly, significant differences between normal students and students at risk for antisocial behavior were found on all HIT-Q scales (F -values ranged from 17.43 to 27.10, $p < .00$, $df = 2$, $n = 1303$). As expected, students within the normal range on the TRF scored lower than students within the borderline and clinical range on the TRF, and students in the borderline range on the TRF scored lower than students within the clinical range on the TRF. Effect sizes were small (η^2 ranging from .03 to .04).

Relevance and practice of the Anomalous Response scale

Firstly, Pearson's correlation showed a significant association between AR and SD. The relationships between the HIT-Q scales and the AR scale were substantially stronger than the relationships between the HIT-Q scales and SD, see Table 2. Secondly, CFA's showed that the six-factor model including an AR factor and a SD factor could statistically be preferred to the higher order 7-factor solution, based on a significant χ^2 difference test, including a general response set factor underlying the AR factor and the SD factor, respectively $\chi^2 (1445) = 4408.05$, $p < .001$, $\chi^2 (1447) = 4441.18$, $p < .001$. This result indicates that AR and SD, although significantly associated, are separate constructs.

Furthermore, the relevance of the AR scale in the present sample was investigated. Based on the American cutoff points, a high percentage of the students in the present study should be considered suspect (29% scoring higher than 4.00) or should even be deleted from the present sample (20% scoring higher than 4.25). Cross-tabulation shows that there were no significant differences based on sex ($\chi^2 = .00$, $p = .96$), ethnicity ($\chi^2 = 1.87$, $p = .17$), and age ($F = .60$, $p = .28$) between students scoring below 4.25 and students scoring above this cutoff score.

The high numbers of ‘suspect’ students cast doubt on the cutoff points for the AR scale, and therefore the six-factor CFA model was tested in six samples comprising different percentages of students depending on their AR scores. Model fit did not improve – RMSEA values remained .037 in all models and CFI values dropped from .86 to .82 – when samples in which 2%, and 5% to 25% (with intervals of 5%) of the students with the highest scores on the AR scale were removed, were compared to the six-factor model including the complete sample.

Profile norms

Firstly, means and standard deviations of the HIT-Q scales (cognitive distortions) derived from the American normative sample, the sample of the study by Nas and colleagues (2008), and the sample of the present study were compared. As shown in Table 3, means for all (sub) scales of the HIT-Q were substantially higher in both the present sample and the delinquent and non-delinquent group by Nas et al. (2008) as compared to the means reported by Barriga et al. (2001). In addition, means for all cognitive distortion subscales were higher in the delinquent group compared to the non-delinquent group by Nas et al. (2008) and the present study. These differences in the prevalence of cognitive distortions between the American norm sample and the different delinquent and non-delinquent Dutch samples underscore the limitations concerning generalizability in the use of the American profile form.

Table 3. *Descriptive statistics for the HIT-Q scales*

Scale	Present study		Nas et al., 2008*				Barriga et al., 2001#	
	Non-delinquents		Delinquents		Non-delinquents		Non-delinquents	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
HIT-Q Total score	2.63	.66	2.83	.72	2.66	.60	2.39	.69
- <i>Self-Centered</i>	2.50	.74	2.79	.86	2.69	.67	2.42	.74
- <i>Blaming Others</i>	2.70	.78	2.92	.81	2.72	.73	2.42	.79
- <i>Minimizing/ Mislabeling</i>	2.70	.78	2.87	.84	2.69	.67	2.31	.78
- <i>Assuming the Worst</i>	2.57	.71	2.78	.76	2.58	.66	2.35	.72
Anomalous Response	3.49	.86	3.37	.89	3.31	.71	3.33	1.07
Social Desirability	3.63	.65	-	-	-	-	-	-

* Delinquent: $N = 329$ males, age 14 – 19 ($M = 16.85$, $SD = 1.57$)

Non-delinquent: $N = 143$ males, age 14 – 19 ($M = 15.73$, $SD = 1.43$)

$N = 412$ (226 male and 186 female), age 14 – 19 ($M = 17.27$, $SD = 1.30$)

Discussion

The present study examined whether the How I Think Questionnaire (Barriga, Gibbs et al., 2001) – measuring self-serving cognitive distortions – can be used as a tool for evaluation of prevention programs and diagnostic purposes.

The adequate reliability and validity of the HIT-Q as found by Barriga et al. (2001) and Nas et al. (2008) was replicated in a mixed gender, non-delinquent sample representing students with diverse ethnic backgrounds. Confirmatory factor analyses (CFA) showed that the cognitive distortions structure (six-factor solution, including four cognitive distortions factors, one AR factor, and one SD factor) was justified in the present sample, in line with previous findings. Fit indexes in the present study were even better compared to the North American version, with a lower RMSEA (respectively, .037 and .055), and fit indexes were almost similar in comparison to the findings of Nas and colleagues (2008), with equal RMSEA values, but a somewhat lower CFI index. A CFI value somewhat below the index (.86 instead of $> .90$) is a common phenomenon in the literature concerning model fit of confirmatory factor analysis; CFA's including a relatively large number of observations (> 30) commonly fail to be associated with acceptable levels of incremental close-fit (Gignac, 2007). The discussion concerning the use of structural equation model fit criteria is still ongoing, and the existence of golden rules, especially concerning the application of CFA's, is being argued (e.g., Markland, 2007; Marsh, Hau, & Wen, 2004). In addition, it is recommended to revise or rephrase items with low factor loadings, in particular item 6 that failed to achieve a minimum factor loading in all three studies.

Internal consistency of the HIT-Q was found to be adequate in the present study. Convergent validity of the HIT-Q was satisfactory, with significant positive associations with both teacher reported antisocial behavior and self reported antisocial behavior. The strength of the relationships was in line with previous findings on both teacher reports (Nas et al., 2008) and self reports (Barriga, Gibbs et al., 2001). In line with the literature (Achenbach, Dumenci, & Rescorla, 2002), correlations between the HIT-Q scales and teacher reported antisocial behavior appeared to be substantially smaller than correlations between the HIT-Q and self report. These findings presumably indicate the limited scope of teachers concerning behaviors of their students. It is likely that antisocial behavior also takes place outside of school grounds (e.g. on the streets and in public places, at home), indicating that future research on antisocial behavior should further include diverse sources (e.g. parent reports, police records).

In line with the findings by Barriga et al. (2001) and Nas et al. (2008), divergent validity was found to be satisfactory, with negative but small associations between the HIT-Q scales and age, social desirability and academic grades in the present study. Furthermore, the significant associations between the HIT-Q scales and social desirability indicate that one should control for social desirability in future (evaluation) studies.

Adequate construct validity of the HIT-Q was indicated by negative associations with moral judgment, and positive associations with perception of community and attitude towards antisocial behavior. Since high levels of cognitive distortions can be construed as reflecting an immature moral development (Gibbs, 2003), the HIT-Q scales should inversely relate to moral judgment, as was found in the present study. These findings were only partly in line

with the study by Nas and colleagues (2008), in which the results concerning the relationship between the HIT-Q and moral judgment were inconclusive. They did establish a negative relationship, but significance depended on the instrument used to assess moral judgment.

Finally, discriminant validity was satisfactory in the present study as well. In line with the literature, boys scored higher on all HIT-Q (sub) scales than girls (Barriga, Morrison, Liao, & Gibbs, 2001), and as hypothesized, students from ethnic minority groups scored higher on several (sub) scales of the HIT-Q than students with a Dutch ethnic background (Harland et al., 2005; Junger-Tas et al., 2003). In line with the literature, students from ethnic minority groups appear to score higher on the scales expressing an external locus of control (Sodowsky, Kuo-Jackson, Richardson, & Corey, 1998). These sex and ethnic differences should be taken into account in future research, since they imply that initial levels of cognitive distortions may differ for diverse groups. Furthermore, it was found that the lower the risk for antisocial behavior, the lower the level of cognitive distortions (based on TRF classification). These findings indicate the importance of cognitive distortions in relation to antisocial behavior.

Meaning of Anomalous Response

In addition, based on contradictory results from previous studies, the present study further investigated the meaning and application of the AR scale in the Dutch population and its relation to the Social Desirability scale. First of all, CFA models including AR and SD as separate constructs showed a significant better fit compared to models including a higher order factor 'Response set', underlying anomalous response and social desirability. In addition, the relationship between AR and the HIT-Q scales appeared to be substantially stronger than the relationship between SD and the HIT-Q scales. These findings indicate that AR and SD are two separate constructs. Future research should first focus more specifically on the differences in meaning and application of both constructs.

Furthermore, according to the manual (Barriga, Gibbs et al., 2001), participants scoring above 4.0 or 4.25 on the AR scale are suspect responders and respectively need to be treated carefully or even need to be eliminated from further analysis. Based on these norms, 29 percent of the participants in the present study appeared to be treated carefully and 20 percent needed to be removed from the sample. These are extremely high percentages compared to other studies reporting exclusion of 10% in delinquent samples and 2% in normal samples of adolescents (Liao, Barriga, & Gibbs, 1998).

Because of the conflicting results between the previous Dutch study (Nas et al., 2008) and the present study versus the North American study by Barriga, Gibbs, et al. (2001), CFA's in which several percentages of high scorers on AR were deleted, were compared. Model fit did not improve when students with high AR scores were removed from the sample. These findings cast doubt on the relevance and function of the AR scale in the Dutch

population. Although AR appeared to be a reliable scale, implementation of the cut off scores as described by Barriga, Gibbs, et al. (2001) was not satisfactory in the present study, and would have led to deletion of a substantial part of the sample. Therefore, the relevance of the AR scale as an indicator for suspect responding does not find support based on these data in terms of internal validity.

A high AR score might partly be an indication of (extreme) prosociality instead of suspect responding, possibly due to a cultural difference in the meaning of anomalous response between both the American validation sample versus the present Dutch sample and ethnic differences within the Dutch sample. Another explanation might possibly be that the meaning of AR in the Dutch population differs from the North American population due to interpretation problems caused by translation or language problems. Suspect response statements might be interpreted differently in diverse cultures or populations.

Generalizability of the Profile form

Finally, the current study investigated the application of the profile norms as set by the North American manual (Barriga, Gibbs et al., 2001). Findings showed that there were substantial differences in levels of cognitive distortions between the North American sample and the Dutch sample, and between the delinquent and non-delinquent groups in the Netherlands. Mean scores on all HIT-Q scales were higher in the Dutch studies (present study; Nas et al., (2008) compared to the North American study by Barriga, Gibbs, et al. (2001). In addition, and in line with the expectations, mean scores were higher in the delinquent group compared to the non-delinquent group in Nas et al. (2008) and in the present study. Furthermore, both a previous Dutch study (Van der Velden et al., in press) and the present study show a small overrepresentation of at risk youth based on the HIT-Q norms when compared to the Dutch TRF norms (respectively, 43% borderline or clinical based on HIT-Q and 34% based on TRF, 42% borderline or clinical based on HIT-Q and 29% based on TRF).

These findings indicate that, for the application of the HIT-Q as an instrument for diagnostic purposes, the profile form and cut off points for classification need to be revised for specific populations. With profile norms suitable for the specific population, the HIT-Q appears to be an adequate instrument for the evaluation of prevention programs. Although the HIT-Q profile form can be a useful starting point, the HIT-Q should always be part of broader set of instruments assessing different aspects of antisocial behavior.

Limitations

There are also some limitations which should be taken into account. First of all, it should be noted that the generalizability of the findings of the present study are limited to adolescents with a low educational level. All participants attended the seventh or eighth grade of prevocational secondary schools, representing the lowest educational level in the Dutch

educational system. To enlarge generalizability, future research should include students from diverse educational levels and diverse grades.

Secondly, selection of participants was not randomized since all recruited schools participated in a longitudinal evaluation study on reducing antisocial behavior. This might have led to an overrepresentation of at risk youth in the present sample, since one could suggest that schools in need for a preventive program targeting antisocial behavior were more likely to volunteer.

Conclusion

Although reliability and validity of the How I Think Questionnaire appeared to be adequate, the present study also showed that the application of the AR scale as a tool to select suspect responders does not find support in this sample. Furthermore, it appears that the original profile form needs to be revised for the Dutch population. Based on these results, it can be concluded that the psychometric quality of the instrument is satisfactory, which indicates that the HIT-Q can be an adequate tool for the evaluation of prevention programs. For the use of the instrument for diagnostic purposes in a wider population of (non-delinquent) adolescents in the Netherlands, a revision of the profile form, including adaptation of the cut off points for classification, is recommended.

In addition, the differences between sexes and diverse ethnic groups imply that future studies should include these variables, and it even might be necessary to develop different profile forms for the diverse groups. Although the AR scale appears to have another meaning in the Dutch population, future research should further investigate its meaning in a more heterogeneous sample, including adolescents with diverse educational levels or levels of cognitive functioning. In addition, since AR and SD appeared to be separate constructs, and SD significantly correlated with all HIT-Q scales, studies using the HIT-Q should control for social desirability. Finally, relatively high levels of self-serving cognitive distortions were found, indicating the need for prevention in a substantial part of the sample. Although results indicate that the revised profile form can be used to classify ‘at risk’ students, the selection should not only be based on the outcomes of the HIT-Q.

Chapter 4

Effects of EQUIP for Educators on students' self-serving cognitive distortions, moral judgment, and antisocial behavior

Abstract

A quasi-experimental pre-test/ post-test study using a control group was conducted to investigate the effects of EQUIP for Educators – implemented as an universal prevention program – on prevalence of antisocial behavior, attitude towards antisocial behavior, self-serving cognitive distortions, and moral judgment of young adolescents. Participants were 764 students (present at both waves: N experimental group = 512, N control group = 110) from nine prevocational secondary schools in the Netherlands (M age = 14.07, SD = .60). Repeated measures were used to examine intervention effects. Students in the experimental group reported a significant reduction of both attitude towards antisocial behavior and self-serving cognitive distortions compared to students within the control group. Ethnicity appeared to play a moderating role; effects on both attitude towards antisocial behavior and self-serving cognitive distortions were strongest for students with a Dutch ethnic background.

Introduction

High rates of violence and aggression among youth stresses the importance to identify effective prevention programs (Flannery et al., 2003). Although the need for early prevention is illustrated by studies showing that antisocial behavior occurs along a developmental continuum of behavioral severity (Schaeffer, Petras, Ialongo, Poduska, & Kellam, 2003; Tremblay, 2006), continuation of support across the transition from childhood to adolescence is of great importance as well. Problem behaviors such as aggression and antisocial behavior increase dramatically during adolescence (Simons-Morton, Haynie, Saylor, Davis Crump, & Chen, 2005).

Since schools play a significant role in shaping the lives of adolescents in their role as social agent (Cho, Hallfors, & Sánchez, 2005) and have access to youth at various levels of risk for extended periods of time, they provide an excellent context to implement empirically based prevention and intervention programs targeting antisocial behavior (Cunningham & Henggeler, 2001). Although many school and community based violence prevention programs have been developed and applied in recent years (Gottfredson & Gottfredson, 2002), relatively few have been rigorously evaluated. Since the need for evidence-based prevention programs is increasing (Boxer, Goldstein, Musher-Eizenman, Dubow, & Heretick, 2005), the goal of the present study is to investigate the effectiveness of a multi component intervention, EQUIP for Educators, targeting antisocial behavior. Antisocial behavior can be conceptualized as outward behavior that either directly or indirectly harms others through the violation of important moral or social norms, and includes aggressive and delinquent acts (Barriga, Morrison, Liau, & Gibbs, 2001; Liu, 2004). In the present study, effects on both prevalence of antisocial behavior and attitude towards antisocial behavior will be investigated.

EQUIP for Educators

EQUIP for Educators is an adapted version of the original treatment program EQUIP for juvenile offenders (Gibbs, Potter, & Goldstein, 1995), and is dedicated to both primary and secondary prevention in an educational context. The program is guided by a theoretical framework based on psycho education, referring to the teaching and training of skills, knowledge, and mature awareness required for competent daily living (DiBiase, Gibbs, Potter, & Spring, 2005). The program is intended to encourage and equip youth to think and act responsibly, using a peer-helping approach (Gibbs et al., 1995). Since a peer group approach is not sufficient to counter negative peer pressure, helping skill limitations of at risk youth must also be addressed (Gibbs, Potter, Barriga, & Liau, 1996). The helping skills curriculum of EQUIP for Educators is based on Aggression Replacement Training (Goldstein & Glick, 2001) and includes social skills training, anger management training, and moral education. EQUIP for Educators aims to remediate developmental delays in moral judgment, self-serving cognitive distortions and deficiencies in social skills.

Moral judgment can hereby be defined in terms of justice and the capability of using moral criteria in the evaluation of behavior (Colby & Kohlberg, 1987). Delay in moral judgment can be seen as ‘the persistence beyond early childhood of a pronounced degree of me-centeredness or egocentric bias’ (Gibbs, 1991, p.96) and implies immature moral judgment. Kohlberg (1984) originally proposed six invariant stages for the development of moral judgment, while Gibbs (2003) revised the stages of this theory and reduced the number of stages to four. In his revision, stage 5 and 6 are not regarded as truly developmental stages, but represent an existential phase. According to Gibbs (2003) the immature level consists of stages 1 and 2, while the mature level comprises stages 3 and 4. Relations between moral judgment and antisocial behavior are generally found to be modest (Barriga, Morrison et al., 2001); a low moral judgment stage does not necessarily lead to antisocial behavior, unless it is combined with a high degree of self-serving cognitive distortions (Gibbs, 2003).

According to social-cognitive theories, people act upon their interpretation of social events and antisocial behavior is based on deficiencies in interpreting these events, i.e. cognitive distortions. Barriga, Gibbs, Potter, and Liao (2001, p.1) define cognitive distortions as ‘inaccurate or biased ways of attending to or conferring meaning upon experiences’. Barriga, Gibbs and colleagues (2001) focus on those cognitive distortions that, at a high level of prevalence, facilitate aggression and other types of antisocial behavior. They distinguish four types of those self-serving cognitive distortions: self-centeredness (according status to one’s own views and needs to such a degree that views of others are scarcely considered), blaming others (misattributing blame to external sources), minimizing/ mislabeling (depicting antisocial behavior as causing no real harm or being acceptable) and assuming the worst (gratuitously attributing hostile intentions to others). Self-centeredness is considered as a primary cognitive distortion that precedes antisocial behavior. Blaming others, minimizing/mislabeling and assuming the worst are secondary cognitive distortions that neutralize feelings of guilt and blame that accompany antisocial behavior (Gibbs, 2003).

Self-serving cognitive distortions are assumed to block moral judgment development because one does not consider oneself to be responsible for one’s delinquent behavior, and fulfill a defensive or neutralizing role (Gibbs, 1991). Consistent with these findings, Barriga, Morrison and colleagues (2001) established a partially mediating effect of self-serving cognitive distortions on the relationship between moral judgment and antisocial behavior in young adults.

Effects of EQUIP

Previous studies investigating the effects of the original EQUIP program for juvenile delinquents (Gibbs et al., 1995) showed improvements in both skills and (attitude towards) behavior. An early study by Leeman and colleagues (1993), using the preliminary version of the program, reported that delinquent adolescents in the EQUIP group evidenced significant

improvements in institutional conduct and recidivism rates compared to the control group. The researchers concluded that EQUIP promoted therapeutic change in juvenile delinquents. Furthermore, a recent Dutch study showed that EQUIP encourages and equips incarcerated delinquent youth to reduce self-serving cognitive distortions and to develop a less positive attitude against delinquent behavior (Nas, Brugman, & Koops, 2005). However, similar to the findings of Leeman et al. (1993), change in level of moral judgment did not differ between the groups. These positive effects of EQUIP in groups of delinquent adolescents are promising and have led to the development of a preventive version of the EQUIP curriculum (DiBiase et al., 2005). So far, there are no evaluation studies published reporting the effects of EQUIP for Educators.

Present study

In the present study, this preventive version of the program is evaluated for the first time in a large sample of students attending prevocational secondary school in the Netherlands. The effects of EQUIP for Educators on prevalence of antisocial behavior, attitude towards antisocial behavior, self-serving cognitive distortions, and moral judgment were examined using a quasi-experimental pre-test/ post-test with control group design. Although the EQUIP for Educators program is in the present study implemented as a universal prevention program - targeting the majority of the school population – a substantial part of the sample can be referred to as ‘at risk’. The Dutch educational system differs from other international educational systems by distinguishing various levels after elementary school, based on students’ cognitive functioning. Prevocational secondary education represents the lowest educational level in the Netherlands, which is associated with a higher risk for antisocial behavior (Boxer et al., 2005; Farrington, 2005). While previous evaluation studies on the original program for juvenile offenders only included small groups of male delinquents in their samples, this study is conducted in a large sample of prevocational secondary education students, including both males and females. In line with the aims of EQUIP for Educators, it is hypothesized that students in the experimental group will show a larger increase in stage of moral judgment, and a larger decrease in levels of self-serving cognitive distortions, in attitude towards antisocial behavior and in prevalence of antisocial behavior in comparison with the control group.

In the analyses, sex differences are taken into account since numerous studies established higher rates of antisocial behavior (Côté, Vaillancourt, Barker, Nagin, & Tremblay, 2007; Van Lier, Vitaro, Wanner, Vuijk, & Crijnen, 2005) and self-serving cognitive distortions (Barriga, Morrison et al., 2001) for males than females. Furthermore, since the sample used in the present study consists of students from diverse ethnic backgrounds, and ethnicity has been found to be an important variable in Dutch studies of antisocial behavior (Boendermaker & Van Yperen, 2003; Van der Laan & Blom, 2006), it is

also investigated whether ethnicity influences the effectiveness of the program. Ethnic status can be defined as an easily identifiable characteristic that implies a common cultural history with others possessing the same characteristic. The most common ethnic ‘identifiers’ are race, religion, country of origin, language and/ or cultural background (Okazaki & Sue, 1998). In the present study, a student is characterized as being part of an ethnic minority group when at least one of his/ her parents is born in another country than the Netherlands (e.g. Morocco, Surinam, Turkey, Netherlands Antilles, or other).

Furthermore, the present study aims to give an overview of the prevalence of different forms of self reported antisocial behavior, and a categorisation of the seriousness of self-serving cognitive distortions and teacher reported antisocial behavior, to find out whether the present sample is suitable for evaluation research. Since higher rates of problem behavior may indicate better opportunities for improvement (Flannery et al., 2003; Stoolmiller, Eddy, & Reid, 2000), it is important to map students’ initial level of antisocial behavior prior to their participation in an intervention. In the present study, it is investigated whether the seriousness of antisocial behavior at pre test influences the effectiveness of the intervention. It is hypothesized that students who can be classified as borderline/ clinical based on TRF norms will show a larger reduction of self-serving cognitive distortions and/ or antisocial behavior compared to the students within the normal range.

Method

Participants

Based on geographic location, about 80 prevocational secondary schools in the Netherlands were approached by letter to participate in the present study. Twelve schools showed interest and were further informed about participation possibilities and preconditions. Schools were able to join the program when they were willing to participate with at least three eight grade classes and were able to provide circumstances to have the teachers well-trained (during a 18 hours course offered by the research project) and implement the program in classroom settings in a period of 16 weeks (two school hours a week). Three schools were interested in the program, but could not take part in the intervention due to logistic problems with school schedules (lack of available hours). Two of these schools participated in the study as control group and their eight grade students filled out the questionnaires at the same data waves as the experimental schools. At first, nine schools were willing and able to implement EQUIP for Educators according to the research requirements and committed to voluntary participation. During the first year of the research, one experimental school dropped out because of a lack of motivation. Another experimental school resigned from participation in the second year, due to management problems.

Eventually a total of nine (7 experimental schools and 2 control schools) preparatory vocational secondary schools located in the western (mostly highly urbanized) part of the

Netherlands, with a total of 764 eight grade students, participated in this study. Since the program was implemented in classroom settings, all students of the attending classes participated in the research. Students in the experimental group filled out self-report questionnaires just before and immediate after the intervention, and students participating in the control group were asked to fill out these questionnaires at the same data waves. A total of 622 students ($n_{\text{exp}} = 512$, $n_{\text{cont}} = 110$) participated in both waves of data collection (see Table 1). Furthermore, teachers filled out a questionnaire about externalizing behavior for each student at the pre test. For generalization purposes, an univariate analysis of variance was conducted to find out whether students who filled out the questionnaires at both the pre and post test differed on the teacher reports of antisocial behavior at pre assessment from students who were absent at one of the two waves. Results showed that there were no significant differences between these groups ($F(1, 693) = .98$, $p = .32$).

Table 1. *Characteristics of the sample, including participants present at both data waves*

	Age		Sex		Ethnicity		SES	
	<i>M</i>	<i>SD</i>	Boys	Girls	Dutch	Minority	Low	High
Total group	14.07	.60	370 59%	252 41%	275 45.3%	332 54.7%	259 83.8%	50 16.2%
Experimental group	14.10	.59	301 58.8%	211 41.2%	209 41.8%	291 58.2%	223 84.5%	41 15.5%
Control group	13.94	.61	69 62.7%	41 37.3%	66 61.7%	41 38.3%	36 80.0%	9 20.0%

Note. Reported percentages are valid percentages

Since it is possible to repeat a class in the Netherlands, or start later in a grade based on cognitive functioning and/ or skills, the age range is rather large: representing students from 12.9 to 16.3 years (ages 12.9-14, $n = 283$; ages 14-15, $n = 288$; ages 15-16, $n = 49$; ages 16 >, $n = 1$; missing, $n = 1$). The mean age at pre test was 14.07 ($SD = .60$) years and only 5.3% of the students was older than 15. Class-size ranged from 9 to 28 students. A majority (54.7%) of the participants belonged to an ethnic minority group. Socioeconomic status was ascertained by asking students about their father's highest completed education and appeared to be largely low, i.e. elementary school or secondary education, or moderate, i.e. non-university higher education (see Table 1). Note that the reported percentages are valid percentages, meaning that missing values were not taken into account. The percentage of students who did not know the educational level of their father was high (50.3%).

There were no significant differences between the experimental group and the control group concerning sex ($\chi^2(1) = .58$, $p = .45$) and socioeconomic status ($\chi^2(1) = .57$, $p = .45$). Further results showed a small but significant difference between the groups on age (F

(1,620) = 7.11, $p = .01$), and a significant difference on ethnicity ($\chi^2(1) = 14.06$, $p = .00$). Students in the experimental group were older ($M = 14.10$, $SD = .59$, in years) than students in the control group ($M = 13.94$, $SD = .61$, in years), and the majority of the experimental group belonged to an ethnic minority group ($n = 291$ versus $n = 209$), while a minority of the control group belonged to an ethnic minority group ($n = 41$ versus $n = 66$). Based on these findings, ethnicity will be included in the analyses as a moderator.

Furthermore, one-way ANOVA's conducted at the pre assessment showed no significant differences between experimental and control groups on prevalence of antisocial behavior ($F(1,621) = 3.45$, $p = .06$), teacher reported antisocial behavior ($F(1, 571) = 1.01$, $p = .32$), attitude towards antisocial behavior ($F(1, 612) = 2.55$, $p = .11$), self-serving cognitive distortions ($F(1,614) = 1.52$, $p = .22$), and moral judgment ($F(1,616) = 2.78$, $p = .10$).

Measures

Antisocial Behavior Questionnaire. The Antisocial Behavior Questionnaire consisted of 12 items selected from self report scales used in previous research (Høst, Brugman, Tavecchio, & Beem, 1998; Leenders & Brugman, 2005). Each item, representing a specific type of antisocial behavior, was followed by a question about the prevalence of their behavior ('did you conduct this behavior?') and a question about their attitude towards antisocial behavior ('how wrong is it to conduct this behavior?'). After the intervention, the questions concerning the prevalence of their behavior were extended with a time period ('did you conduct this behavior in the past three months?'). The questionnaire could be scored on a 4-point scale (prevalence of behavior: 1 = never, 2 = once, 3 = sometimes, 4 = often; attitude towards behavior: 1 = very wrong, 2 = pretty wrong, 3 = not so wrong, 4 = not wrong at all). Scale mean scores for prevalence of antisocial behavior (with scores above 1 indicating that the respondent committed, at least once, one or more forms of antisocial behavior) and attitude towards antisocial behavior (with low scores indicating a more negative attitude towards antisocial behavior) were used in the analysis. For prevalence of antisocial behavior, reliability for the total group at pre test was adequate ($\alpha = .73$) and stability (in the control group, $r(110) = .55$, $p < .00$) was high. For attitude towards antisocial behavior, both reliability for the total group at pre test ($\alpha = .86$) and stability (in the control group, $r(109) = .66$, $p < .00$) were high. Furthermore, reliability for prevalence of antisocial behavior was stable over time, for both the experimental group (α pre = .73, α post = .73) and the control group (α pre = .71, α post = .72). Finally, reliability for attitude towards antisocial behavior was stable over time, for both the experimental group (α pre = .86, α post = .88) and the control group (α pre = .84, α post = .88).

Teacher's Report Form (Achenbach, 1991). The Dutch translation of the externalizing behavior scale from the Teacher's Report Form (Verhulst, Ende, & Koot, 1997) was used to assess antisocial behavior as perceived by teachers. The questionnaire comprised 34 items on

delinquency and aggression such as ‘Fights a lot’ or ‘Threatens other people’, all of which are rated as 0 = not true (as far as I know); 1 = somewhat true; and 2 = very true or often true. Teachers rated problems based on students’ functioning over the preceding two months. Total scores were computed and based on national norms for peers of the same age range and sex. Based on guidelines provided in the Teacher’s Report Form Manual (Achenbach, 1991; Verhulst et al., 1997), students were assigned to normal (scores up to the 67th percentile), borderline (scores between the 67th and 70th percentile) and clinical range (scores above the 70th percentile) groups. The instrument evidenced high validity and reliability (Rescorla et al., 2007). These results were replicated in the present study, with a Cronbach’s alpha of .95.

How I Think Questionnaire. To measure self-serving cognitive distortions the How I Think Questionnaire (Barriga, Gibbs et al., 2001) was used. The HIT comprises 39 items addressing four types of cognitive distortions (i.e. 9 items self-centered; 10 items blaming others; 9 items minimizing; 11 items assuming the worst), 8 anomalous response items (AR) designed to screen for suspect responding (for example: "In the past, I have lied to get myself out of trouble"), and 7 positive filler items to encourage full use of the scale. We replaced the 7 positive fillers by 11 social desirability items based on the Marlowe-Crowne questionnaire (Crowne & Marlowe, 1960). While the positive fillers were not meant to result in a meaningful score, the social desirability items do. Reliability of the social desirability subscale was found to be high (Cronbach’s alpha’s ranging from .75 at pre test to .77 at post test).

Each of the 39 cognitive distortion items refers to one or another of four categories of antisocial behavior derived from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM IV; American Psychiatric Association, 2000): disrespect for rules, laws or authorities (opposition defiance); physical aggression; lying; stealing. For example, the item “I can’t help losing my temper a lot” represents a blaming others cognitive distortion item applied to an opposition defiance behavioral referent. Participants responded along a 6-point Likert scale (from ‘disagree strongly’ to ‘agree strongly’), with responses to individual items of 4 (slightly agree) or higher indicating cognitive distortion and responses of 3 (slightly disagree) or lower indicating absence of distortion. Mean scale and subscale scores were used in the analyses. Based on the American normative sample (comprising referred and non-referred youth), students were assigned to normal (scores up to the 74th percentile), borderline (scores between the 74th and 84th percentile) and clinical range (scores above the 84th percentile) groups (Barriga, Gibbs et al., 2001).

The Dutch translation of the instrument evidenced good validity and reliability (Nas, Brugman, & Koops, 2008). The reliability of the HIT in the present study was high, with Cronbach’s alpha’s of .77 at pre test and .83 at post test for the overall scores. Stability appeared to be high as well (in the control group, $r(109) = .61, p < .00$). Furthermore,

reliability was stable over time, with α pre = .77 and α post = .82 for the experimental group and α pre = .79 and α post = .86 the control group.

Moral Judgment Rating Task. This task assesses stage of moral judgment development as defined by Kohlberg (Colby & Kohlberg, 1987). The Moral Judgment Rating Task (MJRT) consists of two times short dilemmas containing two conflicting moral values, followed by a series of reactions favoring the first value and a series of reactions favoring the second value. The dilemmas and the quotes are derived from the Moral Judgment Interview (MJI) manual (Colby & Kohlberg, 1987) and were used earlier in the Moral Judgment Sorting Task (Boom, Brugman, & Van der Heijden, 2001). Quotes are stage typed representing stages 1 to 4 and in-between stages as defined in the MJI manual. This results in 4 sets of 7 items is 28 items. The task is comparable to (and inspired by) the SROM-sf (Basinger & Gibbs, 1987), however, the response format is different in order to allow a Rating Scale analysis (Wright & Masters, 1982). Students categorized the quotes as ‘for wiser people’ (1), ‘worthy of respect’ (2), or ‘childish’ (3).

Rating scale analysis was used to assess how well the task was performed, because classical reliability (Cronbach’s alpha) seems less appropriate since it assumes all items measure the same construct whereas our items are designed to reflect different stages. Validity was supported in that, for full stages 1 to 3, the ordering in terms of difficulty estimates corresponded with the ordering proposed by Kohlberg (although separation between stages was weak). The higher the stage the quote represents, the more likely it is that it is judged to be wiser, keeping ability constant; and, the higher ability of the respondent is, the more likely an assertion is judged to be childish, keeping stage of quote constant. Since we found that stages higher than stage 3 are out of reach for most participants in our relatively young sample we decided to omit items reflecting these stages from further use in the analysis. In addition, the analysis revealed that the first set of quotes gave mixed results, not contributing positively to the expected ordering; therefore we omitted this set also from further use in the analysis. For further modeling we did not use the rather complex Rasch estimates and instead used the mean score over the items per person.

Semi-structured teacher interviews To gain insight in the perspective of the teacher on the effects of EQUIP for Educators, interviews were conducted just after the intervention. The interview consisted of 17 questions concerning evaluation and effects. For example, teachers were asked whether the behavior of the students changed after the intervention, how they would criticize the different components of the program, whether the atmosphere in the class was influenced by the program, and whether they would like to continue with EQUIP for Educators next college year. Furthermore, there was time for additional remarks and questions. Participation took place on voluntary basis. Teachers from five schools participated in the interviews, which took place in groups (including all teachers from schools who worked with EQUIP). For each school, an evaluation form was filled out.

Students' evaluation forms To gain insight in the way students experienced EQUIP for Educators and to gain insight in the effect of the program on their thinking and behavior, evaluation questionnaires were filled out. Questionnaires included two scales. One scale contained 7 items concerning the evaluation of the different components and methods (e.g. 'How did you like the sessions concerning social skills?' and 'How did you like to work in subgroups?') and was rated on a three point scale ('I liked it', 'I liked it a little', 'I did not like it'). Reliability for this scale was adequate, with a Cronbach's alpha of .64. The second scale consisted of 3 items concerning the effect of the program (e.g. 'Did EQUIP for Educators have influence on your behavior?'), which were rated on a three point scale ('Yes', 'A little', 'No'). Reliability for this scale was adequate, with a Cronbach's alpha of .71. Three experimental schools gave permission to let students fill out evaluation forms, which resulted in participation of 248 students in this evaluation. Anonymity was guaranteed.

Design and Procedures

Intervention The EQUIP for Educators program used in the present study consisted of 32 classroom meetings. Two introduction sessions were added to the original 30 lessons, in which the rules that apply during the EQUIP meetings were being discussed and self-serving cognitive distortions were introduced. Although these distortions appear to be common among all people, previous research shows they are particularly prevalent among antisocial adolescents (Liau, Barriga, & Gibbs, 1998). The other 30 meetings were structured around the three main components; anger management, social skills, and social decision making. Meetings from the different components were alternated to bring out the interrelationships among them and to strengthen the previously learned skills and behaviors (DiBiase et al., 2005). For example: students learn specific skills to stay calm (anger management), then they learn how to act constructively by practicing these skill(s) during role-play sessions (social skills), and finally they learn how to make mature moral decisions (social decision making).

The educational material from the EQUIP for Educators program was translated and expanded with a reader for students, containing an overview of each EQUIP session in addition to the original worksheets. Several examples in the anger management and moral dilemma sections from the original version were replaced by less severe examples, since they were found to be ill-suited for this sample and involved intensively aggressive scenarios (Seroczynski, 2006).

Students were coached by the mentor or a regular teacher who was intensively involved with the class. Teachers were extensively trained before implementation and received several hours of both individual and group wise supervision by a professional EQUIP trainer during the intervention period. All seven experimental schools used the EQUIP for Educators program for the first time and implemented the intervention in eighth grade. Due to practical circumstances and governmental guidelines, schools could only

organize two EQUIP meetings per week instead of three, as advised by the authors of EQUIP for Educators (DiBiase et al., 2005). As a result of this adaptation, the time schedule changed and program duration became 16 weeks in the present study instead of 10 weeks. Meetings of one school hour (50 to 60 minutes, depending on the school rules) were held twice a week in classroom settings, so groups stayed intact during implementation.

Implementation of the program was monitored by the researcher and trainer. There was intensive and structured contact with the participating schools during the intervention. Teachers were asked to fill out evaluation forms every fourth session, random observations took place during EQUIP meetings in the class, and both oral and written evaluations after implementation were organized with teachers and students.

Procedure After a brief oral and written instruction, students completed the self report questionnaires within classroom settings during school hours. A supervisor of the research team and one of their own teachers was present during the assessments and students were encouraged to ask for help when needed. Parents of the participants received written information about the project, participation was granted by means of passive consent. Participants were assured of the confidentiality of their responses. Small rewards were provided after completion of the questionnaires.

The questionnaires were pre-tested in two different pilots and adapted according to feedback provided by students and teachers. The complete program was tested in a pilot intervention of six months ($n = 59$). Based on the results of the pilot and the feedback by the students and teachers, some small adjustments (mostly concerning language) were made.

Results

Descriptive statistics

Prevalence of antisocial behavior and self-serving cognitive distortions

First, descriptive data from self reports on antisocial behavior and self-serving cognitive distortions at pre test will be presented.

Almost half of the respondents (49%) reported to have - at least once - beaten and/ or kicked someone (physical aggression). Vandalism also appeared to be common among the sample; 40 percent of the students reported to have deliberately destructed something, and 30 percent reported raising a fire. Different forms of relational aggression were prevalent; 83% reported gossiping, while 55% excluded someone from the group. Minor stealing was reported by 28 percent of the students, while the most serious form of stealing - robbery - was the least reported crime among the total sample (1%). At pre test, most participants scored within the normal range on self-serving cognitive distortions (67%) according to the USA norms (see Instruments). Although only a small group was assigned to the borderline group (13%), a substantial percentage (20%) of the students showed a high level of self-serving cognitive distortions and was marked as 'clinical' (Barriga, Gibbs et al., 2001). According to

teacher reports at pre test, the majority of the students (75%) scored within the normal range, while 17% could be assigned to the clinical group (Achenbach, 1991). In sum, these figures show that the prevalence of both self-serving cognitive distortions and antisocial behavior is high, confirming that (a substantial part of) this sample is indeed at risk.

Relationships between moral judgment, cognitive distortions and antisocial behavior

Zero-order correlations between the main and control variables were computed for the total group at pre assessment. Prevalence of antisocial behavior appeared to be positively associated with self-serving cognitive distortions ($r(615) = .55, p < .00$) and attitude towards antisocial behavior ($r(613) = .53, p < .00$) as expected, but the correlation with moral judgment was non-significant ($r(617) = -.03, p = .47$). The relationship between attitude towards antisocial behavior and moral judgment was non-significant as well ($r(608) = .01, p = .75$). Self-serving cognitive distortions correlated negatively with moral judgment ($r(610) = -.14, p < .01$), and positively with attitude towards antisocial behavior ($r(606) = .50, p < .00$). Furthermore, prevalence of antisocial behavior, attitude towards antisocial behavior, and self-serving cognitive distortions correlated negatively with social desirability ($r(618) = -.51, p < .00$, $r(609) = -.36, p < .00$, and $r(615) = -.29, p < .00$ respectively), but there was no significant relationship between social desirability and moral judgment ($r(613) = -.06, p < .13$). Social desirability will be included in the analyses on self-serving cognitive distortions and antisocial behavior as a covariate. Finally, there were no significant relationships between chronological age and moral judgment ($r(616) = -.03, p = .45$), chronological age and self-serving cognitive distortions ($r(614) = .01, p = .72$), and chronological age and prevalence of self reported antisocial behavior ($r(621) = -.06, p = .13$). The relationship between chronological age and attitude towards antisocial behavior was significant ($r(613) = -.10, p = .01$). Based on these findings, age was only included as a covariate in further analyses for attitude towards antisocial behavior.

Teacher and student evaluations

After the intervention, six out of seven experimental schools intended on continuing with EQUIP for Educators in the next college year. Open interviews with teachers, which were held directly after the intervention, showed that most teachers (85%) evaluated the program as being successful in its attempt to positively influence thinking and acting of the students. Students from three schools ($n = 248$) filled out evaluation forms anonymously and gave their opinion on the program after completing the intervention. No less than 63 percent of the students stated to have learned something from their classmates, 56 percent of the students reported to think more prosocially after the intervention, and 38 percent of the students reported to behave more prosocially afterwards.

Table 2. Effects of EQUIP for Educators on antisocial behavior, cognitive distortions and moral judgment

	Experimental				Control				F	□ ² _p	
	Pre test		Post test		Pre test		Post test				
	M	SD	M	SD	M	SD	M	SD			
<i>Total group</i>											
Prevalence of antisocial behavior	1.55	.35	1.37	.33	1.62	.37	1.38	.33	(1,615)	1.50	.00
Attitude towards antisocial behavior	1.89	.53	1.80	.54	1.80	.49	1.86	.59	(1,607)	11.07*	.02
Self-serving cognitive distortions	2.52	.64	2.44	.66	2.44	.61	2.53	.67	(1,610)	9.05*	.02
Moral judgment	2.03	.33	2.05	.32	2.09	.30	2.13	.32	(1,609)	.32	.00
<i>Dutch ethnic group</i>											
Prevalence of antisocial behavior	1.64	.37	1.41	.32	1.69	.32	1.45	.29	(1,269)	.50	.00
Attitude towards antisocial behavior	2.04	.52	1.92	.51	1.83	.49	1.97	.57	(1,268)	17.88**	.06
Self-serving cognitive distortions	2.54	.66	2.36	.68	2.38	.51	2.54	.66	(1,269)	22.14**	.08
Moral judgment	2.06	.35	2.09	.29	2.13	.27	2.20	.30	(1,271)	.77	.00
<i>Ethnic minority group</i>											
Prevalence of antisocial behavior	1.50	.32	1.34	.32	1.53	.42	1.30	.37	(1,328)	3.35	.01
Attitude towards antisocial behavior	1.77	.50	1.71	.54	1.73	.46	1.66	.54	(1,323)	.03	.00
Self-serving cognitive distortions	2.50	.62	2.50	.63	2.50	.70	2.52	.67	(1,327)	.01	.00
Moral judgment	2.01	.32	2.03	.33	2.05	.32	2.01	.32	(1,323)	.65	.00

Note. Social desirability was included as a covariate in the analyses of prevalence of antisocial behavior and self-serving cognitive distortions

Age was included as a covariate in the analyses of attitude towards antisocial behavior

* $p < .05$, ** $p < .00$

Effects of EQUIP for Educators

First of all, repeated measures show a reduction of prevalence of antisocial behavior for both the experimental group and the control group, but the difference is not significant, see Table 2. Secondly, the attitude towards antisocial behavior decreased (became more negative) in the experimental group and increased in the control group. Thirdly, the level of self-serving cognitive distortions decreased significantly in the experimental group compared to the control group. Both effect sizes were small according to Cohen (1988), with partial eta squares (η_p^2) of .02. Finally, the increase of moral judgment for both the experimental and control group was non-significant.

Sex and ethnic differences

The influence of sex and ethnic differences on the effect of the intervention was analyzed by including these variables in the repeated measures analyses as a between-subjects factor. There were no 'group X sex' interaction effects on change of prevalence of antisocial behavior ($F(1, 613) = .95, p = .34$), on change of attitude towards antisocial behavior ($F(1, 605) = .70, p = .41$), on change of self-serving cognitive distortions ($F(1,608) = 2.67, p = .10$), and on change of moral judgment ($F(1,607) = .21, p = .65$). Although there was no significant 'group X ethnicity' interaction effect on change of moral judgment ($F(1,594) = 1.40, p = .24$) and prevalence of antisocial behavior ($F(1,598) = 3.19, p = .07$), significant interaction effects for both change in attitude towards antisocial behavior ($F(1,591) = 6.94, p = .01, \eta_p^2 = .01$) and change of self-serving cognitive distortions ($F(1,593) = 8.36, p = .00, \eta_p^2 = .01$) were established. Both effects were small (Cohen, 1988). When comparing the attitude towards antisocial behavior among students on ethnicity, the effect of EQUIP for Educators is stronger for students with a Dutch ethnic background. For these students, the attitude towards antisocial behavior became more negative in the experimental group and more positive in the control group (see Figure 1).

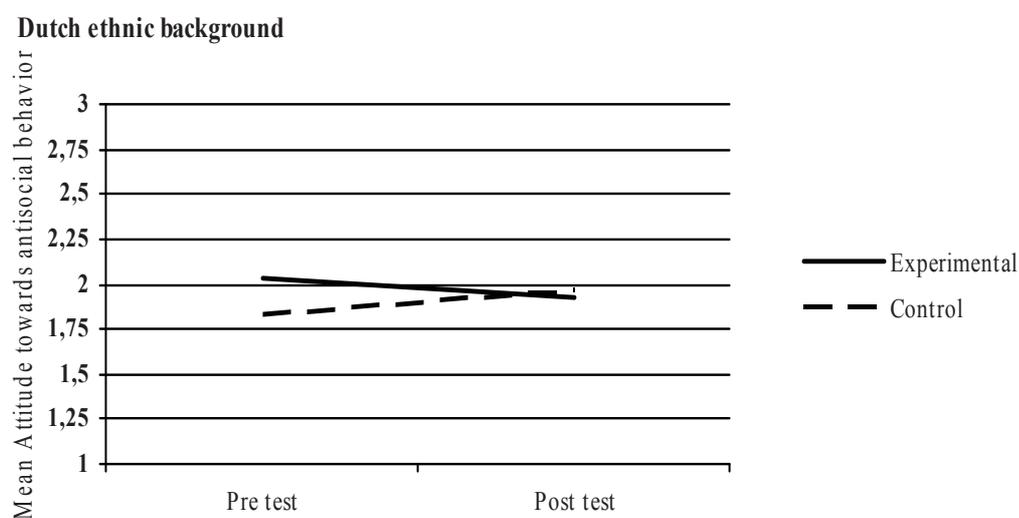


Figure 1. *Interaction of attitude towards antisocial behavior X condition X ethnicity*

In contrast, students from ethnic minority groups in both the experimental and control group showed a decrease in attitude towards antisocial behavior. Furthermore, as shown by the interaction effect mentioned above, the effect of the intervention on self-serving cognitive distortions appears to apply in a higher extent to students with a Dutch ethnic background. For these students, the level of self-serving cognitive distortions decreased in the experimental group and increased in the control group, while in the ethnic minority population the level of self-serving cognitive distortions remained stable for both the experimental and control group (see Figure 2).

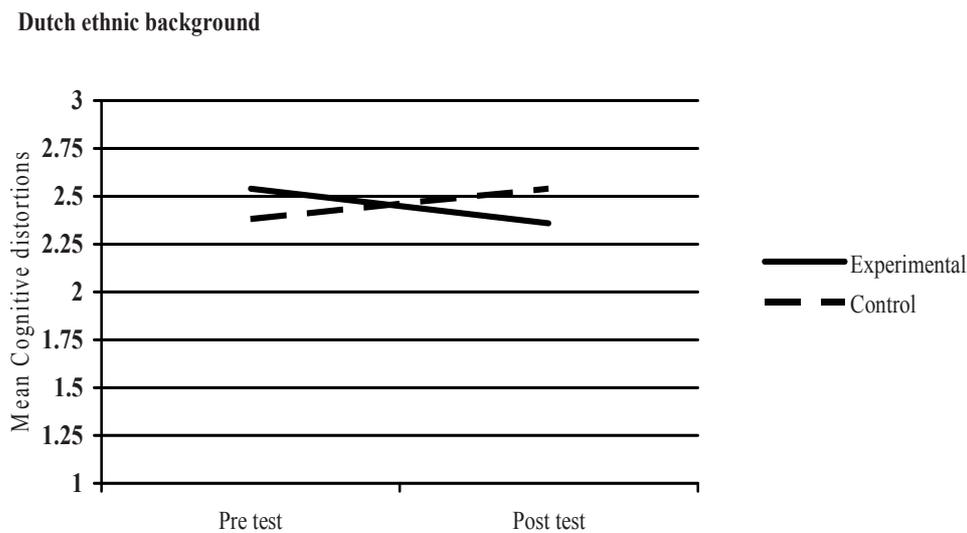


Figure 2. *Interaction of self-serving cognitive distortions X condition X ethnicity*

Since interaction effects on ethnicity were established, separate repeated measures for students with a Dutch ethnic background and students from ethnic minority groups were conducted. This distinction in ethnic groups shows that the effects of the intervention on attitude towards antisocial behavior and self-serving cognitive distortions only apply for students with a Dutch ethnic background. Students with a Dutch ethnic background who followed EQUIP reported a more negative attitude towards antisocial behavior and lower levels of self-serving cognitive distortions after the intervention, while students with a Dutch ethnic background in the control group showed a more positive attitude towards antisocial behavior and higher levels of self-serving cognitive distortions at the post test. Students from ethnic minority groups in both the experimental and control group showed a more negative attitude towards antisocial behavior at the post test and remained stable on the level of self-serving cognitive distortions (see Table 2). These are both medium effects (Cohen, 1988).

Participants' initial level of self-serving cognitive distortions and antisocial behavior

Finally, to investigate the influence of the seriousness of antisocial behavior at pre test on the effects of EQUIP for Educators, TRF (Achenbach, 1991) profiles were added as between-subject factors to the analyses (see Instruments). The TRF provides guidelines for categorization into normal, borderline and clinical groups, but since the numbers of respondents in the control group of the present study appeared to be too small for this division into three groups, the borderline and clinical groups were taken together. In line with TRF norms, students were categorized as normal (up to the 67th percentile) or borderline/ clinical (above the 67th percentile). There were no differences between the experimental and the control group in change of both attitude towards antisocial behavior ($F(1,556) = .02, p = .88$) and self-serving cognitive distortions ($F(1,558) = .45, p = .51$) based on categorization into normal and borderline/ clinical TRF profile groups. In sum, there were no differences between effects of the program based on participants' initial level of antisocial behavior.

In addition, a comparison of HIT profiles (Barriga, Gibbs, et al., 2001) before and after intervention, using cross tabulation, showed a positive change in the experimental condition. The number of students assigned to borderline groups at pretest that could be categorized as normal after intervention, and the number of students assigned to clinical groups at pretest that could be categorized as borderline or normal after intervention was larger than reverse ($n = 95$ versus $n = 65$). Although this result indicates a positive trend, the changes were not significant according to the McNemar-Bowker test ($\chi^2(3) = 6.09, p = .11$).

Furthermore, comparison of HIT profiles showed a negative change in the control group. The number of students assigned to borderline groups at pretest that could be categorized as normal after intervention, and the number of students assigned to clinical groups at pretest that could be categorized as borderline or normal after intervention was smaller than reverse ($n = 11$ versus $n = 18$). However, the changes were not significant according to the McNemar-Bowker test ($\chi^2(3) = 6.07, p = .11$).

Discussion

The present study provides a first evaluation of EQUIP for Educators, implemented as an universal school based intervention program for young adolescents (DiBiase et al., 2005), by using a pre test/ post test design comparing an experimental and control group. This quasi-experimental study focused on the change in prevalence of antisocial behavior, attitude towards antisocial behavior, self-serving cognitive distortions, and moral judgment.

First of all, it was investigated whether this type of intervention was justified, e.g. whether the sample could be defined as an 'at-risk' group for the development of antisocial behavior. Results showed, in line with previous research in the Netherlands (Van der Laan & Blom, 2006; Vollebergh et al., 2006), high rates of various forms of antisocial behavior in the present sample. Next to relational aggression, such as excluding someone from the group and

gossiping, physical aggression and vandalism were reported by substantial percentages of the students. Teacher reports on externalizing behavior confirmed this trend, showing that 17 percent of the students could be assigned to the clinical range (scoring above the 70th percentile) according to the norms of the Teachers' Report Form (Achenbach, 1991). Furthermore, based on mean scores for self-serving cognitive distortions, almost a quarter of the students could be classified as 'clinical' (Barriga, Gibbs et al., 2001). Based on these findings, it can be concluded that a substantial part of the present sample is indeed 'at-risk' for the development of antisocial behavior and could, given the strong relationship between self-serving cognitive distortions and antisocial behavior, benefit from a preventive intervention targeting limitations in the development of moral judgment, self-serving cognitive distortions, and deficiencies in social skills.

So the question remains whether EQUIP for Educators, implemented in classroom settings, is a successful intervention for the young adolescents in this sample (with high mean scores on self-serving cognitive distortions). Firstly, results show that the experimental and control group differed significantly with regard to change in both attitude towards antisocial behavior and self-serving cognitive distortions. A more negative attitude towards antisocial behavior and a lower level of self-serving cognitive distortions after intervention were found for the experimental group when compared to the control group, although effect sizes were small (Cohen, 1988). Comparison of HIT profiles before and after the intervention confirmed this latter finding. The number of EQUIP students in the normal range increased, while their presence in the borderline and clinical group decreased. For students in the control group, reverse effects were found. Although these changes are limited, they do indicate the clinical meaning of the intervention. Evaluations also support these findings, since 85 percent of the teachers evaluated the program as being successful in its attempt to positively influence thinking and acting of students. Furthermore, 56 percent of the students reported to think more prosocially after the intervention and 38 percent of the students reported to act more prosocially after the intervention.

Secondly, prevalence of antisocial behavior decreased on post test. However, in contrast the expectations, this effect appeared to be non-significant since both the experimental and control group showed this reduction. This finding can be attributed to changes in the instrument to assess prevalence of antisocial behavior: at pre test students were asked whether they ever conducted specific types of antisocial behavior, while at post test they were asked whether they conducted these types of behavior in the last three months. It might still be possible that a part of the decrease in prevalence of antisocial behavior is caused by EQUIP for Educators and will become visible on follow up assessment. Longitudinal data are needed to gain insight in the impact of the intervention on the long term.

Thirdly, similar to previous findings (Leeman et al., 1993; Nas et al., 2005), change in moral judgment of the experimental group did not differ from moral judgment in the control

group. Although we could not establish an effect of EQUIP for Educators on moral judgment immediately after intervention, the impact of EQUIP on moral judgment might be a more gradual process. It is possible that a ‘sleeper’ effect occurs, meaning that the effects on moral judgment will become visible and more powerful after a longer period of time (Leeman et al., 1993). Longitudinal data are needed to further investigate the developmental processes and impact of the intervention on the long term.

Since previous studies stress the importance of possible differences between boys and girls (Barriga, Morrison et al., 2001; Côté et al., 2007; Van Lier et al., 2005), and students from diverse ethnic groups (Boendermaker & Van Yperen, 2003; Van der Laan & Blom, 2006), moderation of both sex and ethnicity was investigated. Sex appeared to have no influence on the effects of EQUIP for Educators on prevalence of and attitude towards antisocial behavior, self-serving cognitive distortions, and moral judgment. However, ethnicity does play a moderating role in the effectiveness of the intervention. Results show that, when ethnicity of the student was included in the analyses, the effects on attitude towards antisocial behavior and self-serving cognitive distortions were strongest for students with a Dutch ethnic background. The attitude towards antisocial behavior from these students in the experimental group became more negative over time, while the attitude towards antisocial behavior became more positive in the control group. In contrast, there was no significant difference for students from ethnic minority groups; the attitude towards antisocial behavior in both the experimental and control group became more negative over time.

In line with these results, the level of self-serving cognitive distortions from students with a Dutch ethnic background in the experimental group decreased after the intervention, while the level of self-serving cognitive distortions increased in the control group. There was no significant difference for students from ethnic minority groups; self-serving cognitive distortions remained stable in both the experimental group, and the control group. In sum, the intervention had a greater impact on students with a Dutch ethnic background than on students from ethnic minority groups in reducing both the attitude towards antisocial behavior and self-serving cognitive distortions. These effects for the different ethnic groups require further investigation and seem to be of great importance for implementation of the program in multicultural groups. In contrast to the expectations (Flannery et al., 2003; Stoolmiller et al., 2000), we could not establish larger effects of the program for students with higher initial levels of antisocial behavior. There were no differences in effect of the intervention between students who were assigned to the normal range and students assigned to the borderline/clinical range on behavior based on TRF profiles, indicating that the program can be successful for all students in the sample.

Limitations

Some limitations with regard to the design of the study should be taken into account. First of all, the number of participants in the control group was substantially smaller than in the experimental group, which limited investigation of differences between specific (sub) groups. As a consequence of this limitation, we were not able to differentiate between different ethnic minority groups. The results might vary between diverse ethnic minority groups as well, due to specific cultural differences. In order to optimize EQUIP for Educators for all students, future research should further investigate these possible differences to explain the processes of change. Insight in cultural differences can lead to better implementation and evaluation of the program.

Secondly, students could not be assigned randomly to the experimental group or the control group since the intervention was implemented on class level and all students in the class participated. Furthermore, in line with ethical norms and to prevent possible bias (teachers who teach EQUIP teach other classes as well), all classes of the participating schools were included in the experimental group. Thirdly, the results are based on self-report only, which might have led to a limited scope of problem behaviors. Omitting the observation of and interaction with students by multiple sources in different contexts might have led to a limitation in reported problem behaviors. In addition, changing the self report questionnaire on prevalence of antisocial behavior might have been of importance as well. Future research should use the same questionnaires for both pre and post test assessment.

Furthermore, some limitations of the present study might have contributed to the small effect sizes. First of all, the present study only focussed on the effects of EQUIP for Educators on moral judgment, self-serving cognitive distortions, and both attitude towards and prevalence of antisocial behavior. Future evaluation research should also include other important components of the program, such as social skills and moral atmosphere, and focus on the processes underlying the effects of EQUIP for Educators by investigating the relations between morality, cognition, and behavior. Secondly, the program was implemented for only 2 hours a week as opposed to the advised 3 hours (DiBiase et al., 2005), which made it impossible to follow the original curriculum in which a session of each component could be attended to weekly.

Thirdly, the present study was conducted within the first year of implementation of the program, which can on the one hand lead to strong motivation among teachers, but on the other hand to problems due to a lack of experience. To improve effectiveness, longitudinal research on quality of program implementation - including teacher characteristics- is needed (Durlak & DuPre, 2008; Gottfredson & Gottfredson, 2002).

Conclusion

In sum, it can be concluded that EQUIP for Educators is only partly successful in its aim to equip students to think and act more prosocially. Effects of the intervention on prevalence of antisocial behavior and moral judgment could not be established. However, teacher and student evaluations are promising and the program had a significant impact on self-serving cognitive distortions and attitude towards behavior, especially for students with a Dutch ethnic background. In addition, comparison based on HIT profiles shows a reduction of EQUIP students in clinical and borderline groups and an increase of EQUIP students in the normal group, while the reverse is found for the control group. Although this latter finding was non-significant and the program had only impact on attitude towards behavior and cognition, investing in longitudinal research on the program in this population seems beneficial.

Chapter 5

EQUIP for Educators: the longitudinal effects of a school based universal prevention program on moral cognitions and antisocial behavior in young adolescents

Abstract

EQUIP for Educators (EFE) is an educational multi-component program, designed to motivate and teach students how to think and act responsibly. In the present study, the effects of EFE – implemented as a universal prevention program – on prevalence of and attitude towards antisocial behavior, self-serving cognitive distortions, moral judgment and social skills were evaluated longitudinally in a sample of 863 students attending prevocational secondary education. Latent Growth curve Models (LGM) showed an immediate effect on attitude towards antisocial behavior and both immediate and longer term effects on self-serving cognitive distortions. Ethnicity influenced the effect of EFE on self-serving cognitive distortions: the program appeared to be more effective for students with a Dutch ethnic background, compared to students from ethnic minority groups. In addition, repeated measures showed a delayed positive effect on teacher reported behavior. In sum, EFE was at least partly successful in its aim to equip students to think and act more prosocially.

Introduction

Due to the prevalence, seriousness and persistence of antisocial behavior among youth (Flannery et al., 2003; Wilson & Lipsey, 2007), a growing need for evidence-based prevention programs targeting a large variety of problem behaviors exists (Boxer, Goldstein, Musher-Eizenman, Dubow, & Heretick, 2005). Antisocial behavior can be defined as outward behavior that either directly or indirectly harms others through the violation of important moral or social norms, and includes aggressive and delinquent acts (Barriga, Morrison, Liao, & Gibbs, 2001; Liu, 2004). Evaluation of preventive interventions targeting antisocial behavior in adolescents is increasingly guided by developmental and prevention science (Aber, Brown, & Jones, 2003). The aim of prevention science is a greater understanding of predictors and mechanisms, leading to improvements in the development and implementation of preventive interventions (Aber et al., 2003; Biglan, Mrazek, Carnine, & Flay, 2003; Gottfredson & Gottfredson, 2002). The present study aims at investigating the longitudinal effects of a preventive program on self-serving cognitive distortions, moral judgment, social skills and prevalence of and attitude towards antisocial behavior, and the expected influence of changes in these characteristics on the effect of prevalence of antisocial behavior.

Preventive interventions

Preventive interventions can be divided into three types: *universal* preventive interventions (primary prevention), which focus on an entire population, *selective* or *targeted* interventions (secondary prevention), which focus on at risk subgroups in need for intervention, and *clinical* or *indicated* interventions (tertiary prevention), which focus on youth with detectable symptoms underlying antisocial behavior (Institute of Medicine, 1994). Limbos and colleagues (2007) argued in their review that higher effectiveness of prevention programs focusing on youth violence was reported as the level of intervention increased from universal to indicated. Although selective and indicated preventive interventions might be most efficient because of their focus on at risk youth, they on the other hand rely mainly on correct identification of these adolescents. Studies using teacher reports and parental reports for this screening process both suggest substantial rates of false positives (unnecessarily exposed to the intervention and the risks associated with labeling) or false negatives (not exposed to the intervention they could benefit from) (Offord, Kraemer, Kazdin, Jensen, & Harrington, 1998; Van Lier, Verhulst, & Crijnen, 2003).

Wilson and Lipsey (2007) found in their meta-analysis that the most common and effective approaches are universal interventions delivered to all the students in a classroom or school and targeted interventions for selected students who participate in programs outside their classroom. Since schools play a significant role in shaping the lives of adolescents, they are an important setting for interventions to prevent or reduce antisocial behavior in

adolescents (Cho, Hallfors, & Sánchez, 2005; Wilson, Gottfredson, & Najaka, 2001). They form a setting in which a lot of problem behavior (e.g. relational and physical aggression, vandalism) occurs and it is the only setting with almost universal access to adolescents (Wilson & Lipsey, 2007). School based programs are defined as ‘interventions operated in a school building, by school staff or under school or school system auspices’ (Wilson et al., 2001, p. 251). Universal, school based interventions are delivered in classroom settings to all the students in the class. Schools participating in such interventions often represent students with low socioeconomic background and/ or from higher crime neighborhoods who may be considered at risk. Interventions in this context might be particularly beneficial, because students showing prosocial skills and behavior can positively influence at risk students. Since the risk level of the students receiving the intervention forms a significant factor for the effects of universal interventions, the present study uses a sample comprising a substantial percentage of at risk students.

Although Wilson and Lipsey (2007) found substantial similarity of the mean effect sizes across treatment modalities (e.g. behavioral, cognitive and emotional oriented programs) for all three types of prevention, universal and selected interventions included in the analysis most frequently used cognitive approaches. In line with these findings, Wilson and colleagues (2001) showed that self-control interventions or interventions promoting social competency using cognitive-behavioral and behavioral methods were associated with positive effects. Other factors contributing to the effectiveness of preventive interventions targeting antisocial behavior in adolescents are: program intensity (frequency), focus (Wilson et al., 2001; Wilson & Lipsey, 2007), and implementation quality (Wilson & Lipsey, 2007).

Wilson & Lipsey (2007) reported in their review a mean effect size of .21 (standardized mean differences) for universal prevention programs. It should be noted that outcome measures based on self-report showed smaller effect sizes than measures from other sources or informants, while the most common type of measure used in this review to assess aggressive behavior was a teacher-reported survey. Irrespective program formats and modalities, these effect sizes represent a decrease in antisocial behavior and are not only statistically significant but also likely to be of practical significance to schools, since the numbers of students who act out decreased substantially.

Limitations of at risk youth

Youth who are at risk for the development of antisocial behavior often show limitations which can be defined in terms of the three D’s: sociomoral developmental Delay (that is persistence into adolescence of immaturity in moral judgment and egocentric bias), social cognitive Deficits and Distortions, and social skill Deficiencies. By remediating these delays and deficiencies, a reduction in antisocial behavior is expected (Gibbs, Potter, Barriga, & Liao, 1996).

Sociomoral Developmental Delay

Moral judgment refers to basic patterns of mature or immature thought and develops through a sequence of stages. Kohlberg (1984) originally identified six invariant hierarchically ordered stages of moral development, while Gibbs (2010) reduced this number to four. In his revision, stages 5 and 6 are not regarded as truly developmental stages, but represent an existential phase that universalizes the perspective of the standard stages 3 and 4. Stage 1 and 2 represent immature or superficial moral judgment, while with the transition to stage 3 and 4 moral judgment advances beyond superficiality to a mature understanding of moral norms and values (Gibbs, 2010) and may function as a buffer against antisocial behavior (Stams et al., 2006). Another aspect of the sociomoral developmental delay in at risk youth is the persistence into adolescence of a strong ‘me-centeredness’ or, in other words, a lack of social perspective-taking capacities.

Social Cognitive Deficits and Distortions

Immature moral judgment does not necessarily lead to antisocial behavior, unless it is combined with a high degree of self-serving cognitive distortions (Gibbs, 2003). According to social-cognitive theories (Dodge & Schwartz, 1997), people act upon their interpretation of social events and antisocial behavior is based on deficiencies in interpreting these events, i.e. cognitive distortions. Self-serving cognitive distortions can be defined as ‘inaccurate or biased ways of attending to or conferring meaning upon experiences’ (Barriga, Gibbs, Potter, & Liao, 2001, p.1) and are linked to externalizing behaviors. Barriga and colleagues (2001) developed a four-category typology of self-serving cognitive distortions: self-centeredness (according status to one’s own views and needs to such a degree that views of others are scarcely considered), blaming others (misattributing blame to external sources), minimizing/mislabeling (depicting antisocial behavior as causing no real harm or being acceptable) and assuming the worst (gratuitously attributing hostile intentions to others). Self-centeredness is considered a primary cognitive distortion that precedes antisocial behavior. Blaming others, minimizing/mislabeling and assuming the worst are secondary cognitive distortions that neutralize feelings of guilt and blame that accompany antisocial behavior (Cromwell & Thurman, 2003; Gibbs, 2003).

Social skill Deficiencies

Social skills are crucial for social action and can be defined as ‘balanced and constructive behavior in difficult interpersonal situations’ (Gibbs, Potter, & Goldstein, 1995, p.165). Since at risk youth typically do not know what specific steps are involved in constructive problem solving and have problems with social perspective taking (Dodge & Rabiner, 2004), social skills training is intended to displace out-of-control destructive responses with incompatible constructive ones.

EQUIP for Educators as a prevention program in secondary schools

The heterogeneity and diverse (treatment) needs of secondary school students suggest the need for interventions targeting these different deficiencies at the same time (Leeman, Gibbs, & Fuller, 1993). EQUIP is one of the multi-component interventions for adolescents with antisocial behavior problems, aiming to equip them to think and act more prosocially. Based on earlier research, EQUIP was found to be an effective intervention program for delinquent youth. Results of an American study using a preliminary version of the EQUIP program showed a substantial reduction of recidivism (Leeman et al., 1993), while results of a study with incarcerated boys in the Netherlands showed a significant reduction of cognitive distortions and a more negative attitude towards antisocial behavior in the experimental group as compared to the control group (Nas, Brugman, & Koops, 2005). Based on these promising results, the original treatment program for juvenile offenders has been adapted to a broader based prevention curriculum, EQUIP for Educators for students (grades 5-8) (DiBiase, Gibbs, Potter, & Spring, 2005).

The EFE program combines peer-group with skills-training approaches, and is based on a combination of the Positive Peer Culture (PPC) model and Aggression Replacement Training (ART). The PPC model aims to transform the negative peer group culture into a positive culture, in which individuals feel responsible for and help each other (Vorrath & Brendtro, 1985). ART (Goldstein, Glick, & Gibbs, 1998) aims to reduce helping-skills limitations, by social skills training, anger management, and moral education. EQUIP adapts and assimilates the ART curriculum into a PPC format (Leeman et al., 1993). Firstly, social skills are learned in the context of four phases: modeling, enactment (role play), providing feedback and practicing the skill (DiBiase et al., 2005). Secondly, to remediate sociomoral developmental delay, the moral education component uses group discussions to stimulate age-appropriate moral reasoning. Thirdly, the anger management component equips students with skills, such as self monitoring of emotions and thoughts and thinking ahead, to manage anger and to correct self-serving cognitive distortions.

Present study

The aim of the present study is to investigate the initial and long term effects (up till 9 months after the intervention) of EFE – implemented as a universal prevention program – on students' self-serving cognitive distortions, moral judgment, social skills, attitude towards antisocial behavior and both self reported and teacher reported prevalence of antisocial behavior. It is expected that the levels of self-serving cognitive distortions, attitude towards antisocial behavior, and prevalence of antisocial behavior will be reduced, while social skills and moral judgment are expected to improve after intervention. Furthermore, it is expected that these changes in self-serving cognitive distortions, moral judgment, social skills and prevalence of and attitude towards antisocial behavior occur immediately after the intervention and remain

stable at follow up. To investigate the relationships between the variables over time, a multivariate longitudinal growth model will be tested. It is hypothesized that the expected reduction in prevalence of antisocial behavior after the intervention is related to the increase of moral judgment and social skills, and the decrease of self-serving cognitive distortions and attitude towards antisocial behavior.

Based on research indicating that treatment outcome effects may vary according to students' initial behavior status prior to participation (Flannery et al., 2003; Stoolmiller, Eddy, & Reid, 2000), differential effectiveness of the preventive intervention will be examined. Since a universal prevention program such as EFE targets complete classes or grade levels, the sample will contain many low-risk participants who only rarely exhibit problem behaviors. Therefore, it might be difficult to establish a strong main effect, and differences in behavioral outcomes depending on initial status become of significance (Flannery et al., 2003). It is hypothesized that students' with higher baseline levels of antisocial behavior would show larger effect sizes, since students with higher levels of antisocial behavior have more room for improvement. On the other hand, one may assume that their behavior patterns are more entrenched and therefore more difficult to change.

In addition, sex and ethnic differences are taken into account since numerous studies established higher rates of antisocial behavior (Côté, Vaillancourt, Barker, Nagin, & Tremblay, 2007; Van Lier, Vitaro, Wanner, Vuijk, & Crijnen, 2005) and self-serving cognitive distortions (Barriga, Morrison, Liao, & Gibbs, 2001) for males than females, and ethnicity has been found to be an important variable in Dutch studies of antisocial behavior (Boendermaker & Van Yperen, 2003; Van der Laan & Blom, 2006). Ethnic status can be defined as an easily identifiable characteristic that implies a common cultural history with others possessing the same characteristic. The most common ethnic 'identifiers' are race, religion, country of origin, language and/ or cultural background (Okazaki & Sue, 1998). In the present study, a student is characterized as being a member of an ethnic minority group when at least one of his/ her parents is born in another country than the Netherlands (e.g. Morocco, Surinam, Turkey, Netherlands Antilles, or other).

Method

Based on geographic location – the western (mostly highly urbanized) part of the Netherlands – about 80 prevocational secondary schools in the Netherlands were approached by letter to participate in the present study. Twelve schools showed interest and were further informed about participation possibilities and preconditions. Schools could join the program when they were able to meet research requirements, meaning that they were able to participate with at least three eight grade classes, were willing to train the teachers (during an 18 hours course offered by the research project), and implement EFE in classroom settings in a period of 16 weeks (two school hours a week). Although all twelve schools were enthusiastic and

motivated, only nine schools were able to meet these research requirements and committed to voluntary participation. The other three schools could not take part due to logistic problems with school schedules (lack of available hours). Two of these latter schools participated in the study as non-intervention group and their eight grade students filled out the questionnaires at the same data waves as the intervention schools. Before the start of the implementation, two intervention schools dropped out because of management problems. Both schools were excluded from further participation. In sum, nine schools participated in the present study: seven intervention schools and two non-intervention schools.

Participants

Participants were 863 eight grade students from nine (7 intervention schools and 2 non-intervention schools) preparatory vocational secondary schools. A total of 705 students participated in the intervention group and followed EFE, while 158 students participated in the non-intervention group and followed the normal educational curriculum. The program was implemented in classroom settings and all students of the attending classes in the intervention group participated in the research. Since it is possible to repeat a class in the Netherlands, or start later in a grade based on cognitive functioning and/ or skills, the age range is rather large given all participants attended the same grade: representing students from 12.9 to 16.7 years (ages < 14, $n = 398$; ages 14-15, $n = 385$; ages 15-16, $n = 66$; ages 16 >, $n = 2$; missing, $n = 12$). The mean age at the start of the college year was 14.11 ($SD = .62$) years. Class-size ranged from 9 to 28 students, with smaller classes containing a higher percentage of students that needed additional teacher support. 56 percent of the participating students were boys. A majority (54.9%) of the participants belonged to an ethnic minority group. Socioeconomic status (SES) was ascertained by asking students about their father's highest completed education and appeared to be largely (86%) low, i.e. elementary school or secondary education. Note that the reported SES percentages are valid percentages, meaning that missing values were not taken into account. The percentage of students who did not know the educational level of their father was high (55%).

Table 1. *Overview of project design, data collection, and intervention*

	T1	T2	T3	EFE	T4	T5	T6
	n	n	n		n	n	n
Intervention Cohort 1	-	-	166	X	166	153	135
Intervention Cohort 2	-	381	331	X	308	322	-
Intervention Cohort 3	59	82	81	X	76	-	-
Non-Intervention Cohort	142	130	128	121			

Design and Procedure

This study used an extension of Latent variable Growth curve Modeling (LGM) (Duncan, Duncan, Strycker, Li, & Alpert, 1999; Singer & Willett, 2003). Data collection took place according to various time paths; data waves were spread over three different college years (seventh grade to ninth grade) to gain insight in the trajectories of moral cognitions and behavior, and longitudinal effects of the intervention. Cohorts were formed based on the number of pre- and posttest assessments (e.g. students in Cohort 1 had one pre assessment and three post assessments). According to possibilities based on school schedules, schools were assigned to one of these cohorts. In addition, schools that were not able to meet research requirements were assigned to the non-intervention cohort, which participated only at baseline measurements (Table 1). Students who followed the EFE program filled out one to three self report questionnaires prior to and one to three self report questionnaires after the intervention (four times in total), over a period of sixteen months with time intervals of four months. Students participating in the non-intervention group filled out the same questionnaires at four waves as well, following the timing of the second cohort. Furthermore, teachers filled out a questionnaire about externalizing behavior for each student participating in both the intervention and non-intervention group at pre and post test, with a time interval of eight months. After a brief oral and written instruction, students completed the self report questionnaires within classroom settings during school hours. Students were encouraged to ask the supervisor for help when needed. Parents received written information about the project, participation was granted by means of passive consent. Participants were assured of the confidentiality of their responses. Small rewards were provided.

Pilots educational material and questionnaires

The self report questionnaires were pre-tested in two different pilots and adapted according to feedback. The educational material from the EFE program was translated, adapted (examples used in the original EFE manual were replaced by less violent examples (Seroczynski, 2006), in accordance with the situation at Dutch schools), and expanded with a reader for students (with permission of the authors). The complete program was then tested in a pilot intervention of six months. Based on the results of the pilot and the feedback by the students and teachers, some small adjustments – especially concerning language – were made.

Teacher training and implementation

In each intervention school, teachers received an 18 hour training before implementation and 12 hours of individual and group wise supervision during implementation by a professional and experienced EQUIP trainer. During the training, teachers were provided with information on the background of the program, the content of the program, and teaching styles and they practiced role play, which was taped on video and evaluated afterwards. All schools used the EFE program for the first time. Due to practical circumstances, schools could only organize two EQUIP meetings per week instead of three, as advised by the authors (DiBiase et al.,

2005). The meetings were held in classroom settings, so groups stayed intact during implementation. Implementation of the program was monitored by the researchers and trainer: there was intensive and structured contact with all participating schools during the whole process. Teachers were asked to fill out evaluation forms (included in the original material) every fourth session, random observations took place during EQUIP meetings in the class, and evaluations after intervention with teachers and students were organized. Process evaluation of the implementation of EFE by both teachers and students participating in this study was positive: e.g. 85% of the teachers was positive about the effects of EFE on students' cognitions and behavior (see Chapter 4).

Intervention

The EFE curriculum as used in the present study consists of 32 classroom meetings (two introduction sessions were added to the original 30 lessons), coached by a teacher who was intensively involved with the class. Meetings of one school hour (50 to 60 minutes, depending on the school rules) were held twice a week, over a period of about four months. In the first two meetings, the rules that apply during the EQUIP meetings were being discussed and self-serving cognitive distortions, or 'thinking errors', were introduced. The other 30 meetings were structured around the three main components: anger management, social skills, and social decision making. Meetings from the different components were alternated to bring out the interrelationships among them (DiBiase et al., 2005).

Measures

The Antisocial Behavior Questionnaire. The Antisocial Behavior Questionnaire consisted of 12 items selected from self report scales used in previous research (Høst, Brugman, Tavecchio, & Beem, 1998; Leenders & Brugman, 2005). Each item, representing a specific type of antisocial behavior, was followed by a question about the prevalence of their behavior ('did you conduct this behavior?') and a question about their attitude towards antisocial behavior ('how wrong is it to conduct this behavior?'). After the intervention, the questions concerning the prevalence of their behavior were extended with a time period ('did you conduct this behavior in the past three months?'). The questionnaire could be scored on a 4-point scale (prevalence of behavior: 1 = never, 2 = once, 3 = sometimes, 4 = often; attitude towards behavior: 1 = very wrong, 2 = pretty wrong, 3 = not so wrong, 4 = not wrong at all). Mean scores for the scale 'Prevalence of antisocial behavior' (with scores above 1 indicating that the respondent committed, at least once, one or more forms of antisocial behavior) and for the scale 'Attitude towards antisocial behavior' (with low scores indicating a more negative attitude towards antisocial behavior) were used in the analysis. Reliability of these scales were adequate, with Cronbach's alpha's ranging from .72 to .77 at different assessments (cf. Cortina, 1993).

Teacher's Report Form (Achenbach, 1991). The Dutch translation of the externalizing behavior scale from the Teacher's Report Form (Verhulst, Ende, & Koot, 1997) was used to assess antisocial behavior as perceived by teachers. The questionnaire comprised 34 items on delinquency and aggression such as 'Fights a lot' or 'Threatens other people', all of which are rated as 0 = not true (as far as I know); 1 = somewhat true; and 2 = very true or often true. Teachers rated problems based on students' functioning over the preceding two months. Total scores were computed. Based on guidelines provided in the Teacher's Report Form Manual (Achenbach, 1991; Verhulst et al., 1997), students were assigned to normal (scores up to the 67th percentile), borderline (scores between the 67th and 70th percentile) and clinical range (scores above the 70th percentile) groups. Both international and Dutch studies established high validity and reliability of the instrument (Achenbach, 1991; Rescorla et al., 2007; Verhulst et al., 1997). The high reliability was replicated in the present study, with Cronbach's alpha's of .96 at both waves.

How I Think Questionnaire. To measure self-serving cognitive distortions the How I Think Questionnaire (Barriga, Gibbs et al., 2001), was used. The HIT comprises 39 items addressing four types of cognitive distortions (i.e. 9 self-centered; 10 blaming others; 9 minimizing; 11 assuming the worst), 8 so called anomalous response items (AR) designed to screen for suspect responding (for example: "In the past, I have lied to get myself out of trouble"), and 7 positive filler items to encourage full use of the scale. We replaced the 7 positive fillers by 11 social desirability items based on the Marlowe-Crowne questionnaire (Crowne & Marlowe, 1960). While the positive fillers were not meant to result in a meaningful score, the social desirability items do. Reliability of the social desirability subscale was found to be high (Cronbach's alpha's ranging from .74 to .77).

Each of the 39 cognitive distortion items refers to one or another of four categories of antisocial behavior derived from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM IV; American Psychiatric Association, 2000): disrespect for rules, laws or authorities (opposition defiance); physical aggression; lying; stealing. For example, the item "I can't help loosing my temper a lot" represents a blaming others cognitive distortion item applied to an opposition defiance behavioral referent. Participants responded along a 6-point Likert scale (from 'disagree strongly' to 'agree strongly'), with responses to individual items of 4 (slightly agree) or higher indicating cognitive distortion and responses of 3 (slightly disagree) or lower indicating absence of distortion. Mean scale and subscale scores were used in the analysis. Based on the American normative sample (comprising non-referred youth), students were assigned to normal (scores up to the 74th percentile), borderline (scores between the 74th and 84th percentile) and clinical range (scores above the 84th percentile) groups (Barriga, Gibbs et al., 2001). Nas and colleagues (2008) found that the Dutch translation of the instrument evidenced good validity and reliability. The reliability of the HIT in the present

study was high, with Cronbach's alpha's ranging from .66 to .82 on subscale level and Cronbach's alpha's of .90 to .94 for the overall score.

Moral Judgment Rating Task. This task assesses moral judgment development as defined by Kohlberg (Colby & Kohlberg, 1987). The Moral Judgment Rating Task (MJRT) consists of two times short dilemmas containing two conflicting moral values, followed by a series of reactions favoring the first value and a series of reactions favoring the second value. The dilemmas and the quotes are derived from the Moral Judgment Interview (MJI) manual (Colby & Kohlberg, 1987) and were used earlier in the Moral Judgment Sorting Task (Boom, Brugman, & Van der Heijden, 2001). Quotes are stage typed representing stages 1 to 4 and in-between stages as defined in the MJI manual. This results in 4 sets of 7 items is 28 items. The task is comparable to (and inspired by) the SROM-sf (Basinger & Gibbs, 1987), however, the response format is different in order to allow a Rating Scale analysis (Wright & Masters, 1982). Students categorized the quotes as 'for wiser people' (1), 'worthy of respect' (2), or 'childish' (3).

Rating scale analysis was used to assess how well the task was performed, because classical reliability (Cronbach's alpha) seems less appropriate since it assumes all items measure the same construct whereas our items are designed to reflect different stages of the same construct. Validity was supported in that, for full stages 1 to 3, the ordering in terms of difficulty estimates corresponded with the ordering proposed by Kohlberg (although separation between stages was weak). The higher the stage the quote represents, the more likely it is that it is judged as being wiser, keeping ability constant. The higher the ability of the respondent, the more likely an assertion is judged as being childish (keeping stage of quote constant). Since we found that stages higher than stage 3 are out of reach for most participants in our relatively young sample we decided to omit items reflecting these stages from further use in the analysis. In addition, the analysis revealed that the first set of quotes gave mixed results, not contributing positively to the expected ordering; therefore we omitted this set also from further use in the analysis. For further modeling we did not use the rather complex Rasch estimates and instead used the mean score over the items per person.

Inventory of Adolescent Problems-Short Form. Students' social skills in problematic and stressful interpersonal situations were assessed using an adapted version of the "Inventory of Adolescent Problems-Short Form" (IAP-SF; Gibbs et al., 1995). The original IAP-SF is a 20-item imaginable role play interview, describing difficult real life situations. In this study, a shortened written version of the instrument was used. Based on the situations used in the IAP-SF, four new situations were formulated (such as: "Michel's aunt is going to marry. His brother is wearing old and dirty jeans. Michel thinks this is very inappropriate."). Students were asked about their reaction in these situations, answering with 'yes' or 'no' on five possible (randomised) answers (e.g. 'you observe his dirty jeans, but ignore it and do not comment'). Formulated answers corresponded with the scoring categories of the original

instrument (highly socially skilled, a bit socially skilled, socially skilled, not socially skilled or socially unskilled, highly socially unskilled). A total mean score was computed. Reliability was low to moderate, with Cronbach's alpha ranging from .56 to .71.

Program integrity

Numerous researchers have argued that an essential prerequisite of evaluation research is to ensure integrity of treatment, that is, carrying out procedures as intended (Durlak & DuPre, 2008; Gottfredson & Gottfredson, 2002; Landenberger & Lipsey, 2005). In the present study, several steps were taken to address program integrity. To begin with, all the teachers involved in running the EFE curriculum sessions had at least an undergraduate degree and they received 18 hours of training. Second, the clear specifications of implementation in the EFE manual were used as a guideline, to carry out the intended EFE procedures accurately (DiBiase et al., 2005). Third, once the EFE sessions had begun, several monitoring procedures were initiated. Supervision meetings coached by the EQUIP trainer were organized, in order to provide teachers with feedback regularly. When needed, individual supervision was given, including observation of the teacher during an EFE session by the trainer and/ or co-training by the trainer to support the teacher. Furthermore, teachers were asked to complete checklists (DiBiase et al., 2005) as a review or self-evaluation after each fourth session. These checklists consisted of six general items ("Did you appropriately use the ask, don't tell method?") and seven component specific items (e.g. for social skills: "Did you read and discuss the steps of the social skills on the handout?"). Lastly, there were random observations during the EFE sessions at each participating school, in which the same checklists were completed by the researchers.

Statistical analysis

An extension of Latent variable Growth curve Modeling (LGM) was used to examine the initial and longer-term effects of the preventive program EFE over a three college year period (spanning approximately a year and a half in calendar years). LGM is a powerful and flexible technique, that can be used to model longitudinal change using repeated measures (Duncan et al., 1999). A basic linear LGM presupposes a steady increase or decrease over a small number of equally spaced measurement occasions for each person. The increase is assumed to be linear, with the average intercept and the average slope describing the trend. The individual intercepts are allowed to be different for each respondent and are assumed to be normally distributed in the population with unknown mean and variation. The individual slopes indicating the estimated amount of increase or decrease according to a linear function can also be different for each respondent and are also assumed to be normally distributed in the population with unknown mean and variation. In the present study, the basic LGM was extended for each of the five variables of interest, by modeling the possible effect of the

intervention as a second LGM in addition to a baseline LGM. The challenge is to disentangle the baseline trajectory and intervention effect trajectory with only four measurements available. To meet this challenge three intervention cohorts and one non-intervention cohort were used, which differ in terms of the number of pre and post measurements (see Table 1).

Assuming little or no change for the baseline trajectory, the three intervention cohorts are aligned on the intervention, resulting in a combined model with six virtual assessments (see Figure 1). This is in fact an accelerated cohort-sequential design as described by Duncan, et al. (1999). Each cohort covers another subset of four of the six virtual assessment occasions and has two virtual occasions missing by design. The intervention effect is modeled with one, two, or three of the virtual post intervention assessments. The same model is assumed in each intervention cohort as indicated by the fact that the intercepts and slopes in Figure 1 are constrained to be the same over cohorts. Intercept 1 corresponds to the baseline level of the variable (e.g. initial level of prevalence of antisocial behavior) and slope 1 represents a possible change over subsequent measurement occasions therein. Intercept 2 represents the immediate increase or decrease in the level after intervention, slope 2 indicates the change over measurement occasions after intervention (which represents long term effects).

Because both the intercepts and slopes are allowed to differ between participants (indicated by the variations of intercept 1 and 2, and slope 1 and 2) the model is very sensitive to possible effects. Since the full model is rather complex with 4 latent variables characterizing each of the 5 growth models, several steps were taken to reduce complexity, including the report of only significant and / or theoretically important parameters. Univariate models for the five variables of interest are presented; model 1 is a growth model, with an intercept term before intervention (but no slope term) and an intercept and slope term after intervention (IC1, IC2 and S2). This model is the most advanced model, in which both the intercept and the slope are able to change due to the intervention. This model assumes both changes in level (intercept) as in linearity (slope) after the intervention. Model 2 is a model with two intercepts, before and after intervention (IC1 and IC2), assuming only change in intercept after intervention. Model 3 is a model without intervention terms (but with slope term allowed). To determine statistical significance of the effects of EFE, model 1 and model 2 will be compared to model 3.

In addition, sex and ethnicity will be added to all models as potential predictors for the baseline and growth parameters, and interaction between these predictors will be investigated. Furthermore, the possible impact of initial level of antisocial behavior on the immediate and longer term effects of EFE will be investigated.

Finally, a multivariate LGM, including predictors and possible intervention effects, will be tested. In particular, the relationships between the potential changes in self-serving cognitive distortions, moral judgment, social skills, attitude towards antisocial behavior, and prevalence of antisocial behavior will be investigated.

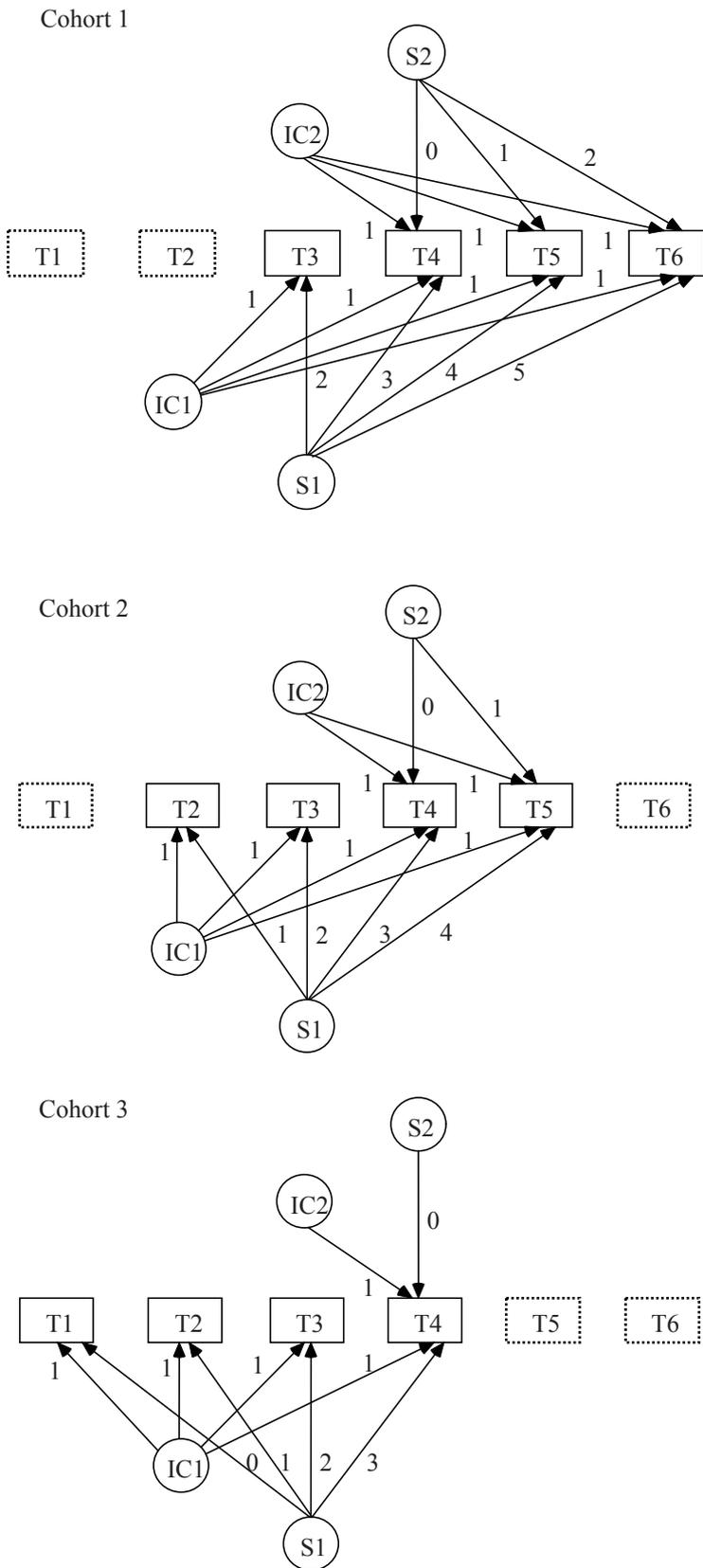


Figure 1. Representation of the cohort-sequential linear growth model

Note. IC1 = Intercept before intervention, S1 = Slope before intervention, IC2 = Intercept after intervention, S2 = Slope after intervention, Tx = Assessment x

For the analysis of all LGM's in the present study Mplus software (Muthén & Muthén, 1998-2007) will be used. Due to the high model complexity, the usual fit indexes are not helpful here; instead the Bayesian Information Criterion (BIC) are used to compare models that show an acceptable fit to the data. This criterion is based on the log-likelihood (log L) to compare different models (Schwartz, 1978). The lower the value of this criterion, the smaller the distance between the model at hand and the true model. The BIC is meant to compare nested models (one can be seen as a reduced version of the other). An absolute difference in BIC of more than 2 is meaningful, irrespective how large the BIC's are. Full Information Maximum Likelihood (FIML) estimation was used to deal with missing data. The models were directly fitted to all available data (Widaman, 2006).

In addition, repeated measures are used to investigate the effect of EFE on teacher reported antisocial behavior; LGM's could not be used because teachers only participated in two assessments (pre and post intervention).

Results

Missing data and attrition

Due to time constraints at assessment, the number of missing values fluctuated on the different measures used in this study. Moreover, for all measures individual scale scores were only computed if at least 80 percent of the items were scored. Attrition of scale scores ranged from 0.4% to 4.8% at first assessment, from 0.9% to 3.8% at second assessment, from 0.8% to 2.9% at third assessment, and from 0.8% to 4.1% at fourth assessment.

Furthermore, 53.1% of the students participated in all four assessments, 25.7% missed one assessment, 11.6% missed two assessments, and 9.6% missed three assessments. Univariate analysis of variance showed that students who were absent at at least one assessment showed the same prevalence of antisocial behavior as students who were present at all four assessments ($F(1, 747) = .04, p = .88$). Lastly, 104 respondents did not fill out the questionnaire including ethnicity, since this questionnaire was only assessed at the first data wave.

Descriptive statistics

Means and standard deviations for the different cohort samples at four assessments are presented in Table 2. Firstly, all cohorts, including the non-intervention cohort, show a decrease in prevalence of antisocial behavior. The decrease might be caused by the adaptation of the questionnaire after intervention, adding a time element to the items, instead of by the intervention itself (see Instruments). In modeling the latent growth of prevalence of antisocial behavior, a distinction between questionnaire effects and intervention effects is made.

Table 2. Descriptive statistics for students' self-reported moral cognitions and behavior

	T1			T2			T3			T4		
	M	SD	n									
<i>Intervention cohort 1</i>												
Prevalence of antisocial behavior	1.65	.38	166	1.40	.33	166	1.42	.35	153	1.43	.36	135
Attitude towards antisocial behavior	2.03	.52	164	1.91	.52	166	2.00	.52	153	2.10	.54	134
Self-serving cognitive distortions	2.55	.67	163	2.32	.67	166	2.34	.65	151	2.47	.69	133
Moral judgment	2.07	.35	166	2.09	.28	165	2.10	.33	147	2.11	.30	136
Social skills	3.47	.44	159	3.50	.48	158	3.47	.48	145	3.40	.50	130
<i>Intervention cohort 2</i>												
Prevalence of antisocial behavior	1.57	.41	381	1.54	.38	331	1.37	.33	308	1.33	.33	322
Attitude towards antisocial behavior	1.85	.55	380	1.86	.53	326	1.76	.55	305	1.78	.57	318
Self-serving cognitive distortions	2.64	.65	381	2.54	.64	329	2.51	.65	306	2.48	.76	322
Moral judgment	2.02	.32	358	2.02	.33	329	2.03	.32	305	2.07	.31	319
Social skills	3.52	.43	363	3.53	.44	319	3.53	.43	300	3.50	.50	308
<i>Intervention cohort 3</i>												
Prevalence of antisocial behavior	1.48	.29	59	1.53	.36	82	1.50	.31	81	1.34	.35	76
Attitude towards antisocial behavior	1.68	.52	59	1.68	.48	82	1.69	.51	79	1.61	.49	74
Self-serving cognitive distortions	2.57	.59	58	2.55	.59	81	2.53	.68	80	2.42	.72	75
Moral judgment	1.98	.32	58	2.01	.29	79	2.05	.27	81	2.05	.34	74
Social skills	3.55	.42	57	3.59	.40	80	3.59	.43	80	3.56	.46	74
<i>Non-intervention cohort</i>												
Prevalence of antisocial behavior	1.58	.36	142	1.63	.37	130	1.39	.32	128	1.37	.32	121
Attitude towards antisocial behavior	1.75	.50	142	1.77	.49	129	1.86	.56	128	1.79	.57	121
Self-serving cognitive distortions	2.47	.56	142	2.43	.60	130	2.54	.65	128	2.50	.75	120
Moral judgment	2.07	.34	142	2.11	.31	131	2.13	.30	128	2.10	.35	121
Social skills	3.49	.39	139	3.50	.46	128	3.41	.51	126	3.41	.52	120

Note. Cohort 1 = 1 pre assessment, 3 post assessments; Cohort 2 = 2 pre assessments, 2 post assessments; Cohort 3 = 3 pre assessments, 1 post assessment.

Secondly, all intervention cohorts show a decrease in attitude towards antisocial behavior and self-serving cognitive distortions after intervention, while the non-intervention cohort stays stable or shows an increase. Thirdly, moral judgment increases in all cohorts – including the non-intervention cohort – over time. Lastly, social skills remain quite stable over the four assessments for all intervention cohorts, while there is a small decrease in social skills after the second assessment for the non-intervention cohort.

Effects of EQUIP for Educators

Univariate LGM's

The immediate and longer term effects of EFE were investigated using accelerated cohort-sequential two-part LGM's for students' prevalence of antisocial behavior, attitude towards antisocial behavior, self-serving cognitive distortions, moral judgment, and social skills. Table 3 contains the fit indices (BIC values), means and variances for the relevant models.

Analysis of univariate LGM's for prevalence of antisocial behavior, in which we corrected for the questionnaire effect, indicated that the model without intervention terms (Model 3) had the lowest BIC value and provided a better fit than the effect models (Model 1 and Model 2). Both the intercept and slope term after intervention appeared to be non significant, see Table 3. Latent growth curves for attitude towards antisocial behavior showed that Model 1, the model with both an intercept and slope term after intervention, fitted best (BIC value of 3338 versus 3350 for Model 3 and 3353 for Model 2), indicating changes in level ($M_{diff} = -.08$) and linearity ($Slope_{post} = .07$) after the intervention. The effect size is small, Cohen's $d = .15$ (Cohen, 1988). For self-serving cognitive distortions, the model including the intercept after intervention (Model 2) showed the best fit. The intercept of self-serving cognitive distortions decreased after intervention ($M_{diff} = -.10$) although the effect size is small, Cohen's $d = .15$ (Cohen, 1988). Finally, analysis of univariate LGM's for both moral judgment and social skills indicated the best fit for Model 3, the model without intervention terms. To control for possible influences of social desirability and age, both variables were added to the models as described above. All models including social desirability and age showed a worse fit, so these predictors were left out of the analyses.

Since LGM's represent linear trajectories at student level and allow for individual variance, the explained variance in these models can be divided into different components (regarding intercept, slope, intervention intercept and intervention slope). This division increases the proportion of explained variance by the model compared to fixed effects models (self-serving cognitive distortions 70%, moral judgment 36%, social skills 56%, prevalence of antisocial behavior 71%, and attitude towards antisocial behavior 70%).

Table 3. Effects of EQUIP for Educators on antisocial behavior, self-serving cognitive distortions, moral judgment, and social skills

	BIC	Intercept 1		Slope 1		Intercept 2		Slope 2		Residual variance			
		M	Var	M	Var	M	Var	M	Var	RV1	RV2	RV3	RV4
Prevalence of antisocial behavior													
Model (IC1, IC2, S2)	1153	1.57	.08	-	-	.03 ns	.00 ns	-.01 ns	-	.06	.04	.03	.04
Model (IC1, IC2)	1148	1.57	.08	-	-	.02 ns	.00 ns	-	-	.06	.04	.03	.04
Model (IC1, S1)	1078	1.58	.10	-.00 ns	-	-	-	-	-	.05	.03	.03	.04
<i>Dutch boys</i>		1.67 ^a											
<i>Dutch girls</i>	1030	1.67	.07	-.01 ns	-	-	-	-	-	.06	.04	.03	.04
<i>Ethnic minority boys</i>		1.54											
<i>Ethnic minority girls</i>		1.54											
<i>Initial level: TRF normal</i>		1.45 ^a											
<i>Initial level: TRF borderline</i>	961	1.54	.08	.00 ns	-	-	-	-	-	.07	.04	.03	.04
<i>Initial level: TRF clinical</i>		1.63											
Attitude towards antisocial behavior													
Model (IC1, IC2, S2)	3338	1.85	.18	-	-	-.08	.05	.07	-	.12	.09	.08	.11
Model (IC1, IC2)	3353	1.85	.18	-	-	-.05	.05	-	-	.12	.09	.08	.11
Model (IC1, S1)	3350	1.83	.15	-.00 ns	.01	-	-	-	-	.13	.10	.08	.09
<i>Dutch boys</i>		2.01 ^a						.13 ^a					
<i>Dutch girls</i>	3022	1.91	.17	-	-	-.11	.05	.03	-	.12	.09	.08	.10
<i>Ethnic minority boys</i>		1.81				-.11		.13					
<i>Ethnic minority girls</i>		1.71				-.11		.03					
<i>Initial level: TRF normal</i>		1.82 ^a				-.08		.08					
<i>Initial level: TRF borderline</i>	3064	1.87	.18	-	-	-.08	.05	.06	-	.12	.08	.08	.11
<i>Initial level: TRF clinical</i>		1.92				-.08		.04					
Self-serving cognitive distortions													
Model (IC1, IC2, S2)	4467	2.56	.28	-	-	-.11	.07	.02 ns	.01 ns	.14	.14	.12	.19
Model (IC1, IC2)	4455	2.56	.28	-	-	-.10	.07	-	-	.14	.14	.12	.20
Model (IC1, S1)	4456	2.59	.23	-.03	.01	-	-	-	-	.16	.14	.14	.23
<i>Dutch boys</i>		2.62 ^a				-.15 ^a							
<i>Dutch girls</i>	4085	2.44	.26	-	-	-.15	.07	-	-	.14	.14	.12	.20
<i>Ethnic minority boys</i>		2.62				-.03							
<i>Ethnic minority girls</i>		2.44				-.03							

Table 3. Effects of EQUIP for Educators on antisocial behavior, self-serving cognitive distortions, moral judgment, and social skills (continued)

	BIC	Intercept 1		Slope 1		Intercept 2		Slope 2		Residual variance			
		M	Var	M	Var	M	Var	M	Var	RV1	RV2	RV3	RV4
<i>Initial level: TRF normal</i>	4141	2.52 ^a	.27	-	-	-.11	.07	-	-	.14	.14	.12	.20
<i>Initial level: TRF borderline</i>		2.59				-.07							
<i>Initial level: TRF clinical</i>		2.66				-.03							
Moral judgment													
Model (IC1, IC2, S2)	1213	2.05	.04	-	-	.01 ns	.01	.03	-	.06	.06	.05	.06
Model (IC1, IC2)	1212	2.05	.04	-	-	.02 ns	.01 ns	-	-	.06	.06	.06	.06
Model (IC1, S1)	1203	2.01	.04	.02	.00 ns	-	-	-	-	.06	.06	.06	.06
<i>Dutch boys</i>		2.01 ^a											
<i>Dutch girls</i>	1093	2.09	.03	.02	-	-	-	-	-	.06	.06	.06	.06
<i>Ethnic minority boys</i>		1.95											
<i>Ethnic minority girls</i>		2.03											
<i>Initial level: TRF normal</i>	1090	2.02	.03	.02	.00 ns	-	-	-	-	.06	.06	.06	.06
<i>Initial level: TRF borderline</i>		2.02											
<i>Initial level: TRF clinical</i>		2.02											
Social skills													
Model (IC1, IC2, S2)	2532	3.51	.12	-	-	.00 ns	.04	-.03	-	.07	.07	.08	.10
Model (IC1, IC2)	2530	3.51	.12	-	-	-.01 ns	.04	-	-	.07	.07	.08	.10
Model (IC1, S1)	2522	3.54	.09	-.02	.01	-	-	-	-	.08	.08	.08	.09
<i>Dutch boys</i>		3.47											
<i>Dutch girls</i>	2309	3.47	.09	.00	.01	-	-	-	-	.07	.07	.08	.10
<i>Ethnic minority boys</i>		3.47											
<i>Ethnic minority girls</i>		3.47											
<i>Initial level: TRF normal</i>	2343	3.56 ^a	.09	-.02	.01	-	-	-	-	.08	.08	.08	.08
<i>Initial level: TRF borderline</i>		3.52											
<i>Initial level: TRF clinical</i>		3.48											

Note. IC1 = Intercept before intervention, S1 = Slope before intervention, IC2 = Intercept after intervention, S2 = Slope after intervention. M = Mean, Var = Variance, RV x = Residual variance at assessment number x, BIC = Bayesian Information Criterion
^a= Difference between groups (sex or ethnic) is significant

Sex and ethnic differences and initial status of antisocial behavior

The influence of sex and ethnic differences on the baseline levels of the main variables (loading on intercept and when possible slope terms at baseline) and their influence on the effect of the intervention (loading on intercept and slope terms after intervention) was analyzed by adding these (dummy) variables to the LGM (with the lowest BIC value) as predictors. The models including ethnicity and sex as predictors showed a better fit than the models without predictors, see Table 3. As shown in Table 3, there are differences between ethnic groups and sex at baseline for all variables, except social skills. Students from the Dutch ethnic group show higher levels of prevalence and attitude towards antisocial behavior and moral judgment at baseline. Further, boys report a more positive attitude towards antisocial behavior, a higher level of self-serving cognitive distortions, and a lower level of moral judgment. With regard to the influence of the predictors on the effect of the intervention, a significant difference between ethnic groups was found on self-serving cognitive distortions. Students from ethnic minority groups show a lower decrease in intercept after intervention than students with a Dutch ethnic background (-.03 versus -.15). Furthermore, sex appears to play a role in the change over measurement occasions after intervention for attitude towards antisocial behavior. While the immediate effect on attitude towards antisocial behavior gradually reduces for boys, the long term effect remains almost stable for girls (.13 versus .03).

The differential effectiveness, based on initial level of antisocial behavior, was investigated by comparing normal, borderline and clinical groups based on TRF profiles (Achenbach, 1991), see categorization as described in the Instruments section. Although the models including TRF as a predictor for effectiveness showed no significant effects on both intercept and slope terms after intervention for all main variables, differences at baseline could be established for most variables. In line with the expectations, the prevalence of antisocial behavior, attitude towards antisocial behavior and self-serving cognitive distortions increased from normal to clinical, while the level of social skills decreased from normal to clinical. There were no baseline differences for moral judgment (see Table 3).

Finally, interaction effects of predictors were investigated (ethnicity X sex, ethnicity X initial level of antisocial behavior, sex X initial level of antisocial behavior) by adding the product scores as predictor terms to the models. None of the interactions appeared to have impact on the effect of the intervention: models including interaction terms did not fit well (high BIC values), and intercept and slope terms after intervention did not show significant differences.

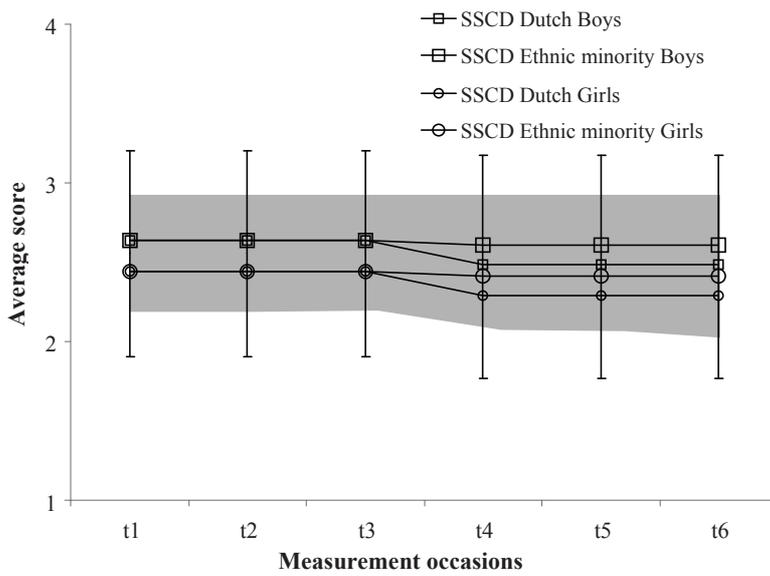


Figure 2. Average effect of EFE on self-serving cognitive distortions (SSCD)

Bivariate model with influence of attitude towards antisocial behavior taken into account (see Figure 4) and with differential effect of predictors (sex and ethnic group) illustrated. Intervention occurred between t3 and t4. Error bars represent modeled standard deviations (one SD up and one SD down from the average score) to visualize effect size. The grey band represents modeled residuals in standard deviation metric as an alternative way to visualize the magnitude of the effects.

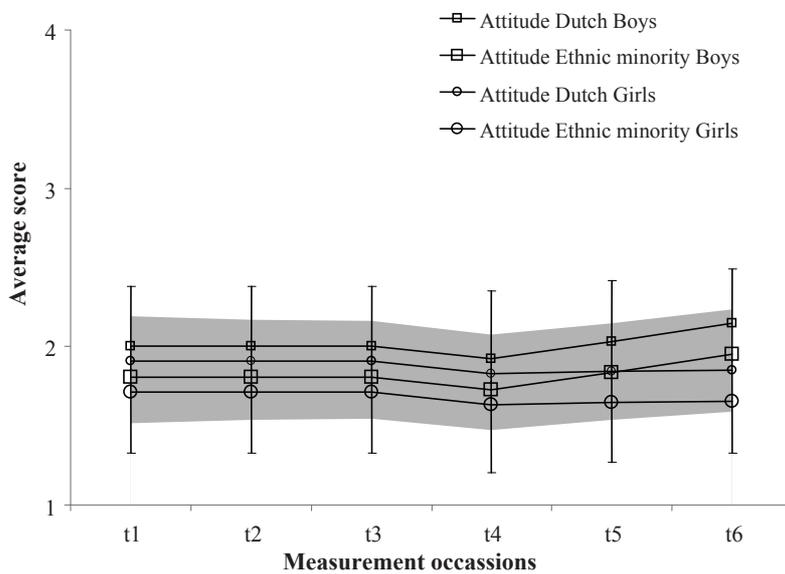


Figure 3. Average effect of EFE on attitude towards antisocial behavior (Attitude)

Bivariate model with influence on self-serving cognitive distortions taken into account (see Figure 4) and with differential effect of predictors (sex and ethnic group) illustrated. Intervention occurred between t3 and t4. Error bars represent modeled standard deviations (one SD up and one SD down from the average score) to visualize effect size. The grey band represents modeled residuals in standard deviation metric as an alternative way to visualize the magnitude of the effects.

Bivariate LGM

The bivariate LGM included the effect intercepts of self-serving cognitive distortions and attitude towards antisocial behavior, the effect slope of attitude towards antisocial behavior, and the predictor terms sex and ethnicity (see Figure 4). Findings replicated results from the univariate models, by confirming the influence of ethnicity on the baseline level of attitude towards antisocial behavior and on the effect of self-serving cognitive distortions (see Figure 2). These findings indicate that the baseline levels of attitude towards antisocial behavior were lower for students from ethnic minority groups (they had a more negative attitude) and that the effect of EFE on self-serving cognitive distortions was smaller for students from ethnic minority groups. Furthermore, in line with the univariate findings, the bivariate model showed lower baseline levels of both self-serving cognitive distortions and attitude towards antisocial behavior for girls (see Figure 2 and 3). In addition, the bivariate model shows that both the baseline levels and the effects of self-serving cognitive distortions and attitude towards antisocial behavior are highly correlated ($r = .63$ and $r = .81$ respectively).

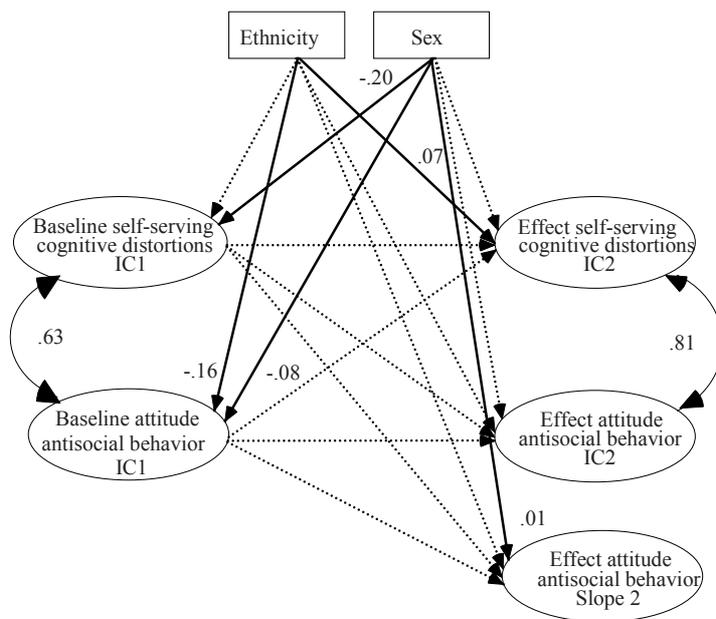


Figure 4. Bivariate latent growth model including ethnicity and sex

Note. Dotted lines represent non-significant paths.

The figure represents a simplified visualization of two growth models (IC1 and IC2 correspond to these terms as used in Figure 1).

Teacher reported antisocial behavior

Since teacher reports of antisocial behavior were only assessed twice (pre and post intervention), repeated measures instead of accelerated growth models were used to investigate the effects of EFE. Results for the total group showed a small increase of teacher reported behavior for both the intervention ($M_{pre} = 50.65$, $SD_{pre} = 11.22$; $M_{post} = 52.42$, SD_{post}

= 12.56) and non-intervention group ($M_{pre} = 50.31$, $SD_{pre} = 11.42$; $M_{post} = 53.03$, $SD_{post} = 10.47$). The difference between groups appeared to be non-significant ($F(1, 555) = 1.28$, $p = .26$). However, a part of the participating schools implemented EFE in the first semester (September to February), while the other part of the schools implemented EFE in the second semester (February to June). Since teacher reports were assessed at the same time for all schools, i.e. at the beginning of September and at the end of June, it was investigated whether there was a difference in immediate effects (directly after the intervention) and delayed effects (four months after the intervention) of EFE on teacher reported antisocial behavior. Repeated measures showed that there was no immediate effect – effect for the schools who implemented EFE in the second semester – on teacher reported antisocial behavior ($F(1, 252) = 2.83$, $p = .09$). Both the intervention ($M_{pre} = 50.51$, $SD_{pre} = 10.07$; $M_{post} = 54.99$, $SD_{post} = 12.26$) and non-intervention group ($M_{pre} = 50.31$, $SD_{pre} = 11.42$; $M_{post} = 53.03$, $SD_{post} = 10.47$) increased in teacher reported antisocial behavior. However, analysis did show that there was a delayed effect of EFE on teacher reported antisocial behavior ($F(1, 430) = 5.95$, $p = .02$, $\eta_p^2 = .01$). The intervention group ($M_{pre} = 50.70$, $SD_{pre} = 11.68$; $M_{post} = 51.36$, $SD_{post} = 12.54$) showed a smaller increase in teacher reported antisocial behavior compared to the non-intervention group ($M_{pre} = 50.31$, $SD_{pre} = 11.42$; $M_{post} = 53.03$, $SD_{post} = 10.47$).

Discussion

This study examined the immediate and longer term effects of EQUIP for Educators implemented as a universal prevention program in prevocational secondary schools. Accelerated LGM's, including six assessments centered on the intervention, were used to evaluate the effects of EFE on prevalence of antisocial behavior, attitude towards antisocial behavior, moral judgment, and social skills of eight grade students in the Netherlands.

First of all, in line with previous research on the effects of EQUIP among delinquent youth (Nas et al., 2005), EFE was found to influence the way students evaluated antisocial behavior: students' attitude towards antisocial behavior became more negative immediate after the intervention. The change to a more negative attitude towards antisocial behavior after intervention might indicate that students relabeled antisocial behavior by correcting minimizing/ mislabeling cognitive distortions, which can be an important step in the process of changing actual behavior. However, the influence of EFE on attitude towards antisocial behavior is only a temporarily effect, which gradually reduces over time. This latter finding mainly concerned boys, which indicates that girls are more susceptible for maintenance of attitudinal changes. Continuation of the intervention or embedding the EQUIP principles in the academic curriculum are possibly needed to maintain the positive impact of attitude towards antisocial behavior.

Furthermore, the present study also established an effect of EFE on self-serving cognitive distortions: students showed a decrease in level of self-serving cognitive distortions

immediate after the intervention, which remained stable over follow up assessments up till nine months after intervention. These results indicate that EFE was successful in changing students' cognitions over an extended period of time. An explanation for the differences in longer term effects of EFE on attitude and cognition might be that self-serving cognitive distortions represent cognitive restructuring. In addition, the bivariate LGM showed a strong relationship between the direct effect of EFE on self-serving cognitive distortions and the direct effect on attitude towards antisocial behavior. Since cognitions and attitudes are presumed to precede and to predict behavior (Leenders & Brugman, 2005; Vitaro, Brendgen, & Tremblay, 2000), these findings are promising for the effects of the intervention on behavior.

LGM's showed no differences in level and rate of change after intervention for both prevalence of antisocial behavior and prosocial behavior, that is social skills. A previous study on the preliminary version of EQUIP for juvenile delinquents did establish a reduction in antisocial behavior after intervention, although it should be noted that these behavioral changes were limited to a correctional facility (Leeman et al., 1993). Leeman and colleagues (1993) also established significant gains in social skills in the EQUIP group after intervention. However, these positive effects of the intervention on both prosocial and antisocial behavior could not be replicated by Nas and colleagues (2005) who used the final version of EQUIP. The findings of the Nas et al. (2005) study on incarcerated juvenile male delinquents are in line with the present study, although Nas and colleagues (2005) only report immediate effects of the intervention. In addition, the present study evaluated EFE over a longer period of time, including four data waves.

In the past decades, a substantial part of the research concerning antisocial behavior focused on the development of and predictors underlying antisocial behavior in order to gain insight in effective ways to change this behavior and to break through the negative spiral (e.g. Moffitt, 1993; Tremblay, 2000). Despite these efforts, reviews of the literature summarizing efforts to intervene with antisocial adolescents invariably lead to negative conclusions. Antisocial behavior appears to be difficult to change and most evaluations of prevention programs present small or at most modest effect sizes (Wilson & Lipsey, 2007). According to the results of the present study, effect sizes for both attitude towards antisocial behavior and self-serving cognitive distortions are small as well (Cohen, 1988). These findings are in line with previous studies reporting effect sizes of diverse universal preventive intervention programs. Limbos and colleagues (2007) argue that effect sizes depend on type of prevention, with increasing success from universal to indicated prevention. Since EFE was implemented as a universal program in the present study – targeting complete classes – the sample included a substantial part of students with normal or low rates of antisocial behavior at pre test. These sample characteristics might have contributed to the small effect sizes (Flannery et al., 2003), but differential effectiveness could not be established in the present study.

Lastly, effects of EFE on moral judgment could not be established in this study, similar to the findings of previous studies evaluating the EQUIP program (Leeman et al., 1993; Nas et al., 2005). Leeman and colleagues (1993) speculated there might be a ‘*sleeper effect*’, in other words a delayed effect on moral judgment, but the longer term results of this study do not support this possibility. Since none of the available studies on EQUIP provided evidence for the positive impact of the intervention on moral judgment, it could be questioned in what way the moral judgment component contributes to the success of the curriculum on other outcome variables. As Palmer (2007) has noted, the effects of the different components within the EQUIP program have not been dismantled so far, and there is no way of knowing “whether the moral reasoning component brings additional value to such interventions” (Palmer, 2007, p. 211). Further research concerning the processes leading to the effects on attitude towards antisocial behavior and self-serving cognitive distortion is needed, including specification of the role and impact of the different components and formats of EQUIP.

Although previous studies stressed the importance of possible differences between boys and girls in antisocial behavior and related processes (Barriga, Morrison et al., 2001; Côté et al., 2007; Van Lier et al., 2005), sex only appeared to be of influence on the longer term effects of EFE on attitude towards antisocial behavior: effects remained stable for girls over time, while they reduced for boys. As expected, ethnicity was found to play a moderating role in the effectiveness of the intervention. According to the LGM’s including ethnicity as a predictor, effects of EFE on self-serving cognitive distortions were stronger for students with a Dutch ethnic background compared to students from ethnic minority groups. These findings are of great importance for future implementation in multicultural groups and require further investigation on possible differences between ethnic subgroups.

It was hypothesized that the reduction of prevalence of antisocial behavior was related to an increase in moral judgment and social skills, and a decrease in self-serving cognitive distortions and attitude towards antisocial behavior. Since reduction on prevalence of antisocial behavior after intervention could not be established, further investigation of the influence of self-serving cognitive distortions, moral judgment, social skills, and attitude towards antisocial behavior was not possible. However, a bivariate LGM was analyzed to gain insight in the relationship between the effects, which appeared to be strong.

Although repeated measures including pre and post assessment of teacher reports showed no immediate effects, a delayed effect (four months after the intervention) on teacher reported antisocial behavior could be established. It appears that changes in behavior as reported by teachers become visible after a longer period of time. An explanation for this finding can be that teachers were asked to assess their students’ behavior over the preceding two months. This might have biased immediate effects, since this period of two months includes a substantial part of the implementation phase of the intervention and covers half of the intervention time. Furthermore, changes in prevalence of antisocial behavior as assessed

by teachers were found to be opposite to the expectations: students showed an increase in antisocial behavior instead of the expected decrease. This increase was significantly smaller in the intervention group, compared to the non-intervention group. Since rates of teacher reported antisocial behavior dropped substantially from immediate to delayed assessment, it might be possible that a 'sleeper' effect occurs in this case, meaning that the effects on teacher reported antisocial behavior will become visible and more powerful after a longer period of time. An extension in number of teacher assessments covering a longer time span is needed to further investigate the impact of the intervention on the long term.

Strengths, limitations and future directions

The findings of the present study contribute to the evaluation of universal prevention programs implemented in secondary schools. First of all, this study used an innovative combination of a longitudinal accelerated cohort-sequential design and an interrupted time-series design with switching replications, in which cohorts followed different time paths and differed in number of assessments before and after intervention (Table 1). Randomization was not regarded as necessary because the intervention cohorts fulfilled a controlling role as well. Since the effect of the intervention is determined for each individual compared to his own baseline level, the effect of bias due to differences in baseline is accounted for, and by implication, bias due to differences between cohorts are accounted for as well.

However, some limitations should be noted. A first limitation pertained to the fact that the program was implemented for the first time in all schools. Although teachers were trained intensively, they had to get used to the new way of teaching and their coaching role. Lack of experience might have contributed to the small effect sizes of cognitions and attitude, and to the absence of effects on behavior and moral judgment. As pointed out in several reviews (Wilson & Lipsey, 2007; McGuire, 2008), program intensity is an important factor contributing to the effectiveness of prevention programs. In contrast, Bennett and Gibbons (2000) state that frequency is not of influence on effectiveness. In the present study, schools were only able to teach EFE twice a week, while the authors of the program recommend three sessions a week (DiBiase et al., 2005). Although there is no consensus concerning the importance and influence of frequency in the literature, this deviation from program standards might have had an impact on the findings concerning the effectiveness of the EQUIP program. Further, the principles of EFE were only applied in the specific EFE sessions, since the program was not embedded in the larger environment of the school. This restriction complicated transfer of what was learned to other situations. Repeating the message is characteristic for the EQUIP philosophy and stresses the importance of expansion of environments in which EQUIP rules and principles are followed and coached.

Lastly, attrition rates were substantial, with 53 percent of the students present at all four assessments. The issue of attrition is a major concern in longitudinal studies (Skara &

Sussman, 2003), especially when adolescents are the study population (Stephens, Thibodeaux, Sloboda, & Tonkin, 2007). Minimization in loss of participants is needed in order to maintain the integrity of the cohort. High attrition rates can negatively influence both internal and external validity of evaluation research. However, missing data analyses in the present study showed that there were no differences in baseline levels of antisocial behavior between students who were absent at all four assessments and students who were absent during at least one assessment.

To gain a more complete insight in the impact of the intervention, future research should also focus on changes perceived by respondents in the broader environment of the students. Parent reports should be added to assess change in behavior at home and observations in classes should be included to gain insight in interactions among students and teachers, and in the teacher characteristics which might contribute to the success of the program.

Lastly, this study also reinforces the need to consider the potential influence of ethnic background on the effectiveness of prevention and intervention programs. Although the findings of the present study show effects of EFE for all students, larger intervention effects on self-serving cognitive distortions were found for students with a Dutch ethnic background. Future research should include subgroups of ethnic minorities.

Conclusion

The present study replicated the immediate effects found in the study among male delinquents by Nas and colleagues (2005), and extended previous research by including a large sample and longitudinal evaluation of the program. Despite the differences in context (correctional facility versus school classes), format (alternative version of EQUIP), composition of the sample (only boys versus boys and girls in the present study), and sample size (small versus large sample in the present study), effects of both studies are identical. This replication and extension of findings in two complete different samples indicates at least that EQUIP is consistent and stable in its influence on cognitions and attitude.

Self reports show that EFE had an effect on attitude towards antisocial behavior and self-serving cognitive distortions, but did not influence prevalence of antisocial behavior, moral judgment and social skills. In addition, a small effect on teacher reported antisocial behavior was established. In sum, the program appeared to be at least partly successful in its attempt to equip students to think and act more prosocially. Furthermore, future research is needed to gain insight in samples which benefit from a cognitive-behavioral prevention program such as EFE.

Chapter 6

General Discussion

Antisocial behavior in adolescents is highly prevalent and often associated with negative outcomes throughout life (Liu, 2004; Redding, Goldstein, & Heilbrun, 2005). Since antisocial behavior occurs along a developmental continuum of behavioral severity (Schaeffer, Petras, Ialongo, Poduska, & Kellam, 2003), effective prevention programs are needed to alter this negative developmental pathway. The current thesis was based on the assumption that antisocial behavior in adolescents can be explained by limitations in the moral, cognitive and social skills domain. Therefore, this thesis aimed at investigating 1) an explanatory model of antisocial behavior based on these limitations, 2) the psychometric quality and the suitability for evaluation of an instrument assessing self-serving cognitive distortions, and 3) the evaluation of EQUIP for Educators, implemented as a school based universal prevention program for prevocational secondary school students.

Summary of main findings

Relationships between moral cognitions and antisocial behavior

The importance of studying antisocial behavior within a developmental context is increasingly emphasized by researchers attempting to untangle the roots and consequences of this problem behavior across the lifespan (Rutter, Moffitt, & Caspi, 2006). In addition, the role of (interaction between) diverse individual and environmental factors is being recognized. Recent longitudinal studies reveal heterogeneity in the development of antisocial behavior (e.g., Odgers et al., 2008), emphasizing its multidimensional nature. Although there is still no absolute consensus regarding the most salient factors contributing to antisocial behavior (Odgers, Vincent, & Corrado, 2002), the second chapter of the current thesis focuses on the important and often studied link between moral cognitions and moral behavior. This study builds on the findings by Barriga, Morrison, Liau and Gibbs (2001), investigating an explanatory model of antisocial behavior based on moral cognitive processes, comprising self-serving cognitive distortions, moral judgment, and perception of community. Follow up data in a sample of 724 students from eight preparatory vocational secondary schools were used to investigate whether moral cognitions predicted antisocial behavior.

In contrast to the expectations (Gibbs, 2003), the present study could not provide evidence for a model in which antisocial behavior could be explained by the proposed moral cognitions. Self-serving cognitive distortions even appeared to be preceded by self reported antisocial behavior. These findings can be explained by the assumption that self-serving cognitive distortions were primarily used as neutralizers in an attempt to reduce feelings of guilt or shame (Bandura, 2002; Festinger, Carlsmith, & Bem, 2007). Nonetheless, the relationship between self-serving cognitive distortions and antisocial behavior was – at least for a substantial part of the population (for girls) – reciprocal, indicating the possible relevance of cognitive-behavioral approaches in prevention programs targeting antisocial behavior. In

addition, models predicting antisocial behavior based on self report were found to be much stronger than these models based on teacher report, possibly reflecting the limited observation possibilities by teachers (Achenbach, Dumenci, & Rescorla, 2002). Furthermore, the findings as described in Chapter 2 were based on a limited time frame – including two assessments with a four month interval – emphasizing the need for longitudinal research including different time sequences and using a more experimental design in order to enlarge knowledge on mediation processes and the causality of cognition and behavior in this context.

Despite the enormous amount of literature on the processes leading to antisocial behavior and on effective preventive interventions targeting these problem behaviors, there is still little known about differences in effects between various subpopulations (Limbos et al., 2007). In the current study, small but significant sex and ethnic differences were established, indicating the existence of differential developmental processes. In this respect, girls' self-serving cognitive distortions appeared to precede moral judgment and antisocial behavior, while boys' perception of community appeared to precede self-serving cognitive distortions. These differences can be explained by the assumption that self-serving cognitive distortions is an individual characteristic, which play a more substantial role for girls, while affiliation with a larger social network and the peer group is of greater influence for boys (Cillissen & Mayeux, 2004). For students from ethnic minority groups, stage of moral judgment appears to be influenced by both self-serving cognitive distortions and perception of community, possibly partly reflecting the idea that their moral cognition is based on collectivist values (De Mey, Baartman, & Schulze, 1999). For students with a Dutch ethnic background, perception of community appeared to precede self-serving cognitive distortions, suggesting a relative strong impact of peers in the native population. The native population is characterized by a more individualistic upbringing, in which emotional independence of parents and bonding with peers are among the most important developmental tasks for adolescents (Stevens & Vollebergh, 2008). Although the sex and ethnic differences as reported in this study were small, they stress the complexity of the processes underlying antisocial behavior. Future studies are required to further specify and explain the differential processes leading to antisocial behavior. This can be facilitated by distinguishing diverse types of antisocial behaviors (e.g. relational aggression, physical aggression, overt and covert antisocial behavior), by making a distinction between ethnic minority groups, and by investigating interaction between sex and ethnicity.

Assessment of self-serving cognitive distortions in the evaluation of prevention programs

Many intervention and prevention programs targeting children or adolescents, use a cognitive-behavioral approach to reduce antisocial behavior (e.g., Bennett & Gibbons, 2000; Wilson, Gottfredson, & Najaka, 2001; Wilson & Lipsey, 2007). Cognitive-behavioral approaches are

based on the assumption that biased thinking patterns need to be changed in order to positively influence behavior. Barriga, Gibbs, Potter, and Liao (2001) refer to these biased or dysfunctional thinking patterns, in which one inaccurately confers meaning upon experiences and daily events, by using the term self-serving cognitive distortions. Although this thesis could not provide evidence for an explanatory model using moral cognitions, self-serving cognitive distortions were found to be 1) strongly associated with antisocial behavior, 2) of influence on antisocial behavior in girls, and 3) highly prevalent in the targeted population (see Chapter 2). In addition, self-serving cognitive distortions form a key element of the EQUIP program, which is evaluated in the current thesis (see Chapter 4 and 5). Based on the importance of self-serving cognitive distortions in relation to antisocial behavior as described above, and in order to facilitate evaluation of the prevention program EQUIP, the psychometric quality and the suitability for evaluation of the How I Think Questionnaire (HIT-Q) – an instrument used to assess self-serving cognitive distortions – was investigated in a mixed gender, non-delinquent sample representing 1534 students with a low educational level and from diverse ethnic backgrounds (see Chapter 3).

The adequate reliability and validity of the HIT-Q as found by Barriga et al. (2001) and Nas et al. (2008) was confirmed in the current thesis and the cognitive distortions structure was supported. Furthermore, the findings concerning the anomalous response scale (AR) – designed to detect disingenuous, incompetent, or otherwise suspect responding – casted doubt on its relevance and function in the Dutch population. Because of these findings, and since anomalous response tendency and social desirability appeared to be separate constructs, social desirability was included as a covariate in the evaluation studies of the current thesis (instead of using the AR cutoff points and removing ‘suspect’ students from the samples). Furthermore, the Dutch samples scored substantially higher compared to the American samples on the HIT-Q scales, and delinquent groups scored substantially higher compared to the non-delinquent groups in the Netherlands. These findings indicate that, for the application of the HIT-Q as an instrument for evaluation, the original profile form and cut off points for classification need to be revised for specific populations. In sum, the HIT-Q is an adequate instrument for evaluation, although it should always be part of a broader set of instruments assessing different aspects of antisocial behavior when used as diagnostic tool.

Evaluation of a universal prevention program in secondary school: EQUIP for Educators

A substantial part of the prevention programs implemented in adolescent populations use the cognitive-behavioral approach. This approach assumes that biased or dysfunctional thinking patterns contribute to the development and persistence of antisocial behavior (Milkman & Wanberg, 2007), and has proven to be most successful in reducing antisocial behavior in student populations (Bennett & Gibbons, 2000; Wilson et al., 2001). The EQUIP program

(Gibbs, Potter, & Goldstein, 1995) is one of these programs which was been found to be promising (Leeman, Gibbs, & Fuller, 1993; Nas, Brugman, & Koops, 2005).

Chapter 4 provided a first evaluation of EQUIP for Educators (EFE), implemented as an universal school based intervention program for young adolescents (DiBiase, Gibbs, Potter, & Spring, 2005). The immediate effects of EFE on prevalence of antisocial behavior, attitude towards antisocial behavior, self-serving cognitive distortions, and moral judgment were examined using a quasi-experimental pre-test/ post-test with control group design. A total of 622 prevocational secondary school students was assessed at both pre and post test. Preceding the evaluation of EFE, it was investigated whether the application of this type of intervention was justified. Since higher rates of problem behavior may indicate a stronger need and create better opportunities for improvement (Flannery et al., 2003; Stoolmiller, Eddy, & Reid, 2000), it is important to map students' initial level of antisocial behavior prior to their participation in an intervention. Findings showing high rates of various forms of antisocial behavior were in line with previous research in the Netherlands (Van der Laan & Blom, 2006; Vollebergh et al., 2006). Results indicated that a substantial part of the present sample could indeed be labeled as 'at-risk' and might, given the strong relationship between self-serving cognitive distortions and antisocial behavior, benefit from a preventive intervention targeting self-serving cognitive distortions, limitations in the development of moral judgment, and deficiencies in social skills. With regard to the evaluation of EFE, findings showed that the experimental group reported a more negative attitude towards antisocial behavior and a lower level of self-serving cognitive distortions after the intervention compared to the control group, although effect sizes were small (Cohen, 1988). These findings were supported by evaluations, since 85 percent of the teachers certified the program as being successful in its attempt to positively influence thinking and acting of students, and 56 percent of the students reported to think more prosocially after the intervention. However, there were no differences in change of prevalence of antisocial behavior and moral judgment between the experimental group and the control group. Although sex and initial level of antisocial behavior appeared to have no influence on the effects of EFE, the intervention did have a greater impact on students with a Dutch ethnic background than on students from ethnic minority groups in changing the attitude towards antisocial behavior and reducing self-serving cognitive distortions.

The evaluation of the immediate effects of EFE (as reported in Chapter 4) was replicated and extended with follow up assessments (Chapter 5). The immediate and longer term (up till 9 months after the intervention) effects of EFE on both self reported and teacher reported prevalence of antisocial behavior, attitude towards antisocial behavior, self-serving cognitive distortions, moral judgment, and social skills were examined, using six assessments. A total

of 705 students participated in the intervention group following EFE, while 158 students participated in the non-intervention group, following the normal educational curriculum. This second evaluation study replicated previous findings (see Chapter 4) and found that EFE influenced both the way students evaluated antisocial behavior and their level of self-serving cognitive distortions. Although the influence of EFE on attitude towards antisocial behavior appeared to be only temporary, the effect of the intervention on self-serving cognitive distortions remained stable over follow up assessments up till nine months after intervention, indicating that EFE was successful in changing students' cognitions over an extended period of time.

A strong relationship between the direct effects of EFE on self-serving cognitive distortions and on attitude towards antisocial behavior was found. No effects on level and rate of change after intervention for self reported prevalence of antisocial behavior, social skills, and moral judgment could be established. However, delayed effects (four months after the intervention) on teacher reported antisocial behavior were established. Several possible explanations for the absence of effects on self reported prevalence of antisocial behavior, moral judgment, and social skills were discussed in Chapter 5; e.g. the presence of a substantial percentage of students with low base rates of antisocial behavior and minimal opportunities for improvement, lack of program intensity, and problems due to the first year of implementation. Although there were no differential effects based on initial level of antisocial behavior (in line with Chapter 4), sex and ethnicity did influence the effects of EFE. Immediate effects of EFE on self-serving cognitive distortions were stronger for students with a Dutch ethnic background compared to students from ethnic minority groups (in line with Chapter 4). A possible explanation might be that a peer approach such as EFE is less successful in minority groups, since for the native population emotional independence of parents and bonding with peers are among the most important developmental tasks (Stevens & Vollebergh, 2008). While immediate effects of the intervention on attitude towards antisocial behavior were established for both sexes, this effect gradually reduces over time for boys, indicating that girls are more susceptible for maintenance of attitudinal changes.

In sum, both evaluation studies (Chapter 4 & Chapter 5) replicated the immediate effects found by Nas and colleagues (2005). The longitudinal study (Chapter 5) extended these findings with longer term effects. Despite the differences in context (correctional facility versus school classes), format (alternative version of EQUIP), composition of the sample (only boys versus boys and girls in the present study), and sample size (small versus large sample in the present study), effects of the studies were identical, indicating that EQUIP is at least consistent and stable in its influence on cognitions and attitude. The program appeared to be at least partly successful in its attempt to equip students to think and act more prosocially,

since self reports show that EFE had an effect on attitude towards antisocial behavior and self-serving cognitive distortions, and teacher reports established an effect on antisocial behavior.

General conclusion

The primary aim of the current thesis was to find out whether EQUIP for Educators (EFE), implemented as a school based universal prevention program in prevocational secondary school students, is effective. A precondition for determining the effectiveness of intervention and prevention programs is the presence of a sound theoretical foundation (Boendermaker, 2008). EFE is based on an extensively investigated theory (cognitive-behavioral therapy, e.g. Milkman & Wanberg, 2007; Sukhodolsky, Kassinove, & Gorman, 2003), and combines diverse theoretically based and empirically evaluated components (Aggression Replacement Training: e.g. Goldstein & Glick, 1994; Hatcher et al., 2008; Moynahan & Stromgren, 2005; Positive Peer Culture: Vorrath & Brendtro, 1985). In order to qualify the effectiveness of EFE, the criteria and guidelines as provided by Van Yperen and Van Bommel (2009) were used.

First of all, the evidential value of the research design (see Table 1, Chapter 1), was found to be *quite strong* for the short term evaluation study (Chapter 4) and *strong* for the long term evaluation study (since follow ups up to nine months were included, see Chapter 5). These findings indicate that, based on evidential value, EFE can be qualified as promising or even evidence-based. An important aspect in determining evidential value is the reliability and validity of the instruments used to establish effects. The psychometric quality of the instruments used in the present thesis to assess antisocial behavior – which was the main outcome variable – was already confirmed in several previous studies (Achenbach et al., 2002; Høst, Brugman, Tavecchio, & Beem, 1998; Leenders & Brugman, 2005; Rescorla et al., 2007; Verhulst, Ende, & Koot, 1997). In addition, the psychometric quality and the suitability for evaluation purposes of an instrument assessing self-serving cognitive distortions – another important outcome variable – could also be established in this thesis (see Chapter 3).

A second criterion is the number of available studies providing evidence for effectiveness of the program (Van Yperen & Van Bommel, 2009). The current thesis presents a first evaluation of EFE implemented as a school based universal prevention program, meaning that the criterion of proven effectiveness of the program could not be met. However, based on the number and evidential value of previous studies who established effects of EQUIP in both incarcerated juveniles (Brugman & Bink, in press; Leeman et al., 1993; Nas et al., 2005) and adults (Devlin, 2008; Liau et al., 2004), the program could be qualified as promising.

The third criterion is the presence of a clear link between the effects and the aim and population of the study. The primary aim of EQUIP is to reduce antisocial behavior, which is assumed to be preceded by reductions in levels of self-serving cognitive distortions and improvements in moral judgment and social skills (Gibbs et al., 1995). The evaluation studies of the present thesis found reductions in self-serving cognitive distortions and a more negative attitude towards antisocial behavior. In addition, the longitudinal study (see Chapter 5) showed that the reductions in self-serving cognitive distortions are stable over a longer period of time (at least nine months after the intervention), and established effects on teacher reported antisocial behavior. A stabilization of antisocial behavior as reported by the teacher in the intervention group was found, while the control group showed an increase. These findings indicate that EFE is at least partly successful in its aims; effects were found for some of the assumed mediating variables (self-serving cognitive distortions and attitude towards antisocial behavior) and partly for the outcome variable (teacher reported antisocial behavior).

The fourth criterion concerns the effect size, which should be in accordance with type of intervention, and relevant in practice. The effect sizes of the both the mediating variables the outcome variable as reported in the evaluation studies of the present thesis were small (Cohen, 1988), but nonetheless in line with previous studies, reporting only small effects on (mostly) outcome variables for unselected general student samples (Limbos et al., 2007; Wilson, Lipsey, & Derzon, 2003). Furthermore, reviews of the literature summarizing efforts to intervene with antisocial adolescents invariably lead to negative conclusions. Antisocial behavior appears to be difficult to change (Simons-Morton, Haynie, Saylor, Davis Crump, & Chen, 2005), and most evaluations of prevention programs in general present small or at most modest effect sizes (Wilson & Lipsey, 2007). Since EFE in the present thesis did show positive effects, and both teacher and student evaluations were positive, the program appears to be relevant for practice.

Based on these conclusions, EFE can be qualified as promising. In order for EFE to acquire the predicate evidence-based, longitudinal studies establishing effects on antisocial behavior are clearly needed.

Implications for practice

The universal implementation of EFE in schools appears to be promising, although further evidence supporting the success of the program in altering the developmental pathway of antisocial behavior on the long term is still needed. Because treatment gains in prevention often become visible after several years, it would be worthwhile to investigate long term effects. Moreover, diverse sources reporting on antisocial behavior of the participants should be included in addition to self and teacher reports (e.g. police records, court mandates).

Furthermore, EFE is intended for both primary and secondary prevention (DiBiase et al., 2005). In the current thesis, EFE was implemented universally in order to: avoid selection procedures with high rates of false positives or false negatives (Offord, Kraemer, Kazdin, Jensen, & Harrington, 1998; Van Lier, Verhulst, & Crijnen, 2003); create a heterogeneous group allowing for positive peer culture to work; and avoid stigmatizing at risk youth (by targeting complete classes instead of only selected students). However, since to date there are no evaluations of EFE implemented as a selected program available, comparison to find out what works best, in terms of effects and cost-effectiveness (Durlak & Wells, 1997), is not possible.

Furthermore, attention should be paid to the differential effectiveness of EFE in various subpopulations, since the current thesis found differences for diverse ethnic groups and for males and females. Insight in specific factors contributing to the successes on the one hand and the lack of effects on the other hand is needed in order to find out what causes these differences. When needed, the program or the implementation of the program should be adapted, in order to extend its reach.

Strengths, limitations, and future directions

Since there is an increasing need for evidence-based programs aiming to reduce antisocial behavior (Boxer, Goldstein, Musher-Eizenman, Dubow, & Heretick, 2005) and a lack of thoroughly evaluated prevention programs, an important strength of the present thesis is the evaluation of EFE in a universal preventive setting. The evaluation studies reported in this thesis assessed immediate effects (Chapter 4 & 5) and longer term effects (Chapter 5). Both studies included a large, mixed-gender sample representing students with diverse ethnic backgrounds. Another strength is the fact that EFE in the present thesis is implemented by teachers, since the use of regular school employees to operate prevention programs has been found to improve implementation quality (Gottfredson & Gottfredson, 2002), and effectiveness (Wilson et al., 2003). Furthermore, implementation of EFE in the current thesis was valid and realistic in a regular school situation. Most empirical interventions are developed and implemented without regard to contextual constraints, such as teacher time, resources, skills, and support operating on teachers and classroom settings. They are often too complex, too costly, and too distinct by bearing little resemblance to regular educational practices (Cunningham & Henggeler, 2001). In the current thesis, schools were selected based on willingness and possibilities for training of the teachers and the implementation of the program in classroom settings. School personnel was extensively informed, training and educational material was financed by the research project, and the researcher visited and contacted schools on a very regular basis during the implementation of the intervention.

Furthermore, the present thesis focuses on the developmental pathways of behavior and moral cognitions. The use of a combination of an accelerated longitudinal design with elements of an interrupted time series design (Chapter 5) made it possible to gain insight into the trajectories of moral cognitions and behavior, and longitudinal effects of the intervention (cf. Flannery et al., 2003). This type of design is more sensitive to trends in individual trajectories and the way these are affected by the intervention than a pretest-posttest design, which is mostly used. Complex latent growth curve models for both the baseline trajectories and the effects of the intervention on these trajectories are combined to assess developmentally relevant effects. An advantage of the design used in this thesis was the inclusion of multiple cohorts, which made correction for possible biases due to age, number of assessments, and time of year possible. Since students in the intervention condition differed in number and time of pre and post assessments, they formed their own control group. This procedure assured a very strong and valid matching, which in turn accounted for correct attribution of effects to the intervention (cf. Shadish, Cook, & Campbell, 2002). In addition, since a non intervention condition was available (see Chapter 4), these students were also included in the longitudinal analyses. Inclusion of this non intervention cohort revealed the effect of change in time perspective in the antisocial behavior questionnaire after intervention (see Chapter 5), and made correction for this bias possible. The type of design as used in the present thesis reveals an important shortcoming in the quality requirements for research as proposed by Van Yperen and Van Bommel (2009). Developmental processes are not taken into account, since randomization in an experimental design is proposed as the sole criterion for strong evidential value (see Table 1, Chapter 1).

When interpreting the findings of the current thesis, there are also some limitations which should be noted. To begin with, there are several limitations concerning the implementation of EFE. Based on the idea of interrelationships between the different components of the program, it is recommended to have three sessions of EFE each week. In the present thesis, schools were only able to organize two EFE sessions a week, due to limited school hours available. Because two introduction lessons were added, EFE was taught in a time period of sixteen weeks instead of ten weeks as in the original EFE program. Gottfredson and Gottfredson (2002) argue that the effectiveness of prevention practices would be improved if schools increased the intensity of the activities they are conducting, and if they incorporated more of the content. On the contrary, Bennett and Gibbons (2000) state that frequency is not of influence on effectiveness. Although there is no consensus concerning the importance and influence of frequency in the literature, this deviation from program standards might have had an impact on the findings concerning the effectiveness of the EQUIP program. On the other hand, one could also argue that the impact of EQUIP increases, because by spreading the lessons over a longer time period more opportunities arise to link the EQUIP lessons to

incidents in everyday behavior. Consequently, more rehearsal is possible of what students had learned. Finally, EFE was only taught in the class. Principles might need to be used in the complete school setting to facilitate transfer and enlarge effectiveness, since better integration of preventive activities into normal school operations and settings is found to predict higher quality programming (Gottfredson & Gottfredson, 2002). Based on these findings, future research should aim at investigating implementation quality of EFE.

There are also some limitations concerning the instruments and samples used in these studies. First of all, the instrument assessing moral judgment was first used in this thesis and showed low stability and moderate reliability. Although most findings based on this instrument were in line with previous research, future research investigating the psychometric quality of the instrument is needed. Secondly, the present thesis mainly used self reports, and although social desirability was included in the analyses in order to correct for possible biases, future research should include more diverse sources (e.g. parental reports, police records). Lastly, when interpreting the findings it should be noted that generalizability is limited, since the sample used in this thesis was rather homogeneous, in terms of educational level and cognitive functioning. Participants were all eight grade students attending prevocational secondary education, mostly living in highly urbanized areas. In addition, the percentage of students from ethnic minority groups is substantially higher in the present study (50 percent and higher) than the percentages as reported by the Central Bureau of Statistics (20 to 25 percent; 2009), which limits findings to similar populations.

Finally, some additional recommendations for future research need to be made. To begin with, previous research indicated an extensive list of factors contributing to the development and persistence of antisocial behavior (e.g. Rutter, Moffitt & Caspi, 2006). Although EQUIP is a multicomponent treatment program, its reach is limited to only some of these dynamic factors, since the program focuses on developmental delays. In order to gain insight in the contribution of these specific delays to effectiveness, future research should specify the influence of the different components included in the program.

Furthermore, the EQUIP program for juvenile delinquents prescribes 3 mutual help meetings per week in addition to the EQUIP meetings. A potential shortcoming of the preventive version of the EQUIP program is the omission of these mutual help meetings (DiBiase et al., 2005). During a mutual help meeting a problem or thinking error of one of the group members is at issue; other group members learn to practice their helping skills by trying to help this person to solve his/her problem or correcting his/her thinking error. Mutual help meetings facilitate transfer of the skills learned at the EQUIP meetings to real-life situations. The omission of mutual help meetings in EFE possibly reduced this transfer and the impact of

the Positive Peer Culture (PPC), which probably influenced the effectiveness of the program. Mutual help sessions can be especially effective in a universal prevention program, since school populations comprise 'regular youth', whose prosocial thinking and acting can be exemplary for at risk youth. Future research should investigate possibilities to add mutual help sessions as included in the original intervention version of EQUIP (Gibbs et al., 1995) and study its effectiveness.

References

- Aber, J. L., Brown, J. L., & Jones, S. M. (2003). Developmental trajectories toward violence in middle childhood: course, demographic differences, and response to school-based intervention. *Developmental Psychology, 39*, 324-348.
- Achenbach, T. M. (1991). *Manual for the Teacher's Report Form and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., Dumenci, L., & Rescorla, L. A. (2002). Ten-year comparisons of problems and competencies for national samples of youth: Self, parent and teacher reports. *Journal of Emotional and Behavioral Disorders, 10*, 194-203.
- Allen, J. P., Porter, M. R., & McFarland, F. C. (2006). Leaders and followers in adolescent close friendships: Susceptibility to peer influence as a predictor of risky behavior, friendship instability, and depression. *Development and Psychopathology, 18*, 155-172.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- Bandura, A. (2001). Social-cognitive theory of moral thought and action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development: Vol. 3. Application* (pp. 45-103). Hillsdale, NJ: Erlbaum.
- Bandura, A. (2002). Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education, 31*, 101-119.
- Barriga, A. Q., Gibbs, J. C., Potter, G. B., & Liao, A. K. (2001). *How I Think (HIT) Questionnaire Manual*. Champaign, Ill.: Research Press.
- Barriga, A. Q., Hawkins, M. A., & Camelia, C. R. (2008). Specificity of cognitive distortions to antisocial behaviours. *Criminal Behaviour and Mental Health, 18*, 104-116.
- Barriga, A. Q., Morrison, E. M., Liao, A. K., & Gibbs, J. C. (2001). Moral cognition: explaining the gender difference in antisocial behavior. *Merrill-Palmer Quarterly, 47*, 532 - 562.
- Basinger, K. S., & Gibbs, J. C. (1987). Validation of the Sociomoral Reflection Objective Measure - Short Form. *Psychological Reports, 61*, 139-146.
- Beem, A. L., Brugman, D., Høst, K., & Tavecchio, L. W. C. (2004). Students' perception of school moral atmosphere: from moral culture to social competence. A generalizability study. *European Journal of Developmental Psychology, 1*, 171-192.
- Bennett, D. S., & Gibbons, T. A. (2000). Efficacy of child cognitive-behavioral interventions for antisocial behavior: A meta-analysis. *Child & Family Behavior Therapy, 22*, 1-15.

- Berndt, T. J. (1996). Transitions in friendship and friends' influence. In J. A. Graber, Brooks-Gunn, J., Petersen, A.C. (Ed.), *Transitions through adolescence: Interpersonal domains and context*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Biglan, A., Mrazek, P. J., Carnine, D., & Flay, B. R. (2003). The integration of research and practice in the prevention of youth problem behaviors. *American Psychologist*, *58*, 433-440.
- Boendermaker, L. (2008). *Wat werkt bij jeugdigen met gedragsproblemen? (What works in antisocial youth?)*. Utrecht: NJi.
- Boendermaker, L., & Van Yperen, T. (2003). *Kansen in de keten. Een gemeenschappelijk referentiekader voor de justitiële jeugdinrichtingen. (Chances in the chain. A shared frame of reference for juvenile correctional institutes)*. Utrecht: NIZW Jeugd.
- Boom, J., Brugman, D., & Van der Heijden, P. G. M. (2001). Hierarchical structure of moral stages assessed by a sorting task. *Child Development*, *72*, 535-548.
- Boxer, P., Goldstein, S. E., Musher-Eizenman, D., Dubow, E. F., & Heretick, D. (2005). Developmental issues in school-based aggression prevention from a social-cognitive perspective. *The Journal of Primary Prevention*, *26*, 383-400.
- Brugman, D., & Bink, M. D. (in press). Effects of the EQUIP peer intervention program on cognitive distortions and recidivism among delinquent adolescent. *Psychology, Crime, and Law*.
- Cho, H., Hallfors, D. D., & Sánchez, V. (2005). Evaluation of a high school peer group intervention for at-risk youth. *Journal of Abnormal Child Psychology*, *33*, 363-374.
- Cillissen, A. H. N., & Mayeux, L. (2004). From Censure to Reinforcement: Developmental Changes in the Association Between Aggression and Social Status. *Child Development*, *75*, 147-163.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2 ed.). Hillsdale, NJ: Erlbaum.
- Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon, Eisenberg, N. (Ed.), *Handbook of child psychology, 5th ed.: Vol 3. Social, emotional, and personality development*. Hoboken, NJ: John Wiley & Sons Inc.
- Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment, Vol. 1: Theoretical foundations and research validation*. New York, NY: Cambridge University Press.
- Colby, A., Kohlberg, L., Gibbs, J. C., & Lieberman, M. (1983). A longitudinal study of moral judgment. *Monographs of the Society for Research in Child Development*, *48*.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98-104.
- Côté, S. M., Vaillancourt, T., Barker, E. D., Nagin, D., & Tremblay, R. E. (2007). The joint development of physical and indirect aggression: Predictors of continuity and change during childhood. *Development and Psychopathology*, *19*, 37-55.

- Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms on reactive and proactive aggression. *Child Development, 67*, 993-1002.
- Cromwell, P., & Thurman, Q. (2003). The devil made me do it: use of neutralizations by shoplifters. *Deviant Behavior, 24*, 535-550.
- Crown, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting and Clinical Psychology, 24*, 349-354.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting and Clinical Psychology, 24*, 349-354.
- Cunningham, P. B., & Henggeler, S. W. (2001). Implementation of an empirically based drug and violence prevention and intervention program in public school settings. *Journal of Clinical Child Psychology, 30*, 221-232.
- De Mey, L., Baartman, H. E. M., & Schulze, H. J. (1999). Ethnic variation and the development of moral judgment of youth in Dutch society. *Youth & Society, 31*, 54-75.
- DeMatteo, D., & Marczyk, G. (2005). Risk factors, protective factors, and the prevention of antisocial behavior among juveniles. In K. Heilbrun, N. E. S. Goldstein & R. E. Redding (Eds.), *Juvenile delinquency. Prevention, assessment, and intervention*. New York, NY: Oxford University Press.
- Devlin, R. (2008). *Responsible Adult Culture (RAC): cognitive and behavioral changes at a community-based correctional facility*. Unpublished Dissertation, Ohio State University.
- DiBiase, A. M., Gibbs, J. C., Potter, G. B., & Spring, B. (2005). *EQUIP for Educators. Teaching youth (grades 5-8) to think and act responsibly*. Champaign, Ill.: Research Press.
- Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*, 755-764.
- Dodge, K. A., & Rabiner, D. L. (2004). Returning to Roots: On Social Information Processing and Moral Development. *Child Development, 75*, 1003-1008.
- Dodge, K. A., & Schwartz, D. (1997). Social information processing mechanisms in aggressive behavior. In D. M. Stoff., J. Breiling & J. D. Maser (Eds.), *Handbook of antisocial behavior*. New York, NY: Wiley & Sons.
- Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling. Concepts, issues, and applications*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Durlak, J. A., & DuPre, E. P. (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*, 327-350.

- Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs: The future is exciting. *American Journal of Community Psychology, 25*, 233-243.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Iver, D. M. (1993). Development during adolescence: the impact of stage–environment fit on young adolescents' experiences in schools and in families. *American Psychologist, 48*, 90-101.
- Farrington, D. P. (2005). Childhood origins of antisocial behavior. *Clinical Psychology and Psychotherapy, 12*, 177-190.
- Festinger, L., Carlsmith, J. M., & Bem, D. J. (2007). Issue 4: Does Cognitive Dissonance Explain Why Behavior Can Change Attitudes? In J. A. Nier (Ed.), *Taking sides: Clashing views in social psychology (2nd ed.)*. (pp. 410). New York, NY, US: McGraw-Hill.
- Flannery, D. J., Liao, A. K., Powell, K. E., Vesterdal, W., Vazsonyi, A. T., Guo, S., Atha, H., & Embry, D. (2003). Initial behavior outcomes for the PeaceBuilders universal school-based violence prevention program. *Developmental Psychology, 39*, 292-308.
- Gest, S. D., Farmer, T. W., Cairns, B. D., & Xie, H. (2003). Identifying children's peer social networks in school classrooms: links between peer reports and observed interactions. *Social Development, 12*, 513-529.
- Gibbs, J. C. (1991). Sociomoral developmental delay and cognitive distortion: implications for the treatment of antisocial youth. In W. M. Kurtines & J. L. Gerwitz (Eds.), *Handbook of moral behavior and development* (Vol. 3, pp. 95-110). Hillsdale, NJ: Erlbaum.
- Gibbs, J. C. (2003). *Moral development and reality. Beyond the Theories of Kohlberg and Hoffman*. Thousand Oaks, CA: Sage.
- Gibbs, J.C. (2010). *Moral development and reality. Beyond the Theories of Kohlberg and Hoffman(2nd edition)*. Boston, MA: Allyn & Bacon
- Gibbs, J. C., Potter, G. B., Barriga, A. Q., & Liao, A. K. (1996). Developing the helping skills and prosocial motivation of aggressive adolescents in peer group programs. *Aggression and Violent Behavior, 1*, 283-305.
- Gibbs, J. C., Potter, G. B., & Goldstein, A. P. (1995). *The EQUIP program: teaching youth to think and act responsibly through a peerhelping approach*. Champaign, Ill.: Research Press.
- Gignac, G. E. (2007). Multi-factor modeling in individual differences research: some recommendations and suggestions. *Personality and Individual Differences, 42*, 37-48.
- Goldstein, A. P., & Glick, B. (1994). Aggression Replacement Training: curriculum and evaluation. *Simulation & Gaming, 25*, 9-26.
- Goldstein, A. P., & Glick, B. (2001). Aggression replacement training: Application and evaluation management. In G. A. Bernfeld, Farrington, D.P., Leschied, A.W. (Ed.),

- Offender rehabilitation in practice: Implementing and evaluating effective programs.* New York, NY: John Wiley & Sons Ltd.
- Goldstein, A. P., Glick, B., & Gibbs, J. C. (1998). *Aggression Replacement Training: A comprehensive intervention for aggressive youth.* Champaign, Ill: Research Press.
- Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: results from a national survey. *Journal of Research in Crime and Delinquency, 39*, 3-35.
- Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime.* Stanford, CA: Stanford University Press.
- Harland, P., Van der Laan, P. H., Smeenk, W., & Weerman, F. M. (2005). *Wangedrag op school en delinquentie elders. Een verkenning onder leerlingen van het VMBO. (Misconduct at school and delinquency elsewhere. An examination among prevocational secondary education students).* Den Haag/ Leiden: Boom Judicial Publishers/ NSCR.
- Hatcher, R. M., Palmer, E. J., McGuire, J., Hounscome, J. C., Bilby, C. A. L., & Hollin, C. R. (2008). Aggression Replacement Training with adult male offenders within community settings: A reconviction analysis. *Journal of Forensic Psychiatry & Psychology, 19*, 517-532.
- Higgins, A., Power, C., & Kohlberg, L. (1994). The relationship of moral atmosphere to judgments of responsibility. In B. Puka (Ed.), *New research in moral development* (pp. 190-222). New York, NY: Garland Publishing.
- Høst, K., Brugman, D., Tavecchio, L., & Beem, L. (1998). Students' perception of the moral atmosphere in secondary school and the relationship between moral competence and moral atmosphere. *Journal of Moral Education, 27*, 47-70.
- <http://statline.cbs.nl/StatWeb/publication/>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.
- Institute of Medicine. (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research.* Washington, DC: National Academy Press.
- Junger-Tas, J., Cruijff, M. J. L. F., Van de Looij-Jansen, P. M., & Reelick, F. (2003). *Etnische minderheden en het belang van bindingen. Een onderzoek naar antisociaal gedrag onder jongeren. (Ethnic minorities and the value of alliances. A study on antisocial behavior in youth).* Houten/Antwerpen: Bohn Stafleu Van Loghum.
- Kazdin, A. E. (2005). *Parent management training: Treatment for oppositional, aggressive, and antisocial behavior in children and adolescents.* New York, NY: Oxford University Press.

- Kline, R. B. (1998). *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford Press.
- Kohlberg, L. (1984). *Essays in moral development: Vol. 2. The psychology of moral development. The nature and validity of moral stages*. San Francisco, CA: Harper & Row.
- Krueger, R. F. (1999). Personality traits in late adolescence predict mental disorders in early adulthood: A prospective-epidemiological study. *Journal of Personality, 67*, 39-65.
- Kuther, T. L., & Higgins - D'Alessandro, A. (2000). Bridging the gap between moral reasoning and adolescent engagement in risky behavior. *Journal of Adolescence, 23*, 409-422.
- Lacourse, E., Côté, S. M., Nagin, D. S., Vitaro, F., Brendgen, M., & Tremblay, R. E. (2002). A longitudinal-experimental approach to testing theories of antisocial behavior development. *Development and Psychopathology, 14*, 909-924.
- Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology, 1*, 451-476.
- Lardén, M., Melin, L., Holst, U., & Långström, N. (2006). Moral judgement, cognitive distortions and empathy in incarcerated delinquent and community control adolescents. *Psychology, Crime and Law, 12*, 453-462.
- Leeman, L. W., Gibbs, J. C., & Fuller, D. (1993). Evaluation of a multicomponent group treatment program for juvenile delinquents. *Aggressive Behavior, 19*, 281-292.
- Leenders, I., & Brugman, D. (2005). Moral/non-moral domain shift in young adolescents in relation to delinquent behavior. *British Journal of Developmental Psychology, 23*, 1-16.
- Liau, A. K., Barriga, A. Q., & Gibbs, J. C. (1998). Relations between self-serving cognitive distortions and overt vs. covert antisocial behavior in adolescents. *Aggressive Behavior, 24*, 335-346.
- Liau, A. K., Shively, R., Horn, M., Landau, J., Barriga, A. Q., & Gibbs, J. C. (2004). Effects of Psychoeducation for Offenders in a Community Correctional Facility. *Journal of Community Psychology, 32*, 543-558.
- Limbos, M. A., Chan, L. S., Warf, C., Schneir, A., Iverson, E., Shekelle, P., & Kipke, M. D. (2007). Effectiveness of interventions to prevent youth violence. A systematic review. *American Journal of Preventive Medicine, 33*, 65-74.
- Liu, J. (2004). Childhood externalizing behavior: theory and implications. *Journal of Child and Adolescent Psychiatric Nursing, 17*, 93-103.
- Loeber, R. (1991). Questions and advances in the study of developmental pathways. In D. T. Cicchetti, S.L. (Ed.), *Rochester Symposium on Developmental Psychopathology, Vol. 3: Models and integrations*. Rochester, NY: University of Rochester Press.

- Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modeling. *Personality and Individual Differences, 42*, 851-858.
- Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320-341.
- Milkman, H., & Wanberg, K. (2007). *Cognitive-behavioral treatment. A review and discussion for corrections professionals*. Washington, DC: National Institute of Corrections.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*, 674-701.
- Moffitt, T. E. (2003). Life-course-persistent and adolescence-limited antisocial behavior: A 10-year research review and a research agenda. In B. B. Lahey, T. E. Moffitt & A. Caspi (Eds.), *Causes of conduct disorder and juvenile delinquency* (pp. 49-75). New York, NY: Guilford Press.
- Moynahan, L., & Stromgren, B. (2005). Preliminary results of aggression replacement training for Norwegian youth with aggressive behaviour and with a different diagnosis. *Psychology, Crime & Law, 11*, 411-419.
- Muthén, L. K., & Muthén, B. O. (1998-2007). Mplus VERSION 5. Los Angeles, CA: Muthén & Muthén.
- Nagin, D., & Tremblay, R. E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and nonviolent juvenile delinquency. *Child Development, 70*, 1181-1196.
- Nas, C. N., Brugman, D., & Koops, W. (2005). Effects of the EQUIP programme on the moral judgement, cognitive distortions, and social skills of juvenile delinquents. *Psychology, Crime and Law, 11*, 421-434.
- Nas, C. N., Brugman, D., & Koops, W. (2008). Measuring self-serving cognitive distortions with the How I Think Questionnaire. *European Journal of Psychological Assessment, 24*, 181-189.
- Odgers, C. L., Moffitt, T. E., Broadbent, J. M., Dickson, N., Hancox, R. J., Harrington, H., Poulton, R., Sears, M. R., Thomson, W. M., & Caspi, A. (2008). Female and male antisocial trajectories: from childhood origins to adult outcomes. *Development and Psychopathology, 20*, 673-716.

- Odgers, C. L., Vincent, G. M., & Corrado, R. R. (2002). A preliminary conceptual framework for the prevention and management of multi-problem youth. In R. R. Corrado, Roesch, R., Hart, S.D., Gierowski, J.K. (Ed.), *Multi-problem violent youth: A foundation for comparative research on needs, interventions and outcomes*. Amsterdam, the Netherlands: IOS Press.
- Offord, D. R., Kraemer, H. C., Kazdin, A. E., Jensen, P. S., & Harrington, R. (1998). Lowering the burden of suffering from child psychiatric disorder: trade-offs among clinical, targeted, and universal interventions. *Journal of the American Academy Child & Adolescence Psychiatry*, 37, 686 - 694.
- Okazaki, S., & Sue, S. (1998). Methodological issues in assessment research with ethnic minorities. In P. Balls Organista, K. M. Chun & G. Marin (Eds.), *Readings in Ethnic Psychology* (pp. 26-40). New York, NY: Routledge.
- Overbeek, G., Vollebergh, W., Meeus, W., Engels, R., & Luijpers, E. (2001). Course, co-occurrence, and longitudinal associations of emotionan disturbance and delinquency from adolescence to young adulthood: A six-year three-wave study. *Journal of Youth and Adolescence*, 30, 401-426.
- Palmer, E. J. (2007). Moral cognition and aggression. In T. A. Gannon, T. Ward, A. R. Beech & D. Fisher (Eds.), *Aggressive offenders' cognition. Theory, research, and practice* (pp. 199-214). Chichester: Wiley.
- Patterson, G. R. (2005). The Next Generation of PMTO Models. *The Behavior Therapist*, 28, 27-33.
- Power, F. C., Higgins, A., & Kohlberg, L. (1989). *Lawrence Kohlberg's approach to moral education*. New York, NY: Columbia University Press.
- Redding, R. E., Goldstein, N. E. S., & Heilbrun, K. (2005). Juvenile delinquency: past and present. In K. Heilbrun, N. E. S. Goldstein & R. E. Redding (Eds.), *Juvenile delinquency. Prevention, assessment, and intervention*. New York, NY: Oxford University Press.
- Rescorla, L. A., Achenbach, T. M., Ginzburg, S., Ivanova, M., Dumenci, L., Almqvist, F., Bathiche, M., Bilenberg, N., Bird, H., Domuta, A., Erol, N., Fombonne, E., Fonseca, A., Frigerio, A., Kanbayashi, Y., Lambert, M. C., Liu, X., Leung, P., Minaei, A., Roussos, A., Simsek, Z., Weintraub, S., Weisz, J., Wolanczyk, T., Zubrick, S. R., Zukauskienė, R., & Verhulst, F. C. (2007). Consistency of teacher-reported problems for students in 21 countries. *School Psychology Review*, 36, 91-110.
- Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene-environment interplay and psychopathology: Multiple varieties but real effects. *Journal of Child Psychology and Psychiatry*, 47, 226-261.
- Ryan, A. M. (2001). The peer group as a context for the development of young adolescent motivation and achievement. *Child Development*, 72, 1135-1150.

- Schaeffer, C. M., Petras, H., Ialongo, N., Poduska, J., & Kellam, S. (2003). Modeling growth in boys' aggressive behavior across elementary school: links to later criminal involvement, conduct disorder, and antisocial personality disorder. *Developmental Psychology, 39*, 1020-1035.
- Schermelleh - Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*, 23-74.
- Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461-464.
- Seroczynski, A. D. (2006). Book review. EQUIP for Educators: Teaching youth (grades 5-8) to think and act responsibly. *Journal of Moral Education, 35*, 124-126.
- Shadish, W.R., Cook, T.D. & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.
- Simons-Morton, B., Haynie, D., Saylor, K., Davis Crump, A., & Chen, R. (2005). The effects of the Going Places program on early adolescent substance use and antisocial behavior. *Prevention Science, 6*, 187-197.
- Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis. Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Skara, S., & Sussman, S. (2003). A review of 25 long-term adolescent tobacco and other drug use prevention program evaluations. *Preventive Medicine, 37*, 451-474.
- Sodowsky, G. R., Kuo-Jackson, P. Y., Richardson, M. F., & Corey, A. T. (1998). Correlates of self reported multicultural competencies: counselor multicultural social desirability, race, social inadequacy, locus of control racial ideology, and multicultural training. *Journal of Counseling Psychology, 45*, 256-264.
- Solomon, D., Watson, M., Battistich, V., Schaps, E., & Delucchi, K. (1996). Creating classrooms that students experience as communities. *American Journal of Community Psychology, 24*, 719-748.
- Stams, G. J. M. M., Brugman, D., Dekovic, M., Rosmalen, L. v., Laan, P. v. d., & Gibbs, J. C. (2006). The moral judgment of juvenile delinquents: A meta-analysis. *Journal of Abnormal Child Psychology, 34*, 697-713.
- Stephens, R. C., Thibodeaux, L., Sloboda, Z., & Tonkin, P. (2007). Research note: an empirical study of adolescent student attrition. *Journal of Drug Issues, 37*, 475-488.
- Stevens, G. W. J. M., & Vollebergh, W. A. M. (2008). Mental health in migrant children. *Journal of Child Psychology and Psychiatry, 49*.
- Stoolmiller, M., Eddy, J. M., & Reid, J. B. (2000). Detecting and describing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior. *Journal of Consulting and Clinical Psychology, 68*, 296-306.

- Sukhodolsky, D. G., Kassinove, H., & Gorman, B. S. (2003). Cognitive-behavioral therapy for anger in children and adolescents: a meta-analysis. *Aggression and Violent Behavior, 9*, 247-269.
- Sykes, G. M., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review, 22*, 664-670.
- Tremblay, R. E. (2000). The development of aggressive behaviour during childhood: What have we learned in the past century? *International Journal of Behavioral Development, 24*, 129-141.
- Tremblay, R. E. (2006). Prevention of youth violence: why not start at the beginning? *Journal of Abnormal Child Psychology, 34*, 481-487.
- Van der Laan, A. M., & Blom, M. (2006). *Jeugddelinquentie: risico's en bescherming (Juvenile delinquency: risks and protection)*. The Hague: WODC.
- Van der Velden, F., Brugman, D., Boom, J., & Koops, W. (in press). Moral cognitive processes explaining antisocial behavior in young adolescents. *International Journal of Behavioral Development*.
- Van Lier, P. A. C., Verhulst, F. C., & Crijnen, A. A. M. (2003). Screening for disruptive behavior syndromes in children: the application of latent class analysis and implications for prevention programs. *Journal of Consulting and Clinical Psychology, 71*, 353 - 363.
- Van Lier, P. A. C., Vitaro, F., Wanner, B., Vuijk, P., & Crijnen, A. A. M. (2005). Gender differences in developmental links among antisocial behavior, friends' antisocial behavior, and peer rejection in childhood: results from two cultures. *Child Development, 76*, 841-855.
- Van Yperen, T., & Van Bommel, M. (2009). *Erkenning interventies: criteria 2008-2009 (Acknowledgement of interventions: criteria 2008-2009)*. Utrecht: NJi/ RIVM.
- Verhulst, F. C. (2008). The fourteen-year prediction of antisocial behavior. In J. J. Hudziak (Ed.), *Developmental psychopathology and wellness: Genetic and environmental influences*. Arlington, VA: American Psychiatric Publishing.
- Verhulst, F. C., Ende, v. d., J., & Koot, H. M. (1997). *Handleiding voor de Teacher's Report Form (Manual for the Teacher's Report Form)*. Rotterdam: Afdeling Kinder- en Jeugdpsychiatrie, Erasmus Ziekenhuis Rotterdam.
- Vitaro, F., Brendgen, M., & Tremblay, R. E. (2000). Influence of deviant friends on delinquency: Searching for moderator variables. *Journal of Abnormal Child Psychology, 28*, 313-325.
- Vollebergh, W. A. M., Van Dorselaer, S., Monshouwer, K., Verdurmen, J., Van der Ende, J., & Ter Bogt, T. (2006). Mental health problems in early adolescents in the Netherlands. Differences between school and household surveys. *Social Psychiatry and Psychiatric Epidemiology, 41*, 156-163.

- Vorrath, H. H., & Brendtro, L. K. (1985). *Positive peer culture* (2 ed.). Hawthorne, NY: Aldine.
- Widaman, K. F. (2006). III. Missing data: What to do with or without them. *Monographs of the Society for Research in Child Development, 71*, 42-64.
- Wilson, D. B., Gottfredson, D. C., & Najaka, S. S. (2001). School-based prevention of problem behaviors: a meta-analysis. *Journal of Quantitative Criminology, 17*, 247-272.
- Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior. Update of a meta-analysis. *American Journal of Preventive Medicine, 33*, 130-143.
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school based intervention programs on aggressive behavior: a meta-analysis. *Journal of Consulting and Clinical Psychology, 71*, 136-149.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Chicago, Ill: MESA press.

Summary

Aggression and delinquency are highly prevalent among adolescents. They form a major social concern, since adolescent problem behavior is associated with immediate and lasting problems throughout life. In response to these high rates of antisocial behavior, there has been a surge of research to investigate preventive strategies aiming to reduce these problem behaviors among adolescents. In this context, schools appear to be an important location, since they form a setting in which a lot of interpersonal aggression occurs and since they have almost universal access to adolescents. School-based prevention programs can be distinguished in *universal* interventions (focusing on entire populations), and *selective* interventions (targeting subgroups of at risk children). Although the literature on selective school-based prevention programs is rich, there is relatively little known about the effects of universal preventive interventions implemented within general populations of students. EQUIP for Educators (EFE) is one of those prevention programs, teaching youth (grades 5-8) to think and act responsibly. The program is based on cognitive-behavioral theory and includes three main components: anger management, social skills, and social decision making.

The aim of the current thesis was to provide a first evaluation of EFE implemented as a universal school-based prevention program in high school students. Additionally, an explanatory model of antisocial behavior was tested, and the psychometric quality and the suitability for evaluation of an instrument assessing self-serving cognitive distortions were investigated. The EFE program was implemented in classroom settings by regular teachers, for two school hours a week over a period of 16 weeks. Students completed questionnaires four times with time intervals of 4 months representing pre and post assessments. In addition, teachers completed antisocial behavior questionnaires twice.

Chapter 2 focused on the link between moral cognitions and moral behavior by presenting a longitudinal model explaining antisocial behavior including moral judgment, self-serving cognitive distortions, and perception of community. More specifically, the aim was to gain a better insight into the concurrent relationships between the moral cognitions and antisocial behavior, and to elaborate the understanding of the relationships over time. Follow-up data of 724 students were used to investigate to what extent moral cognitions and antisocial behavior were stable over time, and whether moral cognitions predicted antisocial behavior or vice versa. In contrast to the expectations, this study could only partly provide evidence for an explanatory model of antisocial behavior by the proposed moral cognitions. The relationship between self-serving cognitive distortions and antisocial behavior was only reciprocal for girls. For boys, antisocial behavior appeared to precede self-serving cognitive distortions. A possible explanation might be that self-serving cognitive distortions were

primarily used as neutralizers in an attempt to reduce feelings of guilt or shame following aggression or delinquency. Additionally, small but significant sex and ethnic differences were established, indicating the existence of differential developmental processes. These findings stress the possible relevance of cognitive-behavioral approaches in preventive strategies targeting antisocial behavior.

In *Chapter 3* – based on the importance of self-serving cognitive distortions in relation to antisocial behavior as described above, and in order to facilitate evaluation of the prevention program EQUIP – the psychometric quality and the suitability for evaluation of the How I Think Questionnaire (HIT-Q) was investigated in a sample representing 1534 students. The adequate reliability and validity of the HIT-Q, an instrument used to assess self-serving cognitive distortions, as found in previous studies was replicated in the current thesis and the cognitive distortions structure was supported. Furthermore, the findings concerning the anomalous response scale (AR) – designed to detect disingenuous, incompetent, or otherwise suspect responding – casted doubt on its relevance and function in the Dutch population. This latter finding led to the inclusion of social desirability in the evaluation studies of this thesis, instead of using the AR cutoff points and removing ‘suspect’ students from the samples. Furthermore, based on substantial differences in means scores between diverse samples, it was recommended to revise the original profile form and cut off points for classification in order to use the HIT-Q as an instrument for evaluation.

Chapter 4 examined the immediate effects of EQUIP for Educators on prevalence of antisocial behavior, attitude towards antisocial behavior, self-serving cognitive distortions, and moral judgment using a quasi-experimental pre-test/ post-test with control group design including 622 students. To begin with, this chapter provided an overview of the prevalence of different forms of self reported antisocial behavior in order to determine the suitability of the present sample for evaluation research. Findings were in line with previous research in the Netherlands, showing high rates of various forms of antisocial behavior. These results indicate that a substantial part of the present sample could indeed be labeled as ‘at-risk’ which justified implementation of EFE in this population. Furthermore, EFE appeared to be partly successful in its aim to equip students to think and act more prosocially. Students in the experimental group reported a more negative attitude towards antisocial behavior and a lower level of thinking errors after the EFE sessions compared to students in the control condition. In addition, effects of EFE were stronger for students with a Dutch ethnic background than for students from ethnic minority groups. These findings were supported by qualitative data, since 85 percent of the teachers evaluated the program as being successful in its attempt to positively influence students, and 56 percent of the students reported positive changes in thinking after the intervention.

Chapter 5 focused on both the initial and long term (up till 9 months after the intervention) effects of EFE on students’ self-serving cognitive distortions, moral judgment,

social skills, attitude towards antisocial behavior, and prevalence of antisocial behavior by using Latent Growth Curve models including four measurements in a sample of 863 students. In addition, effects on teacher reported behavior were investigated using repeated measures. This second evaluation study found that EFE temporarily influenced the way students evaluated antisocial behavior, and extendedly influenced their level of self-serving cognitive distortions. In addition, immediate effects of EFE on self-serving cognitive distortions were stronger for students with a Dutch ethnic background compared to students from ethnic minority groups. Although no effects on level and rate of change after intervention for self reported prevalence of antisocial behavior, social skills, and moral judgment were found, this evaluation did establish delayed effects (four months after the intervention) on teacher reported antisocial behavior.

The main conclusion that can be drawn from the studies as described in the current thesis is that a cognitive-behavioral based prevention program as EFE can be promising. EFE was partly successful, since the evaluation studies could only establish small effects of the intervention on attitude towards antisocial behavior, self-serving cognitive distortions, and teacher reported antisocial behavior. Future research on EFE and adaptation of the curriculum are needed in order to increase effectiveness.

Samenvatting

(Summary in Dutch)

De prevalentie van agressie en delinquentie onder adolescenten is hoog en zorgt voor grote maatschappelijke onrust en bezorgdheid. Probleemgedrag onder adolescenten wijst zowel op actuele problemen als op problemen in de verdere levensloop van ouders en mogelijk ook van slachtoffers. Naar aanleiding van de hoge mate van antisociaal gedrag is er in de afgelopen jaren veel onderzoek gedaan naar de effecten van preventieve programma's. Deze programma's zijn erop gericht om probleemgedragingen van adolescenten te voorkomen of te verminderen. Scholen vormen een belangrijke setting voor deze programma's, aangezien hier veel inter-persoonlijke agressie plaatsvindt en omdat op school grote groepen jongeren tegelijkertijd bereikbaar zijn. Er kan een onderscheid gemaakt worden tussen primaire preventie programma's, die zich richten op algemene populaties jongeren en secundaire preventie programma's, die zich richten op subgroepen bestaande uit jongeren bij wie sprake is van beginnend probleemgedrag. EQUIP voor het Onderwijs is een preventie programma dat zowel primair als secundair ingezet kan worden en zich richt op het stimuleren van verantwoordelijk denken en handelen van leerlingen (vanaf groep zeven van de basisschool tot en met de tweede klas van het voortgezet onderwijs). Het programma is gebaseerd op de cognitieve gedragstheorie en bevat drie hoofd componenten: reductie van zelfbeschermende denkfouten, bevordering van sociale vaardigheden en ontwikkeling van moreel oordelen.

Het doel van de huidige dissertatie was het geven van een eerste evaluatie van EQUIP voor het Onderwijs, geïmplementeerd als een primair preventie programma voor leerlingen van het voortgezet onderwijs. In dit kader werd een verklarend model voor antisociaal gedrag getoetst en werd de psychometrische kwaliteit en de geschiktheid voor evaluatie doeleinden van een instrument onderzocht om denkfouten te meten. In totaal hebben 11 VMBO scholen deelgenomen aan dit onderzoeksproject in een interventie of non-interventie conditie. EQUIP voor het Onderwijs werd geïmplementeerd in de klas en werd gegeven door de reguliere leerkrachten. Er vonden twee bijeenkomsten per week plaats tijdens schooluren, gedurende een periode van 16 weken. Leerlingen vulden vier keer vragenlijsten in (voor- en nametingen), met telkens een tijdsinterval van vier maanden. Daarnaast vulden leerkrachten twee keer vragenlijsten in over antisociaal gedrag.

In *Hoofdstuk 2* werd de relatie tussen morele cognities (moreel oordelen, denkfouten en perceptie van de school als gemeenschap) en antisociaal gedrag onderzocht in een longitudinale studie. Het doel van deze studie was meer inzicht te krijgen in de relaties tussen morele cognities en antisociaal gedrag en het uitbreiden van de kennis van de relaties op de langere termijn. Om na te gaan in welke mate morele cognities en antisociaal gedrag stabiel waren over tijd en om te onderzoeken of morele cognities daadwerkelijk voorafgaan aan

antisociaal gedrag of vice versa, werd data van 724 leerlingen over twee metingen met een tijdsinterval van vier maanden gebruikt. In tegenstelling tot de verwachting kon de huidige studie slechts gedeeltelijk bewijs leveren voor de verklaring van antisociaal gedrag door de genoemde morele cognities. Alleen bij meisjes bleek de relatie tussen denkfouten en antisociaal gedrag wederkerig te zijn. Bij jongens volgden denkfouten op antisociaal gedrag. Een verklaring zou zijn dat denkfouten een neutraliserende rol spelen en beogen om gevoelens van schaamte en schuld te reduceren. Daarnaast werden er kleine, maar significante verschillen gevonden op basis van sekse en etniciteit, wat duidt op het bestaan van differentiële ontwikkelingspaden.

In *Hoofdstuk 3* werd de psychometrische kwaliteit en de geschiktheid van de Hoe Ik Denk Vragenlijst (HID) onderzocht in een steekproef van 1534 leerlingen. De betrouwbaarheid en validiteit van het instrument was adequaat en kwam overeen met resultaten van voorafgaande studies. Wel riepen de bevindingen met betrekking tot de AR schaal, welke bedoeld is om verdachte antwoorden te filteren, twijfel op over de relevantie en de functie van deze schaal voor de Nederlandse populatie. Deze bevindingen leidden ertoe dat leerlingen met ‘verdachte’ scores niet uit de dataset werden verwijderd. In plaats daarvan werd gekozen voor het corrigeren voor sociale wenselijkheid in de evaluatie studies van deze dissertatie. Tot slot leidde substantiële verschillen in gemiddelde scores tussen diverse steekproeven tot de aanbeveling het originele profiel formulier en de daarbijhorende afbreeknormen voor classificatie in normale en klinische groepen nader te bepalen. Dit laatste is nodig om de HID als instrument voor diagnostische doeleinden te kunnen gebruiken.

Hoofdstuk 4 onderzocht de directe effecten van EQUIP voor het Onderwijs op prevalentie van antisociaal gedrag, houding ten opzichte van antisociaal gedrag, denkfouten en moreel oordelen in een quasi-experimentele studie met controle groep ($N = 622$). Dit hoofdstuk geeft een overzicht van de mate van zelf gerapporteerde antisociale gedragingen om te bepalen of de huidige doelgroep geschikt is voor een preventieve interventie. Bevindingen waren in lijn met eerder onderzoek in Nederland, dat liet zien dat verschillende vormen van antisociaal gedrag veel voorkomen bij adolescenten. Op grond van zowel leerling- als leerkrachtscores kan een substantieel gedeelte van de huidige steekproef als ‘risicovol’ worden beschouwd, hetgeen het aanwenden van EQUIP voor het Onderwijs bij deze groep legitimeert. Met betrekking tot de evaluatie van EQUIP voor het Onderwijs kan geconcludeerd worden dat het programma in ieder geval gedeeltelijk succesvol blijkt te zijn in zijn doelstelling leerlingen prosocialer te laten denken en handelen. Leerlingen in de interventie conditie rapporteerden een negatievere houding ten opzichte van antisociaal gedrag en een lagere mate van denkfouten na de EQUIP sessies vergeleken met de leerlingen in de non-interventie conditie. Daarnaast bleken de effecten van EQUIP voor het Onderwijs sterker te zijn voor autochtone leerlingen dan voor allochtone leerlingen. De positieve effecten van deze studie werden ondersteund door evaluaties, waarin 85 procent van de

leerkrachten aangaf dat het programma succesvol was in de poging leerlingen positief te beïnvloeden. In aanvulling hierop gaf 56 procent van de leerlingen aan positief beïnvloed te zijn in hun denken na de interventie.

Hoofdstuk 5 onderzocht zowel de directe als de langere termijn effecten (tot negen maanden na de interventie) van EQUIP voor het Onderwijs op denkfouten, moreel oordelen, sociale vaardigheden, houding ten opzichte van antisociaal gedrag en prevalentie van antisociaal gedrag bij leerlingen. Hiervoor werden latente groei curven met vier metingen gebruikt in een groep van 863 leerlingen. Daarnaast werden effecten van het programma op antisociaal gedrag, zoals door de leerkracht gerapporteerd, onderzocht met behulp van herhaalde metingen analyse. In deze tweede evaluatie studie binnen deze dissertatie werd gevonden dat EQUIP voor het Onderwijs tijdelijk van invloed was op de houding van leerlingen ten opzichte van antisociaal gedrag, en direct en op de langere termijn van invloed op denkfouten. Daarbij waren de directe effecten van het programma op denkfouten sterker voor autochtone leerlingen dan voor allochtone leerlingen. Hoewel er geen effecten van de interventie op zelfgerapporteerd antisociaal gedrag, sociale vaardigheden en moreel oordelen werden gevonden, liet deze evaluatie wel vertraagde positieve effecten zien (vier maanden na de interventie) op leerkracht gerapporteerd antisociaal gedrag.

De belangrijkste conclusie op basis van de studies in deze dissertatie is dat een cognitief gedragsprogramma als EQUIP voor het Onderwijs veelbelovend kan zijn op het gebied van preventie van antisociaal gedrag, ondanks de bevinding dat de relatie tussen cognitie en gedrag slechts gedeeltelijk wederkerig was. EQUIP voor het Onderwijs was ten dele succesvol, aangezien het programma een (klein) positief effect had op houding ten opzichte van antisociaal gedrag, denkfouten en prevalentie van leerkracht gerapporteerd antisociaal gedrag. Nader onderzoek en eventueel aanpassingen van het programma zijn nodig om de effectiviteit van EQUIP voor het Onderwijs te vergroten.

Dankwoord

(Acknowledgements)

En dan het dankwoord. Mijn proefschrift is af en ik wil graag de mensen bedanken die hieraan bijgedragen hebben. Om te beginnen natuurlijk dank aan de scholen die mee hebben gewerkt aan dit onderzoek. Alle leerlingen bedankt voor het invullen van de vragenlijsten (“Oh, nee niet weer die juf van de universiteit met die lange vragenlijst...”). Dank aan alle docenten die de training hebben gevolgd en EQUIP voor het Onderwijs vol enthousiasme hebben gegeven aan hun leerlingen. Dank ook aan alle coördinatoren en directeuren die het mogelijk hebben gemaakt zowel docenten als leerlingen te ‘EQUIP-en’ en die ervoor gezorgd hebben dat we het onderzoek eromheen zo goed mogelijk konden uitvoeren.

Verder wil ik mijn promotoren en co-promotor bedanken voor hun begeleiding tijdens mijn aio-schap. Daan, enorm bedankt voor je steun, je kennis en bovenal je vertrouwen in mij. Wat heb ik veel van je geleerd in de afgelopen vier jaar. Dank voor alle gesprekken over werk en andere zaken, dank voor je altijd snelle reacties op mijn stukken op welke dag van de week en welk tijdstip dan ook, dank dat ik altijd voor alles bij je binnen kon komen vallen (iets waar ik dan ook ruimschoots gebruik van heb gemaakt). Ik had me geen betere promotor en dagelijks begeleider kunnen wensen! Jan, bedankt voor je betrokkenheid bij het project, de wekelijkse besprekingen en je feedback op de stukken. Maar bovenal bedankt dat je als helpdesk fungeerde voor statistische problemen en me wegwijs hebt gemaakt in de wereld van structural equation models en latent growth curve modeling. Willem, bedankt voor je kritische blik, je hielp me de grote lijn te bewaken. De bespreking die me het meest is bijgebleven was tijdens de AME 2007 in New York, met uitzicht op the Empire State Building.

Dank aan de leden van de leescommissie voor hun bereidheid om dit proefschrift te beoordelen en voor hun aanwezigheid tijdens de verdediging ervan. I especially wish to sincerely thank Prof. Dr. John Gibbs for taking part in the reading committee, for attending the defense of my thesis, and for all the helpful comments concerning EQUIP in the past four years.

Daarnaast bijzonder veel dank aan Jan van Westerlaak, EQUIP trainer van dit project. Jan, zonder jou geen geslaagde uitvoering van het programma! Dank voor het trainen van alle docenten, voor je deelname aan ouderavonden, voor alle praktische hulp, voor je interesse in het onderzoek en vooral voor je enorme betrokkenheid.

Dank ook aan JJI Teylingereind en de Stichting EQUIP. Het was heel leerzaam om deel uit te mogen maken van het EQUIP Netwerk en resultaten en ervaringen te delen met alle instellingen in Nederland en België die werken met EQUIP of EQUIP voor het Onderwijs.

Verder dank aan Suzanne van Kessel voor de vertaling van het originele EQUIP for Educators lesmateriaal en daarnaast voor de vertaling van hoofdstuk 3.

Dank ook aan de studenten die mee hebben geholpen met de dataverzameling op de scholen en in het bijzonder student-assistenten Heleen van Huijkelom & Marije de Hoogd en stagiaire Mariëlle Breg. Jullie drie hebben een enorme berg werk van me overgenomen, van observaties in de klassen tot controle van de datasets, mijn dank is groot!

En natuurlijk wil ik mijn collega's van Ontwikkelingspsychologie bedanken voor de gezellige vier jaar in Utrecht. Met name dank aan de Jonkies en de post docs voor de praatjes over allerlei werkgerelateerde en privé zaken en jullie betrokkenheid bij alles wat mij bezighield. In het bijzonder wil ik mijn oud kamergenootjes bedanken. Chantal en Cathelijne, bedankt voor de gezellige tijd in H225. We hebben heel wat af gekletst en veel mooie en minder mooie dingen gedeeld. Dank voor jullie betrokkenheid en jullie steun. Chantal, super dat ik jouw paranimf mocht zijn en dat je nu deze rol in mijn promotie op je wilt nemen. Je was niet alleen een hele lieve kamergenoot, maar bent echt een goede vriendin geworden. Ik vind het heel bijzonder dat we al heel wat mooie life events hebben kunnen delen en hoop dat dit in de toekomst zo zal blijven. Verder dank voor de de gezellige tijd aan collega's van 'beneden', Muriel en Veroni. Muriel, heel bijzonder dat we vriendinnen waren op de middelbare school en deze vriendschap na jaren weer op hebben kunnen pakken toen we collega's werden in Utrecht. Ik vind het heel leuk dat je mijn paranimf wilt zijn.

En dan een bijzonder plek in dit dankwoord voor mijn vader, die mijn aio-periode helaas niet meer mee heeft mogen maken. Papa, jij gaf me mee er altijd uit te halen wat er in zit en ik weet hoe leuk je het had gevonden dat ik dit proefschrift heb geschreven. Dank je wel voor je immense vertrouwen in mij en voor de kracht die je me daarmee gegeven hebt. Ik draag dit proefschrift op aan jou. Mama, bedankt voor je steun en hulp tijdens het schrijven van dit proefschrift. Dank voor het luisteren, voor je geduld en voor de vele dagen oppassen op Frederique zodat ik toch nog even door kon werken.

Lieve Frederique, jouw komst heeft het laatste jaar van mijn proefschrift wel heel bijzonder gemaakt. Door jou kon ik beter relativeren en extra genieten van deze tijd. Zo blijkt dat wetenschap en moederschap toch goed samen kunnen gaan ;-)

Liefste Gideon, jij was mijn allergrootste steun in de afgelopen 4 jaar (en natuurlijk de 7 jaar daarvoor ;-)). Dank je wel voor het luisteren naar mijn verhalen (van SPSS tot moeilijke leerlingen tijdens de dataverzameling), voor je hulp bij mijn eigen 'morele dilemma's' en voor je enorme geduld en je vertrouwen in mij. Wat ben ik blij met jou!

Curriculum Vitae

Floor van der Velden was born in Eindhoven on November 8th, 1978. After completing secondary education (Atheneum, Eckart College Eindhoven) in 1997, she studied Law at Tilburg University (1998). After one year she continued her studies in Human Resource Management and Labour Relations and earned her bachelor degree in 2002. Subsequently, she graduated cum laude from Utrecht University with a master's degree in Pedagogical Sciences in 2005. After graduation she started her PhD project at the Department of Developmental Psychology of Utrecht University, and worked on her dissertation about the effects of EQUIP for Educators on moral cognitions and antisocial behavior in prevocational secondary school students.