Make a Move and Long Live Love+

Citation for published version (APA):

van Lieshout, S. (2017). *Make a Move and Long Live Love+: implementation and evaluation of two Dutch sex education programs*. [Doctoral Thesis, Maastricht University]. Maastricht University.
<https://doi.org/10.26481/dis.20170706svl>

Document status and date:

Published: 01/01/2017

DOI:

[10.26481/dis.20170706svl](https://doi.org/10.26481/dis.20170706svl)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Make a Move and Long Live Love+

implementation and evaluation
of two Dutch sex education programs

Sanne van Lieshout

Make a Move and Long Live Love+

implementation and evaluation
of two Dutch sex education programs

‘Uhhh, I hate it when adults use the term “sexually active”. What does it even mean?
Am I gonna like deactivate some day or is it a permanent state of being?’
Juno MacGuff in *Juno* (2007)

The research presented in this dissertation was conducted at the department of Work and Social Psychology, Maastricht University.

Cover Design and Lay-out: vormvast-ccid.nl

Printing: printsupport4u.nl

ISBN: 9789492597038

©Sanne van Lieshout, Hitchin 2017

This research was funded by grants from ZonMw. Lang Leve de Liefde – Bovenbouw: 124290002. Make a Move (Boy’s Contact): 124330006

Make a Move and Long Live Love+

implementation and evaluation
of two Dutch sex education programs

PROEFSCHRIFT

Ter verkrijging van de graad van doctor aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. dr. Rianne M. Letschert,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen op
donderdag 6 juli 2017 om 14:00 uur

door
Sanne van Lieshout

Promotores

Prof. dr. G. Kok

Prof. dr. R.A.C. Ruiter

Copromotor

Dr. F.E.F. Mevissen

Beoordelingscomissie

Prof. dr. F.R.H. Zijlstra (voorzitter)

Prof. em. dr. H.W. van den Borne

Dr. C.M. Markham (University of Texas)

Dr. R.M. Meertens

Prof. dr. N. Sevdalis (King's College London)

Table of contents

Chapter 1	General introduction	7
Chapter 2	Make a Move: A comprehensive effect evaluation of a sexual harassment prevention program in Dutch residential youth care	23
Chapter 3	Make a Move: Process evaluation of the implementation of a sexual harassment prevention program in Dutch residential youth care	47
Chapter 4	Long Live Love+: Effect evaluation of a Dutch online school-based sex education program for secondary education.	69
Chapter 5	Long Live Love+: Process evaluation of an online school-based sex education program in the Netherlands	91
Chapter 6	General discussion	113
	References	131
	Summary	149
	Samenvatting	155
	Valorization	163
	Acknowledgements	171
	Curriculum Vitae	177

Chapter 1

General introduction

Health behavior change programs must be carefully evaluated. Evaluation is the assessment of the value of events, processes, or, in this case, intervention programs. The purpose of program evaluation is to assess whether program goals have been met, and whether the program has indeed brought about the desired change (Rossi, Freeman, & Lipsey, 1999). These effect evaluation outcomes give feedback to program developers and can help decision-makers to select effective programs that suit the target population. A thorough program evaluation should include, however, not only an effect evaluation but also a process evaluation; i.e. an evaluation describing program implementation and operationalization (Bartholomew Eldredge, Markham, Ruiter, Fernández, Kok, & Parcel, 2016; p552; Moore et al., 2014; Shepherd et al., 2014). A process evaluation looks into program implementation and stakeholders' experiences with the program. Process evaluations can also assess how theoretical methods have been operationalized in the program applications. The outcomes of process evaluations guide interpretation and explanation of the outcomes of an effect evaluation, as program effectiveness is affected by the way in which interventions are implemented (Biggs, Vernberg, Twemlow, Fonagy, & Dill, 2008; Durlak & DuPre, 2008; James, Reddy, Ruiter, McCauley, & van den Borne, 2006; Kalafat, Illback, & Sanders, 2007; Kärnä et al., 2011). Furthermore, process evaluations give feedback to program developers about how the program was received (for instance by teachers and students), which determinants influenced program implementation, and how to potentially enhance the program, improve future implementation, and subsequently increase program effects.

Historically, effect evaluations have received much more attention in scientific literature and research funding, and it is only since the last two decades that assessments of program adoption and implementation by so-called process evaluation studies are gaining increasing attention (Coyle et al., 1996; Dane & Schneider, 1998; Durlak, 1998; Moore et al., 2014). It is, however, the combination of effect evaluation and process evaluation that give the best possible insight into intervention effectiveness (de Visser et al., 2015; Moore et al., 2014).

This dissertation presents the effect and process evaluations of two Dutch sex education programs: one for secondary schools (Long Live Love+) and one for residential youth care (Make a Move). The effect evaluation of both programs was carried out through a Randomized Controlled Trial (RCT), with an intervention and a control group, a baseline measure and a post-intervention measure. The effect evaluation of Make a Move also included a six-month follow up. The process evaluation of Long Live Love was performed through interviews with teachers and focus group discussions with students. The teacher interviews focused both on implementation, as well as on teachers' experiences, including perceived barriers and

facilitators. The process evaluation of Make a Move used trainers' diaries to measure levels of implementation, and included interviews with the trainers for a deeper understanding of their implementation context and the barriers they came across.

This general introduction to the dissertation will first elaborate on program adoption, implementation, and evaluation of school-based health promotion programs. Following this, adolescent sexual health will be set out, and the Dutch school and youth care setting will be explained. From there on, we will introduce the two intervention programs Make a Move and Long Live Love+, of which the evaluations are presented in this dissertation.

Program adoption, implementation, and evaluation

Adoption and implementation research can be performed in the form of determinants' studies, carried out when planning for adoption and implementation of a newly developed intervention. Alternatively, adoption and implementation research can be part of a process evaluation. When programs are not found to be effective, it is often suggested that either the program does not work, or that it has not been implemented with sufficient levels of completeness and fidelity (Biggs et al., 2008; Durlak & DuPre, 2008; Kalafat et al., 2007; Kärnä et al., 2011). Therefore, understanding the factors influencing implementation can help to improve the delivery of health programs and in turn enhance program effectiveness.

Program adoption refers to the decision to work with a certain intervention, and is something that program developers need to plan for (Bartholomew Eldredge et al., 2016). A model frequently used to illustrate the adoption process of interventions is Rogers' Diffusion of Innovation Theory (Rogers, 2003). This theory does not only distinguish categories of program adopters (innovators, early adopters, early majority, late majority and laggards), but also takes into account program characteristics that determine its uptake, such as complexity and compatibility of the intervention. Adoption is not only determined by program and user characteristics, as the environment or context also plays a role (Rogers, 2003). In the school context, for instance, opinions of parents or colleagues, or organizational support and capacity may also affect teachers' decision-making. Research on teachers' adoption decisions regarding sex education showed that adoption was strongly related to outcome beliefs (the student learning outcomes they expected), subjective norms (opinions of pupils, parents and colleagues), and self-efficacy (confidence to create an open atmosphere in the class, to lead role plays, to discuss sexuality). Furthermore, decisions are related to contextual factors such as school policy and teachers' dispositions, i.e. sexual morality and sense of responsibility (Paulussen, Kok, & Schaalma, 1994).

Whereas adoption is the decision to take up a program, implementation refers to the program's actual use. It can be defined as the way in which an intervention is put into practice and is delivered to the target population (Durlak, 2015). Two frequently used indicators of implementation success are completeness and fidelity. Completeness refers to the quantity of the program delivered; to what extent a program has been implemented. Fidelity refers to whether program implementation took place as it was meant by program developers. It refers to whether a program was implemented according to the program guidelines.

Besides completeness (referred to as dosage delivered) and fidelity (referred to as program integrity), Dane and Schneider (1998) have also described participant responsiveness (is the audience attentive and motivated?), and program differentiation (is the program unique, to what extent does it differ from already available programs?). Participant responsiveness is also referred to as dosage received (Bartholomew Eldredge et al., 2016; Linnan & Steckler, 2002). When participants miss out on certain program components because they are absent or because they are not actively engaging, they may not recall these elements. In these cases, dosage received may be lower than the dosage delivered (completeness). Durlak and DuPre (2008) have added more aspects of implementation, namely monitoring of control/comparison conditions, program reach or scope, and adaptation. Monitoring of control or comparison conditions can particularly be important in RCTs in order to be able to know what the intervention is compared with, and to explain possible change over time in the control condition. A control group may or may not have been exposed to a similar intervention during program implementation in the intervention group. Program reach examines what percentage of a defined target population is reached with the intervention as well as the characteristics of participants, as they might not be a representative sample of the population a program is meant for (Bartholomew Eldredge et al., 2016; Durlak & DuPre, 2008; Linnan & Steckler, 2002). The last addition, adaptation, is also referred to as program reinvention. It measures to what extent implementers have made adjustments to the program. Program adaptation or re-invention of programs is mostly done with the intention to tailor the program components to the (perceived) specific needs of the target population. It is also done to create more ownership among program users (Fleuren, Wiefferink, & Paulussen, 2002; van Yperen & Bakker, 2008). According to Durlak and Dupre (2008), 100% fidelity levels are never reached, which allows room for adaptation. This makes that fidelity and adaptation frequently co-occur, meaning that whereas a teacher may implement certain components strictly according to the guidelines, other elements might be adapted.

All these elements play a role in program implementation. Therefore measurement of any of these elements can give insights into program implementation and thus serve as a means to measure how successful behavior change programs have been implemented. In the process evaluation of Make a Move and Long Live Love+ the main focus was on completeness, fidelity, and adaptation. This is in line with the most common implementation measures in literature, as can be seen in the review of Durlak and DuPre (2008). For Make a Move, we aimed to measure dosage received through attendance lists. In Long Live Love+ this was done through focus group discussions with students, but a thorough examination of dosage received was beyond the scope of this research. Monitoring of control condition was done for Long Live Love+, as the control group for LLL+ was a so-called active control group, received their standard sex education program. A meta-analysis of randomized controlled trials in HIV care for example, showed that patients in control groups also received effective care (De Bruijn et al, 2010). This is relevant information when assessing effectiveness of interventions. However, evaluations of neither Make a Move nor Long Live Love+ aimed to compare the program with specific alternative programs.

The indicators of implementation as described above can all be part of a process evaluation, which may guide interpretation of the outcomes of effect evaluations. It is the effect evaluation which assesses whether the intervention indeed brought about the desired change. As health behavior change programs are developed to enhance certain health outcomes among an identified target population, it is of vital importance to establish whether or not interventions were effective indeed. Therefore, program developers also need to consider an evaluation plan. The first requisite for performing a thorough effect evaluation is a acquiring a good understanding of the intervention program and how it aims to reach its objectives (Bartholomew Eldredge et al., 2016; Wholey, Hatry, & Newcomer, 2010). This can for example be schematized in a logic model. Subsequently, effect and process evaluation questions need to be formulated, after which indicators and their measures can be developed. These indicators and measures can for instance be determinants, behavioral outcomes, health outcomes, or changes in the environment. The evaluation design can be specified based on the selected measures and should facilitate the evaluation team to answer the following two questions: how do indicators of program effects compare with what happened before and after program implementation? And can any observed changes be attributed to the intervention evaluated? Although randomized controlled trials (RCTs) are a widely used type of effect evaluation, there are several alternative design for program effect evaluations (Bartholomew Eldredge et al., 2016; p569; Rossi et al., 1999; p309).

Program implementation as key for program effectiveness

Research has shown that correct implementation of programs is a key element in bringing about behavioral change (Biggs et al., 2008; Durlak & DuPre, 2008; James et al., 2006; Kalafat et al., 2007; Kärnä et al., 2011). Several studies in school settings have suggested that completeness and fidelity of program implementation is not self-evident (Ahtola, Haataja, Kärnä, Poskiparta, & Salmivalli, 2013; Bessems, Assema, Martens, & Paulussen, 2011; Wind et al., 2008). A Dutch program on healthy eating, for example, showed average implementation completeness levels of 70% per lesson (Bessems et al, 2011). Level of implementation correlated positively with teachers' appreciation of the program, and was further explained by technical/organizational barriers and time constraints. Similar levels of implementation were found in a Finnish anti-bullying programme in which 80% of the listed topics were implemented and teachers spent on average 11 hours on the program as compared to the 15 hours suggested. Only two-third (67%) of the teachers had implemented all ten lessons of the program (Ahtola et al., 2013). Implementation fidelity was positively related to teacher's experience of program support from the head teacher. Alternatively, their experience of general support did not enhance implementation and was even negatively related to implementation adherence (Ahtola et al., 2013). In another study, Wind and colleagues (2008) compared implementation of a multi-component fruit and vegetable intake school program among three European countries. Lesson implementation by teachers varied both within and among countries, and ranged from two to all 16 lessons. A review by Dusenbury and colleagues (Dusenbury, Brannigan, Falco, & Hansen, 2003) about implementation of school-based drug abuse prevention programs showed that for each program, 56% up to 84% of implementers modified the intervention at least slightly (program adaptation).

As the examples above indicate, in a school context, implementation does not only require specific attributes of the implementer such as attitudes, skills and self-efficacy, but also (technical) support from colleagues or the school director in order to be able to implement these programs (Dusenbury et al., 2003; Paulussen, et al., 1994; Schutte et al., 2014; Wiefferink et al., 2005). Furthermore, teachers' outcome beliefs and perceived instrumentality are predictors of the quantity and quality of program use (Wiefferink et al., 2005). Perceived instrumentality includes perceptions of suitability and feasibility of implementation. Indeed, Kirby, Laris, and Rollieri (2007) identified 17 characteristics of effective implementation of sexuality and HIV education programs, at different socio-ecological levels. Among those are support from appropriate authorities such as the Ministry of Health, school districts, or community organizations, the use of trained educators with desired characteristics, and the provision of monitoring, supervision, and support. Forman and colleagues (2009) interviewed

developers of evidence-based school-based interventions about barriers and facilitators of their programs. Two-third of the respondents indicated that they had made adjustments to the content and/or the format as a result of problems encountered during implementation. Whereas 46% of the problems were pointed out by the primary implementer, a third of the changes were a result of student responses. The authors identified several issues which can either serve as a facilitator or barrier, and classified these as either related to school organization or to implementer characteristics (Forman et al., 2009).

Building upon the model of the adoption of a school-based sex education by Paulussen et al. (1994), Wiefferink, Paulussen and colleagues (2005) developed a process model that gives a schematic view of the determinants of implementation of the school-based sex education program Long Live Love and how these determinants influence program implementation. In this model, curriculum implementation is a result of teachers' curriculum-related beliefs, shaped by the characteristics of the program strategy and of the interactive context. The influence of the program strategy on teachers' curriculum-related beliefs is both direct, and indirect through the interactive context (see Figure 1).

Figure 1: Model for explaining teachers' implementation of Long Live Love. Wiefferink et al. (2005)

Adolescent sexual health

Sexual health is an important element of general health and well-being (Douglas & Fenton, 2013). According to the World Health Organization (WHO) “*sexual health requires a positive and respectful approach to sexuality and sexual relationships, as well as the possibility of having pleasurable and safe sexual experiences, free of coercion, discrimination and violence.*” (World Health Organization, 2006). One way of achieving sexual health and well-being is through comprehensive sex education (CSE).

Throughout the world, there are several systematically developed school-based sex education programs (Coyle et al., 2001; Schaalma et al., 1996; Tortolero, Markham, & Peskin, 2010; Wight & Dixon, 2004). Besides descriptions of the interventions and effect evaluations, several studies have also looked into aspects of implementation of sex education programs (Buston, Wight, Hart, & Scott, 2002; Byers, Sears, & Foster, 2013; Ollis, 2010, 2014; Poobalan et al., 2009; Schutte et al., 2014).

The Netherlands has a long standing tradition of CSE, and worldwide, the Dutch approach for CSE is appraised (Ferguson, Vanwesenbeeck, & Knijn, 2008; Lewis & Knijn, 2001; van Teijlingen et al., 2015). Indeed, teenage pregnancy rates and STI infection rates among young people in the Netherlands are among the lowest in Europe and elsewhere (Cook & Cameron, 2015; Part et al., 2013). As Dutch teenagers become sexually active at a similar age as their American or European counterparts (mean age being 16.7), authors have suggested that the low rates of STI's and teenage pregnancies can be largely attributed to Dutch adolescents being more likely to correctly and consistently use contraceptives and a culture of openness regarding discussing adolescents' sexuality, including CSE (Cook & Cameron, 2015; Part et al., 2013; van Teijlingen et al., 2015).

The aims of Dutch sex education programs have been described as enhancing self-reliance and mutual respect, rather than self-esteem (Lewis & Knijn, 2001). Furthermore, when Lewis and Knijn (2001) compared Dutch and English sex education materials, they concluded that although the topics were largely the same, classroom discussions in the Netherlands were more open, materials were more comprehensive, and the tone of the materials towards sex were more positive than in England. Mass media in the Netherlands have historically played a significant role in breaking the general taboo regarding sexuality, as well as discussing adolescent sexuality. In line with this, the Dutch researchers Visser and Van Bilsen already stated more than 20 years ago: “*If sexuality cannot be discussed outside school, the effect of the education provided will be limited.*” (Visser & Van Bilsen, 1994; see also Ketting and Visser, 1994).

Despite these praises for Dutch CSE, a culture of openness, and the relatively low rates of teenage pregnancy and STI infections in the Netherlands, there is still room for improvement regarding the sexual health and well-being of Dutch adolescents. For example, a recent study among Dutch 12-25 year olds showed that 40.6% of young woman and 20.4% of young men have ever experienced a form of non-volitional sex, ranging from an unwanted touch or kiss to rape, and the figures are even higher for those with same-sex experiences (Dukers-Muijrers, Somers, de Graaf, Meijer, & Hoebe, 2015). In 2012, 6.264 girls under twenty became pregnant, of whom 4061 chose to have an abortion (Rutgers, 2014). Chlamydia is the most common STI among Dutch adolescents and was diagnosed in 2.5% of women and in 1.5 of men aged between 15 and 30 (van Bergen et al., 2005). However, a more recent study suggests that national prevalence among 6-30 year olds must be around 3.0% (van den Broek et al., 2012).

Another concern is the acceptance of sexual minorities, or those who identify as gay, lesbian, bisexual, and/or transgender (LGBT). In line with the more liberal legislation, Dutch adolescents appear to have more tolerant attitudes towards homosexuality than their American peers (Collier, Horn, Bos, & Sandfort, 2015). Yet, Dutch homosexual adolescents report more incidents of sexual harassment and lower levels of mental health than their heterosexual peers (Baams, Beek, Hille, Zevenbergen, & Bos, 2013; Dukers-Muijrers et al., 2015). Therefore, it is important to address safe sex practices, while also addressing relationships and wishes and boundaries, and discussing and enhancing the acceptance of sexual diversity in sex education.

At-risk populations

Overall, Dutch adolescents have better sexual and reproductive health outcomes than their peers elsewhere in Europe (Cook & Cameron, 2015; Part et al., 2013). However, there are several risk groups among Dutch adolescents, related to various aspects of sexual health. For example, young girls with a Surinamese or Antillean background have a higher chance of unplanned pregnancy, and sexual coercion is more often reported by adolescents who had their first sexual intercourse at a young age (De Graaf, Kuijer, Van Acker, & Meijer, 2012).

A specific at-risk group for becoming involved in sexual harassment (as a perpetrator and/or victim), are boys in youth care. Growing up in an unstable social environment with severe disturbances, and having been the victim of or exposed to sexual and/or physical and emotional abuse during childhood, are factors that have been associated with perpetrators of sexual harassment (Abbey, McAuslan, Zawacki, Clinton, & Buck, 2001; Barbaree &

Marshall, 2006; Burton, Miller, & Shill, 2002; P. de Bruijn, Burrie, & van Wel, 2006; Salter et al., 2003; Wekerle et al., 2009; Wolfe, Wekerle, Scott, Straatman, & Grasley, 2004).

Childhood maltreatment significantly contributes to trauma symptoms, attitudes justifying dating violence, and actual dating violence (Wolfe et al., 2004). In addition, studies show that many of these adolescents have grown up with emotional, sexual and/or physical abuse or neglect, and have witnessed domestic violence or had substance abusing parents (Connor, Doerfler, Toscano, Volungis, & Steingard, 2004).

Besides the behavioral and emotional problems of these boys, the atmosphere in the group homes might also contribute to developing positive attitudes towards deviant sexual behavior (Murnen, 2000). Group homes, particularly male-only homes, typically have a macho culture, sometimes referred to as an atmosphere of hyper-masculinity (Colton, 2002; Gault, 2012). Sexual aggression, being related to social recognition and power struggles, often defines manhood in such group settings (Murnen, 2000). Furthermore, bullying (both with and without sexual overtones) is an everyday practice in adolescent residential care groups, (Gruber & Fineran, 2008; Sekol, 2013; Sekol & Farrington, 2009; Smith, 2006).

Together, the experiences boys have gone through at home and their living conditions in youth care, make boys in residential youth care a specific risk group for perpetrating (as well as becoming a victim of) sexual harassment and sexual violence (Boendermaker, van Rooijen, Berg, & Bartelink, 2013; Colton, 2002). In response to this, the Dutch organization Rutgers, Center for Sexual and Reproductive Health and Rights, developed an intervention called Make a Move.

Make a Move

Make a Move is a sexual harassment prevention program for boys in residential youth care. In this dissertation we refer to sexual harassment, meaning any form of sexually offensive behaviour that can take place either in general or within a dating relationship. Interventions specifically directed at young people who are at-risk for sexual harassment perpetration, such as those in youth in care, are scarce (Antle, Sullivan, Dryden, Karam, & Barbee, 2011; Ball et al., 2012). Around the world, there are several teen dating violence prevention programs available (Cornelius & Resseguie, 2007; De Koker, Mathews, Zuch, Bastien, & Mason-Jones, 2014; Taylor, Stein, & Burden, 2010; Taylor, Stein, Mumford, & Woods, 2013). However, to our knowledge, only two interventions have targeted at-risk youth in care-related settings specifically (Ball, Kerig, & Rosenbluth, 2009; Wolfe et al., 2003), and

none have been implemented in the Netherlands. Make a Move is the first intervention in the Netherlands aiming to prevent sexual harassment.

Make a Move was developed for boys in residential youth care. The Netherlands has a complex system of youth care services, most of it organized on the municipality or provincial level. Low-threshold services include ‘youth and family centers’ and care through the school system. When more intensive support is necessary, the youth care as described in this dissertation comes into play. Youth care services include ambulant care, placement in foster care or residential care (group homes). There is ‘regular’ residential care, care with a specific focus on mental health services, care tailored to youth with intellectual disabilities, and youth correctional institutions. Make a Move was implemented in regular residential care. Residential youth care institutions (group homes) accommodate around 12,000 boys and girls on an annual basis in 52 institutes across the country (Boendermaker et al., 2013). Children may be placed in residential care if their parents are unable to take care of their children and if other forms of (ambulant) youth care prove insufficient. This can be the case if the child has severe emotional (internalizing) or behavioral (externalizing) problems, or when family circumstances prevent parents from taking care of their children. Placement can be either on a voluntary basis or on a juridical measure. Placement in youth care on a juridical measure is, however, not the same as a youth correctional institute, as the latter is meant for juvenile offenders. Purposes of a placement are very personal and can thus vary, but usually, the main aim is to facilitate a safe place to grow up, while working the family and/or individual problems (Boendermaker et al., 2013). In most cases, the child and staff together work toward either moving back to living with their family or other relatives or train self-reliance in preparation for (assisted) independent living. The average length of stay is currently approximately eleven months (Boendermaker et al., 2013).

At the time Make a Move was developed and institutions were being recruited for its trial implementation, a report about sexual abuse in youth care had just been released. In 2010, the Ministers for Youth and Family and of Justice decided that there should be an investigation into signs of the sexual abuse of children and adolescents who had been placed in institutions or foster families on governmental authority. For this purpose the Commissie Samson (in English: Samson Committee) was established. The committee investigated signs of and responses to the sexual abuse of children in care during the period between 1945 and 2010, and looked into current mechanisms for identifying abuse. The report ‘Surrounded by care, still unsafe’ revealed that a substantial amount of children in care is sexually abused, both by care takers and by their peers (Commissie Samson, 2012). Furthermore, professionals were only aware of 2% of incidents that minors reported to the Commission. On average, children

in care reported victimization of sexual abuse almost twice as often (143 per 1000) as average among Dutch children (74 per 1000 children) in 2010. Children in residential institutions reported the highest prevalence (194 per 1000) and girls and children with intellectual disabilities are most at risk (Commissie Samson, 2012). Following this report, several actions were taken. For Rutgers and their Make a Move program, these developments and the timing of the program becoming available made that some youth care institutions were more open towards participation in the trial implementation of Make a Move. Some institutions already took different actions, but others were still searching for ways to address sexuality and for them Make a Move came just at the right time.

Sex education in secondary education: Long Live Love+

The Dutch school system distinguishes primary and secondary schools. Children go to primary school aged 4 to 12, after which they continue to secondary education. Based on the advice of their teacher and a national examination score during the last year of primary school, they can attend preparatory secondary vocational education (VMBO, four years, age 12-16), higher general secondary education (HAVO, five years, age 12-17) or pre-university education (VWO, six years, age 12-18). Graduation from these streams allows students to continue vocational training (MBO), higher professional education (HBO) or university, respectively. Graduation from VMBO (level 4 out of 4 VMBO levels) also allows entrance to the last two years of HAVO education, and a HAVO diploma allows entrance to the last two years of VWO education, as an alternative route to higher levels of education. For more information on the Dutch school system, see EP Nuffic (2015).

Long Live Love+ is a spin-off of the widely used Long Live Love program. The original Long Live Love program was developed in the early 1990s and has since then been updated several times (Schaalma et al., 1996; Schutte et al., 2014; Vanwesenbeeck, Poelman, Reinders, & Schaalma, 2002). It is a school-based and teacher-led sex education program for 13- and 14 year olds aiming to provide students with communication and negotiation skills to enable safe sex practices (Schutte et al., 2014). However, not all students become sexually active at the same age.

Surveys among adolescents in the Netherlands show that, although the average age for first time intercourse is 16.7 years, Dutch adolescents with higher educational levels have their first sexual experiences at a later age than their peers with a lower level of education: Whereas 22% of lower educated 14 and 15 year olds have had sexual intercourse, the percentage for HAVO and VWO students at that age is only 7% (De Graaf et al., 2012).

However, health education results in more effective change when the information has a high personal relevance for the receiver (Kreuter & Wray, 2003; Petty & Cacioppo, 1986; Petty, Cacioppo, & Goldman, 1981; Rosenstock, Strecher, & Becker, 1988; Rothman & Schwarz, 1998). Furthermore, repetition fosters learning (Emier, Windchey, Zaino, Feldman, Stephen, & Scheetz, 1980; Lee, Swanson, & Hall, 1991; Rock, 1957). It thus seems unlikely that sex education at one single time point throughout multiple years of secondary school suffices to address most adolescents in a timely manner. Not only is the repetition beneficial in itself, but a repeated sex education curriculum also has the potential to be relevant for students who have become sexually active at a later stage.

Students differ in their mental and learning capacities, as reflected in the Dutch school system. This has implications for how to address students with different levels of intellectual abilities, using different methods and strategies in health promotion programs. We can thus conclude that, to enhance the effectiveness of sex education, additional programs are needed that are age-appropriate, adapted to the educational level of the target population, the sexual health situation of the specific target population, and preferably repeated over time.

With these insights, the original Long Live Love program for young adolescents was not only updated several times since the early 1990s, but also expanded with two additional programs for adolescents aged 15 and over. One is Long Live Love upper classes, the other spin-off is Long Live Love – Vocational training (LLL-MBO). Two chapters in this dissertation describe the effect and process evaluation of Long Live Love upper classes (in Dutch: Lang Leve de Liefde - Bovenbouw), further referred to as LLL+. The LLL+ curriculum targets the upper classes of the higher levels in secondary schools: *HAVO* ('higher general continued education' - 12 to 17 yrs) and *VWO* ('preparatory scholarly education' or 'pre-university training' – 12 to 18 yrs). The four themes addressed in LLL+ are sexual wishes and boundaries, pregnancy prevention, STI prevention, and sexual diversity.

Aims and outline of this dissertation

The aims of this dissertation are as follows: to evaluate the effectiveness of both Make a Move and Long Live Love+ and to evaluate how successfully the programs were implemented, i.e. to what extent the programs were implemented as planned, what factors determined the implementation, and how the implementers reflected upon the program and its implementation. **Chapter 2** describes the results of the effect evaluation of Make a Move. This was done through a RCT with 10 institutions in the intervention group, and 10 in the waiting-list control group (N = 177 at baseline). Surveys were administered before the first

Make a Move meeting, at the end of the eighth meeting, and six months later. Mixed linear modelling analysis was used to deal with the nested design and the large dropout to follow-up. Surveys were administered before the first Make a Move meeting, at the end of the eighth meeting, and six months later. Main outcome measures were psycho-social determinants, including attitudes, self-efficacy, and intentions. **Chapter 3** reports on the quality of the implementation of the program, based on the diaries that Make a Move trainers kept (in 10 of the 13 groups) and interviews with the trainers (N = 7). The interviews were analyzed using thematic content analysis. **Chapter 4** reports on the findings of the effect evaluation of LLL+. In this RCT, a total of n=999 students from 17 schools participated. After the schools were matched in pairs, one school was randomly assigned to the intervention group and the other to the waiting-list control group. Outcome measures included determinants such as attitudes, self-efficacy and intentions, as well as behavioral outcomes, related to sexual relationships, pregnancy prevention, STI prevention, and the acceptance of sexual diversity. **Chapter 5** describes the implementation of LLL+, and how teachers and students reflected on the program. Teachers were mostly interviewed individually, whereas students participated in focus group discussions. These qualitative findings were analyzed using thematic content analysis. Finally, **chapter 6** summarizes and discusses the major findings of the four studies, and proposes a more extensive process model for explaining the implementation of sex education programs. Moreover, it discusses the limitations, recommendations and conclusions of this dissertation.

Chapter 2

Make a Move: A comprehensive effect evaluation of a sexual harassment prevention program in Dutch residential youth care

This chapter has been published as:

Van Lieshout, S., Mevissen, F.E.F., van Breukelen, G., Jonker, M., Ruiter, R.A.C. (2016). Make a Move: A Comprehensive Effect Evaluation of a Sexual Harassment Prevention Program in Dutch Residential Youth Care. *Journal of Interpersonal Violence*
DOI: 10.1177/0886260516654932

Abstract

Sexual harassment - unwanted sexual comments, advances or behaviors – and sexual violence are still prevalent worldwide, leading to a variety of physical, cognitive and emotional problems among those being harassed. In particular, youth in care are at risk of becoming perpetrators (and victims) of sexual harassment. However, in general, there are very few interventions targeting this at-risk group, and no such programs exist in the Netherlands. To this end, a group intervention program - Make a Move - targeting determinants of sexual harassment was developed. This program was implemented and evaluated among boys ($N = 177$) in Dutch residential youth care (20 institutions). A pre-test, post-test, and six-month follow-up design including an intervention and a waiting-list control group with randomized assignment of institutions (cluster randomized trial) was used to measure the effects of the intervention on determinants of sexual harassment. Multilevel (mixed) regression analysis with Bonferroni correction for multiple testing ($\alpha = .005$) showed no significant effects of Make a Move on determinants of sexual harassment (p 's $> .03$; Cohen's d 's $< .44$). Results are discussed in light of a 3-way explanatory model focusing on intervention content, evaluation, and implementation as potential explanations for not finding any measurable intervention effects.

Introduction

Sexual harassment – in this paper conceptualized as unwanted sexual comments, advances or behaviors that cause harm to the victim – and sexual violence towards women are major problems worldwide. According to the World Health Organization, globally, 35.6% of women experience sexual violence at least once in their lifetime (World Health Organization, 2013). Although sexual harassment is more prevalent in less developed countries with higher gender inequalities (Agardh, Odberg-Pettersson, & Östergren, 2011; Devries et al., 2013; Peterman, Palermo, & Bredenkamp, 2011), Western countries also show worrisome rates. A surveillance study by the American Centers for Disease Control and Prevention (CDC), for example, showed that 10.4% of American high school students (girls but also boys) who had been out with or dated someone in the previous year had been the victim of some type of sexual harassment in the twelve months preceding the survey (Centers for Disease Control and Prevention, 2014). In a recent British study, on average, 45% of 13- and 14-year-old girls and boys reported having been a victim of sexual harassment by someone they had ever dated or been with in a relationship. The authors reported that emotional abuse and controlling behaviors (38%), physical abuse (17%), and sexual victimization (14%) were the three most reported types of harassment (Fox, Corr, Gadd, & Butler, 2014). Moreover, Dutch studies suggest that 4%-23% of boys and 18%-36% of girls have experienced some form of physically violent sexual behavior (De Bruijn et al., 2006; De Graaf, Meijer, Poelman, & Vanwesenbeeck, 2005; Dukers-Muijters et al., 2015). The experience of sexual harassment can lead to a variety of physical, cognitive and emotional problems, such as post-traumatic stress disorder (PTSD), substance abuse, depression, and anxiety (Holmes & Sher, 2013; Kilpatrick et al., 2000, 2003; Van Berlo & Mooren, 2009), and as such, there is a strong need for prevention efforts. Furthermore, adolescence is an important period in terms of developing (healthy) sexual attitudes and behaviors, a further reason why interventions targeted at this particular group may help to prevent these individuals from becoming perpetrators of sexual harassment (Foshee, Benefield, Ennett, Bauman, & Suchindran, 2004; White & Smith, 2004).

Understanding the factors that are linked to an individual becoming or being a perpetrator of sexual harassment is important for designing prevention programs (Bartholomew Eldredge et al., 2016). Meta-analyses and other reports on prevalence rates show that the vast majority of perpetrators of sexual harassment are male, and a significant proportion of them are underage (Barbaree & Marshall, 2006; Fago, 2003; Hall & Barongan, 1997). Psychiatric and personality disorders have also been identified as determinants of sexual harassment behaviors (Kafka & Hennen, 2002; Zakireh, Ronis, & Knight, 2008). Furthermore, growing up in an unstable social environment with severe disturbances,

and having been the victim of or exposed to sexual and/or physical and emotional abuse during childhood, are factors that have been associated with sexual harassers (Abbey et al., 2001; Barbaree & Marshall, 2006; Burton et al., 2002; De Bruijn et al., 2006; Salter et al., 2003; Wekerle et al., 2009; Wolfe et al., 2004). Others studies have shown that perpetrators of sexual harassment generally have an earlier sexual debut and a greater number of sex partners as compared to their non-offending peers (Abbey et al., 2001; Barbaree & Marshall, 2006; Wekerle et al., 2009). Wolfe and Wekerle (2004) found that childhood maltreatment significantly contributed to trauma symptoms, attitudes justifying dating violence, and actual dating violence. Furthermore, some studies have reported a relationship between committing acts of sexual harassment and having friends who approve of forced sex (Abbey et al., 2001; De Bruijn et al., 2006).

The findings described above suggest that in particular youth in residential care (youth living in 24-hour group living facilities that are not licensed as hospitals or youth correctional institutions and offer (mental health) care for children and adolescents with serious emotional and/or behavioral problems, including law violations) are highly vulnerable in terms of involvement in sexual harassment and violence, both as a victim and/or a perpetrator, as these adolescents often have behavioral and emotional disturbances (Connor et al., 2004; Jonson-Reid, Scott, McMillen, & Edmond, 2007). In addition, studies show that many of these individuals have grown up with emotional, sexual and/or physical abuse or neglect, and have witnessed domestic violence or had substance abusing parents (Connor et al., 2004). Besides the behavioral and emotional problems of these boys, the atmosphere in the group homes might contribute to developing positive attitudes towards deviant sexual behavior (Murnen, 2000). Group homes, particularly male-only homes, typically have a macho culture, sometimes referred to as an atmosphere of hyper-masculinity (Colton, 2002; Gault, 2012). Sexual aggression, being related to social recognition and power struggles, often defines manhood in such group settings (Murnen, 2000). Furthermore, bullying (both with and without sexual overtones) is an everyday practice in adolescent residential care groups, (Gruber & Fineran, 2008; Sekol, 2013; Sekol & Farrington, 2009; Smith, 2006). Together, the experiences boys have gone through at home and their living conditions in youth care, make boys in residential youth care a specific risk group for perpetrating (as well as becoming a victim of) sexual harassment and sexual violence (Boendermaker et al., 2013; Colton, 2002).

Over the years, many sexual harassment prevention programs have been developed to empower women and men against sexual victimization and to prevent potential perpetrators from sexually harassing others (see e.g., De Graaf, de Haas, Zaagsma, & Wijsen, 2015;

De Koker, Mathews, Zuch, Bastien, & Mason-Jones, 2014; DeGue et al., 2014; Taylor, Stein, Mumford, & Woods, 2013). However, evidence regarding the effectiveness of these programs is not straightforward, as the interventions differ along multiple dimensions such as intervention goals, methods, settings, design, and target groups, making it difficult to draw general conclusions. In addition, interventions specifically directed at young people who are at-risk for sexual harassment perpetration (such as those in youth in care) are scarce (Antle et al., 2011; Ball et al., 2012).

To our knowledge, only two interventions have targeted at-risk youth in care-related settings (Ball et al., 2009; Wolfe et al., 2003), and none has been implemented in the Netherlands. Ball and colleagues' *Expect Respect* dating violence prevention program for pupils who have experienced domestic violence or sexual abuse in the United States showed improved communication skills as well as anger control, knowledge of healthy relationships, and awareness of abusive behaviors (Ball et al., 2009). Wolfe and colleagues (2003) evaluated a Canadian teen dating violence prevention program targeting adolescents who were recruited through Child Protection Services. Participants showed decreased emotional distress and reduced incidence of both physical and emotional abuse after participating in the program. Veerman and Van Yperen (2007) point out the difficulties of performing a randomized controlled trial (RCT) in youth care practices. These difficulties, alongside possible ethical considerations, may contribute to the lack of intervention evaluations found in scientific literature targeting this target group.

In response to the need for a sexual harassment prevention program specifically targeting at-risk boys in Dutch youth care, the Dutch organization Rutgers --Center for Sexual and Reproductive Health and Rights -- developed an intervention called Make a Move. This intervention was designed to be delivered in a group setting using groups of six to eight boys. In the Netherlands, youth care services include ambulant care, placement in foster care or residential care (group homes), residential care with a specific focus on mental health services or disabilities, and youth correctional institutions. Make a Move targets boys in regular residential care. Residential youth care institutions (group homes) accommodate around 12,000 boys and girls on an annual basis (Boendermaker et al., 2013). Children may be placed in residential care if their parents are unable to take care of their children and if other forms of (ambulant) youth care prove insufficient. This can be the case if the child has severe emotional or behavioral problems, or when family circumstances prevent parents from taking care of their children. Placement can be either on a voluntary basis or a juridical measure although residential care should not be mistaken for youth correctional institutes. The main purposes of a placement vary as these are very personal, but usually

the main aim is to facilitate a safe place to grow up combined with family and/or individual interventions (Boendermaker et al., 2013). In most cases the child and staff together work towards either moving back to living with their family or other relatives or train self-reliance in preparation for (assisted) independent living. The average length of stay has in recent years decreased from 24 months to approximately eleven months (Boendermaker et al., 2013).

Make a Move was implemented by freelance trainers working for Rutgers. The main goal of Make a Move is to promote respectful (sexual) relationships and to prevent sexual harassment. Subgoals include, for example, increasing awareness of differences in relationship expectations and teaching the boys how to control their wishes and desires if they are not in line with their partners' wishes. Make a Move intends to be a prevention program, not a treatment program. Make a Move has a cognitive-behavioral approach and addresses determinants related to sexual harassment such as attitudes and social-relational skills. As part of a trial implementation of the program, an effect and process evaluation was performed to assess the effectiveness of Make a Move in influencing determinants of sexual harassment. Outcome measures were based on the objectives of the program and taken from previous studies exploring determinants of sexual harassment. Previously identified determinants include gender attitudes (Abbey et al., 2001; McCauley et al., 2013), adverse sexual beliefs, attitudes about gender roles, acceptance of verbal pressure in obtaining sex (Abbey et al., 2001), rape myth acceptance, acceptance of interpersonal violence, adverse sexual beliefs, rape empathy, and self-efficacy regarding dealing with potentially abusive situations (O'Donohue, Yeater, & Fanetti, 2003). This paper describes the effect evaluation of a trial implementation of Make a Move in Dutch residential youth care. The process evaluation is described elsewhere (Van Lieshout, Mevissen, Jonker, & Ruiter, under revision).

Methods

Recruitment, randomization, and study design

All (N = 49) Dutch youth care institutions that offer residential care to boys aged 12-18 were contacted by e-mail and telephone and were invited to participate (with one or more groups of boys) in the trial implementation of Make a Move. Twenty (40.8%) institutions agreed to participate in the pilot; the remaining 29 (59.2%) institutions had no interest in participating at that moment. Reasons mentioned for non-participation varied and ranged from already partaking in another research project, lack of time, to mainly taking care of girls

at that time. To prevent large baseline differences with regard to background characteristics between the intervention and control groups, institutions were matched in pairs based on characteristics such as geographical setting, type of care offered, and number of participating groups of boys within the institutions. Within each matching pair, one institution was randomly assigned to the intervention condition (the Make a Move program) and the other was assigned to the waiting-list control condition (no prevention program). In the end, each treatment condition included 10 institutions. The effectiveness of the Make a Move program in changing determinants of sexual harassment was evaluated using a baseline measure (T0), a direct post-measure (T1), and a six-month follow-up (T2).

The 20 participating institutions were requested to form one to three groups of boys, each consisting of six to eight boys aged 12-17 years. Important exclusion criteria for participation were as follows: having severe Autism Spectrum Disorders (ASD), being unable to function in group meetings (e.g., due to aggression regulation problems), having an IQ level below 80, and being known by the institution to suffer from post-traumatic stress due to sexual abuse or have been convicted as perpetrator of sexual assault. The reason for the latter exclusion criteria is that Make a Move is a preventive program, not a treatment component. In addition, at the start of the program, boys had to have a remaining stay of at least eight months within the institution in order to be able to finish the full program including the follow-up survey. In the end, 29 groups were formed (14 in the intervention condition, 15 in the control condition) with number of groups per institution ranging from one to three. The number of boys per group at the start of the program ranged from three to nine.

Figure 1: Timeline Make a Move trial implementation

At baseline, 177 boys (44% in the intervention group) started the pilot and filled out the first questionnaire (T0). In the end, 139 boys (46% in the intervention group) completed the direct post measure (T1), and 68 (32.4% in the intervention group) completed the six-month follow-up (T2). Individual time measures were matched on date of birth, and mother's initials. Not all boys consistently filled out these matching criteria. Matching of all three questionnaires was possible for 63 boys while matching of two questionnaires was possible for 64 boys (59 for T0 and T1, one for T0 and T2 and four based on T1 and T2).

Participants

With regard to the 177 boys starting the pilot at T0, mean age was $M = 14.8$ ($SD = 1.64$, range 12 to 18 years). The majority were of Dutch origin (61.6%). A few participants had

a (partly) Antillean-Surinam (10.7%), or Turkish or Moroccan (5.1%) background, and the remaining boys (22.6%) were from a very diverse range of (non-) Western backgrounds. Most boys (53.4%) had no religious orientation. Of the remaining boys, 27.6% had a Christian orientation (Catholic, Protestant or similar), 7.5% an Islamic, and 11.5% another sometimes unclearly defined orientation (Agnostic, God, or an 'I don't know' response). The educational level of most boys (95.5%) did not exceed lower vocational education, which is the lowest level of secondary education in the Netherlands.

Of these 177 boys, most (71%) had had sexual intercourse, and 19% reported having had sexual experience without intercourse (i.e. touching, oral sex). Age at first intercourse ranged from 6 to 17 years old ($M=13.3$; $SD=1.85$, median 13 years old). Reported number of sexual partners ranged from 1 to 30 (median = 5). Ninety-two percent had had a relationship in the past two years. Regarding reported sexual harassment behaviors, 40.9% reported to have sometimes or regularly tried to persuade somebody into having sex and 23.8% had sometimes or regularly taken advantage of another person's drunkenness to get sex, 23.5% had sometimes or regularly used nagging in order to get sex and 17.2% had sometimes or regularly lied to get sex. One in ten (9.7%) had sometimes or regularly become angry towards a partner, 4.0% had sometimes or regularly used violence, and 4.0% had sometimes or regularly threatened their partner with violence in order to get sex.

Procedure

All participating institutions received an informed consent form to sign. In addition, they received a letter that they were asked to use to inform parents or legal guardians about Make a Move, which included information about the option to withdraw the child by requesting an objection form (i.e. passive consent). No requests were received for an objection form. After the institutional consent forms had been returned, all participating boys received the first paper-pencil questionnaire, which included information about the research, their privacy, and the fact that they could withdraw at any time, as well as instructions on how to fill out the questionnaire. Their consent was given by ticking the box indicating that they understood and agreed with the terms of agreement. Prior to the start of the intervention group, the contact persons of the institutions were asked to invite the potential participating boys for a brief introduction talk with the aim to inform and motivate them to participate. After the baseline measurement was completed, seven schooled trainers from Rutgers (one male, six female, aged 51-58, $M=54.6$), often assisted by an employee from the youth care institution, delivered the Make a Move program to boys in the intervention group, during eight group meetings over a time period of about eight weeks. Some trainers visited several institutions,

and some institutions were visited by several trainers. Within groups, one visiting trainer consistently led all meetings, but trainers were not always consistently assisted by one and the same youth care employee. No group meetings were organized for the boys in the control group during those eight weeks. Eight weeks after T0, the boys were asked to fill out the post-test survey (T1). The T0 and T1 surveys were filled in during a group meeting under the supervision of a trainer (intervention group) or a researcher (control group) to assure a quiet and private surrounding for the boys and in case the boys needed help with the survey. For the six-month follow-up (T2), all institutions were contacted by phone to make arrangements for administering the six-month follow-up questionnaires to the boys. As most boys had by then left the institutes, T2 was not administered in a similar group setting as T0 and T1. Instead, institutes were requested to forward the questionnaires to the boys who had left the institute. After the T2 measurement, waiting-list control group institutions were assigned a trainer and also ran the intervention.

Make a Move program

Make a Move is a group intervention for boys aged 12-17 years about sexuality and relationships, designed to prevent sexual harassment and dating violence. The program consists of eight weekly meetings of 90 minutes each. The themes covered by the eight meetings are, in order: *men, image, girls, sex, flirting, dating, pleasurable sex, and the future*. Each meeting includes several exercises such as role play, discussion, and watching short movie clips. See Table 1 for an overview of the content of the program. A competitive element runs through the program. Boys can earn credits for doing the exercises as well as for active participation and good manners, both individually and as a group. At each meeting, a new score-keeper is appointed among the boys, and the final winner receives an incentive. This competitive element of the intervention was added after a pre-test of the program, in order to increase the boys' engagement in the tasks.

Table 1: Make a Move Program

Lesson	Assignment	Determinant(s) addressed	Goals (summary)	Elaboration on assignments (methods and application)
1. Manhood	1. Getting to know each other		Boys: <ul style="list-style-type: none"> get to know each other 	Boys are asked beforehand to bring along a personal item which they can use to introduce themselves
	2. Ground rules and safety	Attitudes, emotions (feelings of safety and trust)	Boys: <ul style="list-style-type: none"> acknowledge the importance of respecting each other acknowledge the importance of group rules regarding safety and trust mention that they are entitled to set their own boundaries in this program are willing to ask questions about girls, flirting, dating, intimacy and sexuality 	Trainer creates a safe group environment by stimulating the boys to formulate group rules that will be written down for everybody to read. Boys sign a contract
	3. Typical men	Attitudes, norms	Boys: <ul style="list-style-type: none"> mention different aspects of manliness are aware that perceptions on manliness differ among different people, genders and cultures acknowledge that one's ideas about manliness are influenced by multiple factors such as group norms, family, and culture 	Group discussion on worksheet with images of six different men
2. Friends and image	1. My web profile	Self-esteem, knowledge, perceptions and attitudes, skills	Boys: <ul style="list-style-type: none"> can list their own positive traits (looks and behavioral) can give examples how they can use their positive traits to make contact with others in a positive way 	Individual assignment with worksheet. Boys choose from a list descriptions that fit their looks, character and things they like to do. Followed by group discussion
	2. What do you do together with a close friend?	Perceptions and attitudes, skills	Boys: <ul style="list-style-type: none"> know how to name some common/basic emotions can indicate their wishes and boundaries when interacting with others, both one-on-one and in groups 	Watch short movie clip with funny and not-so-funny situations, individual assignment with worksheet, group discussion
	3. Friends among each other	Perceptions and attitudes, skills	Boys: <ul style="list-style-type: none"> can indicate their wishes and boundaries when interacting with others, both one-on-one and in groups can give examples of positive and negative ways in which friends can behave towards each other in group settings can list situations in which boys might experience peer pressure can mention ways how to deal with peer pressure 	Team assignment in which cards indicate type of question/assignment: 1 true/false questions, 2 sharing opinions, 3 role play

Lesson	Assignment	Determinant(s) addressed	Goals (summary)	Elaboration on assignments (methods and application)
3. Girls	1. Typical girls	Knowledge, perceptions and attitudes	Boys: <ul style="list-style-type: none"> • have more knowledge and less misconceptions about girls, intimacy and sexuality • know that there is not one type of boy and one type of girl 	Individual assignment with true/false statements, group discussion
	2. Boys want sex, girls want love	Perceptions and attitudes	Boys: <ul style="list-style-type: none"> • are aware that some ideas about boys, girls and sexuality are stereotyped • are aware that stereotype ideas may lead to misunderstandings 	Watch short movie clip in which 3 boys and 3 girls react to the statement 'Boys want sex, girls want love', followed by group discussion
	3. Your ideal girlfriend	Perceptions and attitudes	Boys: <ul style="list-style-type: none"> • can list qualities and traits they find important in a (future) partner 	Individual assignment in which boys receive 4 different colors of sticky notes to write down what they find important in an ideal girlfriend (or boyfriend) under the headings: looks, character, background, and what the ideal relationship would look like
4. Sex	1. Word association task 'sex'	Knowledge, perceptions and attitudes	Boys: <ul style="list-style-type: none"> • can list different aspects of sexuality, such as: feelings, sexual behavior, sexual preferences and orientation • are aware that sex is something very personal and can thus be different for each individual 	Brainstorm in two teams, followed by group discussion
	2. Sex in the picture	Perceptions and attitudes	Boys: <ul style="list-style-type: none"> • recognize their own wishes and boundaries concerning sexuality • are aware that stereotypical ideas may lead to misunderstandings 	Watch short clip in which three boys and three girls discuss 'the first time' and being a virgin
	3. Sex quiz	Knowledge	Boys: <ul style="list-style-type: none"> • have more knowledge and less misconceptions about boys and sexuality • can recognize (their own) misconceptions about sexuality 	Individual assignment, group post-discussion
	4. Towards the top	Knowledge, perceptions and attitudes	Boys: <ul style="list-style-type: none"> • know that girls can experience having sex in a different way than boys • can mention consequences of the differences in which boys and girls experience sex • are aware there are different ways to communicate sexual wishes and boundaries 	Answer questions in teams, psycho-education, group post-discussion

Lesson	Assignment	Determinant(s) addressed	Goals (summary)	Elaboration on assignments (methods and application)
5. Flirting	1. Flow master	Perceptions and attitudes, skills and self-efficacy	Boys: <ul style="list-style-type: none"> can mention and recognize signals with which girls may give the feedback that she is having a good time can mention ways in which they can get in touch with a girl in a positive, welcome, and respectful manner 	Watch short clip, group post-discussion, individual assignment, group post-discussion
	2. Respect	Perceptions and attitudes, skills and self-efficacy	Boys: <ul style="list-style-type: none"> acknowledge the importance of being aware of feelings, both their own and the feelings of the other person (when flirting) now that a feeling can change over time, for example a yes (wish) can change into doubt or no when circumstances change know how to react when a flirting situation changes 	Group session, use of emoticon signs
	3. The right move	Skills	Boys: <ul style="list-style-type: none"> can mention and recognize indirect, (non-) verbal and subtle ways in which girls may give feedback can apply different ways of making contact with girls in a positive and respectful manner, both offline and online know what to do when they interact with girls and start doubting what they want themselves, or think the other person is unsure acknowledge that friends can assist in discovering one's wishes and boundaries, but can also encourage crossing boundaries 	Watch short video clip, group discussion, role play, group post-discussion
	4. Chat contact	Perceptions and attitudes, skills	Boys: <ul style="list-style-type: none"> can mention and recognize indirect, (non-) verbal and subtle ways in which girls may give feedback can think of ways to react when a girl seems to show doubt or to be uncomfortable in intimate interaction (eg take a break, ask her what she wants) are aware that boys and girls can experience the same situation in different ways 	Watch short video clip, individual assignment, group post-discussion
6. Dating	1. The date	Skills	Boys: <ul style="list-style-type: none"> show ability in dealing with expectations or pressure from others (friends) regarding flirting and dating can deal with situations in which there's a difference in how a boy and a girl experience (sexual) contact 	Watch short movie clip, team role play assignment. Boys receive a card with 'Golden Dating Rules'

Lesson	Assignment	Determinant(s) addressed	Goals (summary)	Elaboration on assignments (methods and application)
7. Sex is fun	2. Dating skills	Knowledge, perceptions and attitudes, skills	Boys: <ul style="list-style-type: none">are aware of their own ideas, attitudes and misperceptions regarding love and datingknow what to do / how to clarify when they are not sure about the wishes and boundaries of another personknow how to deal with rejection and no-signals when interacting with a girl	Role play in teams
	1. Body mapping	Knowledge, perceptions and attitudes, skills and self-efficacy	Boys: <ul style="list-style-type: none">acknowledge everyone has their own personal wishes and boundaries in physical contactrecognize the importance of knowing what the other person likes and does not like when being intimate	Individual body drawing assignment, group discussion
	2. Boundaries in the picture	Perceptions and attitudes	Boys: <ul style="list-style-type: none">know there are different (religious, cultural, social, and personal) rules that may affect one's behaviorare aware of misconceptions regarding boundaries when flirting and in sexual relationshipsknow criteria to assess when flirting/dating/sexual behavior is harassment	Watch short movie clip, followed by group discussion
	3. Do you or don't you want to? Can you fix it?	Perceptions and attitudes	Boys: <ul style="list-style-type: none">know the difference between acceptable and unacceptable sexual behaviorcan apply (in a training setting) adequate and respectful rules for behavior and interaction with regard to sexualityknow criteria to recognize risky situationscan effectively regulate impulsive behavior and thus adequately handle risky situations	Watch short movie clip, group discussion
8. The future				
8. The future	1. Warming up: sexual career	No program goal, ice-breaker		Group discussion
	2. Timeline of the future	Intention	Boys: <ul style="list-style-type: none">think about their future and the choices they want to make (in terms of relationships and sex)are aware that choices they make now have an influence on the possibilities later	Individual timeline assignment, group post-discussion
	3. Good fathers		Boys: <ul style="list-style-type: none">list what they have learned about boys, girls, dating, sex and internet	Individual assignment and role play
	Overarching	Competition	Motivation <ul style="list-style-type: none">are motivated and stimulated to actively participate	At each meeting two boys are appointed to be a referee and a score-keeper. At the end of the last meeting, the one with the highest score receives a prize

Measures

The questionnaire included measures on background characteristics: age, ethnicity, religion, educational level (very low, low, high, unknown), sexual history (relationship experience over the past 2 years (yes vs. no), sexual experience with vs. without penetration (yes vs. no), age first time sex, lifetime number of sex partners, and sexual harassment behaviors such as ‘have you ever tried to persuade somebody into having sex?’ (never – sometimes – regularly – very often). This last response scale was recoded into three response categories: never, sometimes, and regularly / very often.

In addition, the questionnaire included six validated scales previously used in empirical research to measure antecedents of sexually offensive behavior, as well as eight measures based on elements taken from the Reasoned Action Approach (Fishbein & Ajzen, 2010) that were adapted to the outcome goals of the intervention. From the six validated scales, we selected a subset of items with the highest inter-item correlations, in order to keep the survey short and maximize the chance that the boys would be able to concentrate on it. All psychosocial measures used 7-point Likert scales. Higher scores indicate more of that attribute. We calculated mean scores for scales at each time point for a participant only if more than half of the items that made up one measure had been answered. The outcome measures and their mean scores are showed in Table 2.

Self-esteem was measured with five items ($\alpha = .64$) taken from Heatherton and Polivy (1991); e.g., ‘I am confident about my talents’ (1 = totally disagree, 7 = totally agree). Our measure of **empathy** was based on Bos, Dijkster, and Koomen (2007) and Van Alphen, Dijkster, Bos, Van den Borne, and Curfs (2011). Boys were asked to read a short scenario in which a boy assaulted a girl on the street and then had to indicate by use of two items the degree to which they experienced empathy (‘I feel empathy / compassion for the girl’, 1 = totally disagree, 7 = totally agree; $r = .71$). **Adverse sexual beliefs** were measured using three items ($\alpha = .68$); e.g., ‘I think women mostly date men as to make use of them.’ (1 = totally disagree, 7 = totally agree) and adapted from Burt (1980). **Attitude towards dating violence** was measured by nine items ($\alpha = .81$) e.g., ‘I think that a girl who cheats on her boyfriend deserves to be beaten.’ (1 = totally disagree, 7 = totally agree) and was based on Price and Byers (1999). **Rape attitude** included eight items ($\alpha = .91$) e.g., ‘I think a boy is allowed to force a girl into having intercourse when she has allowed him to take off her pants’ (1 = totally disagree, 7 = totally agree), and was based on Maxwell, Robinson, and Post (2014). The scale **outcome expectancies** ($\alpha = .89$) was based on O’Donohue, Yeater, and Fanetti (2003). The scale measured to what extent the participant would expect seven outcomes that could occur as a consequence of persuading a girlfriend who

is not in the mood to have sex (e.g., I would enjoy, she would enjoy, I would feel guilty; 1 = no, definitely not, 7 = yes, absolutely).

Attitudes towards stating boundaries when with friends ('To say no to my friends when I don't agree with them, seems to me...', $\alpha = .60$), **attitudes towards communicating with a partner** ('To ask my girlfriend what she does and does not want during sex, seems to me...', $\alpha = .88$) and **attitudes towards self-control** ('To only have sex with my girlfriend when she also wants it, seems to me...', $\alpha = .88$) were each measured by one statement with three semantic differentials (very bad -very good; very unimportant - very important; very uncomfortable- very comfortable. **Self-efficacy towards self-control** was measured with the statement: 'Imagine you have a girlfriend. You want to have sex but she does not want to.' followed by four response questions, such as: 'would you manage to not become angry (1 = no, absolutely not; 7 = yes, definitely; $\alpha = .85$). **Self-efficacy towards communication** was measured by a scale which included 5 items ($\alpha = .87$), such as: 'I'm able to communicate with my sex partner about what we do and do not want regarding sex' (1 = no, absolutely not; 7 = yes, definitely). **Social norms** were measured with 4 items ($\alpha = .61$) e.g., 'my friends believe it's okay to force a date into having sex when you spent a lot of money on her' (1 = no, absolutely not; 7 = yes, definitely). **Self-efficacy to deal with peer pressure** was measured with 3 items ($\alpha = .60$), e.g., 'I'm able to say no to my friends when I don't agree with them' (1 = no, absolutely not; 7 = yes, definitely). Finally, **intention** was measured with 3 items ($\alpha = .78$) and assessed to what extent subjects expect to become angry, start persuading, or start whining when they want to have sex but their girlfriend does not (1 = no, absolutely not; 7 = yes, definitely).

Data analysis

The data were analyzed in IBM SPSS 21 using mixed (multilevel) regression because of the nested structure of the design (repeated measurements nested within boys, who are nested within institutions). The predictors were: condition (0 = control, 1 = intervention) and time of measurement (with dummy-indicators for post-test and follow-up, respectively t1 and t2, and baseline as reference category). To capture unexplained outcome differences between institutions, a random intercept was included. To capture inter-individual differences over time, an unstructured covariance matrix for the repeated measures was selected. In line with the Consort 2010 guidelines for reporting randomized trials, we did not check for baseline differences on outcome measures between intervention and control group, as any potential differences are caused by chance (Gruijters, 2016; Moher et al., 2010). Potential differences have, however, been adjusted for with the random intercept. Due to the small sample size no

distinctions have been made with regard to the different cultural backgrounds or religions. Possible outcome differences between groups within institutions could not be modeled as most institutions contributed only one group, and those contributing more than one group did not (always) include a group identifier in the returned data. Similarly, possible outcome differences based on trainers could not be modeled as some trainers visited several institutions, and some institutions were visited by several trainers.

Histograms were used to check for normality and possible outliers in the distribution of the random intercept, as well as of the residuals per time point. Two outcome measures (social norm and rape attitude) showed severe non-normality (e.g., a large floor or ceiling effect). The scores on these outcome measures were recoded in order to better approach normality or at least to reduce severe skewness and outliers¹. The effect analysis for these outcome measures were then rerun to check the robustness of the untransformed results. Running the effect analyses on the transformed outcome measures, did not change the results in any relevant way. We therefore report the original (untransformed) data. Results were considered significant at $p \leq .005$ (two-tailed) to correct for multiple testing². Effect sizes were calculated as $\beta / \sqrt{\text{var}}$, where β is the estimated group difference at T1 or at T2, and var is the residual outcome variance at that same time point. Note that this effect size comes down to Cohen's d per time point, except that it is not based on the student t-test for two unpaired means, but on multilevel analysis taking into account the baseline T0 and the nesting of boys within institutions.

Results

Table 2 provides the mean scores (observed) on the outcome measures at T0, T1, and T2. The aim of this research was to investigate whether there are program effects at posttest and follow-up, as reflected by a significant condition by time interaction (represented by the product terms *condition*t1dummy* and *condition*t2dummy*). No significant support for a program effect was found on any of the 14 outcome measures (p 's $> .03$; see Table 3) but some small ($d < .2$) and small to medium ($.2 < d < .5$) effect sizes, both at short-term (T1) and long-term (T2), were observed.

Table 2: Mean scores (SD) on outcome measures

1 Since the non-normal outcomes were discrete rather than continuous the best recoding was the pooling of adjacent categories, such that the outcome became much less skewed.

2 Applying Bonferroni correction by dividing the p-value of .05 by the number of outcome measures (14), leads to a p-value of $.05/14=.0036$. However, as the Bonferroni correction is an overcorrection, we rounded it off to .005.

Outcome	T0 baseline Mean (SD)		T1 posttest Mean (SD)		T2 follow-up Mean (SD)	
	I-group	c-group	I-group	c-group	I-group	c-group
Attitude communication	5.57 (1.40)	5.54 (1.53)	5.80 (1.30)	5.71 (1.39)	6.11 (1.02)	6.09 (1.16)
Attitude self-control	6.01 (1.39)	6.04 (1.28)	6.33 (1.09)	6.14 (1.32)	6.54 (1.05)	6.21 (1.07)
Attitude boundaries	5.68 (1.28)	5.48 (1.16)	5.63 (1.24)	5.63 (1.28)	5.81 (.87)	5.86 (1.08)
Empathy	5.82 (1.56)	5.79 (1.61)	6.09 (1.26)	5.74 (1.55)	5.80 (1.59)	6.17 (1.23)
Adverse sexual beliefs	2.72 (1.26)	3.15 (1.44)	2.58 (1.25)	3.00 (1.48)	2.27 (1.48)	2.78 (1.57)
Attitude dating violence	2.20 (.91)	2.35 (1.02)	2.11 (.88)	2.17 (1.02)	1.86 (.91)	2.20 (1.27)
Rape attitude	1.81 (.97)	2.20 (1.39)	1.79 (.99)	2.10 (1.32)	1.75 (1.02)	1.85 (1.21)
Self-efficacy self-control	5.35 (1.64)	5.25 (1.59)	5.65 (1.48)	5.37 (1.78)	5.20 (1.63)	5.39 (1.71)
Self-efficacy communication	5.46 (1.37)	5.66 (1.20)	5.53 (1.31)	5.28 (1.44)	5.83 (1.35)	5.67 (1.06)
Self-efficacy peer pressure	5.09 (1.61))	5.37 (1.27)	5.57 (1.46)	5.39 (1.61)	5.89 (1.20)	5.69 (1.35)
Social norm	5.62 (1.19)	5.73 (1.18)	5.81 (1.14)	5.47 (1.26)	6.05 (1.14)	5.61 (1.33)
Intention	5.74 (1.32)	5.46 (1.51)	6.10 (1.08)	5.83 (1.49)	5.80 (1.41)	5.79 (1.57)
Outcome expectancies	4.42 (1.45)	4.09 (1.58)	5.29 (1.36)	4.63 (1.66)	5.11 (.32)	5.08 (1.51)
Self-esteem	5.69 (1.03)	5.80 (.95)	5.48 (1.08)	5.64 (1.15)	5.64 (.85)	5.68 (1.01)

Table 3: *p*-values and effect sizes program effects

Outcome	p-values	effect sizes
	Condition*t1 Condition*t2	
Attitude communication	Condition*t1 p = .90	.03
	Condition*t2 p = .92	-.01
Attitude self-control	Condition*t1 p = .36	.22
	Condition*t2 p = .48	.22
Attitude boundaries	Condition*t1 p = .39	-.15
	Condition*t2 p = .28	-.27
Empathy	Condition*t1 p = .36	.19
	Condition*t2 p = .13	-.44
Adverse sexual beliefs	Condition*t1 p = .66	-.06
	Condition*t2 p = .63	.10
Attitude dating violence	Condition*t1 p = .97	-.00
	Condition*t2 p = .74	-.08
Rape attitude	Condition*t1 p = .64	-.06
	Condition*t2 p = .22	.32
Self-efficacy self-control	Condition*t1 p = .34	.10
	Condition*t2 p = .46	-.20
Self-efficacy communication	Condition*t1 p = .06	.31
	Condition*t2 p = .35	.23
Self-efficacy peer pressure	Condition*t1 p = .04*	.35
	Condition*t2 p = .26	.33
Social norm	Condition*t1 p = .03*	.38
	Condition*t2 p = .03*	.24
Intention	Condition*t1 p = .54	.11
	Condition*t2 p = .63	-.12
Outcome expectancies	Condition*t1 p = .08	.25
	Condition*t2 p = .51	-.16
Self-esteem	Condition*t1 p = .91	.03
	Condition*t2 p = .48	.14

*.005 ≤ p ≤ .05

To explore the possibility that program effects were dependent on whether participants had sexual experience (0 = no, 1 = yes), we reran the analyses with sexual experience and its two- and three-way interactions with condition and time in the model. When taking into account the critical values of the chi-squares, as in Johnson (1988), no statistical support was found for 3-way interactions. The outcome measure rape attitude ($p = .004$) was the only measure that showed a significant 3-way interaction effect with sex experience. However, its chi-square was below the critical value (chi-square = 9.06; $df=2$; $p=.004$). We therefore rejected the model and thus concluded that sexual experience did not influence program effects. We

also looked into age (recoded as a dichotomous variable, 0= under 15; 1 = 15 and older) as a potential effect moderator. No significant tests of the condition*time*age effects emerged (all chi-squares ≤ 5.39 , $df = 2$; all p 's $> .022$).

Discussion

The present study describes the effect evaluation of a Dutch sexual harassment prevention program, Make a Move, which addresses determinants of sexual harassment among adolescent boys in residential youth care. The program has a cognitive-behavioral approach and includes eight group meetings guided by a trainer. Overall, results showed that after program delivery, boys in the intervention group did not differ significantly from boys in the control group, neither at direct post-test nor follow-up, on any of the 14 outcome measures.

As compared to the adolescent population in general, it seems that the boys in youth care did, on the whole, show higher rates of sexual harassment behaviors. For example, 40.9% reported having tried to persuade a partner into having sex, and 23.8% reported having taken advantage of a person's drunkenness in order to get sex, as compared to 33% and 16.5% respectively in the report of Kuyper and colleagues (2010). These rates illustrate that boys in youth care are indeed a target group specifically at-risk of being or becoming perpetrators of sexual harassment and are thus very much in need of a prevention intervention.

The main conclusion of the effect evaluation is that, even though effect sizes were small to medium, no significant changes were observed in the intervention group. It often is difficult to prove the effectiveness of interventions, even if they have a solid grounding in theory and evidence. Many different factors can play a role in this, and often there is more than one explanation for the lack of results found. Such explanations can generally be linked to the intervention itself, to its implementation, or to the evaluation.

First, the intervention itself might not be well designed, and may not address the exact needs of the target population. In the case of Make a Move, it is difficult to address the exact needs of the target population, as it is a population with a diverse range of problems and negative experiences that often result in emotional and behavioral disturbances (Connor et al., 2004; Jonson-Reid et al., 2007; Veerman & van Yperen, 2007). An extensive range of risk factors, in combination with a complex target population, made it difficult to identify which determinants may contribute most to preventing sexual harassment. In relation to juvenile sexual offenders, Seto and Lalumière (2010) did not find significant differences in terms of

social skills or social competencies between adolescent sex offenders and adolescent non-sex offenders. They argue that in order to understand why sexual offenses are committed, and to prevent re-offending, it is important to discuss sexuality, including sexual fantasies, sexual arousal, and early exposure to sex and pornography. However, according to Seto and Lalumière (2010), interventions often focus on social skills and social competency, which would explain the lack of effects. Make a Move can be positioned somewhat in the middle. It includes communication skills training, but also explicitly addresses aspects of sexuality. Furthermore, discussions about sexuality and sexuality-related attitudes and norms are an important part of the program. At the same time, it is difficult to assess whether the program adequately meets the needs of this particular target group, with its complex and diverse problems and experiences.

Meta-analyses on the effectiveness of youth care institutions in contributing to a positive development of youth with serious behavioral and/or emotional problems also show mixed results (De Swart et al., 2012; Knorth, Harder, Zandberg, & Kendrick, 2008; Souverein, Van der Helm, & Stams, 2013). De Swart et al's review (2012) on the effectiveness of institutional youth care confirms the findings of Seto and Lalumière (2010), who showed that (social) skills training had no positive effects. A review by Knorth and colleagues (2008) suggests that programs which also involve the family show the most promising results in improving the behavioral development of youth in care. All in all, these findings show that more research is needed to establish what may work best in the context of youth care.

A second factor influencing program effectiveness relates to the level of intervention implementation. When a program is not implemented as intended, but rather with lower levels of completeness and fidelity, its potential will not be reached (Brownson, Glanz, Tabak, & Stamatakis, 2015). The implementation of Make a Move appeared to be very challenging and needed to take into account several factors such as the interaction with and between the boys, time constraints, and the fact that the resources and facilities allocated were not always suitable (see process evaluation in next chapter). Differences between the trainers (styles, skills, motivation) are also likely to have played a role in the implementation of the program. Closely linked to adequate implementation (dose delivered), is the dose received by the target group; this also turned out to be very low, in large part because groups were unstable, with many boys missing meetings. Furthermore, issues such as the problematic backgrounds and social power struggles among some boys made it difficult for them to open up and actively participate in the group meetings. This explanation seems to be in line with that of Veerman and VanYperen (2007), who propose that the difficulties of performing RCTs in youth care practices can contribute to the lack of carefully controlled effect evaluations.

Finally, it is worth noting that a program might be effective, but the evaluation of it may not have been optimal. Sufficient participants are needed when using a randomized controlled trial, and the measures included should match the objectives targeted by the intervention (Bartholomew Eldredge et al., 2016). Although the design of our study was based on an RCT and included a baseline measure and two post-tests (direct post-treatment and six-month follow-up), the number of participants that completed all of the questionnaires was rather low. Moreover, it is difficult to estimate how seriously the boys may have answered the questionnaires. It is possible that they might have answered some questions while anticipating what would be socially desirable. Furthermore, trainers reported back that many boys had problems concentrating long enough to answer the full questionnaire, and that some also had difficulties in fully comprehending the questions. Furthermore, the drop-out rate in the intervention group was higher than in the (waiting list) control group, which could have biased our results.

Conclusions and implications

There is a considerable lack of systematically described and evaluated programs addressing sexual harassment which focus on at-risk youth. This paper is among the first to describe an effect evaluation of a program targeting sexual harassment among the specific at-risk group of adolescent boys in youth care. Even though no significant intervention effects were found, the present study contributes to the literature by providing a systematic overview of the intervention content and an in-depth example of how such a program can be evaluated in the complex setting of youth care.

Chapter 3

Make a Move: process evaluation of the implementation of a sexual harassment prevention program in Dutch residential youth care

Chapter in preparation for resubmission as:

Van Lieshout, S., Mevissen, F.E.F., Jonker, M., Ruiter, R.A.C.. Make a Move: process evaluation of the implementation of a sexual harassment prevention program in Dutch residential youth care (revision).

Abstract

Background: It is not self-evident that prevention programs are implemented to the full extent (completeness) and according to its guidelines (fidelity). Incorrect implementation can greatly influence a programs' effectiveness. Exploring the factors that influence implementation can help improving future program implementation and additionally help in explaining a lack of program effects. This paper describes the process evaluation of the implementation of Make a Move (MaM), a sexual harassment prevention program developed for boys in Dutch residential youth care. **Methods:** MaM was planned to be implemented in 10 residential youth care institutions including a total of 78 boys divided over 14 groups. Seven freelance trainers visited the institutions to deliver the eight meetings of the MaM intervention assisted by a co-trainer working at the institution. Trainers' diaries (kept during implementation) and semi-structured interviews with the trainers (after implementation) were used to collect data focusing on trainers' experiences with implementing MaM, their levels of implementation, and factors influencing completeness and fidelity of implementation. **Results:** Completeness levels ranged from 24 to 84% ($M = 56\%$) of the exercises. Fidelity levels were difficult to determine. Organizational problems with(in) the institutions, program characteristics, and the boys' individual problems and behaviors interacted in negatively influencing complete and correct implementation. **Conclusions:** The findings indicated that trainers experienced high time pressure in their attempt to fully implement the program while dealing with the boys' difficult behaviors and the complex, often unsupportive situation of the institutes. It is concluded that especially in more complicated settings such as residential youth care institutions, implementers should receive full organizational support and need a thorough understanding of the group dynamics of the audience to optimize program implementation.

Introduction

Effectiveness of health prevention programs does not only depend on the content, but also on its delivery. Programs that are not implemented completely (the full content) and with sufficient levels of fidelity (according to the program guidelines) will have reduced program effects (Bartholomew Eldredge et al., 2016; Durlak & DuPre, 2008; James et al., 2006; Kalafat et al., 2007). A literature review on implementation of health promotion programs however, showed that only few programs reached implementation levels of 80% or more (Durlak & DuPre, 2008). Also, there were differences up to 87% between lowest and highest implementation for the same intervention program. Durlak and DuPre (2008) further showed that perfect fidelity never occurred but that instead fidelity of some elements frequently co-occurred with adaptation of other parts. More recently, the evaluation of a Finnish anti-bullying program showed average implementation levels of 80% (Ahtola et al., 2013), and a Dutch healthy diet program showed slightly lower figures with an average completeness of 70% per lesson (Bessems et al., 2011). In addition, Dusenbury and colleagues (2013) reviewed the implementation of school-based drug abuse prevention programs and found that 56-84% of the implementers had made modifications to the program.

Implementation is clearly not self-evident and not an automatic process. Understanding the factors influencing implementation can help in improving the delivery of health programs and in turn enhance program effectiveness. Within the context of teacher-delivered sex education, Schutte and colleagues (Schutte et al., 2014) found that self-efficacy, being familiar with the program, having had official program training, and evaluating the program as practical and useful were all related to fidelity of implementation. Others (Wiefferink et al., 2005) stressed that outcome beliefs and perceived instrumentality play a major role. The review by Durlak and DuPre (2008) identified 23 contextual factors influencing program implementation, classified as community level factors, provider characteristics, features of the innovation, and organizational capacity relevant to the delivery system. In line with this, Shepherd and colleagues (2014) reviewed implementation of school-based STI prevention programs and categorized factors related to implementation including school culture, school organization, and teacher expertise.

This paper describes the process evaluation of the try-out implementation of a sexual harassment prevention program called Make a Move (MaM). MaM is a group intervention targeting boys in residential youth care. Most of these boys have psychosocial and behavioral problems and are growing, or have been growing up in an unstable social environment, making them more at risk of becoming a perpetrator or victim of sexual harassment (Abbey et al.,

2001; Barbaree & Marshall, 2006; Burton et al., 2002; P. de Bruijn et al., 2006; Salter et al., 2003; Wekerle et al., 2009; Wolfe et al., 2004). They need specific care for a healthy sexual development. MaM was therefore developed by Rutgers, the Dutch sexuality expertise centre (www.rutgers.nl) and implemented by free-lance trainers. A randomized controlled trial including a pre and a post-test evaluated the programs' effectiveness in positively changing determinants of sexual harassment, but no significant effects were found (Van Lieshout, Mevissen, van Breukelen, Jonker, & Ruiter, 2016). The process evaluation described in this paper assessed (factors influencing) completeness and fidelity of program implementation and the experiences of the implementers with MaM. As not only the extent of dose delivered (i.e., level of implementation) is important for program effectiveness, but also the dose received (i.e. to which extent was the target group present and actively involved), we also aimed to collect information on the interventions' target group engagement and attendance.

Methods

Participants

The seven trainers (one male, six female; aged between 51 and 58, mean age $M = 54.6$) who implemented MaM all participated in the process evaluation. Two worked in a pair and were evaluated simultaneously. Trainers' expertise ranged from working as a sexologist, a behavioral scientist, a mental health youth therapist, a sexuality coach, to working for youth care as a family guardian. They worked freelance for the program developer Rutgers. All had experience working with boys in youth care, in groups or on an individual basis, and most were used to discussing sexuality.

Make a Move

Make a Move is a group intervention for boys aged 12 to 17 years in residential youth care including eight 90 minute meetings. The program focusses on promoting (aspects of) sexual health with the aim of preventing sexual harassment. The themes of the eight meetings are, in respective order, *manhood*, *friends and image*, *girls*, *sex*, *flirting*, *dating*, *pleasurable sex*, and *the future*. Each meeting includes several exercises such as role plays, discussions and watching short movie clips. The program has a positive approach and a competitive element to challenge boys and motivate them to participate actively. The boys can earn points during each meeting for carrying out exercises and various positive behaviors. For an overview of the content of the program, see the effect evaluation in the previous chapter.

Procedure

All 49 youth care institutions in the Netherlands delivering care to adolescent boys were contacted and invited to participate in the try-out implementation and evaluation of MaM. Twenty (41%) agreed to participate, of which ten were randomly assigned to the intervention group and the other 10 to a waiting-list control group. The program MaM was implemented by trainers in fourteen groups from the ten youth care institutions in the intervention group. Each trainer implemented the program in one to four groups, some with more than one group in an institute. In all groups, the meetings were consistently led by the same trainer. The trainers were assisted by a co-trainer who worked in the institution, usually as a group supervisor or as a therapist. Thirteen out of fourteen intervention groups finished MaM; one trainer quit one group after the second meeting as the boys severely

lacked motivation which resulted in impossible conditions to further deliver the program. In the end, 13 groups divided over 9 institutes finished the program and were included in the evaluation. See Table 1 for an overview of how trainers were distributed over the institutions.

Table 1: Trainers and number of groups divided over institutions

Trainer number	Institution									Total number of groups per trainer
	1	2	3	4	5	6	7	8	9	
1	1			1			1	1		4
2		2	1							3
3					3					3
4				1						1
5 and 6						1				1
7									1	1
Total number of groups per institution	1	2	1	2	3	1	1	1	1	13

Before implementing the program, the trainers attended a train-the-trainer workshop in which they were introduced to the content of MaM and informed about the effect and process evaluation of the program. The trainers received the program manual, a diary in which to log if and how they performed and evaluated each MaM exercise, and an attendance list to note the boys' presence during each of the group meetings. Aim of the latter was to investigate the intervention dose received at the individual level.

While trainers attended the train-the-trainer workshop, the institutions were instructed to appoint a contact person from within their organization to assist the trainer. Tasks requested from this contact person were to form boys into groups of six to eight boys, to arrange a meeting room, laptop, and digital projector, to ensure that during every MaM meeting an assistant from the institution was present to assist the trainer (co-trainer), and to organize a one-to-one introductory meeting with each of the boys who were recruited for participation. In some institutions, the contact person functioned as a co-trainer, in other institutions another colleague was appointed for this role. The co-trainer's role was twofold: the first task was to support the trainer by co-leading the MaM meetings and exercises, and second, to support the trainers finding their way in the organization and with the boys. At this point, co-trainers did not receive any training. One-on-one introductory meetings with the boys were included in order to inform them about and motivate them for participating in MaM. Trainers contacted the institutions to arrange when and where the program would be implemented.

Interview procedure and ethical approval

Shortly after program implementation for a group was finished, an interview with the trainer was scheduled. All trainers were individually interviewed, except for the pair that had implemented the program together. Trainers signed consent, agreed with the recording of the interviews and were informed that the data obtained from the interviews would anonymously be communicated with the program developers, reported, and possibly published in scientific journals. As some trainers implemented the program in more than one group, they were planned to be interviewed more than once to evaluate the different experiences for each group separately. However, in the end, the trainers regularly combined the evaluation of their experiences in the different groups during one interview. Eventually, nine interviews with all seven trainers were done. In these nine interviews all thirteen groups were discussed. The length of the interviews varied from thirty minutes to two and a half hours. Ethical approval was obtained from the Ethical Review Committee Psychology and Neuroscience (ERCPN) of Maastricht University.

Instruments: diary and interview

Trainers were provided with a printed diary to describe the general course of each meeting right after its implementation. Descriptions should include the number of boys that attended that meeting, a reflection upon the meeting in general (*‘Overall, how did this session go?’*), and on the specific implementation of each of the individual exercises. For each exercise, the two questions posed were *‘Did you execute the exercise as described in the manual, and if not, why?’* and *‘What was your experience with this activity?’* Trainers had filled out this diary for ten out of thirteen groups. The attendance list had only been used in three groups. Therefore, no calculations can be made about individual dose received.

The semi-structured interviews reflected on the trainers’ diaries and were guided by the interview protocol. Determinants of implementation derived from theories and empirical studies on the implementation of prevention programs shaped the interview protocol, among which the work of Wiefferink and colleagues (9). Trainers were asked about their experiences with MaM and the specific exercises, with a focus on completeness and fidelity, and implementation barriers. The diaries were used to stimulate the trainers’ memory and subsequently their narratives. This method is called data prompting and is increasingly being used in qualitative research (Kwasnicka, Dombrowski, White, & Sniehotta, 2015).

Data analysis

The interviews were recorded and transcribed verbatim. The transcriptions and diary notes were then exported to Nvivo 10. Although theoretical models were used to guide the themes of the semi-structured interview, we deliberately neglected these in the first steps of the coding to ensure a broad and open perspective (inductive). In this stage transcripts were examined line-by-line to identify factors that played a role in implementation (Thomas, 2006). Subsequently, through axial coding, the open codes were examined to understand the roles and relationships of these factors. In step three, selective coding, the theoretical models were used to assist in shaping the final codebook with branched subthemes (deductive). The first and second author independently coded the first two interviews, after which discrepancies were discussed. The first author then coded the remaining interviews and refined the codebook. The second author checked the final codebook, after which differences in interpretation were discussed until consensus was reached. Finally, based on theoretical models and the codebook, five main themes were identified, under which the results have been described.

Results

Implementation

Level of implementation completeness was measured quantitatively and matrices were created to record completeness, based on the diaries that trainers kept in ten out of the thirteen groups. Categories for implementation were fully implemented, partially implemented, or not implemented. As the program contained 25 exercises, each exercise accounted for 4% of the program. This quantification gave insight into completeness levels per group.

Most exercises were fully or partially implemented. On average, trainers implemented 54% of the exercises fully, 28% partially, and left out 16% of the exercises. Levels of implementation, however, varied between groups. Full implementation of the exercises as reported by the trainers varied from 24 to 84% ($M = 54\%$) of the exercises per group, whereas 12 to 52% ($M = 28\%$) of the exercises were reported as partially implemented. Trainers had skipped between 4 and 28% ($M = 16\%$) of the exercises per group. Some exercises were done in every group, and no exercises were skipped in all groups. See Table 2 for the full overview of implementation completeness. Fidelity levels could not be quantified because the diary reporting was not accurate enough and the theme was overshadowed by other topics during the interviews.

Table 2: Implementation completeness levels

Exercise	Group										Total per exercise
	1	2	3	4	5	6	7	8	9	10	
1.2	C	C	C	C	C	C	C	C	C	C	100% C
1.3	N	C	C	C	C	C	C	C	C	P	80% C 10% P 10% N
1.4	N	C	P	C	N	N	C	C	C	N	50% C 10% P 40% N
2.2	C	C	N	C	C	C	C	C	N	N	70% C 30% N
2.3	C	P	N	C	P	P	C	N	P	C	40% C 40% P 20% N
2.4	N	P	N	C	N	N	C	C	C	C	50% C 10% P 40% N
3.2	C	C	C	C	C	C	C	C	C	C	100% C
3.3	C	P	P	C	P	P	C	C	C	N	50% C 40% P 10% N
3.4	C	P	C	C	N	P	C	C	N	C	60% C 20% P 20% N
4.2	C	C	C	C	P	P	C	C	C	C	80% C 20% P
4.3	N	P	P	C	N	N	P	P	C	C	30% C 40% P 30% N
4.4	C	C	C	P	N	N	C	N	C	C	60% C 10% P 30% N
4.5	P	P	C	P	N	N	C	C	P	C	40% C 40% P 20% N

Exercise	Group										Total per exercise
	1	2	3	4	5	6	7	8	9	10	
5.2	P	C	C	C	P	P	C	P	P	P	40% C 60% P
5.3	P	P	P	C	N	P	C	P	N	P	20% C 60% P 20% N
5.4	P	P	C	C	P	P	C	P	C	P	40% C 60% P
5.5	C	N	N	N	P	P	C	C	C	C	50% C 20% P 30% N
6.2	P	P	P	C	P	P	N	N	N	C	20% C 50% P 30% N
6.3	P	P	P	C	P	P	P	C	C	P	30% C 70% P
7.2	C	C	C	C	P	P	C	C	C	C	80% C 20% P
7.3	N	C	N	C	C	C	N	C	C	N	60% C 40% N
7.4	P	P	N	C	C	C	P	P	N	C	40% C 40% P 20% N
8.2	C	C	C	C	P	P	N	C	N	N	50% C 20% P 30% N
8.3	C	P	C	P	P	P	N	C	N	-	30% C 40% P 20% N 10% unknown
8.4	C	P	C	C	P	P	C	C	C	C	70% C 30% P
Total per group	C 52% P 28% N 20%	C 44% P 52% N 4%	C 52% P 24% N 24%	C 84% P 12% N 4%	C 24% P 48% N 28%	C 24% P 56% N 20%	C 72% P 12% N 16%	C 68% P 20% N 12%	C 60% P 12% N 28%	C 56% P 20% N 20% unknown 4%	Average: Completely 54% Partially 28%; Not implemented 18%

*C = completely implemented; P = partially implemented; N = not implemented

** Only in 10 out of the 13 groups, trainers kept a diary, which this data was retrieved from

In general, the trainers stated they had had the intention to adhere to the program as much as possible, knowing they were partaking in the pilot implementation and its effect evaluation. However, all trainers had experienced time pressure in their attempt to fulfill this task. In the end, all trainers (though some more than others) had skipped certain exercises, adjusted them, left out certain components of exercises with multiple elements, or combined exercises. In one institution, the trainer stated that she had to change a lot in the program for two out of the three groups, because of the boys' severe behavioral problems which influenced their attention levels as well as the group atmosphere (see also the result section on vulnerability, safety, and group atmosphere). Meetings in these groups were reduced to 45-60 minutes and organized in a much more structured way: Boys each got turns in which they could respond and share their thoughts and no dialogues, role plays or discussions between boys were allowed.

All trainers experienced time pressure during program delivery, of whom two explicitly pointed out that at times this prevented them from elaborating on certain discussions and asking deepening questions as they had to rush on to the next exercise. One trainer explained that the time pressure made her feel insecure:

"The meeting is again too fully planned, no time scheduled for a break. This makes me feel insecure. Which exercises do I pick, how can I make sure to do everything? Show the movie clip, but shorten it?" (Female trainer, 58)

Besides time pressure, other factors that seemed to have affected completeness and fidelity of program implementation can be differentiated as characteristics related to 1) the program, 2) the institute, and 3) the boys and will be further outlined below.

Program characteristics

All trainers but one appreciated the program and, despite big and small challenges, enjoyed implementing it. The only negative trainer found it difficult to work with a program that she said did not suit her personal style of working, while at the same time having a group of boys who lacked motivation, and co-trainers who critiqued that the program by saying some exercises were childish, and her implementation as they found she went over some exercises too quickly. All others positively reflected on their experience with MaM. They reported that the program had a logical structure and good exercises.

“But, so I really enjoy doing this with these boys. Also with the hyperactive boys, when they’re hyperactive. It is, the nice part is the conversations. What you can get out of the conversations and the questions that are asked then. And yes, there are good elements in the program.” (Female trainer, 52)

Trainers were unanimously positive in their judgment of some exercises. All trainers, e.g., reported that formulating group rules was a good starting exercise. In most groups this went smoothly as *‘these boys are used to rules’*. Regarding the videos, trainers reported that most boys were very eager to watch the videos, although they also critiqued the looks or acting skills of the actors. The trainers themselves, in turn, commented on dysfunctional DVD’s, as some DVDs did not have sound.

There were, however, also differences between how trainers evaluated exercises. Whereas in some groups, for example, the role plays were a big success, in other groups this did not work at all, or only when the trainer played a role too. Similar contradicting responses were found for exercises that required writing or discussions: they went very well in some groups while causing resistance in others.

“Ohh and this is so much fun!! Also in the other institution. This is great, it’s really great! And they always want to play the role of the girl and ask me whether they can use my handbag. This is really so much fun.” (Female trainer, 58)

“The role plays, that is really too much to ask from these boys. They can only do it if I play a role too, and they can respond to me” (Male trainer, 57)

Similarly the trainers’ experiences with the ‘competition-element’ in the program were found to be mixed. Whereas in some groups this enhanced motivation and active participation, in other groups it was regarded as *not cool* to do well or, worse, the competition stimulated an already existing competitive atmosphere which the trainers regarded as undesirable. Trainers indicated that the scoring for the competition was not always clearly described in the trainer manual. Two trainers stated that this led to confusion and made them also less enthusiastic to stimulate the boys to get over their initial resistance against the competition.

According to the trainers, the manual did not always give enough guidance and for some exercises there was too much text, which made it difficult to have a quick glance while performing an exercise. Furthermore, trainers pointed out that many exercises required more time than indicated in the manual. Finally, the program was heterosexually oriented only.

Trainers had to make adjustments to the program to make exercises also suitable for boys who were openly gay, for instance by inviting them to describe their ideal boyfriend instead of girlfriend.

Trainers' attitude towards certain exercises played a large role in completeness and fidelity of implementation.

"Yes, I did this exercise. It's a good one, it's fun and educating. I did, however, not use the pictures of the emoticons (...) because of the lay-out. These are big cards and I'm not working with preschoolers." (Female trainer, 52)

The trainers' attitude towards program components also seemed to be related to the boys' engagement and liking of these components. Among trainers who implemented the program in more than one group, previous experience with exercises, both related to the boys' responses and trainers' attitudes towards the exercises, also affected implementation.

"But the boys in the current group really liked this exercise, so I will also do it in that third group" (Female trainer, 58)

Role of the institution (incl. organization, communication, co-trainer)

Being visitors to the institutions, all trainers relied fully on the cooperation of the organization and the support of the staff. In most institutions, trainers positively evaluated the collaboration with the contact persons and co-trainers, and the way the meetings were scheduled and organized. In multiple groups there had been some form of miscommunication about meetings or disagreement on how to enforce the rules between the visiting trainers and institution staff, but this was often quickly solved.

"The communication with the contact person went very well. There was one misunderstanding as she thought I needed the DVD player and digital projector for the first meeting only. But besides that, all praise and good communication (...)" (Female trainer, 58)

In a few groups, however, it seemed that the coordination and rapport between trainer and institution (contact person and/or co-trainer) as well as the organization within the institution was very chaotic.

“The initial contact person handed it over to a colleague, and it seemed as if she couldn’t get it together. That’s the impression I got. And because of that it took quite a while before I could start (...) No, she wasn’t the co-trainer. That [allocating of co-trainer] really was a total mess. The way it was arranged was absolutely wretched” (Female trainer, 52)

The trainers with such negative experiences reported on, for example, appointed to different locations for each meeting or to rooms that were inappropriate for the MaM meetings (e.g., too small room or with insufficient seating), lack of being notified in advance regarding drop in or drop out of boys or even the sudden absence of an entire group because of holidays, being confronted with certain institutional rules that conflicted with the execution of the program or being confronted by the co-trainer with incidents of sexual harassment that had taken place previously and which the trainer was supposed to solve on the spot during the MaM meeting.

Despite the efforts of the co-trainers, they could not always assist the trainer in overcoming these problems as it often seemed to be beyond their control as well. As one of the trainers stated related to the boys’ attendance or the boys coming in late:

“And no offense to the group leaders [trainer refers to co-trainer here] because they really tried. One of them even sent around an email one day in advance, to all the group homes. Something like: Don’t forget, and please sent them off in time.’ Well, that also didn’t work.” (Female trainer, 58)

A wide range of reasons caused drop out of and/or missing sessions, such as being too aggressive or overstimulated, transfer to another institution, the program interfering with school or internship, having ran away, serving penalties (isolation), attending court, personal or family circumstances, holidays, camps, etc.

Despite their dependence on the collaboration of the institution in implementing the program, most trainers took control and shared their disagreement whenever possible. For example, one trainer overruled the institution’s policy that boys could only go for a cigarette break one by one: *“I simply said that that’s impossible”*. The trainer duo turned around the lack of control regarding what time the boys came in by integrating coming on time as a program element that was rewarded.

“But, if such a cumbersome institution doesn’t work along, in the sense that the boys come in 20 minutes late, what do you do? Are you going to start, or not? Do you make it a program element? That’s what we did. We decided: we make this a program component and the boys will be rewarded if they come on time.” (Female trainer, 53)

In nine out of the thirteen groups, the same co-trainer attended all sessions of the group. In the other four groups co-trainers varied as work schedules and other factors did not allow having the same co-trainer throughout the eight meetings. Having a different co-trainer each meeting bothered the trainers, primarily because the different co-trainers could not go along in the structure and build-up of the group dynamics and some co-trainers even had attitudes about girls and sexuality that contradicted the message of MaM, according to the trainers.

“Well, in general I liked the way they handled the boys, but it really fluctuated and every week I just had to wait and see who would be there. And it’s really a shame that they therefore don’t go along in the structure and the build-up of the training. You’re building something up with the group and then suddenly there is this new group leader who says things like: well, but the way girls behave nowadays...’ I find that very regrettable.” (Female trainer, 58)

However, some trainers had more positive experiences and mentioned how the co-trainers became more comfortable themselves talking about sexuality. Trainers appreciated co-trainers who had a positive attitude towards the program and discussing sexuality and had a natural authority over the groups.

“But now at the other institution, the one who joins our sessions, he already easily talks with the boys about relationships, girls and sex. And that makes such a difference!” (Female trainer, 58)

Although the program used the term co-trainers, most trainers did not perceive the co-trainers as such. Rather they referred to them as ‘a group leader or other employee who is present who know the institute and the boys and can correct them on their behavior’.

“I wouldn’t say ‘co-trainer’, they were present. And that’s how they regarded it themselves too.” (Male trainer, 57)

Role of the Boys

Group composition

The variations in group composition seemed to have been an important factor influencing the group atmosphere which in turn influenced program implementation. The group compositions not only varied in age, but also in boys being or not being acquainted with each other, and in the number of boys being present. The boys' age ranged from twelve to eighteen, which some trainers pointed out was a too wide range. Overall, the younger boys appeared more motivated, while older boys showed more resistance and found certain exercises childish, which the trainers agreed with to some extent. Trainers said that an age range from thirteen to sixteen would be most suitable.

In some groups, boys were already acquainted before MaM started. This could be positive (no time needed to get used to each other and thus more quickly feeling comfortable enough to speak out) but also negative (a history of conflict or hierarchal structures could be brought into the group meetings). Groups in which boys did not know each other beforehand could start fresh and would not bring previous tensions or conflicts into the room but required time to feel comfortable with each other.

Due to high levels of missed meetings, drop out and drop in, the group composition as well as the group size was hardly ever the same during the course of the eight meetings. This fluctuation in group composition and group size seemed to have greatly influenced the atmosphere in each meeting and in turn the implementation of exercises. At the first meeting, group sizes varied from four to nine. Throughout the subsequent meetings, number of boys present varied from one to thirteen. Trainers agreed that the ideal group size would be six to eight. Groups of three were regarded too small for implementing certain exercises (e.g., role plays, team exercises), while eleven was said to be too large to handle. These different compositions influenced whether certain exercises could be performed and caused some turmoil as it affected the atmosphere in the group.

Vulnerability, safety, and group atmosphere

Group atmosphere varied, both between groups and between meetings within groups.

“One group is very different from another group, and then within groups, the atmosphere also changes over time. One day you think: this is such a great group, we’re doing great. And then when you come back three weeks later, things have happened and the atmosphere has changes 180 degrees.” (Male trainer, 57)

These variations were partly due to the instable group composition and partly because of the boys’ behavioral problems, the later sometimes resulting in boys fighting with each other. Trainers pointed out the importance of creating a safe environment during the MaM meetings in which boys were comfortable enough to open up and share their thoughts and experiences. As one trainer formulated it, this was opposite to the atmosphere the boys are used to in the residential groups, because *“in the residential groups, you hide your vulnerabilities.”* (Male trainer, 57). According to the trainers, in some groups getting to know and trust each other took more time than in others, but most boys eventually felt comfortable to share their thoughts, experiences, and the questions they had.

“The boys in this group have serious behavioral problems and were all in care as a forced measure. They didn’t know each other, so they were also a bit afraid of each other. We needed to invest in a safe atmosphere (...) I found the boys quite motivated and enthusiastic. And later on they would also tell us if they had a bad day. Well, that’s very good. And that’s an indicator that the group is safe, otherwise they wouldn’t say that” (Female trainer, 53)

In some groups the boys found it difficult to listen to each other, whereas in other groups boys had been very open and shared a lot. Some trainers reported how they had complimented the boys, saying they had also dared to share their opinions when these differed from the rest of the groups. In other groups, boys were more hesitant to deviate from the group norm, to act in a role play, etcetera. In multiple groups, boys had reported back to the trainers that they enjoyed working with the program, had learned numerous things, and were disappointed when it finished after eight meetings.

Discussion

In this paper we reported on the implementation of Make a Move, a sexual harassment prevention program in residential youth care. The study examined how and to what extent the program's exercises were performed and how the trainers reflected on working with the program. Findings showed that on average, per group, 54% of the exercises had been fully implemented, 28% partially and 16% had not been implemented. However, implementation varied widely, with reported full implementation of exercises ranging from 24 to 84%. The determinants that negatively affected the implementation of MaM can be categorized as factors related to the youth care organization, the program itself, to the characteristics and behavior of the target population, and how the trainer related to and acted upon these factors. Together, these determinants caused time pressure or otherwise hindered completeness and fidelity of implementation and are likely to, at least partially, explain the non-findings of the effect evaluation of Make a Move. In this effect evaluation, the lack of outcomes was discussed in a three factor model, addressing the program itself, its implementation, and its evaluation. This process evaluation gave a more in-depth insight in the challenges trainers faced while implementing MaM. Diary recording and individual interviews showed that the MaM trainers had varying experiences with organizational support, including interaction with the contact persons, co-trainers, as well as general awareness and support from other institution employees. Support for implementation and the allocation of the required resources is key in successful intervention implementation, but this was not always present. Dusenbury and colleagues (2003) state that reinvention or adaptation of a program may be necessary, not only to tailor programs to specific needs in a particular setting, but also to promote ownership and institutionalization of a program. Ownership and dedication need to be considered from the early stages of program development and generating ownership may require additional interventions (Forman et al., 2009). One explanation for the lack of ownership in some institutions may be that it was regarded as a top-down imposed program. Institutions had been actively approached and invited to participate in the RCT, rather than implementation being in response to a commonly identified need. Although individuals within the organizations identified a need for the program and supported implementation, there was no institutionalized ownership. Second, external trainers implementing the program rather than the institutes themselves might have contributed to this lack of ownership and support. Finally, MaM was implemented during a time of transition of the Dutch youth care system, in which employees lost their job or were assigned different roles and positions. Because of the insecurities that this brought along, employees might have had other priorities than the implementation of the program.

Previous studies have already shown the difficulties of program implementation and that this is not always done to the full extent, which affects program effectiveness (Durlak & DuPre, 2008; James et al., 2006; Kalafat et al., 2007). This challenge has also been identified in youth care settings (Aarons & Palinkas, 2007; Daamen, 2015; Stals, 2012; Veerman & van Yperen, 2007), particularly due to the specific needs and characteristics of the target population. A well-known challenge is the tension between implementing an intervention as intended (fidelity), and delivering needs-based care (flexibility). This dilemma is particularly relevant in group interventions. Similar to the findings of this study, Aarons and Palinkas (2007) acknowledge that working with evidence-based programs in child welfare is often challenged by unforeseen events, conflicts, and solutions.

Youth can be in care because of problems in the family household and/or because they show a range of externalizing (behavioral) or internalizing (emotional) problems. These problems, but also the dynamics among the participants, can affect their acceptability and engagement of the intervention. Implementation researchers have pointed out that high levels of student misconduct or conflicts (both peer conflict and student – teacher conflict) may result in teachers spending more time on class management rather than program implementation (Domitrovich et al., 2008). Engagement, dependent on class climate factors such as trust among students, willingness to help each other, and how well students got along with each other, has previously been found to affect the effectiveness of a school-based program (Low, Van Ryzin, Brown, Smith, & Haggerty, 2014). Indeed, Veerman and van Yperen (2007) argued that the constant need to adapt to and meet the needs of certain participants makes interventions in youth care settings unsuitable for rigorous empirical evaluation such as RCTs. We believe, however, that it is always important to perform a rigorous evaluation before rolling out implementation on a wider scale, and with the combination of the effect and process evaluation we derived some very valuable insights that need to be taken into account before further steps are taken.

Limitations

This study had some limitations. Of the thirteen groups that finished the program, trainers kept a diary in ten groups and an attendance list in three. This means there was a 23% data loss in the diaries, which made it impossible to draw any conclusions about the boys' individual dosage of the program received. Furthermore, there might be a social desirability bias in the self-reported completeness levels, causing over-reporting of implementation. This has also been found in other studies (Lillehoj, Griffin, & Spoth, 2004). Finally, trainers with multiple groups did not always make clear distinctions between these groups in the interview, or

would not be able to recall which memories related to which group. This made some of the results difficult to interpret.

Implications and conclusions

The outcomes underline the importance of process evaluations of (health) intervention programs. Besides organizational constraints and characteristics of the intervention and the implementers, the target population also affected program implementation. The boys' engagement played a large role in program implementation. Future implementation of MaM requires embedding in the institution and extensive support for the implementers as well as the promotion of program acceptance among the target population.

Chapter 4

Long Live Love+: Effect evaluation of a Dutch online school-based sex education program for secondary education

Chapter in preparation for submission as:

Van Lieshout, S., Mevissen, F.E.F., van Duin, G. & Kok, G. Long Live Love+: Effect evaluation of a Dutch online school-based sex education program for secondary education.

Abstract

Introduction: Sex education plays an important role in the sexual development and the sexual health and well-being of adolescents. In the Netherlands, Long Live Love is the most widely used school-based sexual education program for young adolescents. Long Live Love+ (LLL+) is its spin-off for older adolescents aged 15+ in higher levels of secondary education. The program includes online and offline components and is freely available via internet. The topics are relationships and wishes and boundaries, pregnancy prevention, STI preventions, and sexual diversity. This paper describes the effect evaluation of LLL+.

Methods: A RCT was performed with a pre-test and direct post-test including 17 secondary schools (nine in the intervention group and eight in the waiting-list control group) from different locations in the Netherlands with $N = 999$ students at baseline. Psycho-social outcome measures included knowledge, attitudes, risk-perception, self-efficacy, norms and intentions regarding consensual safe sex practices and sexual diversity. Behavioral outcomes included carrying condoms, unsafe sex, STI-testing and sexuality-related communication with sex partners.

Results: Multilevel analysis showed a significant time*condition interaction effect for STI-related risk perception ($p < .000$; $d = -.29$) and attribution of homosexuality ($p < .000$; $d = .25$): students in the intervention condition scored higher on STI risk perception and more strongly believe homosexuality is nature rather than nurture or a personal choice, after the LLL+ lessons. There were no effects on any other outcomes (p 's $> .03$).

Conclusion: significant change was only found for 2 out of the 31 psycho-social outcome measures and none on any behavioral outcomes. The limited results are discussed in light of the program and its development process, program implementation, and evaluation.

Introduction

Sexual health is an important element of general health and well-being (Douglas & Fenton, 2013). According to the World Health Organization (WHO) sexual health requires a positive and respectful approach to sexuality and sexual relationships, as well as the possibility of having pleasurable and safe sexual experiences, free of coercion, discrimination and violence.” (World Health Organization, 2006). One way of achieving sexual health and well-being is through comprehensive sex education (CSE) as several reviews have shown (Braeken & Cardinal, 2008; Kirby, 2008; Kirby et al., 2007; UNFPA, 2015). Especially in a school context CSE can be an effective way to reach multiple adolescents simultaneously. The Netherlands has a long standing tradition of CSE, Worldwide, the Dutch approach for CSE is appraised and has been linked to the relative good sexual health of Dutch adolescents as reflected e.g., in the low teenage pregnancy and STI infection rates (Cook & Cameron, 2015; Ferguson et al., 2008; Lewis & Knijn, 2001; Part et al., 2013; van Teijlingen et al., 2015). One of such programs that is widely used and proven to be effective is the Long Live Love program (LLL) for younger adolescents (age 13-14) which was developed in the ‘90s and has since been updated several times (Schaalma et al., 1996; Schutte et al., 2014; Vanwesenbeeck et al., 2002).

In addition, students differ in their mental and learning capacities, as reflected in the Dutch school system (see EP Nuffic, 2015). This has implications for how to address students with different levels of intellectual abilities, using different methods and strategies in health promotion programs. We can thus conclude that, to enhance the effectiveness of sex education, additional programs are needed that are age-appropriate, adapted to the educational level of the target population, the sexual health situation of the specific target population, and preferably repeated over time. To serve this need among adolescents with a higher level of education, the original Long Live Love program for young adolescents (LLL) was expanded with a version for older adolescents: LLL+. This paper reports on the effect evaluation of the LLL+ program.

Long Live Love+

The LLL+ curriculum targets the upper classes of the higher levels in secondary schools: HAVO (*‘higher general continued education’ - 12 to 17 yrs*) and VWO (*‘preparatory scholarly education’ or ‘pre-university training’ – 12 to 18 yrs*). LLL+ is developed following Intervention Mapping (IM); a protocol for theory- and evidence-based intervention development (Bartholomew Eldredge et al., 2016). The program covers four themes

(relationships; (un)safe sex and contraception; (un)safe sex and STI; and sexual diversity) with two 45-minute lessons each. Exercises are interactive and include discussions, (online) quizzes, narratives, and videos. All components and materials are freely available via the program website (www.langlevedeliefde.nl), with separate environments for teachers and students. The exercises can be performed both online and offline, and materials and teacher manuals can be downloaded from the website. A linkage group including teachers was consulted throughout the development (see Mevissen, Van Empelen, Watzeels, Van Duin, Meijer, Van Lieshout, and Kok (under review) and Mevissen, Kok, Watzeels, Van Duin, & Bos (under review), for more details on the development of LLL+). Table 1 gives an overview of the content of the LLL+ . This paper reports on the Randomized Controlled Trial (RCT) performed to evaluate the effectiveness of the LLL+ program.

Table 1. Program overview Long Live Love+

Theme	Exercise	Summary
1 Relationships	1. introduction to relationships	Using sticky notes, students anonymously write down three aspects they like or appreciate in (sexual) relationships. Teacher collects and sticks them on the blackboard, after which 2 statements (provided in manual) will be discussed. The teacher will guide the (group) discussion. In the majority of the cases it will show that boys and girls are not that different from each other. They will conclude that both 'safe' as well as 'unsafe' information can be found on internet. The teacher will provide the students with a list of trustworthy websites.
	2. Desires	An (online) video shows by way of a cartoon-like graph how the sexual arousal curve differs for boys vs. girls but also between boys/girls. The cartoon alternates with a nurse further explaining arousal as well as factors influencing arousal.
	3. What do you want?	Students read two scenarios describing different 'crossing sexual borders' situations and discuss in small groups or in a class setting. Based on both stories students discuss items like 'how would both persons feel?', 'what goes wrong', 'how could the boy/girl have acted differently to avoid this' and 'what if the situation was opposite?'
	1 – 4. Reproductive body parts, menstrual cycle and conception	Short video animations show schematic representations of the male and the female sexual body parts, the different components, and their names and how they function. A third and fourth video animation explain the menstrual cycle and conception. Information includes hygienic advice, facts and myths regarding hymen, and stresses that shapes and looks of reproductive body parts can be very diverse.
2 (Un)safe sex and contraception	5. Safe vs. unsafe sex	Different sexual behaviors (kissing, touching, penetration etc.) are discussed in light of what is (un)safe for pregnancy and/or STI prevention.
	6. Contraception	Students (in pairs, groups or individually) collect information on one or more types of contraception focusing on how it works, but also on what the (dis)advantages of certain types are, how to deal with disadvantages, how to get and how to use the different types, and how to deal with potential emotions (e.g., embarrassment) involved in discussing and organizing contraception use. Finally, different types will be discussed and compared in a group session. The teacher will show a website where students (girls) can get tailored contraception advice.
	7. Hormonal contraception	Short video animation shows how hormonal contraception works in preventing pregnancy.
	1. Quiz	Online quiz on STIs and (un)safe sex including ten questions with multiple choice answer options.
3 (Un)safe sex and STI	2. STI's	Students (in pairs, groups or individually) collect information on one or more types of STI. They answer questions on what it is, potential symptoms, consequences, and how to cope with negative consequences. After all information is collected, the teacher will summarize the results with the class.
	3. Chlamydia	Short video animation shows how chlamydia infects the body and affects the reproductive system

Theme	Exercise	Summary
4 Sexual diversity	4. STI testing	Video shows a young couple discussing to stop condom use and to get an STI test (including standard excuses that are in the end tackled by the girl). The boy in the end agrees and both visit the STI clinic. The video shows the type of conversation they have with the nurses, the questions asked, and the tests done (urine test for boy, vaginal swap for girl, and blood-test for both). After the video, the students receive several questions for a more in-depth discussion of topics such as how would you deal with a resistant partner.
	5. What's your personal STI risk?	Online survey on personal sexual history. A safe sex advice will be provided tailored to the answers provided.
	1. Circle exercise	A circle is drawn on the blackboard. Students are stimulated by teacher to mention homosexual men and women they know within the different levels of their environment (family, school, media). The teacher makes clear that it is not the goal to 'expose the gay' in this exercise or give private information if not wanted on personal family situations. Teacher guides conclusion towards the wide diversity in homosexuals just like with heterosexuals.
	2. Sexual Diversity: terminology	The different terms that students may relate to sexual diversity will be explained. To round off, a video can be shown about a transgender girl. In the video, the girl explains her struggle with being in a female-body while feeling male and talks about the support she received from her family.
	3. Nature or nurture?	A brochure will be handed out to the students by the teacher at the end of the lesson. The brochure includes a summary of the discussed terms. In addition, it includes a list of websites where students can find more information on sexual diversity or homosexual support groups.
	4. Coming-out	Four statements (in teacher manual) will be discussed by the students. If the situation allows, the discussion can first take place in small groups before summing up with the whole class. Preferably mixing homo-negative and homo-positive students and mixed cultures so different opinions can be heard. Teacher will guide the discussion towards certain conclusions (nature) as instructed in the manual. Teacher explains what 'coming out' means. After that, four questions need to be discussed and answered by the students. Teacher will guide the discussion such that also solutions on how to deal with difficult coming out situations will be discussed. In the end the teacher will guide the conclusions that it is important for everybody to be able to be him/herself and that it is brave for young people to get out of the closet and this should be respected. A video (as warming up or wrap-up) can be shown in which a girl talks about her positive (at school) and negative (in the church) experience with coming out.
	5. Coming-out at school	Four questions/statements need to be discussed and answered by the students (in teacher manual). The questions/statements can first be answered/discussed in small groups before summing up with the whole class. Teacher will guide the discussion such that also solutions on how to deal with difficult situations will be discussed. Teacher will ask more questions to stimulate students to elaborate

Methods

Recruitment, randomization and study design

Teachers at secondary education likely to be involved in sex education (teachers of Biology, Social Sciences, mentoring teachers) throughout the Netherlands were contacted by email and phone and invited to participate in the trial implementation of LLL+. This was through formal and informal networks and lists of schools. Reasons not to participate were e.g., already being involved in other research projects, not having time, and not being interested in teaching this new method. In total, 33 teachers from 17 schools participated in the RCT, with a total of 46 classes including 999 students, see Table 2.

Table 2: participating schools

School	Condition	Number of teachers	Number of classes	Number of students (baseline)	City, Province
1	Intervention	2	3	68	Maastricht, Limburg
2	Control	1	1	19	Arnhem, Gelderland
3	Intervention	2	3	58	Amsterdam, Noord Holland
4	Control	2	3	54	Maastricht, Limburg
5	Control	1	1	18	Amsterdam, Noord Holland
6	Intervention	2	3	68	Roermond, Limburg
7	Intervention	2	3	67	Hoorn, Noord Holland
8	Intervention	1	3	64	Steenwijk, Overijssel
9	Intervention	4	5	132	Haren, Groningen
10	Intervention	1	1	32	Gulpen, Limburg
11	Intervention	1	1	5	Amsterdam, Noord Holland
12	Intervention	2	3	62	Horst, Limburg
13	Control	2	3	69	Geldrop, Noord-Brabant
14	Control	3	4	78	Echt, Limburg
15	Control	2	4	82	Breda, Noord Brabant
16	Control	2	2	30	Den Haag, Zuid Holland
17	Control	3	4	93	Mijdrecht, Utrecht
Total	17 schools	33 teachers	47 classes	999	

Although we aimed to include a more general sample of teachers (not only those teaching Biology), the final sample consisted of Biology teachers only. To prevent large baseline differences with regard to background characteristics between the intervention and control groups and an uneven distribution of number of classes and students, before randomization

schools were matched in pairs based on similarities in characteristics such as geographical setting, type of school, and numbers of teachers and classes participating. Of these pairs, one school was randomly assigned to either intervention group or waiting-list control group, and its matching school assigned to the other group. In the end, nine schools (25 classes, N = 556 students) were assigned to the intervention group and eight schools (22 classes, N = 443 students) to the waiting-list control group. Schools in the intervention condition would receive access to the LLL+ website and teaching materials, while schools in the control condition were requested to use their regular Biology teaching methods to cover the theme. All schools worked with either '*Nectar*' or '*Biologie voor jou*' as their regular teaching method.

The effectiveness of LLL+ was evaluated using a baseline measure (T0) and a direct post-test measure (T1). At baseline, 999 students (55.6% in the intervention group; 54.0% female) filled out the questionnaire (T0) and 830 students (58.1% in the intervention group) completed the direct post measure (T1). T0 and T1 surveys for each individual participant were matched using date of birth, and mother's initials, which was possible for 735 students (73.6% of baseline, 88.6% of post-test respondents). Inability to match T0 and T1 for some participants was partly because of drop-out, and partly because students did not consistently fill out matching criteria details. One school in the control group (including 3 classes with 93 participants) completely dropped out after the baseline measure due to personal circumstances of the teacher.

A follow up measure was planned, but this could not be administered through the schools. Therefore, in the post-test questionnaire, students were requested to fill out their email address on a separate piece of paper if they wanted to be contacted for a third (six-month online follow-up) measurement. These email addresses were then sent back to us together with the questionnaires. Because of the very low response (N=25 of out 830 post-test), this follow-up had been cancelled and the students who signed up were informed about this via email.

Procedure

After randomization, all schools (intervention and waiting-list control) received T0 questionnaires for the students by post approximately two weeks before LLL+ implementation was scheduled. The package further included a stamped return envelope, a consent form for the teacher, information and instructions for the teachers on the time frame of the LLL+ implementation (intervention condition only) and evaluation, and on how to safeguard students' privacy during administering of the surveys.

The questionnaire included information about the goal of the survey (i.e. we would like to hear your opinion on and experience with different sexuality-related issues), how privacy was safeguarded (i.e. setting desks apart, similar to when administering exams), that the students could withdraw at any time, as well as instructions on how to fill out the questionnaire. Students were asked to give consent on the introductory page of the questionnaire by ticking the box indicating that they understood and agreed with the terms of agreement. No informed consent from the parents was necessary as teachers in the Netherlands are free to choose which educational materials they use for sex education and LLL+ was used as a substitution of the regular program. Upon completion of T0, teachers in the intervention group could start implementing LLL+ while teachers in the control group were requested to run their regular Biology lessons on sexual reproduction. Teachers were given the freedom to start implementation according to their convenience within the boundaries of the semester, provided the curriculum was taught in a timeframe of approximately 8 weeks. A password was installed on the website and sent via email to all teachers in the intervention group, as to ensure that the waiting-list control group could not be exposed to the intervention until after the evaluation was completed. A reminder was sent after 2 weeks if schools did not return their T0 survey. Approximately 8 weeks after completing T0, all teachers received T1 by post according to the same procedures as described for T0. Ethical approval was received from Maastricht University's Ethical Review Committee Psychology and Neuroscience (ERCPM).

Participants

At baseline, mean age of the students was $M = 15.5$ ($SD = 0.8$, range 13 to 20 years). The majority of students were of Dutch origin (87.4%). A few participants had a (partly) Antillean or Surinam (2.5%), Moroccan (1.3%) or Turkish (1.0%) background, and the remaining students (7.9%) were (partly) from a diverse range of (non-)Western backgrounds. A third (332 students; 33.5%) had been raised religiously (mostly Catholic or Protestant) and among those, 44.0% rated religion as not important at all (41.6% somewhat important; 14.5% very important). Half of the students (512; 51.3%) had had a relationship in the past two years. Four hundred and five (40.5%) reported having had sexual experience without intercourse (i.e. kissing, touching, oral sex) of whom 174 (17.5% of the total sample) had had sexual intercourse. Among those experienced with sexual intercourse, 91 (52.3%) had carried a condom once or more often in the past three months while 83 (47.7%) had never done so. Almost all (98.3%) students with intercourse experience had used one or more (e.g., both condom and pill) forms of contraception the last time they had sexual intercourse. Among those with sexual experience, 19 students (10.9%) had ever undertaken

an STI test. See Table 3 for more details on demographics as well as distribution among the two study conditions.

Table 3: demographics and characteristics at baseline

Variable	Full sample		Intervention group		Control Group	
	M (SD)	%	M (SD)	%	M (SD)	%
Age	15.5 (0.8)		15.4 (.7)		15.5 (.8)	
Cultural background						
	Dutch	88.8		89.9		87.4
	Antillian/Surinam	1.9		1.8		2.0
	Moroccan	.9		.7		1.1
	Turkish	1.1		1.3		.9
	Other	7.0		5.9		8.4
Religious		33.5		33.8		33.1
	“not important”	44.0		45.2		42.5
	“somewhat important”	41.6		41.9		41.1
	“very important”	14.5		12.9		16.4
Relationship in past two years		51.3		49.3		53.7
Sexual experience						
	No	41.4		43.7		38.6
	Somewhat (kissing, touching)	58.0		55.6		60.9
	Intercourse	17.4		15.5		19.9
Intercourse						
	< 3 months ago	71.7		72.1		71.3
	> 3 months ago	28.3		27.9		28.7
Number of sex partners	2.5 (3.3)		2.8 (3.0)		2.2 (3.6)	
Condom use last time sex		55.7		48.8		62.5
STI test		1.9		2.0		1.9
STI		.3		.2		.5
Sexual orientation						
	Straight	93.2		94.6		91.4
	Mostly straight	1.0		.9		1.1
	Bisexual equally	.6		.4		.9
	Mostly same-sex	.3		-		.7
	Same-sex	2.0		1.3		2.9
	Don't know yet	2.3		2.2		2.5
	No answer	.6		.7		.5

Measures

The questionnaire included measures on background characteristics (age, ethnicity, religion, sexual orientation, relationship experience in the past two years) and outcome measures of (psycho-social determinants of) sexuality-related behavior. The construction of the outcome measures was based on the change- and performance objectives as formulated in the program goals (see Mevissen et al., and Mevissen et al., both under review) as well as guided by the Reasoned Action Approach (Fishbein & Ajzen, 2010). All psycho-social measures (except knowledge) used 5-point Likert scales. We calculated mean scores for scales per time point for a participant if more than half of the items that made up the measure had been answered. Higher scores indicate more of that (positive) attribute. The reliability scores refer to the baseline measure (T0).

Psycho-social determinants

Knowledge regarding sex and STI's was measured with nine items; e.g., 'A girl will always bleed the first time she has intercourse'. Correctly answered items were scored as 1, whereas a wrong answer and the option 'I don't know' were scored with a 0. The scale constituted of the sum of the number of correct answers, leading to a maximum score of 9.

STI risk perception was measured with three items ($\alpha = .65$); e.g., 'It is possible that one day I might get infected with an STI' (1 = totally disagree, 5 = totally agree). **Pregnancy risk perception** was measured with two items ($r = .44$); e.g., 'It would never happen to me that I become pregnant/make a girl pregnant unplanned (1 = totally agree, 5 = totally disagree).

Attitudes towards condom use e.g., 'To always use a condom myself when I have sex with someone, seems to me...' ($\alpha = .70$) and **attitudes towards communicating with a sex partner**, e.g., 'To discuss with a partner what we do and do not want during sex, seems to me...' ($\alpha = .75$) were both measured using one statement with three semantic differentials (very bad - very good; very unimportant - very important; very uncomfortable-very comfortable). **Attitude towards pressure methods to have sex** ($r = .53$) was measured with two items e.g., 'to become angry when my partner doesn't want to have sex, seems to me...' (very good – very bad).

Attitude towards contraception methods ($\alpha = .71$) was measured with one statement with three semantic differentials ("To use the contraceptive pill or other contraceptive methods to prevent pregnancy, I find..." (very bad -very good; very unimportant - very important; very uncomfortable- very comfortable).

Attitude towards homosexuality was measured with six items ($\alpha = .91$) e.g., “I think homosexual may work wherever they want to” (1 = totally disagree, 5 = totally agree), based on Larsen, Reed, & Hoffman (1980). **Attribution of homosexuality** included three items ($\alpha = .71$) e.g., “I think being homosexual is a personal choice” (1 = totally disagree, 5 = totally agree), based on Haider-Markel & Joslyn (2008). **Empathy towards homosexuals** was measured with three items ($\alpha = .70$) e.g., “I feel sympathy for homosexuals” (1 = totally disagree, 5 = totally agree), adapted from Batson et al., (1997).

Self-efficacy condom use ($\alpha = .78$) was measured with 7 items e.g., “I am able to correctly put on a condom, also when I am very aroused” (1 = totally disagree, 5 = totally agree). **Self-efficacy condom use preparation** ($r = .38$) was measured with 2 items regarding carrying and buying condoms, e.g., “When necessary, I am always able to buy condoms” (1 = totally disagree, 5 = totally agree). **Self-efficacy for STI testing** was tested with the single item “If I would want to get an STI test, I would definitely be able to” (1 = totally disagree, 5 = totally agree). **Self-efficacy regarding the use of the contraceptive pill** was measured with two items ($r = .65$, e.g., ‘If I (my girlfriend) would use the pill, I’d be able to remember (remind her) to take it every day, also when I’m drunk’ (1 = absolutely not, 5 = yes, definitely).

Self-efficacy regarding refusal of sex was measured with 4 items ($\alpha = .83$), e.g., “I am able to refuse sex if I don’t want it, even if I have flirted with someone throughout the whole evening” (1 = totally disagree, 5 = totally agree). **Self-efficacy regarding communication of wishes and boundaries** was measured with the single item ‘I am able to talk with a partner about what we do and do not like regarding sex’ (1 = totally disagree, 5 = totally agree). **Self-efficacy regarding self-regulation** when the other person does not want to have sex was measured with 2 items ($r = .68$), e.g., “I am able to let go rather than insist on having sex when my partner doesn’t want to” (1 = absolutely not, 5 = yes, definitely).

Norms regarding responsibilities were measured with a Likert Scale (1 = only the girl, 2 = mainly the girl, 3 = boy and girl equally, 4 = mainly the boy, 5 = only the boy) which were later recoded as (-2 = only the girl, -1 = mainly the girl, 0 = boy and girl equally, 1 = mainly the boy, 2 = only the boy). **Perceived responsibility regarding carrying condoms** was measured with the question “According to you, who is responsible for carrying condoms?” **Perceived responsibility for condom use** was questioned as: “According to you, who is responsible for the use of condoms?” **Perceived responsibility regarding arranging contraception** was phrased as: “According to you, who is responsible for arranging contraception?” **Perceived responsibility regarding the contraceptive pill** was measured with two items ($r = .34$), e.g., “According to you, who is responsible when the girl forgets

the pill?” **Perceived responsibility regarding unplanned pregnancy** was questioned as: “According to you, who is responsible if the girl accidentally becomes pregnant?” **Perceived responsibility regarding wishes and boundaries** was measured with three separate items. The first measure was questioned as: “According to you, who decides what happens during sex?” The second outcome was measured with the question: “According to you, who is responsible when a girl does something during the sex against her will?” The final measure was questioned as: “According to you, who is responsible when a boy does something during the sex against his will?”

Intentions for condom use preparation was measured with two items ($r = .48$), e.g., “I am planning to buy condoms when I need them” (1 = totally disagree, 5 = totally agree). **Intention for actual condom use** was measured with five items ($\alpha = .84$), e.g., “Imagine you have a partner. You want to have sex with each other but neither of you have taken an STI test. Are you planning to use a condom?” (1= absolutely not, 5 = yes, definitely). **Intention to communicate** wishes and boundaries was measured with one item: ‘Imagine you have a partner. Are you planning to discuss with him/her what you do and don’t like to do, sexually?’ **Intentions regarding refusal** in a relationship was measured with three situational items ($\alpha = .88$), e.g., ‘Imagine you have a partner. Your partner becomes angry because (s)he want to have sex, but you are not in the mood. Are you planning to refuse the sex?’ **Self-regulation intention** was again measured with one item: ‘Imagine you want to have sex but your partner doesn’t. Are you then going to try to persuade your partner into having sex?’ All measures used the same scale (1= absolutely not, 5 = yes, definitely). For girls, an additional item measured the **intention to use a form of contraception** to prevent pregnancy if she would have a boyfriend and wanted to sleep with him (1= absolutely not, 5 = yes, definitely).

Behavioral measures

Carrying condoms was measured with the question: “In the past 3 months (post-test: 6 weeks), have you ever carried condoms? (0=never, 1=sometimes, 2=regularly, 3=always). Since a histogram showed a severe skewness, the measure was recoded into (0=never, 1=at least once). **Carry condom at last intercourse** was questioned as: “the last time you had sexual intercourse in the past 3 months (post-test: 6 weeks), did you bring along a condom?” **Safe sex practices** were measured with a multiple choice question (multiple answers possible): “the last time you had sexual intercourse, did you do anything to prevent pregnancy or STI transmission?” Answer options were no; yes, condom; yes, the pill; or yes, other, namely ... , which was recoded into 0 = unsafe sex, 1 = pregnancy and/or STI prevention. **STI testing behavior** was measured with the question: “Have you ever (post-test: in the previous six weeks) been tested for an STI?” which was recoded into the binary

(0=no, 1=yes). Two behavioral outcome measures were related to communication with last relationship or sex partner. **Communication wishes and boundaries** consisted of two items ($r = .81$) e.g., “Have you talked about things that you like regarding sex?” **Communication safe sex** consisted of four items ($\alpha = .82$), e.g., “Have you talked about prevention of STI’s?” Answer options were as follows: (1 = never talked about, 2 = talked about it a bit, 3 = extensively talked about).

Data analysis

Psycho-social determinants

The data were analyzed in IBM SPSS 24 using mixed (multilevel) regression because of the nested structure of the design (repeated measurements nested within students, who are nested within classes and schools). An additional advantage of the mixed regression is that it takes into account all participants, including those with only one measure. The predictors were: condition (0 = control, 1 = intervention) and time of measurement (T0 = baseline, T1 = post-test). The aim of this research was to investigate whether there are program effects, as reflected by a significant condition by time interaction (represented by the product terms condition*T). Gender (1 = boy, 0 = girl) and sexual experience intercourse (1 = yes, 0 = no) were taken into account as covariates. To capture unexplained outcome differences between schools, a random intercept was included. To capture inter-individual differences over time, an unstructured covariance matrix for the repeated measures was selected. In line with the Consort 2010 guidelines for reporting randomized trials, we did not check for baseline differences on outcome measures between intervention and control group, as any potential differences are caused by chance (Gruijters, 2016; Moher et al., 2010). Potential differences have, however, been adjusted for with the random intercept.

Results were considered significant at $p \leq .005$ (two-tailed) to correct for multiple testing³. Effect sizes were calculated as $\beta / \sqrt{\text{var}}$, where β is the estimated group difference at T1, and var is the residual outcome variance. This effect size is Cohen’s d , based on multilevel analysis taking into account the baseline T0 and the nesting of the students within schools rather than the student t-test for two unpaired means.

Behavioral outcomes

The analyses for the behavioral outcomes were performed among the subpopulation of those with sexual experience (at baseline and/or post-test, in order to take into account those

3 Applying Bonferroni correction by dividing the p-value of .05 by the number of outcome measures (31), leads to a p-value of $.05/31 = .0016$. However, as the Bonferroni correction is an overcorrection, we rounded it off to .005.

who had recently become sexually active). The exception is communication behavior, as for these two measures all students who had had a relationship / sex partner in the past two years were taken into account. The outcomes related to carrying condoms (last 3 months and during last intercourse) and safe sex practices are dichotomous measures and were therefore analyzed with multilevel logistic regression. The continuous measures were analyzed with mixed regression, in a similar fashion as the psychosocial determinants.

Results

Psycho-social determinants

Table 4 provides the mean scores (observed), p-values, and effect sizes on the psychosocial outcome measures at baseline and post-intervention. Significant time*condition interaction effects were found for STI risk perception ($p < .000$; Cohen's $d = -.29$) and attribution of homosexuality ($p < .000$; Cohen's $d = .25$), both with a medium effect size. When looking at 3-way interactions with sexual experience, the significant effects for STI risk perception and attribution of homosexuality diminished. Model reduction led back to the simplified time*condition model. After the intervention, students in the intervention group had a higher risk perception for STI's ($M = 2.89$, $SD = 0.67$) as compared to students in the control condition ($M = 2.75$, $SD = 0.72$). Furthermore, students in the intervention group scored higher on attribution of homosexuality after the intervention ($M = 3.66$, $SD = 0.88$) as compared to students in the control condition ($M = 3.30$, $SD = 0.85$), meaning that they better understand (acknowledge) that homosexuality is something one is born with (nature), rather than how one is raised (nurture) or a personal choice. None of the other outcome measure showed significant effects (p 's $> .03$, d 's $< -.29$).

Table 4. Psycho-social outcome measures: Mean scores (SD) at baseline and posttest, significance levels, and effect sizes (Cohen's d)

	T0 baseline (N=999)		T1 posttest (N=830)		<i>p</i>	Cohen's <i>d</i>
	Mean (SD)		Mean (SD)			
Outcome	I-group	c-group	I-group	c-group		
Knowledge	7.53 (1.41)	7.70 (1.29)	8.30 (1.06)	8.19 (1.02)	.03	-.20
Risk perception STI's	2.75 (.72)	2.79 (.74)	2.89 (.67)	2.75 (.72)	.000*	-.29
Risk perception pregnancy	2.22 (.78)	2.26 (.81)	2.18 (.74)	2.29 (.79)	.13	.11
Attitude condoms	4.07 (.67)	4.08 (.68)	4.14 (.69)	4.11 (.70)	.84	-.01
Attitude communication	4.12 (.74)	4.16 (.74)	4.22 (.68)	4.21 (.69)	.57	-.04
Attitude pressure	4.46 (.68)	4.57 (.54)	4.51 (.65)	4.58 (.57)	.37	-.06
Attitude contraceptives	4.51 (.58)	4.44 (.62)	4.59 (.53)	4.56 (.57)	.73	.02
Attitude homosexuality	4.40 (.74)	4.30 (.78)	4.40 (.74)	4.32 (.71)	.51	.03
Attribution homosexuality	3.45 (.93)	3.36 (.85)	3.66 (.88)	3.30 (.85)	.000*	.25
Empathy homosexuality	2.89 (.83)	2.80 (.83)	2.86 (.86)	2.75 (.86)	.92	.01
Self-efficacy preparation condom use	3.21 (.83)	3.11 (.84)	3.41 (.81)	3.34 (.78)	.34	.06
Self-efficacy condom use	3.88 (.48)	3.89 (.51)	4.00 (.50)	4.00 (.51)	.30	.06
Self-efficacy STI test	3.90 (.88)	3.83 (.89)	4.14 (.77)	3.97 (.80)	.16	-.11
Self-efficacy contraceptive pill	3.58 (.93)	3.49 (.96)	3.52 (.93)	3.50 (.93)	.07	.12
Self-efficacy boundaries	3.84 (.76)	3.81 (.73)	3.93 (.72)	3.94 (.72)	.40	.03
Self-efficacy communication	3.83 (.84)	3.86 (.81)	3.99 (.74)	3.88 (.80)	.03	-.15
Self-efficacy self regulation	4.28 (.76)	4.36 (.69)	4.28 (.75)	4.34 (.74)	.85	-.01
Intention condom use preparation	3.40 (.85)	3.46 (.84)	3.56 (.80)	3.56 (.80)	.54	-.04
Intention condom use	3.92 (.77)	3.93 (.76)	4.03 (.74)	3.97 (.77)	.14	-.08
Intention communication	3.98 (.81)	3.99 (.81)	4.05 (.79)	4.05 (.81)	.94	.01
Intention refusal	3.87 (.97)	3.83 (.90)	3.97 (.89)	4.04 (.87)	.09	-.10
Intention self-regulation	3.49 (1.05)	3.60 (1.02)	3.57 (1.08)	3.76 (1.00)	.21	.08
Intention contraceptive pill (girls)	4.61 (.65)	4.65 (.58)	4.60 (.74)	4.68 (.59)	.46	.07
Perceived responsibility carrying condoms*	.69 (.58)	.62 (.58)	.60 (.58)	.50 (.57)	.78	-.02
Perceived responsibility condom use*	.48 (.61)	.44 (.62)	.34 (.58)	.36 (.56)	.29	.08
Perceived responsibility contraceptives*	-.55 (.70)	-.56 (.71)	-.40 (.63)	-.43 (.62)	.96	.00
Perceived responsibility pill*	-1.11 (.57)	-1.10 (.55)	-.99 (.56)	-1.04 (.56)	.22	-.08
Perceived responsibility unplanned pregnancy*	-.01 (.28)	-.01 (.22)	.00 (.20)	.00 (.26)	.84	-.02
Perceived responsibility sex*	.04 (.28)	-.01 (.23)	.02 (.28)	.02 (.26)	.04	.16
Perceived responsibility boundaries girl*	.39 (.90)	.41 (.88)	.25 (.83)	.32 (.92)	.46	.06
Perceived responsibility boundaries boy*	-.31 (.91)	-1.10 (.55)	-.22 (.84)	-.26 (.91)	.78	-.02

Note. I = Intervention group, C = Control group. P-value and Cohen's *d* based on the treatment group by time interaction in mixed regression of all available data from all randomized schools and participants (intention to treat analysis), which adjusts for bias due to missing data (drop-out). Comparisons between observed means between time points may be biased due to dropout. *norm items range from -2 to 2 with rates more close to 0 signifying more shared responsibility perceptions.

Behavioral outcomes

Table 5 provides the absolute numbers (and percentage) for the dichotomous behavioral outcomes and mean scores (observed) on the continuous behavioral outcomes, as well as the significance levels (p-values). No significant time*condition interaction effects were found on any of the behavioral outcome measures (p 's > .46).

Table 5: behavioral outcomes

	T0 baseline (N=174)				T1 posttest (N=121)				<i>p</i>
Outcome	I-group		C-group		I-group		C-group		
	number (% yes)	Mean (SD)	number (% yes)	Mean (SD)	number (% yes)	Mean (SD)	number (% yes)	Mean (SD)	
Carried condoms	43 (50%)		48 (54.5%)		31 (47.7%)		27 (41.5%)		.51 ^A
Brought condoms last time intercourse	40 (46.5)%		42 (47.7%)		27 (41.5%)		25 (44.6%)		.96 ^A
Unsafe sex last time intercourse	0		1 (1.1%)		1 (1.5%)		1 (1.8%)		--
done STI test	10 (11.6%)		8 (9.1%)		2 (3.1%)		2 (3.6%)		.70 ^A
Communication wishes and boundaries		1.78 (.74)		1.83 (.77)		1.85 (.74)		1.80 (.76)	.46 ^B
Communication safe sex practices		1.51 (.57)		1.52 (.57)		1.59 (.61)		1.57 (.59)	.70 ^B

A: significance level multilevel binary logistic regression; B: significance level multiple regression

P-value based on the treatment group by time interaction in mixed regression of all available data from all randomized schools and sexually active participants (intention to treat analysis), which adjusts for bias due to missing data (drop-out). Comparisons between observed means and percentages between time points may be biased due to dropout

Discussion

This paper reports on the effect evaluation of a school-based sex education program for higher educated adolescents aged 15+ in the Netherlands called Long Live Love + (LLL+). The program was developed using the Intervention Mapping approach and had the following objectives: 1) adolescents have healthy, happy (sexual) relationships, free of coercion, 2) adolescents prevent pregnancies, 3) adolescents prevent STIs, and 4) adolescents are not prejudiced towards sexual minorities (Mevissen et al., under review). The effect evaluation focused on (determinants of) condom use, STI testing, healthy relationships, and sexual prejudice. After the intervention, students in the intervention group scored higher on risk perception for STI and attribution of homosexuality, meaning that these students perceive it

as more likely that they might contract an STI and better understand that homosexuality is something one is born with, rather than nurture or a personal choice, as compared to students in the control group. However, no other significant changes were found on the determinant level, nor for the behavioral outcomes. This is despite rather high levels of implementation (average 75%, ranging from 22% to 100%) as was reported in the process evaluation (Van Lieshout, Mevissen, De Waal, & Kok, 2017). It can be argued that when there are effects on only two out of the 37 outcome measures these can be regarded as false positive (type I) errors.

Teenagers with higher levels of education report fewer sex partners, experience less sexual victimization and have more knowledge than their lower educated peers regarding sexuality (De Graaf et al., 2012). At baseline, only one student among those with sex experience had ever had sex without any form of protection. There thus seems to be less room for improvement than in a group with more problematic outcomes. Furthermore, the observed baseline means in our study showed that overall this group already scores rather positive when it comes to knowledge, attitudes, self-efficacy, and intentions, with no ratings below 3.11 on a scale up to 5. Only the risk-perception scales and empathy towards homosexuals scores were not particularly high. In a previous study on determinants of consistent condom use among Dutch adolescents, Schaalma, Kok, and Peters (1993) found that intercourse experience has a negative impact on attitudes and perceived social norms regarding condom use. Although there were no significant three-way interactions condition*time*sexual experience in our study, it can be suggested that it may be more suitable to address these determinants of condom use at a time where most participants are sexually active than the 17.5% at baseline in this study.

The limited effects that were found can further be related to three aspects: the content of the program, the implementation, and the evaluation. We will reflect upon these three elements, while also taking into consideration potential limitations. First, health promotion programs need to be theory- and evidence-based and suit the target population. LLL+ was developed through Intervention Mapping, a protocol for the step-wise development of theory- and evidence based interventions (Bartholomew Eldredge et al., 2016). This was rigorously done in extensive collaboration with other parties, and teachers were consulted in planning and linkage groups in various stages of development (Mevissen et al., under review). In general, theory-based interventions which have been developed in a systematic way are more effective than other interventions (Albarracín et al., 2005; Downing, Jones, Cook, & Bellis, 2006; Gehrman & Hovell, 2003). The original Long Live Love program for 13 and 14 year olds showed positive effects on the level of determinants, but effects related

to behavioral outcomes were limited (SOA Aids Nederland & Rutgers, 2015; Vogels et al., 2002). However, systematic development and theory-based by no means guarantee program effectiveness (Gourlan et al., 2016; Prestwich et al., 2014).

The limited effects may indicate that this process of intervention development has not been optimal. In all steps of the development process, suboptimal decisions may have been made. For example, related to the observed needs of the population as defined in step one: the topics included in the LLL+ program may still not be relevant for the almost 60% of students who haven't had their first kiss yet. Maybe a program like LLL+ would better suit these adolescents later, when a larger proportion of the target population is sexually active. However, in the development process, teachers in the linkage group had clearly indicated that there would not be time for sex education in the last school year, due to (preparation for) the final exams (Mevisen et al., under review). LLL+ was therefore developed for year 4 (in which most students turn 16) and 5 of secondary school, but all teachers in this trial implementation used LLL+ in year four, because of the order in which Biology themes are covered in their regular teaching methods. Among 19 year olds, 80% has experience with sexual intercourse, and this percentage increases to 90% among 24 year olds (De Graaf et al., 2012). This suggests that for this target population, an intervention after school, for instance, when they go to higher education (HBO) or university, would be more suitable. An intervention at this stage in life, e.g., an online program, would also overcome the current issue that only students who have Biology classes are reached with LLL+.

Second, in order to be effective a program needs to be implemented with high levels of completeness and fidelity (Biggs et al., 2008; Durlak & DuPre, 2008; Kalafat et al., 2007). The process evaluation of LLL+ (see Van Lieshout et al., 2017) showed rather high levels of implementation ranging from 22 to 100% with an average of 75%. These figures are based on interviews with teachers, which have in previous research found to be more accurate than teacher self-report questionnaires, as the latter seem to yield overestimations (Resnicow et al., 1998). However, these ratings were relatively arbitrary (a simple 'done vs. not done' without much room for detail). Also, a study by Lillehoj, Griffin, and Spoth (2004) suggest a social desirability bias (i.e. higher implementation rates) in self-reported ratings of implementation from program providers as compared to the ratings of independent observers. Although implementation of LLL+ seemed quite good, it is difficult to be sure on whether the reported level of implementation in our study fully matches actual implementation rates. Although a very time-intensive method, classroom observations may help clarifying the actual levels of implementation. The two outcome measures that showed program effects were addressed with a video clip and through class discussions. In the video clip in module three, a young

couple both get an STI test and the boy, despite not having any symptoms, is diagnosed with chlamydia. In module four, students have discussions as they are invited to respond to several statements about homosexuality. Both the video clips and discussions were very positively evaluated by both teachers and students (Van Lieshout et al., 2017), and may thus have been effective methods and modes of delivery.

Third, the evaluation must be rigorously done and measure the right concepts. The outcome measures were based on the program objectives formulated as part of program development and should therefore be a suiting representation, covering exactly what the program aimed to achieve. Furthermore, measuring these objectives and evaluating the program was rigorously done in an RCT with 999 students at baseline. In hindsight, a 7-point Likert Scale instead of a 5-point Likert Scale might have allowed the participants to be more nuanced and may thus have led to more effects (Fishbein & Ajzen, 2010). Furthermore, it is possible that students gave answers which they perceived as socially desired and/or overestimated their abilities. Young people's overestimation of their own skills has previously been found, .e.g., in relation to driving a car (Dunning, Johnson, Ehrlinger, & Kruger, 2003; Gregersen, 1996). Although the outcome measures matched with the program objectives, the number of outcomes had to be restricted as to not make the questionnaires too long. We may thus have missed some effects. Also, it would be advisable to measure the effectiveness of the individual components of the program, preferably in a controlled setting. This would allow more detailed investigation of the effectiveness of the individual components while ensuring that the different program components cannot influence each other.

Finally, a long-term follow up measure may have set light on long-term effects. Previous research has shown that some intervention effects only appear later in time (Resnicow et al., 1998). For HIV/STI prevention, it is indeed the long-term behavioral effects that can impact prevalence rates of infections (El-bassel et al., 2005; Kirby et al., 2007). Also, the interviews performed with teachers and students as part of the process evaluation suggested positive program outcomes (Van Lieshout et al, 2017) that somehow could not be captured by the survey. We intended to add a long-term follow up measure, but because of the very low response (N=25 of out 830 post-test), this follow-up had been cancelled and the students who signed up were informed about this via email.

Conclusion

The effect evaluation of LLL+ showed limited program effects, despite theory and evidence-based program development, moderate to high self-reported implementation, positive program evaluations by teachers and students, and a thorough effect evaluation which matched the program objectives. The low percentage of sexually active students in this sample suggests that this target population may benefit more from an intervention at a later stage in life, for instance while at university. Finally, findings from the process evaluation suggest it is worth investigating the effects of Long Live Love + in more detail, e.g., by evaluating the individual modules separately.

Chapter 5

Long Live Love+. Process evaluation of an online school-based sex education program in the Netherlands

This chapter has been published in Health Education Research as:

Van Lieshout, S., Mevissen, F.E.F., de Waal, E., & Kok, G. (2017). Long Live Love+. Evaluation of the implementation of an online school-based sex education program in the Netherlands. *Health Education Research*, 32(3), 244-257

Abstract

Schools are a common setting for adolescents to receive health education, but implementation of these programs with completeness and fidelity is not self-evident. Programs that are only partially implemented (completeness) or not implemented as instructed (fidelity) are unlikely to be effective. Therefore, it is important to identify which determinants affect completeness and fidelity of program implementation.

As part of the launch of Long Live Love+ (LLL+), an online school-based sexuality education program for adolescents aged 15-17, we performed a process evaluation among teachers and students *to measure the levels of completeness and fidelity, identify factors influencing teachers' implementation, and to evaluate the students' response*. Sixteen Biology teachers from nine secondary schools throughout the Netherlands who implemented LLL+ were interviewed and 60 students participated in 13 focus group discussions.

Results showed that teachers' completeness ranged between 22-100% ($M = 75\%$). Fidelity was high, but many teachers added elements. Teachers and students enjoyed LLL+, particularly the diversity in the exercises and its interactive character. The most important factors that influenced implementation were time and organizational constraints, lack of awareness on the impact of completeness and fidelity, and student response. These factors should be taken into account when developing (online) school-based interventions.

Introduction

In the Netherlands, schools are a primary place for adolescents to receive health education, with mostly teachers deciding which programs to use (adopt) and implement in their classroom (De Graaf et al., 2012). Programs focus for example on healthy eating (Martens, van Assema, Paulussen, Van Breukelen, & Brug, 2008), physical activity (De Meij et al., 2011), or smoking prevention (Ausems, Mesters, Van Breukelen, & De Vries, 2004; R. E. Thomas, McLellan, & Perera, 2013). For such programs to be effective, not only the content is important but also the level of implementation. Only programs delivered with sufficient completeness (quantity of the program) and fidelity (implementation according to the program guidelines) may result in positive health outcomes (Biggs et al., 2008; Durlak & DuPre, 2008; Kalafat et al., 2007; Kärnä et al., 2011).

Completeness and fidelity of implementation are, however, not self-evident (Ahtola et al., 2013; Bessems et al., 2011; Wind et al., 2008). A Dutch program on healthy eating, for example, showed an average implementation rate of 70% per lesson (Bessems et al., 2011). Similar levels of incomplete implementation (80%) were found in a Finnish anti-bullying programme (Ahtola et al., 2013) and by Wind and colleagues (Wind et al., 2008) who compared implementation of a multi-component fruit and vegetable intake school program among three European countries (lesson implementation ranged from two to all 16 lessons). A review (Dusenbury et al., 2003) on implementation of school-based drug abuse prevention programs showed that 56-84% of program implementers modified interventions. Durlak and DuPre (2008) reviewed the implementation of health promotion programs for children and adolescents, and concluded that fidelity levels never reach 100% and adaptation frequently co-occurs.

Ideally, program adoption and implementation are considered during program development (Bartholomew Eldredge et al., 2016). However, both are often overlooked. Implementation not only requires specific attributes of the teacher such as knowledge, attitudes and skills but also (technical) support from colleagues or appropriate authorities such as the school director (Dusenbury et al., 2003; Paulussen; Kok; & Schaalma, 1994; Schutte et al., 2014; Wiefferink et al., 2005). In a Dutch study, teachers' sex education adoption decisions were strongly related to outcome beliefs, subjective norms, and self-efficacy (Paulussen; Kok; & Schaalma, 1994). Furthermore, decisions were related to contextual factors such as school policy and teachers' dispositions, i.e. sexual morality and sense of responsibility. Regarding implementation, teachers' outcome beliefs and perceived instrumentality best predicted program use (Wiefferink et al., 2005). Perceived instrumentality included suitability, clear structure, and feasibility of implementation.

Schutte and colleagues (2014) looked into barriers and facilitators for the implementation of Long Live Love (LLL), a widely used school-based sex education program for young adolescents (aged 13-15) in the Netherlands (Paulussen; Kok; & Schaalma, 1994; Schutte et al., 2014). Whereas adoption was predominantly influenced by teachers' curriculum-related beliefs; completeness, fidelity and continued use were also shaped by contextual factors i.e. school policy and student response. Completeness was correlated with perceived teacher-benefits, evaluating the program as practical, social support, and teachers' implementation self-efficacy. Fidelity was related to finding the program practical and useful, self-efficacy, having followed a LLL training, and being familiar with the program (Schutte et al., 2014) developing and evaluating interventions. With the focus being predominantly on the adoption stage, little is known about the factors influencing the implementation and continuation stages of the diffusion process. This study contributes to the understanding of factors that promote or impede each stage of the diffusion process in the school setting using the sex education program Long Live Love (LLL).

The present study

Based on an observed need for sex education targeting older adolescents (15+) (De Graaf et al., 2005), the original LLL program was expanded with two additional programs to target the different higher educational levels in the Netherlands: LLL+ for higher general continued education and pre-university training (resp. Hoger Algemeen Voortgezet Onderwijs (HAVO) and Voortgezet Wetenschappelijk Onderwijs (VWO)) and LLL-MBO for Middelbaar Beroeps Onderwijs (MBO), or vocational training. The present study investigated LLL+'s completeness and fidelity, factors influencing implementation, and student response.

Long Live Love+

LLL+ is developed following Intervention Mapping (IM); a protocol for theory- and evidence-based intervention development (Bartholomew Eldredge et al., 2016). To promote easy reach and widespread dissemination, LLL+ is accessible online; this overcomes barriers related to costs, access of materials, time, and geography (Griffiths, Lindenmeyer, Powell, Lowe, & Thorogood, 2006; Strecher, 2007). In addition, interactive, personalized and tailored messages and feedback can easily be included (Griffiths et al., 2006).

LLL+ covers four themes (relationships; (un)safe sex and contraception; (un)safe sex and STI; and sexual diversity) with two 45-minute lessons each. Exercises are interactive

and include discussions, (online) quizzes, narratives, and videos. All components and materials are freely available via the program website (www.langlevedeliefde.nl), with separate environments for teachers and students, and downloadable materials and teacher manuals. A linkage group including teachers was consulted throughout the development (see Mevissen et al., under review, for publication, for more details on LLL+). Table 1 gives an overview of the program. In this paper we report the process evaluation of the pilot implementation of LLL+ in nine Dutch secondary schools, using interviews with teachers and students.

Methods

Participants

Seventeen biology teachers from nine secondary schools implemented Long Live Love+, of whom 16 (56% female; aged 24-59 ($M=39.1$, $SD=12.9$)) agreed to be interviewed after program implementation. Together, they implemented LLL+ in a total of 24 classes, of which 15 in HAVO and 9 in VWO. They had 0-27 years ($M = 9.0$, $SD = 9.5$) of teaching experience.

The teachers recruited 4 to 5 students from their class to participate in a focus group discussion (FGD). In total, 13 FDG were formed with a total of 60 pupils (3 -7 per FDG, 57% girls; aged 15-17). One FGD was girls-only, one boys-only.

Procedure

Nine teachers were individually interviewed, 4 in two pairs of 2, and 3 (from the same school) together in a FGD. Students and teachers were informed about the research objectives, the anonymous reporting, and the possibility to skip questions or withdraw from the interview at all times. They gave permission for recording and signed for consent. Teachers and students received gift vouchers of €20 and €5 respectively. Interviews and FGDs lasted 20-90 minutes. Approval was received from Maastricht University's Ethical Review Committee Psychology and Neuroscience (ERCPM).

Instruments

The focus of the teachers' interviews was on completeness and fidelity of implementation and reasons for suboptimal implementation. The interview protocol was based on the models

by Paulussen et al. (1994) and Wiefferink et al. (2005). Teachers were asked to make notes in the teacher manuals, which were used as a reference during the interviews. The topic list for the students covered what exercises they remembered, how these had been implemented, and their opinions about these exercises.

Data analysis

The interviews and FGDs were recorded, transcribed, and exported to QSR NVivo 10. Although theoretical models were used for the interview protocol, we deliberately ignored these in the first coding to ensure a broad and open perspective (inductive). First open coding was conducted, in which the transcripts were examined to identify factors influencing implementation (Thomas, 2006). Subsequently, through axial coding with the theoretical framework, the open codes were examined to understand the roles and relationships of these factors. Thematic content analysis derived a framework for the results section. The first author coded all interviews and FGDs, formulating the codebook. The second author coded transcripts from four teacher interviews and four student FGDs; both researchers discussed and agreed on the coding. From there, axial coding was done by both coders independently, after which main themes and findings were discussed until consensus was reached.

For calculating completeness of implementation, coding per exercise was 1 (complete) or 0 (partially, not, or unknown). These codes were based on self-reported figures from the teachers, derived from the interviews. Their reports were never falsified during the focus group discussions with the students, suggesting correct reporting. The number of exercises fully implemented was divided by the number of exercises available, providing a percentage.

Results

Teachers

Implementation

Reasons for teachers to try out the new LLL+ program varied. Teachers were curious about new teaching materials, felt a need to devote more attention to certain sexuality-related topics, or were interested in LLL+ because of the match with the upcoming new requirements for the Biology curriculum that included a focus on interactivity and problem-based learning.

Implementation behavior of teachers showed moderate levels of completeness and differed between teachers and classes. Some tried to implement all exercises with as much completeness and fidelity as possible; others selected exercises and/or combined LLL+ with their regular teaching methods. Completeness ranged between 22-100% with an average of 75%, see Table 2.

Table 1. Completeness of implementation

Teacher	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	total
Module 1 ex. 1	0	1	0	1	1	1	1	1	1	1	1	1	-	1	1	1	1	88%
Ex. 2	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	0	94%
Ex. 3	0	1	1	1	1	1	1	1	0	1	1	1	-	1	1	1	1	88%
Module 2 ex. 1	0	1	1	1	1	1	1	0	1	1	1	1	-	1	1	1	0	81%
Ex. 2	0	1	1	1	1	1	1	0	1	1	1	1	-	1	1	1	0	81%
Ex. 3	1	1	1	1	1	1	0	0	1	1	1	1	-	1	1	1	0	81%
Ex. 4	0	1	0	0	1	1	0	0	1	1	1	1	-	1	1	1	0	63%
Ex. 5	1	1	0	1	1	1	1	1	0	0	1	1	-	1	1	1	0	75%
Ex.6	1	1	1	1	1	1	1	1	1	0	1	1	-	1	1	1	0	88%
Module 3 ex. 1	1	1	1	0	1	1	1	1	0	0	1	1	-	1	1	0	1	75%
Ex. 2	0	1	1	1	0	1	1	1	1	1	1	1	-	1	1	1	0	81%
Ex. 4	1	0	1	1	0	1	1	1	0	0	1	1	-	1	1	1	0	69%
Ex. 5	1	1	0	1	0	1	0	0	0	1	0	0	-	0	1	0	0	38%
Module 4 ex. 1	1	1	1	1	0	0	1	1	0	0	1	1	-	1	1	1	1	75%
Ex. 2	1	1	1	0	1	1	1	1	0	1	1	1	-	1	1	1	0	81%
Ex. 3	1	1	0	0	1	1	1	1	0	0	1	1	-	1	1	1	0	69%
Ex. 4	1	1	1	1	0	0	1	1	0	0	1	1	-	1	1	1	0	69%
Ex. 5	1	1	1	0	0	0	1	0	0	0	1	1	-	1	1	1	0	56%
Total	67%	94%	72%	72%	67%	83%	83%	67%	44%	56%	94%	94%	-	94%	100%	89%	22%	75%

*exercise 3 of module 3 was not available; teacher 13 was not interviewed.

Two main reasons for incomplete implementation were limited amount of time (limited teaching hours per week) and competing curriculum priorities (students need to work towards the national Biology exams). This concern seemed especially the case for theme 2 related to contraception that was perceived as the most important topic for the Biology exam. Most teachers took more time to discuss the menstruation cycle, which then resulted in time constraints and subsequently skipping other LLL+ exercises.

Other factors influencing implementation were perceived importance of a topic/exercise, curiosity, barriers (e.g., unavailability of computer facilities), or the atmosphere in the class. The latter was shaped by teacher-student interactions, students' responses, and students' attention span:

"I found discussing both situations with the whole group rather taxing, both on the teacher and on the pupils. Also with regard to pupils' concentration. Because it's quite a number of questions to go through." (male, 30).

Especially class atmosphere contributed to implementation being different for teachers within the same school and between classes.

“Oh yes, the manual was very clear (...) But I think you always have to adjust it a bit to the group you’re teaching and the dynamics within that group, the room in which you’re teaching, whether it’s the first hour on Monday morning or the last hour on Friday.”
(female, 27)

The exercises were mostly performed as instructed. However, virtually all teachers added components, ranging from fun to fear-appealing; sometimes driven by a limited concentration span of the students:

“I showed them 20 minutes of a Southpark episode about the discussion whether or not to give sex education, because after one and a half hour of discussing contraceptive methods they are all like: blehbleh“ (female, 27)

And sometimes because teachers thought the extra elements were more effective:

“They find it super interesting, these infections. And sometimes I also show scary pictures. That really helps a lot.” (male, 34)

A third reason for adding elements was teachers’ concerns about lack of (biological) depth in LLL+.

“But I just added extra lessons to it in order to, yes, to also cover the technical aspects. Because I just found that necessary.” (female, 49)

Attitudes towards LLL+

Teachers were very positive about LLL+. They valued the diversity and interactive character of the exercises, praising the combination of discussions, clips, a quiz, and exercises that motivated students to find their own information and answers. They appreciated the off- and online options, the links to other websites, and the focus on skills and emotional learning. They were especially positive about modules 1 and 4 as these topics receive little attention in their regular teaching materials.

“In general I find the program very good in regard to feelings, sexuality and how to deal with it. It has fun exercises and I’m happy with it, because there’s not a lot of good material on that matter.” (male, 52)

The main critique was the limited time allocated to core knowledge for the Biology exams.

“Yes, so if it is really a sex education program, it is very good. But if it is meant to substitute the theme of reproduction, then some parts are lacking.” (male, 35)

Some teachers noted that students could have been more challenged and called LLL+ “easy”.

Teachers’ opinions sometimes varied. Two teachers indicated that sexuality may be more positively framed: sex should be fun. Another praised how LLL+ discusses sexuality in a casual and ‘light’ manner. One found LLL+ a hassle, having to shift between the students’ web-environment and the teachers’. Most teachers, however, found LLL+ a very convenient and low-threshold program.

Attitudes towards the specific themes

Most teachers indicate module 1, relationships, as a good introduction to the topic of sexuality and reproduction, although some found it too abrupt or had to adjust to interactive teaching. Positive comments related to the interactivity, the discussions triggered by shifting perspectives (e.g., boys’ vs. girls’ views), and the convenient way of transferring more sensitive information by use of a video (i.e. sexual arousal film).

“This module I found good, and why? Because here the students really needed to do something. They had to think for themselves, write things down, and discuss in the class or in their groups, so they were active themselves.” (female, 35)

Attitudes regarding module 2, (un)safe sex and pregnancy, were mixed. Teachers were positive about the exercises and valued discussing different contraceptive methods, but also stressed that the menstrual cycle was not covered with enough depth in LLL+ for the Biology exam. They often added information resulting in this module taking more time than allocated.

“So this I extended to make the technical story. I really found this, I thought: no, this is much too short. We hurtle over this.” (female, 49)

Teachers were enthusiastic about module 3, (un)safe sex and STI's, which was perceived an important theme. The clip in which an adolescent couple goes to a clinic for an STI test was particularly well-received.

"There are many misconceptions about a cotton swab in the penis (...) Here they say: that procedure is not necessary if there are no symptoms. The students see that in the clip. I can say it 300 times but they won't believe me." (male, 52)

Teachers had different perceptions of the core message resulting in some focusing on fear, others on reassuring (e.g., that getting tested is anonymous and not painful, and that most STI's can be cured).

Teachers were impressed by module 4, sexual diversity, for its added value, because their other methods hardly attend to this topic.

"Yes I think that if there is really something with an added value, it is particularly this last part, sexual diversity. That is really something that is lacking in the regular methods." (male, 52).

Teachers liked the circle exercise, in which students explored which gay and lesbian people they know. In most classes, students were able to mention lesbians, gay men, and bisexuals in their social environment. Two teachers indicated that their students knew very few, which may have hampered the effect of the exercise.

All teachers but one ("too stereotypical") were lyrical about the two clips:

"Although, that clip also made a big impression on me personally. It really gives me goosebumps every time, also when I think back about it now. I found it really, uhm, yes I think it also made an impression on them [the students]. Yes, definitely." (female, 49)

The exercises on 'nature vs. nurture' and 'coming out' often led to in-depth discussions.

"The clip about Linda, that's what we started with. And then we had a pretty long discussion about coming out in the class room. (...) It was a nice discussion, yes, it was lively. It also became more fun throughout the lessons, these discussions. That's also the atmosphere that this program creates." (male, 26)

Two teachers from one school left out module 4, as they did not see the added value for the subject Biology.

Website and online components

Teachers were very positive about the website describing it as attractive, cheerful, and mostly easy to navigate. A few teachers disliked that the teacher environment was accessible for students, or had difficulties shifting between students vs. teacher environment or online vs. offline implementation of program elements. Computer-rooms were difficult to book and do not facilitate group discussions. Most online components (video's, quizzes) were therefore carried out on the smartboards, with the whole group. Some used the printed options.

Teacher manual

All teachers indicated that the manual had been very helpful, clear, sufficient, and supportive particularly for modules 1 and 4. Most beneficial were: the support for introducing relationships and diversity, suggestions for discussions, and how to ask deepening questions.

“This is very thorough and clear, and it gives a lot of support. In particular with regard to the discussions, as it already outlines all the directions in which the discussions can go. So at times I experienced that as very helpful.” (male, 58)

All teachers indicated not needing additional training as they are used to discussing sexuality with students. Some noted that other teachers might need additional support or would like to exchange experience with other colleagues.

Barriers, self-efficacy, and perceived conditions for implementation

Teachers indicated no difficulties teaching the topic of sexuality. A few initially had felt insecure about the new teaching method (leading group discussions and talking about feelings and opinions related to sexuality) and some of the topics (relationships and homosexuality). Furthermore, the importance and difficulty of ensuring a safe class environment was stressed. Three teachers had limited teaching experience; one lost control over the class several times.

Teachers encountered organizational constraints, such as limited availability of computers, cancelled classes, shifting group compositions, changing schedules and problems accessing the website due to capacity problems or blocked sex-related websites. Also, some of the

online exercises were not (yet) available. Different solutions were found such as working in groups, giving homework assignments, or using printed materials.

Teachers' main complaint was time pressure however this was seen as a general problem rather than related to LLL+. Most time pressure was felt in module 2 (see above). Modules 1 and 4 are topics that are normally not, or only briefly, touched upon which makes LLL+ take additional time.

Student response and compatibility

Almost all teachers were positive about the students' responses. Students actively participated and were very involved and interested:

"Yes, and when we watched that clip it became very silent in the class (...), that is usually an indicator. It remains silent and you see them watch and think." (female, 49).

Teachers' attitudes towards LLL+ were strongly related to students' responses which seem to have strongly affected implementation:

"Yes they really liked this exercise, so I was quite enthusiastic about it." (female, 37).

"We didn't do a summing up afterwards. Because in previous lessons there had been little feedback when I did that, I now talked a bit about it myself, pointed out a few things." (male, 30)

Some teachers mentioned students enjoyed the discussions but boys sometimes got tired of these social aspects, others noted students' uncertainties regarding examination on the topic.

Several comments were made about feasibility of LLL+ as HAVO students participated more actively, speaking their mind, while VWO students were more scientific-minded, asking for evidence or in-depth insights. Other differences were based on personal development and experiences of individual students:

"And I thought: how mature are you yourself in this discussion? Have you ever experienced something like this? If not, it's indeed a bit like learning swimming outside the water. And then it's easy to have a strong opinion." (female, 49)

The atmosphere in the class could differ. Sometimes, students were very engaged; sometimes the atmosphere was giggly.

“As I just said, the two groups are very different. One is very quiet and reserved, and the students wouldn’t easily get in a discussion with each other. And in the other group the response was: this is ridiculous!’ and there was much more of a discussion.” (female, 35)

Intentions future use

All teachers stated they planned to use the program, or elements of it, again:

“I have seen very good elements of which I think: maybe I won’t fully adopt this, but I have seen, yes, I have seen things of which I think these are very useful. For example the clip about Linda, and the circle exercise I will for sure do (...) So I don’t think that in the way it is buildup I will fully use it, but that I will take issues that I find important.” (female, 49).

Teachers noted that LLL+ should also be implemented in other (non-Biology) classes as it is essential for all students.

Students

The atmosphere during the focus group discussions with students was serious in most, but not all groups. In some groups, the interaction was a little giggly or students were boasting about their sex-related experiences. For some, it seemed their first sex education, while others reported they already knew everything. The majority of students were familiar with some topics in LLL+, however, at a more superficial level:

“I think it was pretty much what we already knew but it was more in detail than I knew.” (boy, 15)

General attitude regarding LLL+ and atmosphere during lessons

Students liked LLL+ and mentioned that it was very different and more fun than their regular Biology classes. They appreciated the interactive aspects and classified the program as fun, good, easy and interactive. It grasped their attention more than their regular teaching method.

“Playful, yes that sounds super silly again. But you learn it in a fun way. It’s not like you have to cram facts, but by talking about it and watching clips you learn and that is, I like that more than, then you are also more involved.” (boy, 15)

Despite some critique on amateurish acting in the clips, students were very positive about the movies and said it was a good way to learn; the messages were clear and valuable. Students noted how the focus on feelings impacted the atmosphere in class.

“It was more personal, not just the facts that are told. You are more involved yourself” (girl, 16)

Most students said they liked the open atmosphere.

“I think that because the program was very open, students in the class can also be more open. Because if you just have the book with the facts, you don’t really dare to talk about it in class either.” (girl, 16)

According to most students, the timing for the program was good. There is more interest and the atmosphere was less giggly than in previous years. Most students assumed the vast majority of their class was not sexually active yet.

“I think 95% hasn’t had sex yet” (boy, 16)

“But then, is this the right moment to learn these things?” (interviewer)

“Hmm, well, it’s good to learn. Yes, it is good to learn because then you know beforehand” (boy, 16)

In two FGDs students said that the information came a bit late.

Attitudes regarding specific modules and exercises

Students liked the exercises in module 1 and found it a good start. They enjoyed the interactive character and learning about others' viewpoints. Opinions regarding the arousal curve clip differed from informative to funny:

"That clip was rather something to laugh about than to seriously listen to it" (girl, 17)

"It was not that bad, it was the kind of information you further talk about during the break"
(girl, 16)

Modules 2 (contraceptives) and 3 (STI's) were perceived as more knowledge-geared. Students pointed out that they had learned the most from these modules. However, some felt overwhelmed with the abundance of information. A leaflet summarizing the information would have been helpful. The acting in the clip about the STI test was regarded as amateurish, but overall students appreciated the clip.

"The clip showed that you only have to give urine" (boy, 15)

"Yes, and that these people really are, sometimes, confidential" (boy, 16)

"Yes, because it was very different than what you hear" (boy, 15)

Module 4 on diversity was generally well received. Students liked the discussions in class. However, some felt that it took too much time, because their schools already regularly organize activities for acceptance of LGBT (Lesbian, Gay, Bisexual, and Transgender). In one school students regretted that their teachers had left out module 4. Students praised the two clips about the LGTB youth, evaluating these as gripping as the characters were well-chosen, and their own age.

Students who recalled visiting the LLL+ website evaluated the site as attractive, nice, cheerful, good, and pretty. Regarding usability, they mentioned the site was clear, well organized and professional. It was easy to find the information needed. A few reported that their teacher had difficulties finding the way, or that some of the clips did not work. They appreciated the helpful links to other websites. Only a few students stated they would visit the LLL+ website in private.

LLL+ vs Biology curriculum: skills and self-efficacy vs. knowledge

Despite instructions in the teacher manual to cover certain ‘biology’ elements through the regular teaching method and to finish discussion exercises with a summing up, feedback from students suggests that this was not always done. Filling out the schemes about contraceptive methods and types of STI’s sometimes went too fast. Some students wanted more in-depth biological knowledge and were worried about passing their Biology exam.

Students often compared the program with their regular teaching methods and indicated that LLL+ had a very different approach as it was more about feelings and social aspects rather than knowledge. Students also tended to call LLL+ easy because they did not have to learn so much.

“And it was mainly about each other’s opinions and not really about facts or so.” (boy, 15)

“Which is also good to discuss, because then you understand the other persons character” (girl, 16)

“Yes, that’s what I mean with not clearly including knowledge questions, it is really opinion questions.” (boy, 15)

Yes, and that’s why it is a bit awkward and unclear for us what to expect from the exam.” (girl, 15)

Module 1 and 4 in particular were regarded as very different from regular teaching methods.

“You don’t learn this from a study book. That became clear through the website and the movies we got to see. And, well, you could see the examples, for example what it is like to be homosexual. (...) In the book you read a bit about it [being homosexual] but it doesn’t go into the personal part, about how it feels for such a person.” (girl, 16)

Although the discussions were much appreciated, students questioned whether they had gained enough in-depth knowledge about sexual reproduction to pass their biology exams.

Learning aspects

Both boys and girls expressed to have learned about personal differences, and differences between boys and girls in what they want in relationships, while at the same time indicated to have learned that all in all, boys and girls are not so different from each other on what they perceive important. Many students mentioned they had never before heard about types of contraceptives other than the pill and the condom but also indicated that knowing about the full range of options was mostly relevant for girls. Students also showed more knowledge on STI transmission, and more positive attitudes and confidence towards testing and treatment. Furthermore, they felt more confident to discuss condom use and STI testing with an unwilling partner.

Students also showed more understanding about and empathy towards LGBT, for example that there is diversity among LGBT and that life can be difficult for LGBT.

“There was a lot of discussion. That was interesting. When everyone has different standpoints, you can also adjust yours. (...) I used to find transgenders just weird people....” (girl, 15)

Some students elaborated on how they had never before understood how transgenders can be so very unhappy in their body:

“I think it’s very good that this clip is included because it enhances empathy or so. I thought the clip was good. Because I mean, what if it would happen to you? I mean, you would also want understanding from other people for your situation I think.” (girl, 15)

“But you see that that’s the way he was born, in the clip” (girl, 16)

“Yes, exactly, that he had always wanted that [to be a boy]” (girl, 15)

Discussion

The present study describes the process evaluation of an online school-based sex education program called Long Live Love+ (LLL+), focusing on (factors influencing) implementation of the program. The outcomes showed that the program was not fully implemented (levels per teacher ranged from 22-100%, with an average of 75%) but most teachers implemented the exercises with high levels of fidelity. At the same time, virtually all teachers had added extra elements.

Incomplete implementation could mostly be explained by time pressure, organizational constraints, lack of awareness on the importance of full implementation, and students' response. The time pressure was largely due to teachers feeling a need to provide students with sufficient in-depth Biology information for their exams but also because program components such as group discussions frequently took more time than anticipated. Organizational constraints were mostly related to lack of computer facilities and to (changes in) the teaching schedule that limited the number of classes that could be devoted to LLL+. During program development, teachers in the linkage group had already warned for time and organizational constraints (Mevisen et al., under review) and the program thus included alternative suggestions for completing/implementing exercises, such as letting students do exercises collectively instead of individually, ideas for assigning (parts of) exercises as homework, and providing online vs offline options for exercises, as to suit a variety of circumstances. This seemed to have helped in reaching relative high levels of completeness but full implementation is still hard to achieve. This is in line with previous work showing moderate to high levels of implementation of school-based health education programs, yet with large varieties between schools and with time as a universal barrier (Ahtola et al., 2013; Bessems et al., 2011; Buston et al., 2002; Wind et al., 2008).

As found previously (Schutte et al., 2014), teachers showed high tendencies to 'pick and choose' and to add components to the program, e.g., fun or fear elements. This may have influenced program outcomes negatively; e.g., fear appeals are only effective when self-efficacy for the advised behavior is high (Peters, Ruiter, & Kok, 2013; Ruiter, Kessels, Peters, & Kok, 2014). Previous research has shown that teachers' incomprehension of the theoretical basis for behavior change negative influences implementation (Buston et al., 2002; Shepherd et al., 2014). To avoid 'pick and choose', a clearer justification of the exercises may enhance fidelity and completeness, for example by elaborating on methods (not) used. Finally, the more positive students responded to a theme or exercise, the more enthusiastic the teacher was about these as well, similar as in previous studies (Buston et al., 2002; Paulussen et al., 1994; Schutte et al., 2014; Wiefferink et al., 2005; Wind et al., 2008). Improving students' appreciation of health education programs can thus contribute to implementation completeness and subsequently to program effectiveness.

The flexibility offered in LLL+ allows for a balance between fidelity of implementation and adaptation to circumstances. Program developers' wish for fidelity often conflicts with implementers' desire to alter and adapt a program (Dusenbury et al., 2003), while it is often unknown whether alterations have positive or adverse effects. It is therefore

important to guide adaptation where possible. Fidelity and adaptation often co-exist, as teachers may adhere to certain exercises while modifying others (Durlak & DuPre, 2008).

Teachers' evaluation of the program mostly matched the students'. Both enjoyed the open, interactive and positive approach of LLL+ and the variation in the exercises. Teachers particularly appreciated the modules on relationships and sexual diversity and considered these as very different from their regular teaching. The modules on contraception and STI were considered nice but lacking in-depth biological information. Only a few times teachers had devoted more or less attention to a topic than the students liked. In general, teachers had positive intentions towards using (components of) LLL+ in the future. Finally, students' responses related to relationships and sexual diversity suggest increased empathy for other people's opinions and preferences, while their responses regarding contraception and STIs suggest improved knowledge and attitudes towards STI testing and improved skills and self-efficacy for communicating with a partner on safe-sex related issues after being exposed to LLL+.

Both students and teachers called LLL+ 'easy', likely referring to the relatively limited amount of factual knowledge in the program. This suggests an underestimation of the importance of social skills as opposed to knowledge. A possible explanation is that the program was implemented by Biology teachers during regular Biology classes, where teachers are expected to deliver facts and students to gain knowledge rather than enhance self-efficacy and social skills. To avoid that such a perception negatively influences implementation, the importance of social skills should be communicated to both teachers and students. The importance –and challenges– of social and emotional learning in sex education (Schaalma, Abraham, Gillmore, & Kok, 2004) and in other domains have been described previously (Greenberg et al., 2003; Payton et al., 2000). Dutch schools and teachers are starting to recognize the importance of application in learning and are slowly shifting towards the so-called concept-context approach (SLO, 2013; Sol, 2013).

Limitations

By coincidence, one boys-only and one girls-only focus group were formed. Here, students seemed more open. Although this could just have been two talkative groups, it may also suggest that students might have been more confident speaking their mind about LLL+ and their own sexuality when they are with same-gender peers.

By trying out a new program, teachers in this study can be regarded as early adopters (Rogers, 2003). It may be that these teachers have higher levels of self-efficacy and/or more positive attitudes towards sex education than their colleagues. Therefore, findings from this study may not be generalizable to Dutch (Biology) teachers in general.

Some teachers may have felt controlled in their implementation due to the research setting. In addition, the results are based on self-reports by the teachers, which might -although in line with the data from students- be optimistic. Not all teachers had consistently made notes in their teacher manual regarding level of implementation and their experience with the exercises which may have biased their recall during the interviews. Yet, we have taken a conservative approach for calculating the completeness levels by only considering the implementation of an exercise as complete when a teacher clearly and convincingly stated to have performed the exercise completely. We therefore expect our results to be reliable.

Implications and conclusions

Both teachers and students enjoyed LLL+. Yet, both seemed to underestimate the importance of acquiring social skills. Completeness was moderate to high with high fidelity levels, but most teachers added elements to the program. Students' response was important for teachers' implementation. In order to enhance implementation of health education programs in schools, the most important issues for consideration are student liking and response and teachers' awareness of the importance of social skills and implementation with completeness and fidelity.

Chapter 6

General discussion

Sexual health is an important element of general health and well-being (Douglas & Fenton, 2013). The World Health Organization (WHO) defines sexual health as a state of physical, emotional, mental and social well-being in relation to sexuality, and adds that “*sexual health requires a positive and respectful approach to sexuality and sexual relationships, as well as the possibility of having pleasurable and safe sexual experiences, free of coercion, discrimination and violence.*” (World Health Organization, 2006). Over the years many health promotion programs have been developed that aim to enhance the sexual health and well-being of adolescents, and an increasing number of these are theory- and evidence-based. Meta-analyses have shown that theory- and evidence-based behaviour change programs are more effective than those without a theory or evidence base (Albarracin et al., 2005; Bos, Schaalma, & Pryor, 2008; Downing et al., 2006; Gehrman & Hovell, 2003; Hobbs et al., 2013; Peters et al., 2009). There are, however, certain conditions or parameters for program effectiveness. For instance, programs need to be tailored to the specific target population and they must be personally relevant (Kreuter & Wray, 2003).

In this dissertation we examined the effectiveness of two different sex education programs. Make a Move is a sexual harassment prevention program targeting boys in residential youth care. Long Live Love+ (LLL+) is a school-based sex education program for adolescents aged 15+ who are enrolled in higher level secondary education. Besides the effect evaluation, we also studied how the programs were implemented and evaluated by their users. This chapter will provide an overview of the key findings, reflect on these findings, and compare the two programs. A process model for the implementation of sex education will then be proposed before discussing limitations and recommendations and drawing conclusions.

The interventions and the main findings

Chapter 2 reports on the effect evaluation of Make a Move. The program was developed by Rutgers, the Dutch *center of expertise on sexual and reproductive health and rights*. Make a Move is a group intervention consisting of eight group meetings. The programs goals were formulated as follows: to promote positive sexual behavior and to prevent sexual harassment. In this trial, Make a Move was implemented by freelance trainers who visited the institutions for the meetings. The program was evaluated through a semi-Randomized Controlled Trial (RCT). A total of twenty institutions participated in the evaluation; ten in the intervention group (N = 78 boys, 13 groups) and ten in the waiting-list control group (N = 99, 15 groups). Surveys were administered at the start of the first meeting, at the end of the eighth meeting (control group: eight weeks later), and six months later. Outcome measures were psychosocial determinants of sexual behavior, including attitudes, self-efficacy, and intentions. No effects were found on any of the outcome measures at posttest nor follow-up.

Chapter 3, the process evaluation of Make a Move, describes the implementation of the program. Trainers from 10 out of the 13 groups kept a diary in which they logged implementation of the exercises. On average, trainers rated to have implemented 54% of the exercises fully, 28% partially, and 18% of the exercises were left out. However, these figures are self-reported and might thus be optimistic. Resnicow and colleagues (1998) compared three different methods to measure levels of implementation, namely observation, teacher self-reporting in a diary, and through interviews. In their study, implementation rates based on teachers' self-reports were higher, suggesting teachers over-reported in their diaries. Similarly, Lillehoj, Griffin, and Spoth (2004) found that program providers tended to report higher implementation than independent observers did. We thus have to take into account that the reporting of the Make a Move trainers may have been optimistic. The outcomes of the interviews further showed a plethora of implementation barriers and other challenges that the trainers had come across, many of which were unique in every institute and group, but can be categorized as barriers related to the organization of the institute, the program itself, or the responses of the boys. Such barriers were for instance not being allocated a suitable room, time pressure in the program, and boys' unwillingness or incapability to actively participate in a constructive manner. The results also showed how much variation there had been in implementation among the trainers and groups. Based on the feedback of the trainers, the program and trainer manual have been updated, before Rutgers started offering training workshops for youth care employees to start working with Make a Move themselves.

Chapter 4 describes the effect evaluation of Long Live Love+ (LLL+), a spin-off of the widely used Long Live Love program. Long Live Love is a Dutch sex education program for younger adolescents, which was developed in the 1990s and has since then been updated several times (Schaalma et al., 1996; Schutte et al., 2014; Vanwesenbeeck et al., 2002). LLL+ is designed for students aged 15 and over in higher levels of secondary education. The program was systematically developed according to the Intervention Mapping protocol (Bartholomew Eldredge, Markham, Ruiters, Fernández, Kok, & Parcel, 2016; Mevissen et al., under review; Mevissen et al., under review). The IM protocol guides the stepwise, yet iterative, process of systematic development of theory- and evidence-based interventions. LLL+ can be used online or offline and consists of four modules of two lessons of 45 minutes each. The program goals are 1) adolescents have healthy, happy (sexual) relationships, free of coercion, 2) adolescents prevent pregnancies, 3) adolescents prevent STIs, and 4) adolescents are not prejudiced towards sexual minorities (Mevissen et al., under review; Mevissen et al., under review). Like Make a Move, this intervention was also evaluated through a quasi-experimental design with schools matched in pairs. Pairs were formed based on characteristics such as geographical area and number of classes participating. Of the outcome measures, 6 were behavioral outcomes and 31 were on the level of determinants, such as attitudes, self-efficacy, and perceived norms. Program effects were found on the outcomes STI risk perception and attribution of homosexuality, indicating that after the program, students in the intervention group found it more likely that they ever contract an STI. Furthermore, after the program students in the intervention group acknowledge more that homosexuality is nurture, rather than nature or a personal choice. No effects were found on any of the other outcome measures.

Chapter 5 reports on the process evaluation performed on LLL+, in which sixteen teachers were interviewed about the level of implementation of LLL+ and how they experienced working with the program. Furthermore, 60 students participated in focus group discussions to share their experiences with the program. The process evaluation showed rather high levels of implementation (average 75%, range 22-100%). The main barrier for implementation was time constraints. Both teachers and students shared their enthusiasm about the program. The findings showed that the flexibility offered, including the options to perform exercises online or offline, with a large group discussion or in pairs, helped teachers to implement the program. Furthermore, the teachers' positive evaluations of the program seemed partly shaped by their students' positive responses, which is in line with previous studies (Buston et al., 2002; Paulussen, Kok, Schaalma, 1994; Schutte et al., 2014; Wiefferink et al., 2005; Wind et al., 2008).

Systematic development of an adoption brochure

After this trial implementation of LLL+, the program and the teacher manual were updated, based on the feedback of the teachers. Furthermore, the password was taken off the website. Since 2013 the LLL+ program is freely available for the teachers in the waiting-list control group and all other teachers interested in LLL+ via www.langlevedeliefde.nl. We then developed an adoption brochure. Using the Intervention Mapping protocol, we formulated a matrix in which we specified the following three performance objectives related to the behavioral outcome '*teacher adopts LLL+*':

- PO 1 Teacher decides to teach relationship and sex education
- PO 2 Teacher researches available programs and comes across LLL+
- PO 3 Teacher decides to work with LLL+

After a literature search (Hoekstra et al., 2009; Paulussen et al., 1994; Wiefferink et al., 2005) and based on the data gathered in the process evaluation of LLL+, we identified determinants of these performance objectives and formulated change objectives related to the specific performance objectives and determinants. From this list of change objectives, we identified the change objectives related to teaching LLL+ specifically (as opposed to other available sex education programs) and subsequently selected those change objectives that we ought we could influence within the scope of a brochure. These eleven change objectives were then clustered based on determinant. Methods were selected, ensuring that all change objectives were covered, and the practical application meant we formulated written text and used teacher quotes from the process evaluation that together, with bits of overlap, covered all the change objectives. A graphic designer was involved for the lay-out and a first edition consisted of 2000 brochures. This brochure is now being used by SOA Aids Nederland (STI Aids Netherlands), the current owner of LLL+.

Comparing the programs

Based on the effect evaluations, it can be concluded that neither of the two programs could be proven to be effective in changing the various determinants of sexual health. In both chapter 2 and 4, the lack of findings were discussed in light of program development, implementation, and evaluation. For both programs all three aspects may have contributed,

in various degrees, to the explanation of the limited outcomes. These three elements can also be used as a framework for comparing the two programs.

One aspect in which light the programs can be compared is program development. In the development of effective interventions, it is vital that the program addresses the exact needs of the target population (Bartholomew Eldredge et al., 2016; p. 211). However, addressing the identified needs is very difficult in a youth care population with a diverse range of problems and negative experiences that often result in emotional and behavioral disturbances (Connor et al., 2004; Jonson-Reid et al., 2007; Veerman & van Yperen, 2007). According to Seto and Lalumiere (2010), there are no significant differences in terms of social skills or social competencies between adolescent sex offenders and adolescent non-sex offenders. They argue that to understand why sexual offenses are committed, and to prevent re-offending, it is important to discuss sexuality, including sexual fantasies, sexual arousal, and early exposure to sex and pornography. Instead, re-offending prevention programs often focus on social skills and social competency, which would explain the lack of effects of such secondary prevention programs. Make a Move, albeit being a primary prevention program, has a combination of these components, addressing both social skills as well as explicitly addresses aspects of sexuality. Moreover, discussions about sexuality and sexuality-related attitudes and norms are an important part of the program. Regarding youth care in general, meta-analyses show mixed results of the effectiveness of placement and treatment in youth care institutions on the positive development of youth with serious behavioral and/or emotional problems (De Swart et al., 2012; Knorth, Harder, Zandberg, & Kendrick, 2008; Souverein, Van der Helm, & Stams, 2013). A review of De Swart and colleagues (2012) on the effectiveness of institutional youth care confirms the findings of Seto and Lalumiere (2010), showing that as opposed to cognitive behavioral therapy, (social) skills training had no positive effects on the various program objectives. These previous findings suggest that it remains questionable whether a group intervention like Make a Move can adequately meet the needs of each individual in this particular target group, due to their complex and diverse problems and experiences.

LLL+ targets a very different subpopulation among adolescents than Make a Move. It is developed for 15+ aged adolescents in higher levels of secondary education; a target population that is not regarded as an at-risk group. Participants in LLL+ already scored quite high, especially on attitudes and self-efficacy outcomes, and only 17.5% of participants had ever had sexual intercourse. In a previous study on determinants of consistent condom use among Dutch adolescents, Schaalma, Kok, and Peters (1993) found that intercourse experience has a negative impact on attitudes and perceived social norms regarding condom

use. As there were no significant three-way interactions condition*time*sexual experience in our study, this does not seem to have been in the case in our sample.

However, given that such a low percentage of the target population is already sexually active, LLL+ may yield long-term effects (when they become sexually active), which have not been measured in the current study. Alternatively, it can be suggested that this target population may benefit more from a sex education program at a later stage in life, e.g., when they are at university and a larger proportion of them is sexually active.

The programs can also be compared by looking at their implementation. Levels of program implementation were moderate to high in both programs: average implementation was 75% in LLL+, and Make a Move trainers had implemented 54% of the exercises fully and 28% partially. Some Make a Move exercises were very extensive as they contained a video clip, a post-discussion, and a role play exercise. It appeared that this required taking into account partial implementation rather than a simple yes or no, as partial implementation of such exercises was common. Both LLL+ teachers and Make a Move trainers were requested to record which exercises they did and did not do. Make a Move trainers were given a diary for this, and trainers in 10 out of the 13 groups kept this diary. LLL+ teachers, on the other hand, did not keep track of their implementation in this way. The completeness rates in the process evaluation of LLL+ are thus based on the teacher interviews.

Previous work on reporting on implementation levels has shown that whereas observations and interviews give accurate completeness rates, self-reported (diary) figures are more likely to be overestimations (Lillehoj et al., 2004). This suggests that the figures for Make a Move need to be interpreted with a bit more caution as is derived from self-reported log books. For both programs, observations may have been a more reliable method to measure implementation (Lillehoj et al., 2004; Resnicow et al., 1998). However, this was impossible within the timeframe due to the time-consuming nature of observations.

Both programs were evaluated in a similar fashion: an RCT in which schools (LLL+) or institutions (Make a Move) were matched in pairs before randomization. Both evaluations had a baseline and a post-test measure, Make a Move also had a six-month follow-up. For LLL+, only 25 students indicated at post-test that they were willing to participate in an online follow-up, and therefore they were informed that the third measurement would not take place. Both effect evaluations used mixed multilevel regression to be able to take into account the level of school (LLL+) or institution (Make a Move) and to deal with the drop-out and inability to match all questionnaires. As opposed to the effect evaluation of Make a

Move, the effect evaluation of LLL+ also included behavioral outcome measures. Although for Make a Move the change objectives were not as clear as for LLL+ and although both evaluations had to be a trade-off between examining the objectives and questionnaire length, the evaluations were developed systematically.

Process evaluations

Process evaluations can give insight into how programs were delivered and can reveal how implementers and intervention audience experienced the intervention (De Visser et al., 2015; Durlak & DuPre, 2008; Moore et al., 2014; Shepherd et al., 2014). Findings from process evaluations can therefore serve multiple functions. First, their outcomes can help interpret findings from effect evaluations. Second, they give feedback to program developers, which can be used to make adjustments to a program. These adjustments can support future implementation and in turn enhance program effectiveness. Finally, process evaluations can capture program effects that may not have been captured with effect evaluations. The process evaluations of Make a Move and LLL+ have given an abundance of additional insight into the dynamics of the implementation of these two sex education programs.

The process evaluations of Make a Move showed the many difficulties in program implementation, and thus partially explained the non-findings of the effect evaluation. Mostly, it showed the variety of barriers trainers had come across, and the large role the boys played in program implementation (study 2, chapter 3). The process evaluation of LLL+ showed high levels of implementation, how the teacher manual and flexibility enabled teachers to perform the exercises, and appreciation of the program by both teachers and students. Moreover, although they have not been captured in statistically significant changes, quotes of the students showed things they have learned from the program. In a similar fashion, Make a Move trainers reflected upon certain attitudinal changes they witnessed among some of the boys throughout the course of program implementation. These findings of the process evaluations of both programs provide valuable information beyond the insights of the effect evaluations, showing the relevance and importance of both process evaluations. The outcomes have been a starting point for the program developers to make adjustments to the program and are also a good starting point for further re-assessment of the program, including the full development process.

Complex interventions

Both Long Live Love + and Make a Move can be regarded as so-called ‘complex interventions’. Complex interventions are commonly defined as interventions with multiple interacting components, meaning they have multiple components that target various objectives (Medical Research Council, 2006). Dimensions of complexity as identified by the Medical Research Council (MRC) are related to the number of components of the intervention (i.e., scope; Bartholomew Eldredge et al., 2016), the number and difficulty of behaviors that are required by implementers and receivers of the intervention, the number of groups or organizational levels that are targeted with the intervention, the number and variability of intervention outcomes, and the degree of flexibility that the intervention offers (Medical Research Council, 2006).

There are several explanations to the complexity of both programs. The LLL+ program is an extensive program, covering four themes (wishes and boundaries, pregnancy prevention, STI prevention, and sexual diversity) and addressing students’ knowledge, risk perception, attitudes, skills and self-efficacy, norms, intentions, and behaviors, most of which in relation to more than one of the themes. Although Make a Move has one common theme, namely the prevention of sexual harassment, it also has many different components in its eight meetings of 90 minutes each and addresses self-esteem, attitudes, self-efficacy, intentions, as well as outcome expectations.

Furthermore, sexuality being a sensitive topic, adds to the complexity of both programs. As sexuality is something very personal, both teachers and students can feel uncomfortable discussing the topic. Previous research has found that, among other determinants, students’ perceived discomfort in discussing their sexuality with teachers as well as teachers’ discomfort in discussing issues (including a lack of training) were barriers to the implementation of an intervention to reduce homophobic and transphobic bullying (O’Donoghue & Guerin, 2016). LLL+ tackled potential discomfort in various ways. The buildup of LLL+, the teacher manual, offering alternative ways of doing exercises and pointing out different directions in which discussions could go, and the referrals to the website Teaching Love (www.lesgevenindiefde.nl) reduced complexity as it lowered the threshold for discussions, particularly for modules one (wishes and boundaries, healthy relationships) and four (sexual diversity). This was also reflected in the teacher interviews and focus group discussions with students. Both experienced and inexperienced teachers indicated that the manual had enabled them to talk about sexual diversity in an open manner, which also stimulated students to be open and ask questions. Students indicated

that the curriculum, using personal stories, invited them to be more actively involved and share their thoughts.

Most of the Make a Move trainers frequently discuss sexuality in their work and indicated that this was not uncomfortable in any way. This suggests that the topic in itself, sexuality, did not contribute to the complexity for trainers to implement Make a Move. However, a program element which proved particularly challenging was the role plays. It is a vital method for modeling skills, but has also proven difficult to implement in other sex education programs (Buston et al., 2002). Whereas some Make a Move trainers shared their enthusiasm about these exercises, others reported back that it had caused a lot of resistance, that boys started ‘overacting’ or were only able to do the role plays when the trainer actively participated too (study 2, chapter 3). What further contributed to the complexity of Make a Move was that the trainers were visitors to the institutions. This made that they did not know the boys beforehand and thus needed to spend time on getting to know the boys and creating a positive group atmosphere. Furthermore, it made that they were reliant on their contact persons and co-trainers and could not always make autonomous decisions but were dependent on others instead. The LLL+ teachers, on the other hand, were more autonomous and already familiar with their students. This made it easier for them to implement the curriculum as they did not rely on colleagues and did not need to spend time on getting to know their students.

Finally, the trainer manual of Make a Move added to the complexity. The instructions to the exercises were long, no suggestions for alterations were given, and trainers experienced a lot of time constraints. These factors all made it more difficult for the trainers to implement the program with high levels of completeness and fidelity.

Process model of implementation

At the start of this work, we used the process model of Wiefferink et al. (2005) to guide our research on the implementation of Long Live Love+. Other researchers who have modelled implementation of evidence-based school programs are Domitrovich and colleagues (2008). They have designed an elaborate model, which has a socio-ecological approach with an individual teacher level, school level, and macro level. The model distinguishes the intervention itself and its support system, which are independent from each other, but interrelated. Furthermore, implementation quality is affected by the programs’ core elements, its delivery, and standardization. However, both the models of Wiefferink and Domitrovich miss an important aspect: a feedback loop. I would therefore propose the following model (see Figure 1):

Figure 1: process model implementation sex education

Similar to the model of Wiefferink (2005), I propose in my model that the curriculum-related beliefs of the implementer are determined by 1) the intervention (i.e. the content of the program) and its implementation strategy (i.e. teacher/trainer manual and potential further support from the program developers, such as the Teaching Love website www.lesgevenindiefde.nl, which is referred to on the Long Live Love+ website), and 2) the interactive context. The organization in which implementers work, and different individuals within, can all influence the curriculum-related beliefs of program implementers. However, based on the findings of the studies in this dissertation, I have distinguished the interactive context into two categories. The first, category A, comprises of the more distant levels in the socio-ecological model. It includes organizational capacity and support, and beliefs and opinions of important others, such as head teachers and colleagues (LLL+) or personnel of the youth care institution (Make a Move). The influence of these stakeholders has also been found in previous studies (Durlak & DuPre, 2008; Dusenbury, Brannigan, Falco, & Hansen, 2003; Paulussen, Kok, Schaalma, 1994; Shepherd et al., 2014; Wiefferink et al., 2005).

The second box of interactive context comprises of the program audience: the students. Their response also contributes to the implementers' beliefs about the curriculum and subsequently

program implementation. Whereas implementation of Long Live Love+ was mostly reinforced by the students, implementation of Make a Move was both positively and negatively affected by the response of the boys. For example, the process evaluation of Long Live Love+ showed that teachers' appreciation of the program was largely a result of the students' enthusiasm and curiosity to hear what their classmates thought. In Make a Move, some program components caused a lot of resistance in the majority of the groups. As a result, a trainer reported back about a specific exercise that it was a great exercise in the manual, but that it did not work in practice (study 2, chapter 3). When exercises did go well, trainers evaluated the exercise much more positively which in turn influenced both their implementation of somewhat similar exercises in the program within that same group, as well as the implementation of that specific exercise in subsequent groups. The importance of student response for completeness and fidelity of implementation has also been found in several other studies (Buston et al., 2002; Dusenbury et al., 2003; Paulussen, Kok, & Schaalma, 1994; Schutte et al., 2014; Shepherd et al., 2014; Wiefferink et al., 2005).

The model shows an arrow from program implementation to audience response, rather than from the intervention to the audience response. This is because the response of the audience is not a response to the program *as it is developed and intended to be delivered*, but rather a response to the program *as it was delivered to them*. The audience may not have a clear picture of the intervention program if it is delivered to them in a way that is different from how the program was intended. The individual characteristics of participants affect the audience response, both directly, and through the group dynamics. In LLL+, teachers indicated that whereas in some classes discussions easily flowed, in some other classes there were barely any students who had a firm opinion which they dared to share. In Make a Move, group dynamics played a much bigger role than in LLL+ and trainers indicated that they had needed to spend a lot of time to create a good atmosphere. In most groups, trainers were eventually able to do so. Due to the complexity of the boy's problems, it was very challenging for some boys to actively and constructively participate. Some authors have claimed that the kind of group homes where Make a Move was implemented, have a so-called hyper-masculinity culture (Murnen, 2000). In such an environment, boys cannot let their guard down and share their personal opinions and experiences. Creating a safe atmosphere is therefore a vital aspect in program implementation.

Sometimes the program was adapted as a response to personal characteristics of the audience; e.g., trainers making the Make a Move program less hetero-normative when a boy is openly gay. At other times the group dynamics (shaped by the combination of the individual participants and their characteristics) determined the group's response which in

turn influences curriculum-related beliefs of the implementer and subsequently program implementation. Similarly, LLL+ teachers made small adaptations based on the group response; e.g., extending or shortening group discussions or showing a *Southpark* clip about condoms when the group was too tired to participate in a meaningful way (study 4, chapter 5).

Fidelity and adaptation

The feedback loop in the model visualises the fidelity-adaptation tension. Several reviews have found program adaptation taking place in program implementation (Durlak & DuPre, 2008; Dusenbury et al., 2003; Shepherd et al., 2014). Researchers are increasingly looking into the complex mix of fidelity and adaptation, suggesting that program adaptation may be inevitable, and sometimes even beneficial for program outcomes (Durlak & DuPre, 2008; Lendrum & Humphrey, 2012). Program effectiveness may be a result of the combination of fidelity and adaptation, rather than of fidelity only. Currently, however, there is limited understanding of how program, implementer and organizational level factors interact and together shape implementation and a unique balance between fidelity and adaptation (Durlak & DuPre, 2008). This process model is our interpretation of the fidelity-adaptation tension, hypothesising that adaptation is largely a result of the audience response and how this influences implementers' curriculum-related beliefs. In *Make a Move*, there was also a direct link from audience response to implementation, as in some groups their behaviour (fighting, sexual harassment) simply made it impossible to implement a program. In one group where a boy sexually harassed a trainer, the institutional staff overruled the trainer and decided to end the meeting. As these incidents are regarded as exceptions, the direct link between audience response and implementation has not been taken up in the model.

Another tension that researchers have previously identified in youth care specifically is the tension between providing care through interventions that have proven to be effective, and responding to needs and requests of individuals and their families (Stals, 2012). In the implementation of *Make a Move*, responding to individual needs and problems was one of the challenges that forced trainers to improvise and adapt the program. For example, in one of the groups a boy came out as gay during one of the meetings, something the trainer needed to respond to ad-hoc. In another group, a participant appeared to prostitute himself, which explained a few things regarding his sometimes odd attitudes and responses during the meetings. A step further from program adaptation (and going back to fidelity whenever possible) is program reinvention. This means that implementers re-invent the program, e.g., by simplifying it or making any other adjustments (Rogers, 2003). One of the reasons

for doing so is to adjust the program to the target population. Another reason to re-invent a program is to create more ownership among the users (Fleuren et al., 2002; van Yperen & Bakker, 2008). Whereas adaptation can be regarded as adjustments with intentions to go back to the original program, program re-invention suggests adaptations that are meant to remain.

Limitations

The limitations of the individual studies have been described in the different chapters. There is, however, an overarching limitation. The process evaluations give insights into the implementation process and how stakeholders experienced the program. These insights support the explanation and interpretation of the outcomes of the effect evaluations. However, the level of implementation has not been taken into account as a moderator or mediator in the effect evaluation. With the accumulated knowledge of the four studies, it can be questioned whether the self-reported levels of completeness are accurate enough to take them into account. For such a role in the effect evaluation, it may be required to work with completeness levels obtained through observation studies, as these have been found to be more accurate (Lillehoj et al., 2004; Resnicow et al., 1998).

Another point that can be regarded as a limitation is the double role (and thus potential conflict of interest) of program developers and evaluators. This is previously been pointed out by Goossens and colleagues (Goossens, 2017; pp.111). The authors found that in the Dutch national database of effective youth interventions, in 84% of the evaluation studies at least one of the co-authors was not independent. Make a Move was developed by Rutgers and evaluated by a team of Maastricht University. The studies have been co-authored by one Rutgers employee (MJ). For LLL+, program developers (FM and GK) were also involved in the evaluation. Given the fact that both effect evaluations did not prove program effectiveness, it can be assumed that this has not been an issue in the studies in this dissertation. Rather, when program effectiveness is shown, the double role of authors as program developer and evaluator may raise an eyebrow. However, this point of concern also highlights once more the importance of transparent systematic program development. Without clear program goals and objectives, methods and practical applications, it is a difficult task, even more so for external evaluators, to properly evaluate an intervention. This has indeed been a challenge in relation to the evaluation of Make a Move.

Recommendations

Neither of the two interventions proved to be effective. However, rather than dismissing the programs, it is recommended to critically reassess the different steps taken in the development process. Have the needs assessments been done in the right way, or have certain needs not been identified and as a result of that not been addressed? Or were maybe some suboptimal choices made when the program objectives were formulated, and then had to be reduced to fit in a program of eight lessons (LLL+) or eight meetings (Make a Move)? Furthermore, it is recommended to reassess the potential effectiveness of the individual program components. This allows evaluating the separate components in more detail, both with an effect evaluation and while monitoring implementation.

Make a Move and Long Live Love+ were both implemented for the first time. In this dissertation we have not researched sustained, or so called standardized ‘real life’ implementation. This would, however, be very relevant to research. What happens after lessons learnt from an RCT, when the intervention developer ‘lets go’ of the intervention and implementers are using an intervention more than once? A recommendation would be to perform more implementation research in such a ‘real life’ setting, rather than only during trial implementation with active recruitment of participants, who might not be a representative sample of program implementers. We had aimed to do so and developed an extensive questionnaire in order to investigate the real-life implementation of LLL+ after the RCT, as it was now available online for all teachers, without a password and free of charge. In a larger and more representative sample of teachers, we now wanted to investigate in a quantitative way to what extent the program was delivered with completeness and fidelity, and the reasons why teachers made certain alterations. The questionnaire was programmed through Qualtrics and a banner on the LLL+ website provided teachers with a link to the questionnaire. Furthermore, the brochures as described above, were sent out to 1200 schools throughout the Netherlands with an inset referring to the website and inviting teachers to fill out the questionnaire. Unfortunately, after nine months, not a single teacher had filled out the questionnaire and we decided to take the survey offline.

An alternative way to look into implementation that does not require teachers’ willingness and effort to fill out a questionnaire, would be to track the use of the Long Live Love website. This type of research into the use of online interventions has previously been done by Crutzen, Roosjen, and Poelman (2012) and by Kirk, Morgan, Tonkin, McDonald, and Skirton (2012). Google Analytics data can yield how many unique visitors the website attracts, how they navigate around the website, and what elements of the content they are exposed to. For

LLL+, this could indicate how many visitors the website attracts, how both teachers and students navigate around the website, and which modules and exercises are covered.

Another recommendation for program developers is to plan more for implementers' understanding of the program content. Previous findings have shown that teachers' incomprehension of the theoretical basis of interventions can undermine implementation and subsequently program effectiveness (Buston et al., 2002; Shepherd et al., 2014). This seems to have occurred in both LLL+ and Make a Move too. In LLL+, teachers added fear-appeal messages, which can be counter-effective (Ruiter et al., 2014), and some Make a Move trainers had to improvise significantly. It can be assumed that implementers make these adaptations with good intentions. More research is needed in order to understand the fidelity-adaptation tension and how adaptation can be guided in such a way that it does not undermine program effectiveness.

It has already been established that teacher training enhances implementation (Ollis, 2010, 2014; Woo, Soon, Thomas, & Kaneshiro, 2011). During LLL+ program development, teachers in the linkage group indicated that their colleagues would not have time to attend official teacher training (Mevisen et al., under review). Therefore, teacher guidance had to take place through the teacher manual and teachers were referred to a support website called Teaching Love (www.lesgevenindiefde.nl). However, regardless the shape and form of teaching training and support, it is important for teacher training to not only enhance teachers' self-efficacy regarding discussing sexuality, letting students actively participate in role plays and discussions, but also to give insights into the programs' critical components and the underlying theoretical basis, while elaborating on why certain decisions were made (e.g., not using fear appeals). Teachers need to understand 1) what the critical components, their methods and practical application are; 2) the significant role (power) they have in the delivery, and thus translation of evidence-based programs; and 3) their responsibility in maintaining the core components. Ideally, this will prevent implementers from program adaptation in ways which undermines the theoretical methods and their practical applications.

Conclusions and implications

The studies in this dissertation describe the effect and process evaluations of Make a Move and Long Live Love+, when both were implemented for the first time. Neither of the programs were proven effective. The process evaluations of both programs gave additional insights. The process evaluation of LLL+ showed teachers' appreciation of the program, which was partly shaped by the enthusiasm of their students. The process evaluation of Make

a Move showed the array of barriers trainers had come across, and how they overcame these. A process model has been proposed, which schematizes program implementation, including a feedback loop around audience response.

More funding is needed for evaluation research, also beyond the stage of initial implementation. The increasing attention for mixed methods is an important improvement in the field of evaluation research. However, the next step is to follow-up on the lessons learnt from these evaluations. Particularly for LLL+, there are no indications that the program would not work. Although mistakes may have been made in any of the steps of program development, it is a theory- and evidence-based program, valued by both teachers and students. Therefore, it is important to use the feedback to revise and optimize the intervention (content) and take away barriers for implementation that have been identified. Rather than rejecting these two interventions or arguing that the target population is unteachable, we should go back and revise the different steps in program development. The outcomes of the process evaluations give guidance to this reassessment. Furthermore, evaluation of individual program components allows for assessment in more detail. Currently, however, this is where the funding stops, leaving youth care institutions and schools in doubt about the effectiveness of these two programs.

References

-
- Aarons, G. A., & Palinkas, L. A. (2007). Implementation of Evidence-based Practice in Child Welfare : Service Provider Perspectives, *34*, 411–419. <http://doi.org/10.1007/s10488-007-0121-3>
- Abbey, A., McAuslan, P., Zawacki, T., Clinton, A. M., & Buck, P. O. (2001). Attitudinal, experiential, and situational predictors of sexual assault perpetration. *Journal of Interpersonal Violence*, *16*(8), 784–807. <http://doi.org/0803973233>
- Agardh, A., Odberg-Pettersson, K., & Östergren, P.-O. (2011). Experience of sexual coercion and risky sexual behavior among Ugandan university students. *BMC Public Health*, *11*(1), 527. <http://doi.org/10.1186/1471-2458-11-527>
- Ahtola, A., Haataja, A., Kärnä, A., Poskiparta, E., & Salmivalli, C. (2013). Implementation of anti-bullying lessons in primary classrooms : how important is head teacher support ? *Educational Research*, *55*(4), 376–392. <http://doi.org/10.1080/00131881.2013.844941>
- Albarracín, D., Gillette, J. C., Earl, A. N., Glasman, L. R., Durantini, M. R., & Ho, M.-H. (2005). A Test of Major Assumptions About Behavior Change: A Comprehensive Look at the Effects of Passive and Active HIV-Prevention Interventions Since the Beginning of the Epidemic. *Psychological Bulletin*, *131*(6), 856–897. <http://doi.org/10.1037/0033-2909.131.6.856>
- Albarracín, D., Kumkale, G. T., & Poyner-Del Vento, P. (2017). How people can become persuaded by weak messages presented by credible communicators: Not all sleeper effects are created equal. *Journal of Experimental Social Psychology*, *68*, 171–180. <http://doi.org/10.1016/j.jesp.2016.06.009>
- Antle, B. F., Sullivan, D. J., Dryden, A., Karam, E. A., & Barbee, A. P. (2011). Healthy relationship education for dating violence prevention among high-risk youth. *Children and Youth Services Review*, *33*(1), 173–179. <http://doi.org/10.1016/j.childyouth.2010.08.031>
- Ausems, M., Mesters, I., Van Breukelen, G., & De Vries, H. (2004). Effects of in-school and tailored out-of-school smoking prevention among Dutch vocational school students. *Health Education Research*, *19*(1), 51–63. <http://doi.org/10.1093/her/cyg001>
- Baams, L., Beek, T., Hille, H., Zevenbergen, F. C., & Bos, H. M. W. (2013). Gender Nonconformity , Perceived Stigmatization , and Psychological Well-Being in Dutch Sexual Minority Youth and Young Adults : A Mediation Analysis. *Archives of Sexual Behavior*, *42*, 765–773. <http://doi.org/10.1007/s10508-012-0055-z>
- Ball, B., Kerig, P. K., & Rosenbluth, B. (2009). “Like a family but better because you can actually trust each other”: The Expect Respect dating violence prevention program for at-risk youth. *Health Promotion Practice*, *10*(1 suppl), 45S–58S. <http://doi.org/10.1177/1524839908322115>
- Ball, B., Tharp, A. T., Noonan, R. K., Valle, L. A., Hamburger, M. E., & Rosenbluth, B.

- (2012). Expect Respect support groups: Preliminary evaluation of a dating violence prevention program for at-risk youth. *Violence Against Women*, 18(7), 746–762. <http://doi.org/10.1177/1077801212455188>
- Barbaree, H. E., & Marshall, W. L. (2006). *The Juvenile Sex Offender*. (H. E. Barbaree & W. L. Marshall, Eds.) (second). New York: The Guilford Press.
- Bartholomew Eldredge, L. K., Markham, C., Ruiter, Robert A.C. Fernández, M.E. Kok, G., & Parcel, G. S. (2016). *Planning Health Promotion Programs: An Intervention Mapping Approach*. (Jossey-Bass, Ed.) (4th ed.). San Francisco.
- Batson, C. D., Polycarpou, M. R., Harmon-jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., ... Highberger, L. (1997). Empathy and Attitudes : Can Feeling for a Member of a Stigmatized Group Improve Feelings Toward the Group ? *Journal of Personality and Social Psychology*, 72(1), 105–118.
- Bessem, K. M. H. H., Assema, P. Van, Martens, M. K., & Paulussen, T. G. W. M. (2011). Appreciation and implementation of the Krachtvoer healthy diet promotion programme for 12- to 14- year-old students of prevocational schools. *BMC Public Health*, 11(1), 909. <http://doi.org/10.1186/1471-2458-11-909>
- Biggs, B. K., Vernberg, E. M., Twemlow, S. W., Fonagy, P., & Dill, E. J. (2008). Teacher Adherence and Its Relation to Teacher Attitudes and Student Outcomes in an Elementary School-Based Violence Prevention Program. *School Psychology Review*, 37(4), 533–549.
- Boendermaker, L., van Rooijen, K., Berg, T., & Bartelink, C. (2013). *Residentiële jeugdzorg: wat werkt? / Residential youth care: what works?* Utrecht.
- Bos, A. E. R., Dijker, A. J. M., & Koomen, W. (2007). Sex differences in emotional and behavioral responses to HIV + individuals' expression of distress. *Psychology & Health*, 22(4), 493–511. <http://doi.org/10.1080/14768320600976257>
- Bos, A. E. R., Schaalma, H. P., & Pryor, J. B. (2008). Reducing AIDS-related stigma in developing countries: The importance of theory- and evidence-based interventions. *Psychology, Health & Medicine*, 13, 450–460. <http://doi.org/10.1080/13548500701687171>
- Braeken, D., & Cardinal, M. (2008). Comprehensive Sexuality Education as a Means of Promoting Sexual Health. *International Journal of Sexual Health*, 20. <http://doi.org/10.1080/19317610802157051>
- Brownson, R., Glanz, K., Tabak, R., & Stamatakis, K. (2015). Implementation, dissemination and diffusion of public health interventions. In K. Glanz, B. K. Rimer, & K. Viswanath (Eds.), *Health behavior: Theory, research and practice* (5th edition, pp. 544–592). San Francisco: Jossey-Bass.
- Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, 38(2), 217–230. <http://doi.org/10.1037/0022-3514.38.2.217>

-
- Burton, D. L., Miller, D. L., & Shill, C. T. (2002). A social learning theory comparison of the sexual victimization of adolescent sexual offenders and nonsexual offending male delinquents. *Child Abuse and Neglect*, 26(9), 893–907. [http://doi.org/10.1016/S0145-2134\(02\)00360-5](http://doi.org/10.1016/S0145-2134(02)00360-5)
- Buston, K., Wight, D., Hart, G., & Scott, S. (2002). Implementation of a teacher-delivered sex education programme: obstacles and facilitating factors. *Health Education Research*, 17(1), 59–72. <http://doi.org/10.1093/her/17.1.59>
- Byers, E. S., Sears, H. A., & Foster, L. R. (2013). Factors associated with middle school students' perceptions of the quality of school-based sexual health education. *Sex Education*, 13(2), 214–227. <http://doi.org/10.1080/14681811.2012.727083>
- Centers for Disease Control and Prevention. (2014). Prevalence and Characteristics of Sexual Violence, Stalking, and Intimate Partner Violence Victimization — National Intimate Partner and Sexual Violence Survey, United States, 2011. *Surveillance Summaries*, 63(8).
- Collier, K. L., Horn, S. S., Bos, H. M. W., & Sandfort, T. G. M. (2015). Attitudes Toward Lesbians and Gays Among American and Dutch Adolescents. *The Journal of Sex Research*, 52(2), 140–150. <http://doi.org/10.1080/00224499.2013.858306>
- Colton, M. (2002). Factors associated with abuse in residential child care institutions. *Children & Society*, 16(1), 33–44. <http://doi.org/10.1002/chi.683>
- Commissie-Samson. (2012). *Omringd door zorg, toch niet veilig. Seksueel misbruik van door de overheid uit huis geplaatste kinderen, 1945 tot heden*. Amsterdam.
- Connor, D. F., Doerfler, L. A., Toscano, J. P. F., Volungis, A. M., & Steingard, R. J. (2004). Characteristics of children and adolescents admitted to a residential treatment center. *Journal of Child & Family Studies*, 13(4), 497–510. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=cja&AN=14700820&site=ehost-live>
- Cook, S. M., & Cameron, S. T. (2015). Social issues of teenage pregnancy. *Obstetrics, Gynaecology & Reproductive Medicine*, 25(9), 243–248.
- Cornelius, T. L., & Resseguie, N. (2007). Primary and secondary prevention programs for dating violence: A review of the literature. *Aggression and Violent Behavior*, 12(3), 364–375. <http://doi.org/10.1016/j.avb.2006.09.006>
- Coyle, K., Basen-Engquist, K., Kirby, D., Parcel, G., Banspach, S., Collins, J., ... Harrist, R. (2001). Safer choices: Reducing teen pregnancy, HIV, and STDs. *Public Health Reports*, 116(SUPPL. 1), 82–93. <http://doi.org/10.1093/phr/116.S1.82>
- Coyle, K., Kirby, D. B., Parcel, G., Basen-Engquist, K., Banspach, S., Rugg, D., & Weil, M. (1996). Safer choices: A multicomponent school-based HIV / STD and pregnancy prevention program for adolescent. *The Journal of School Health*, 66(3).
- Crutzen, R., Roosjen, J. L., & Poelman, J. (2012). Using Google Analytics as a process

- evaluation method for Internet-delivered interventions : an example on sexual health. *Health Promotion International*, 28(1), 36–42. <http://doi.org/10.1093/heapro/das008>
- Daamen, W. (2015). *Wat werkt bij het implementeren van jeugdinterventies?*
- Dane, A. U., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control?, 18(1), 23–45.
- de Bruijn, M., Viechtbauer, W., Schaalma, H. P., Kok, G., Abraham, C., & Hospers, H. J. (2010). Standard Care Impact on Effects of Highly Active Antiretroviral Therapy Adherence Interventions. *Arch. Intern. Med.*, 170(3), 240–250.
- de Bruijn, P., Burrie, I., & van Wel, F. (2006). A risky boundary: Unwanted sexual behaviour among youth. *Journal of Sexual Aggression*, 12(2), 81–96. <http://doi.org/10.1080/13552600600841631>
- de Graaf, H., Kuijer, H., Van Acker, J., & Meijer, S. (2012). *Seks onder je 25e. Seksuele gezondheid van jongeren in Nederland anno 2012*. Delft.
- de Graaf, H., Meijer, S., Poelman, J., & Vanwesenbeeck, I. (2005). *Seks onder je 25e. Seksuele gezondheid van jongeren in Nederland anno 2005 / Sex under the age of 25. A population study of Dutch adolescent's sexual health*. Delft.
- de Graaf, I., de Haas, S., Zaagsma, M., & Wijsen, C. (2015). Effects of Rock and Water: an intervention to prevent sexual aggression. *Journal of Sexual Aggression*, 1–16. <http://doi.org/10.1080/13552600.2015.1023375>
- De Koker, P., Mathews, C., Zuch, M., Bastien, S., & Mason-Jones, A. J. (2014). A systematic review of interventions for preventing adolescent intimate partner violence. *Journal of Adolescent Health*, 54(1), 3–13. <http://doi.org/10.1016/j.jadohealth.2013.08.008>
- de Meij, J. S. B., Chinapaw, M. J. M., van Stralen, M. M., van der Wal, M. F., van Dieren, L., & van Mechelen, W. (2011). Effectiveness of JUMP-in, a Dutch primary school-based community intervention aimed at the promotion of physical activity. *British Journal of Sports Medicine*, 45(13), 1052–7. <http://doi.org/10.1136/bjsm.2010.075531>
- De Swart, J. J. W., Van den Broek, H., Stams, G. J. J. M., Asscher, J. J., Van der Laan, P. H., Holsbrink-Engels, G. a., & Van der Helm, G. H. P. (2012). The effectiveness of institutional youth care over the past three decades: A meta-analysis. *Children and Youth Services Review*, 34(9), 1818–1824. <http://doi.org/10.1016/j.childyouth.2012.05.015>
- de Visser, R. O., Graber, R., Hart, A., Abraham, C., Scanlon, T., Watten, P., & Memon, A. (2015). Using Qualitative Methods Within a Mixed-Methods Approach to Developing and Evaluating Interventions to Address Harmful Alcohol Use Among Young People. *Health Psychology*, 34(4), 349–360.
- DeGue, S., Valle, L. A., Holt, M. K., Massetti, G. M., Matjasko, J. L., & Tharp, A. T. (2014). A systematic review of primary prevention strategies for sexual violence

-
- perpetration. *Aggression and Violent Behavior*, 19(4), 346–362. <http://doi.org/10.1016/j.avb.2014.05.004>
- Devries, K. M., Mak, J. Y. T., Garcia-Moreno, C., Petzold, M., Child, J. C., Falder, G., ... Watts, C. H. (2013). The Global Prevalence of Intimate Partner Violence Against Women. *Science*, 340, 27–28. <http://doi.org/10.1126/science.1240937>
- Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. a., Olin, S., ... Ialongo, N. S. (2008). Maximizing the Implementation Quality of Evidence-Based Preventive Interventions in Schools: A Conceptual Framework. *Advances in School Mental Health Promotion*, 1(3), 6–28. <http://doi.org/10.1080/1754730X.2008.9715730>
- Douglas, J. M., & Fenton, K. A. (2013). Understanding sexual health and its role in more effective prevention programs. *Public Health Reports*, 128, 1–4.
- Downing, J., Jones, L., Cook, P. A., & Bellis, M. A. (2006). *Prevention of sexually transmitted infections (STIs): a review of reviews into the effectiveness of non-clinical interventions. Evidence Briefing Update.*
- Dukers-Muijters, N. H. T. M., Somers, C., de Graaf, H., Meijer, S., & Hoebe, C. J. P. A. (2015). Prevalence of non-volitional sex types and associated factors: A national sample of young people. *Plos One*, 10(7). <http://doi.org/10.1371/journal.pone.0132847>
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), 83–87.
- Durlak, J. A. (1998). Why Program Implementation Is Important. *Journal of Prevention & Intervention in the Community*, 17(2), 5–18. <http://doi.org/10.1300/J005v17n02>
- Durlak, J. A. (2015). Studying Program Implementation Is Not Easy but It Is Essential. *Prevention Science*, 1123–1127. <http://doi.org/10.1007/s11121-015-0606-3>
- 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(3–4), 327–350. <http://doi.org/10.1007/s10464-008-9165-0>
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation : implications for drug abuse prevention in school settings. *Health Education Research: Theory and Practice*, 18(2), 237–256.
- El-bassel, N., Witte, S. S., Gilbert, L., Wu, E., Chang, M., Hill, J., & Steinglass, P. (2005). Long-Term Effects of an HIV / STI Sexual Risk Reduction Intervention for Heterosexual Couples. *AIDS and Behavior*, 9(1). <http://doi.org/10.1007/s10461-005-1677-0>
- Emier, B. F., Windch, A. M., Zaino, S. W., Feldman, Stephen, M., & Scheetz, J. P. (1980). The Value of Repetition and Reinforcement in Improving Oral Hygiene Performance. *Journal of Periodontology*, 51(4), 228–234.
- EP Nuffic. (2015). *Education system The Netherlands. The Dutch education system described.*

- Fago, D. P. (2003). Evaluation and Treatment of Neurodevelopmental Deficits in Sexually Aggressive Children and Adolescents. *Professional Psychology: Research and Practice*, 34(3), 248–257. <http://doi.org/10.1037/0735-7028.34.3.248>
- Ferguson, R. M., Vanwesenbeeck, I., & Knijn, T. (2008). A matter of facts ... and more : an exploratory analysis of the content of sexuality education in The Netherlands. *Sex Education*, 8, 93–106. <http://doi.org/10.1080/14681810701811878>
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. (M. Fishbein & I. Ajzen, Eds.). New York: Psychology Press.
- Fleuren, M. A. H., Wiefferink, C. H., & Paulussen, T. G. W. M. (2002). *Belemmerende en bevorderende factoren bij de implementatie van zorgvernieuwingen in organisaties*. Leiden.
- Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-Based Interventions in Schools: Developers' Views of Implementation Barriers and Facilitators. *School Mental Health*, 1(1), 26–36. <http://doi.org/10.1007/s12310-008-9002-5>
- Foshee, V. A., Benefield, T. S., Ennett, S. T., Bauman, K. E., & Suchindran, C. (2004). Longitudinal predictors of serious physical and sexual dating violence victimization during adolescence. *Preventive Medicine*, 39(5), 1007–1016. <http://doi.org/10.1016/j.ypmed.2004.04.014>
- Fox, C. L., Corr, M.-L., Gadd, D., & Butler, I. (2014). Young teenagers' experiences of domestic abuse. *Journal of Youth Studies*, 17(4), 510–526. <http://doi.org/10.1080/13676261.2013.780125>
- Gault, I. (2012). Macho man: hyper-masculinity and adolescent dating violence (pp. 1–8).
- Gehrman, C. A., & Hovell, M. F. (2003). Protecting children from environmental tobacco smoke (ETS) exposure: A critical review. *Nicotine & Tobacco Research*, 5, 289–301. <http://doi.org/10.1080/1462220031000094231>
- Goossens, F. X. (2017). *Taking snapshots of preventive interventions: on the effectiveness of preventive interventions for youth and how it relates to implementation and conflict of interest*.
- Gourlan, M., Bernard, P., Bortolon, C., Romain, A. J., Lareyre, O., Carayol, M., ... Boiché, J. (2016). Efficacy of theory-based interventions to promote physical activity. A meta-analysis of randomised controlled trials. *Health Psychology Review*, 10.
- Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., ... Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist*, 58(6–7), 466–474. <http://doi.org/10.1037/0003-066X.58.6-7.466>
- Gregersen, N. P. (1996). Young drivers' overestimation of their own skill - an experiment

-
- on the relation between training strategy and skill. *Accident; Analysis and Prevention*, 28(2), 243–250.
- Griffiths, F., Lindenmeyer, A., Powell, J., Lowe, P., & Thorogood, M. (2006). Why are health care interventions delivered over the internet? A systematic review of the published literature. *Journal of Medical Internet Research*, 8(2). <http://doi.org/10.2196/jmir.8.2.e10>
- Gruber, J. E., & Fineran, S. (2008). Comparing the impact of bullying and sexual harassment victimization on the mental and physical health of adolescents. *Sex Roles*, 59(1–2), 1–13. <http://doi.org/10.1007/s11199-008-9431-5>
- Gruijters, S. L. K. (2016). Baseline comparisons and covariate fishing: Bad statistical habits we should have broken yesterday. *The European Health Psychologist*, 18(5), 205–209.
- Haider-Markel, D. P., & Joslyn, M. R. (2008). Beliefs about the Origins of Homosexuality and Support for Gay Rights: An Empirical Test of Attribution Theory. *The Public Opinion Quarterly*, 72(2), 291–310. <http://doi.org/10.1093/poq/nfn015>
- Hall, G. C. N., & Barongan, C. (1997). Prevention of Sexual Aggression. *American Psychologist*, 52(1), 5–14.
- Hobbs, N., Godfrey, A., Lara, J., Errington, L., Meyer, T. D., Rochester, L., ... Sniehotta, F. F. (2013). Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. *BMC Medicine*, 11(75), 1–12.
- Hoekstra, F., Klein Velderman, M., Wiefferink, K., van Overveld, C., ter Beek, J., Buijs, G., ... Paulussen, T. (2009). De adoptiefase van schoolgerichte preventieprogramma's: Van boodschap naar beslissing. *Tijdschrift Sociale Geneeskunde*, 87(1), 27–36.
- Holmes, K., & Sher, L. (2013). Dating violence and suicidal behavior in adolescents. *International Journal of Adolescent Medicine and Health*, 25(3), 257–261.
- James, S., Reddy, P., Ruiter, R. A. C., McCauley, A., & van den Borne, B. (2006). The impact of an HIV and AIDS life skills program on secondary school students in KwaZulu-Natal, South Africa. *AIDS Education and Prevention*, 18(4), 281–294. <http://doi.org/10.1521/aeap.2006.18.4.281>
- Johnson, R. (1988). *Elementary Statistics*. (R. Johnson, Ed.) (fifth edit). Boston: PWS-KENT Publishing Company.
- Jonson-Reid, M., Scott, L. D., McMillen, J. C., & Edmond, T. (2007). Dating violence among emancipating foster youth. *Children and Youth Services Review*, 29(5), 557–571. <http://doi.org/10.1016/j.childyouth.2006.12.008>
- Kafka, M. P., & Hennen, J. (2002). A DSM-IV Axis I comorbidity study of males (n = 120) with paraphilias and paraphilia-related disorders. *Sexual Abuse : A Journal of Research and Treatment*, 14(4), 349–366. <http://doi.org/10.1177/107906320201400405>

- Kalafat, J., Illback, R. J., & Sanders, D. (2007). The relationship between implementation fidelity and educational outcomes in a school-based family support program : Development of a model for evaluating multidimensional full-service programs. *Evaluation and Program Planning*, 30, 136–148. <http://doi.org/10.1016/j.evalprogplan.2007.01.004>
- Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Alanen, E., & Salmivalli, C. (2011). Going to Scale: A Nonrandomized Nationwide Trial of the KiVa Antibullying Program for Grades 1–9. *Journal of Consulting and Clinical Psychology*, 79(6), 796–805. <http://doi.org/10.1037/a0025740>
- Kilpatrick, D. G., Acierno, R., Saunders, B., Resnick, H. S., Best, C. L., & Schnurr, P. P. (2000). Risk factors for adolescent substance abuse and dependence: data from a national sample. *Journal of Consulting and Clinical Psychology*, 68(1), 19–30. <http://doi.org/10.1037//0022-006X.68.1.19>
- Kilpatrick, D. G., Ruggiero, K. J., Acierno, R., Saunders, B. E., Resnick, H. S., & Best, C. L. (2003). Violence and risk of PTSD, major depression, substance abuse/dependence, and comorbidity: Results from the National Survey of Adolescents. *Journal of Consulting and Clinical Psychology*, 71(4), 692–700. <http://doi.org/10.1037/0022-006X.71.4.692>
- Kirby, D. B. (2008). The Impact of Abstinence and Comprehensive Sex and STD / HIV Education Programs on Adolescent Sexual Behavior. *Sexuality Research and Social Policy*, 5(3), 18–27.
- Kirby, D. B., Laris, B. A., & Roller, L. A. (2007). Sex and HIV Education Programs: Their Impact on Sexual Behaviors of Young People Throughout the World. *Journal of Adolescent Health*, 40, 206–217. <http://doi.org/10.1016/j.jadohealth.2006.11.143>
- Kirk, M., Morgan, R., Tonkin, E., McDonald, K., & Skirton, H. (2012). An objective approach to evaluating an internet-delivered genetics education resource developed for nurses: using Google Analytics to monitor global visitor engagement. *Journal of Research in Nursing*, 16(6).
- Knorth, E. J., Harder, A. T., Zandberg, T., & Kendrick, A. J. (2008). Under one roof: A review and selective meta-analysis on the outcomes of residential child and youth care. *Children and Youth Services Review*, 30(2), 123–140. <http://doi.org/10.1016/j.childyouth.2007.09.001>
- Kreuter, M., & Wray, R. J. (2003). Tailored and Targeted Health Communication : Strategies for Enhancing Information Relevance. *American Journal of Health Behavior*, 23 (Suppl, S227–S232. <http://doi.org/10.5993/AJHB.27.1.s3.6>
- Kumkale, G. T., & Albarracín, D. (2004). The Sleeper Effect in Persuasion: A Meta-Analytic Review. *Psychological Bulletin*, 130(1), 143–172. <http://doi.org/10.1037/0033-2909.130.1.143>

-
- Kuyper, L., de Wit, J., Adam, P., Woertman, L., & van Berlo, W. (2010). *Laat je nu horen: Een onderzoek naar grensoverschrijdende seksuele ervaringen en gedragingen onder jongeren / Speak up now! Unwanted sexual experiences and behaviors among young people*. Utrecht.
- Kwasnicka, D., Dombrowski, S. U., White, M., & Sniehotta, F. F. (2015). Data-Prompted Interviews: Using Individual Ecological Data to Stimulate Narratives and Explore Meanings. *Health Psychology, 34*(12), 1191–1194. <http://doi.org/10.1037/hea0000234>
- Larsen, K. S., Reed, M., & Hoffman, S. (1980). Attitudes of Heterosexuals toward Homosexuality: A Likert-Type Scale and Construct Validity. *The Journal of Sex Research, 16*(3), 245–257.
- Lee, T. D., Swanson, L. R., & Hall, A. L. (1991). What Is Repeated in a Repetition ? Effects of Practice Conditions on Motor Skill Acquisition. *Physical Therapy, 71*(2).
- Lendrum, A., & Humphrey, N. (2012). The importance of studying the implementation of interventions in school settings. *Oxford Review of Education, 38*(5), 635–652. <http://doi.org/10.1080/03054985.2012.734800>
- Lewis, J., & Knijn, T. (2001). A Comparison of English and Dutch Sex Education in the Classroom The English continue to focus on prevention. *Education and Health, 19*(4), 59–64. Retrieved from <http://sheu.org.uk/sites/sheu.org.uk/files/imagepicker/1/eh194jl.pdf>
- Lillehoj, C. J., Griffin, K. W., & Spoth, R. (2004). Program Provider and Observer Ratings of School-Based Preventive Intervention Implementation: Agreement and Relation to Youth Outcomes. *Health Education & Behavior, 31*(2), 242–257. <http://doi.org/10.1177/1090198103260514>
- Linnan, L., & Steckler, A. (2002). Process Evaluation for Public Health Interventions and Research: an overview. In A. Steckler & L. Linnan (Eds.), *Process Evaluation for Public Health Interventions and Research* (pp. 1–24). San Francisco: Jossey-Bass.
- Low, S., Van Ryzin, M. J., Brown, E. C., Smith, B. H., & Haggerty, K. P. (2014). Engagement Matters: Lessons from Assessing Classroom Implementation of Steps to Respect: A Bullying Prevention Program Over a One-year Period. *Prevention Science, 15*(2), 165–176. <http://doi.org/10.1007/s11121-012-0359-1>
- Martens, M., van Assema, P., Paulussen, T., van Breukelen, G. J. P., & Brug, J. (2008). Krachtvoer : Effect evaluation of a Dutch healthful diet promotion curriculum for lower vocational schools. *Public Health Nutrition, 11*(3). <http://doi.org/10.1017/S1368980007000298>
- Maxwell, C. D., Robinson, A. L., & Post, L. A. (2014). The Nature and Predictors of Sexual Victimization and Offending Among Adolescents. *Journal of Youth and Adolescence, 32*(6), 465–477. <http://doi.org/10.1023/A:1025942503285>

- McCauley, H. L., Tancredi, D. J., Silverman, J. G., Decker, M. R., Austin, S. B., McCormick, M. C., ... Miller, E. (2013). Gender-equitable attitudes, bystander behavior, and recent abuse perpetration against heterosexual dating partners of male high school athletes. *American Journal of Public Health, 103*(10), 1882–1887. <http://doi.org/10.2105/AJPH.2013.301443>
- Medical Research Council. (2006). *Developing and evaluating complex interventions: new guidance*.
- Mevissen, F. E. F., Kok, G., Watzeels, A., Van Duin, G., & Bos, A. E. R. (n.d.). Systematic Development of a Dutch School-Based Sexual Prejudice Reduction Program: an Intervention Mapping approach. *Under review*
- Mevissen, F. E. F., Van Empelen, P., Watzeels, A., Van Duin, G., Meijer, S., van Lieshout, S., & Kok, G. (n.d.). Systematic Development of a Dutch School-based Online Sexual Health Program: Long Live Love+. An Intervention Mapping Approach. *Under review*
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., ... Altman, D. G. (2010). CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology, 63*(8), e1–e37. <http://doi.org/10.1016/j.jclinepi.2010.03.004>
- Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Cooper, C., ... Baird, J. (2014). Process evaluation in complex public health intervention studies: the need for guidance. *Journal of Epidemiology and Community Health, 68*(2), 101–2. <http://doi.org/10.1136/jech-2013-202869>
- Murnen, S. K. (2000). Gender and the use of sexually degrading language. *Psychology of Women Quarterly, 24*.
- O'Donoghue, K., & Guerin, S. (2016). Homophobic and transphobic bullying: barriers and supports to school intervention. *Sex Education, 18*11(December), 1–15. <http://doi.org/10.1080/14681811.2016.1267003>
- O'Donohue, W., Yeater, E. A., & Fanetti, M. (2003). Rape Prevention With College Males: The Roles of Rape Myth Acceptance, Victim Empathy, and Outcome Expectancies. *Journal of Interpersonal Violence, 18*(5), 513–531. <http://doi.org/10.1177/0886260503251070>
- Ollis, D. (2010). Sex Education: Sexuality, Society and Learning “ I haven’t changed bigots but ...”: reflections on the impact of teacher professional learning in sexuality education. *Sex Education, 10*(2), 217–230. <http://doi.org/10.1080/14681811003666523>
- Ollis, D. (2014). The role of teachers in delivering education about respectful relationships : exploring teacher and student perspectives. *Health Education Research, 29*(4), 702–713. <http://doi.org/10.1093/her/cyu032>
- Part, K., Moreau, C., Donati, S., Gissler, M., Fronteira, I., & Karro, H. (2013). Teenage

-
- pregnancies in the European Union in the context of legislation and youth sexual and reproductive health services. *Acta Obstet Gynecol Scand*, 92, 1395–1406. <http://doi.org/10.1111/aogs.12253>
- Paulussen, Theo; Kok, Gerjo; Schaalma, H. (1994). Antecedents to adoption of classroom-based AIDS education in secondary schools. *Health Education Research: Theory and Practice*, 9(4), 485–496. <http://doi.org/10.1093/her/9.4.485>
- Payton, J. W., Wardlaw, D. M., Graczyk, P. A., Bloodworth, M. R., Tompsett, C. J., & Weissberg, R. P. (2000). Social and Emotional Learning: A Framework for Promoting Mental Health and Reducing Risk Behavior in Children and Youth. *J Sch Health*, 70(5), 179–185. <http://doi.org/10.1111/j.1746-1561.2000.tb06468.x>
- Peterman, A., Palermo, T., & Bredenkamp, C. (2011). Estimates and determinants of sexual violence against women in the Democratic Republic of Congo. *American Journal of Public Health*, 101(6), 1060–1067. <http://doi.org/10.2105/AJPH.2010.300070>
- Peters, G.-J. Y., Ruiter, R. a. C., & Kok, G. (2013). Threatening communication: a critical re-analysis and a revised meta-analytic test of fear appeal theory. *Health Psychology Review*, 7(June 2013), 1–24. <http://doi.org/10.1080/17437199.2012.703527>
- Peters, L. W. H., Kok, G., Ten Dam, G.T.M., Buijs, G. J., & Paulussen, T. G. W. M. (2009). Effective elements of school health promotion across behavioral domains : a systematic review of reviews. *BMC Public Health*, 9. <http://doi.org/10.1186/1471-2458-9-182>
- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205.
- Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal Involvement as a Determinant of Argument-Based Persuasion. *Journal of Personality and Social Psychology*, 41(5), 847–855.
- Poobalan, A. S., Pitchforth, E., Imamura, M., Tucker, J. S., Philip, K., Spratt, J., ... van Teijlingen, E. (2009). Characteristics of effective interventions in improving young people's sexual health: a review of reviews. *Sex Education*, 9(3), 319–336. <http://doi.org/10.1080/14681810903059185>
- Prestwich, A., Sniehotta, F. F., Whittington, C., Dombrowski, S. U., Rogers, L., & Michie, S. (2014). Does Theory Influence the Effectiveness of Health Behavior Interventions? *Health Psychology*, 33(5), 465–474.
- Price, L. P., Byers, E. S., and the Dating Violence Research Team (1999). The attitudes towards dating violence scales: Development and initial validation. *Journal of Family Violence*, 14 (4), 351-375
- Resnicow, K., Davis, M., Smith, M., Lazarus-Yaroch, A., Baranowski, T., Baranowski, J., ... Wang, D. T. (1998). How best to measure implementation of school health curricula:

- A comparison of three measures. *Health Education Research*, 13(2), 239–250. <http://doi.org/10.1093/her/13.2.239>
- Rock, I. (1957). The Role of Repetition in Associative. *The American Journal of Psychology*, 70(2), 186–193.
- Rogers, R. (2003). *Diffusion of innovations* (fifth edit). New York: Free Press.
- Rosenstock, I. M., Strecher, V. J., & Becker, M. H. (1988). Social Learning Theory and the Health Belief Model. *Health Education Quarterly*, 15(2), 175–183.
- Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: a systematic approach* (6th edition). California: Thousand Oaks.
- Rothman, A. J., & Schwarz, N. (1998). Constructing perceptions of vulnerability: personal relevance and the use of experiential information in health judgements. *Personality and Social Psychology Bulletin*, 24(10), 1053–1064.
- Ruiter, R. A. C., Kessels, L. T. E., Peters, G. J. Y., & Kok, G. (2014). Sixty years of fear appeal research: current state of the evidence. *International Journal of Psychology : Journal Internationale de Psychologie*, 49(2), 63–70. <http://doi.org/10.1002/ijop.12042>
- Rutgers. (2014). Factsheet Tienerzwangerschappen in Nederland.
- Salter, D., McMillan, D., Richards, M., Talbot, T., Hodges, J., Bentovim, A., ... Skuse, D. (2003). Development of sexually abusive behaviour in sexually victimised males: A longitudinal study. *Lancet*, 361(9356), 471–476. [http://doi.org/10.1016/S0140-6736\(03\)12466-X](http://doi.org/10.1016/S0140-6736(03)12466-X)
- Schaalma, H. P., Abraham, C., Gillmore, M. R., & Kok, G. (2004). Sex education as health promotion: What does it take? *Archives of Sexual Behavior*, 33(3), 259–269. <http://doi.org/10.1023/B:ASEB.0000026625.65171.1d>
- Schaalma, H. P., Kok, G., Bosker, R. J., Parcel, G. S., Peters, L., Poelman, J., & Reinders, J. (1996). Planned Development and Evaluation of AIDS / STD Education for Secondary School Students in the Netherlands: Short-Term Effects. *Health Education Quarterly*, 23(3), 469–487. <http://doi.org/10.1177/109019819602300407>
- Schutte, L., Meertens, R. M., Mevissen, F. E. F., Schaalma, H., Meijer, S., & Kok, G. (2014). Long Live Love. the implementation of a school-based sex-education program in the Netherlands. *Health Education Research*, 29(4), 583–597. <http://doi.org/10.1093/her/cyu021>
- Sekol, I. (2013). Peer violence in adolescent residential care: A qualitative examination of contextual and peer factors. *Children and Youth Services Review*, 35(12), 1901–1912. <http://doi.org/10.1016/j.childyouth.2013.09.006>
- Sekol, I., & Farrington, D. P. (2009). The nature and prevalence of bullying among boys and girls in Croatian care institutions: A descriptive analysis of children's homes and correctional homes. *Kriminologija & Socijalna Integracija*, 17(2), 15–34. Retrieved from

-
- <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=psyc6&AN=2010-03948-002%5Cnhttp://nhs4073201.resolver.library.nhs.uk/?sid=OVID:psycdb&id=pmid:&id=doi:&issn=1330-2604&isbn=&volume=17&issue=2&spage=15&pages=15-34&date=2009&title=Kriminolo>
- Seto, M. C., & Lalumière, M. L. (2010). What is so special about male adolescent sexual offending? A review and test of explanations through meta-analysis. *Psychological Bulletin*, 136(4), 526–575. <http://doi.org/10.1037/a0019700>
- Shepherd, J., Harden, a., Barnett-Page, E., Kavanagh, J., Picot, J., Frampton, G. K., ... Clegg, a. (2014). Using process data to understand outcomes in sexual health promotion: An example from a review of school-based programmes to prevent sexually transmitted infections. *Health Education Research*, 29(4), 566–582. <http://doi.org/10.1093/her/cyt155>
- SLO. (2013). Concept- contextvenster, 1–62.
- Smith, D. M. (2006). *The relationship between emotional childhood maltreatment and bullying*.
- SOA Aids Nederland, & Rutgers. (2015). *Lang leve de Liefde Onderbouw*. Utrecht.
- Sol, I. (2013). De implementatie van de Concept-context Methode op een middelbare school, (september 2012). Retrieved from <http://dspace.library.uu.nl/handle/1874/275615>
- Souverein, F. A., Van der Helm, G. H. P., & Stams, G. J. J. M. (2013). “Nothing works” in secure residential youth care? *Children and Youth Services Review*, 35(12), 1941–1945. <http://doi.org/10.1016/j.childyouth.2013.09.010>
- Stals, K. (2012). *De cirkel is rond. Onderzoek naar succesvolle implementatie van interventies in de jeugdzorg*.
- Strecher, V. (2007). Internet methods for delivering behavioral and health-related interventions (eHealth). *Annual Review of Clinical Psychology*, 3, 53–76. <http://doi.org/10.1146/annurev.clinpsy.3.022806.091428>
- Taylor, B. G., Stein, N., & Burden, F. F. (2010). Exploring gender differences in dating violence/harassment prevention programming in middle schools: results from a randomized experiment. *Journal of Experimental Criminology*, 6(4), 419–445. <http://doi.org/10.1007/s11292-010-9103-7>
- Taylor, B. G., Stein, N. D., Mumford, E. a., & Woods, D. (2013). Shifting Boundaries: An Experimental Evaluation of a Dating Violence Prevention Program in Middle Schools. *Prevention Science*, 14(1), 64–76. <http://doi.org/10.1007/s11121-012-0293-2>
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237–246. <http://doi.org/10.1177/1098214005283748>
- Thomas, R. E., McLellan, J., & Perera, R. (2013). School-based programmes for preventing smoking (Review). *The Cochrane Library*, (4).

- Tortolero, S. R., Markham, C. M., & Peskin, M. F. (2010). It's your game. Keep it Real: Delaying Sexual Behavior with an Effective Middle School Program. *Journal of Adolescent Health, 46*(2). <http://doi.org/10.1016/j.jadohealth.2009.06.008>.It
- UNFPA. (2015). *The Evaluation of Comprehensive Sexuality Education Programmes: A focus on the gender and empowerment outcomes*. New York.
- van Alphen, L. M., Dijkster, A. J. M., Bos, A. E. R., van den Borne, B. H. W., & Curfs, L. M. G. (2011). Explaining Not-in-My-Backyard Responses to Different Social Groups: The Role of Group Characteristics and Emotions. *Social Psychological and Personality Science, 2*(3), 245–252. <http://doi.org/10.1177/1948550610386807>
- van Bergen, J., Gotz, H. M., Richardus, J. H., Hoebe, C. J. P. A., Broer, J., & Coenen, A. J. T. (2005). Prevalence of urogenital Chlamydia trachomatis increases significantly with level of urbanisation and suggests targeted screening approaches: results from the first national population based study in the Netherlands. *Sex Transm Infect, 81*, 17–23. <http://doi.org/10.1136/sti.2004.010173>
- van Berlo, W., & Mooren, T. (2009). Seksueel geweld: gevolgen en behandeling / sexual violence: consequences and treatment. In L. Gijs, W. Gianotten, I. Vanwesenbeeck, & P. Weijnenborg (Eds.), *Seksuologie / Sexuology* (second, pp. 419–435). Houten.
- van den Broek, I. V. F., van Bergen, J. E. A. M., Brouwers, E. E. H. G., Fennema, J. S. A., Hoebe, C. J. P. A., Koekenbier, R. H., ... Op de Coul, E. L. M. (2012). Effectiveness of yearly, register based screening for chlamydia in the Netherlands: controlled trial with randomised stepped wedge implementation. *BMJ, 1*–12. <http://doi.org/10.1136/bmj.e4316>
- van Lieshout, S., Mevissen, F. E. F., De Waal, E., & Kok, G. (2017). Long Live Love+. Evaluation of the implementation of an online school-based sexuality education program in the Netherlands. *Health Education Research, 32*(3), 244–257 <https://doi.org/10.1093/her/cyx041>
- van Lieshout, S., Mevissen, F. E.F., Jonker, Marianne, Ruiter, R.A.C. (n.d.). Process evaluation of Make a Move. *Under review*
- van Lieshout, S., Mevissen, F. E. F., van Breukelen, G., Jonker, M., & Ruiter, R. A. C. (2016). Make a Move: A Comprehensive Effect Evaluation of a Sexual Harassment Prevention Program in Dutch Residential Youth Care. *Journal of Interpersonal Violence*. <http://doi.org/10.1177/0886260516654932>
- van Teijlingen, E., Benoit, C., Bourgeault, I., DeVries, R., Sandall, J., & Wrede, S. (2015). Learning from health care in other countries: the prospect of comparative research. *Health Prospects: Journal of Public Health, 14*(1), 8–12.
- van Yperen, T. A., & Bakker, K. (2008). *Ontwikkeling en borging jeugdinterventies*. Utrecht.
- Vanwesenbeeck, I., Poelman, J., Reinders, J., & Schaalma, H. (2002). Rationele herziening lespakket “Lang Leve de Liefde.”

-
- Veerman, J. W., & van Yperen, T. A. (2007). Degrees of freedom and degrees of certainty: A developmental model for the establishment of evidence-based youth care. *Evaluation and Program Planning*, 30(2), 212–221. <http://doi.org/10.1016/j.evalprogplan.2007.01.011>
- Visser, A. P., & van Bilsen, P. (1994). Effectiveness of sex education provided to adolescents. *Patient Education and Counseling*, 23, 147–160.
- Vogels, T., Buitenhuis, S. E., Bruil, J., Dijkstra, N. S., & Paulussen, T. G. W. M. (2002). *Jongeren, seksualiteit, preventie en hulpverlening. Een verkenning van de situatie in 2002*. Leiden.
- Wekerle, C., Leung, E., Wall, A. M., MacMillan, H., Boyle, M., Trocme, N., & Waechter, R. (2009). The contribution of childhood emotional abuse to teen dating violence among child protective services-involved youth. *Child Abuse and Neglect*, 33(1), 45–58. <http://doi.org/10.1016/j.chiabu.2008.12.006>
- White, J. W., & Smith, P. H. (2004). Sexual Assault Perpetration and Reperpetration: From Adolescence to Young Adulthood. *Criminal Justice and Behavior*, 31(2), 182–202. <http://doi.org/10.1177/0093854803261342>
- Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of practical program evaluation*. (J. S. Wholey, H. P. Hatry, & K. E. Newcomer, Eds.) (third). San Francisco: Jossey-Bass.
- Wiefferink, C. H., Poelman, J., Linthorst, M., Vanwesenbeeck, I., Van Wijngaarden, J. C. M., & Paulussen, T. G. W. (2005). Outcomes of a systematically designed strategy for the implementation of sex education in Dutch secondary schools. *Health Education Research*, 20(3), 323–333. <http://doi.org/10.1093/her/cyg120>
- Wight, D., & Dixon, H. (2004). SHARE - Sexual Health And Relationships: Safe, Happy And REsponsible. *Education and Health*, 22(1), 3–7.
- Wind, M., Bjelland, M., Pérez-Rodrigo, C., te Velde, S. J., Hildonen, C., Bere, E., ... Brug, J. (2008). Appreciation and implementation of a school-based intervention are associated with changes in fruit and vegetable intake in 10- to 13-year old schoolchildren — the Pro Children study. *Health Education Research*, 23(6), 997–1007. <http://doi.org/10.1093/her/cym078>
- Wolfe, D. A., Wekerle, C., Scott, K., Straatman, A.-L., & Grasley, C. (2004). Predicting abuse in adolescent dating relationships over 1 year: the role of child maltreatment and trauma. *Journal of Abnormal Psychology*, 113(3), 406–415. <http://doi.org/10.1037/0021-843X.113.3.406>
- Wolfe, D. A., Wekerle, C., Scott, K., Straatman, A.-L., Grasley, C., & Reitzel-Jaffe, D. (2003). Dating violence prevention with at-risk youth: a controlled outcome evaluation. *Journal of Consulting and Clinical Psychology*, 71(2), 279–291. <http://doi.org/10.1037/0022-006X.71.2.279>

- Woo, G. W., Soon, R., Thomas, J. M., & Kaneshiro, B. (2011). Factors Affecting Sex Education in the School System. *Journal of Pediatric and Adolescent Gynecology*, 24(3), 142–146. <http://doi.org/10.1016/j.jpbg.2010.12.005>
- World Health Organization. (2006). *Defining sexual health. Report of a technical consultation on sexual health 28-31 January 2002, Geneva*. Geneva.
- World Health Organization. (2013). *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*. Geneva: WHO Press.
- Zakireh, B., Ronis, S. T., & Knight, R. A. (2008). Individual beliefs, attitudes, and victimization histories of male juvenile sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 20(3), 323–351. <http://doi.org/10.1177/1079063208322424>

Summary

After the systematic development of health behavior change interventions, it is also very important that programs are extensively evaluated. Evaluation outcomes give feedback to program developers and can help decision-makers to select effective programs that suit the target population. A comprehensive program evaluation does not only include an effect evaluation but also a process evaluation. An effect evaluation aims to measure whether the defined program goals were met and whether the program brought about the desired change, a process evaluation looks into the process of use. Process evaluations can include opinions about the program and experiences of program users, for example what teachers and students thought of a program. Process evaluations can also measure how and to what extent a program was implemented. Two aspects of implementation that are frequently used by evaluators are completeness and fidelity of implementation. Completeness refers to the extent of implementation, whereas fidelity measures whether implementation was done as intended and instructed by program developers.

The outcomes of process evaluations can give valuable additional information about a health promotion program beyond program effectiveness. First, insights into how a program was implemented can help interpret and explain the outcomes of an effect evaluation. For example, a program might not prove effective, but maybe it was implemented with low levels of completeness and fidelity thus lowering the likelihood of success. Second, it is important to know how program users and the target population experienced working with the program. This information is valuable feedback for program developers, which can help program developers in optimizing the program and supporting program implementation with high levels of completeness and fidelity.

Program implementation with high levels of completeness and fidelity is not self-evident. Literature has shown that many school-based health promotion programs are implemented with sub-optimal levels of completeness and fidelity, and many implementers also make adaptations to the program. Wiefferink and colleagues (2005) developed a model which schematizes the implementation of the Dutch Long Live Love program. We used this model for the process evaluations of the interventions evaluated in this dissertation.

The programs evaluated in this dissertation are Make a Move (MaM) and Long Live Love+ (LLL+). Make a Move is a sexual harassment prevention program for adolescent boys in residential youth care. It is a group intervention consisting of eight meetings of 90 minutes each. The program is developed by Rutgers, the Dutch center of expertise regarding sexual and reproductive health and rights (SRHR). The goals of Make a Move have been formulated by Rutgers as promoting positive aspects of sexuality and preventing sexual harassment.

Long Live Love+ is an online school-based sex education program for adolescents aged 15 and over in higher levels of secondary education. The program is developed by Maastricht University and consists of four modules, each with two lessons of 45 minutes. The modules are wishes and boundaries regarding sex and relationships, pregnancy prevention, STI prevention, and sexual diversity. The program goals are 1) adolescents have healthy, happy (sexual) relationships, free of coercion, 2) adolescents prevent pregnancies, 3) adolescents prevent STIs, and 4) adolescents are not prejudiced towards sexual minorities.

In chapter 2 and 3 we reported on the effect and process evaluation of Make a Move, respectively. Chapter 2 describes the content of the sexual harassment prevention program, the outcome measures used, and the outcomes of the effect evaluation. Effects were evaluated in a semi-randomized controlled trial (RCT) among 177 boys in 20 youth care institutions. No effects were found on any of the psycho-social determinants of sexual offensive behavior neither at direct post-test nor at 6-month follow-up, and so the program did not prove to be effective. This was discussed in a three-way explanatory model. These three pillars are the intervention itself, its implementation, and the evaluation. Weaknesses in all three pillars, to a varying degree, may have contributed to not finding significant effects.

Chapter 3, the process evaluation of Make a Move, describes the implementation of the program and the trainers' experiences with using the program. Program implementation was measured through the trainers' diaries, and in individual interviews the trainers reflected on their experiences. On average, trainers had implemented 54% of the exercises fully, 28% partially, and 18% of the exercises were left out. However, these figures are self-reported and might therefore be optimistic. The outcomes also showed a plethora of implementation barriers and other challenges that the trainers had come across, many of which were unique in every institute and group. These barriers were related to the organization of the institute, the program itself, and the responses of the boys. The results also showed the large variation in implementation completeness and fidelity among the trainers and groups. In fact, as the trainers faced so many challenges and improvised a substantial amount, the effect evaluation may not have been able to capture and evaluate the program as it is meant to be. Instead, it showed that the extent of the program that the trainers were able to deliver, and in the improvised way they did so, was without effect.

The outcomes underline the importance of process evaluations of (health) intervention programs. Besides organizational constraints and characteristics of the intervention and the implementers, the behaviors and responses to the program of the target population also affected program implementation. The boys' engagement played a large role in program im-

plementation. Future implementation of Make a Move requires embedding in the institution and extensive support for the implementers as well as the promotion of program acceptance among the target population.

Chapter 4 describes the effect evaluation of Long Live Love+, a spin-off of the widely used Long Live Love program. Whereas the original Long Live Love program targets young adolescents aged 13-14, LLL+ is specifically designed for the upper classes of the higher levels of secondary education. The objectives of LLL+ are: 1) adolescents have healthy, happy (sexual) relationships, free of coercion, 2) adolescents prevent pregnancies, 3) adolescents prevent STIs, and 4) adolescents are not prejudiced towards sexual minorities. Each of the four themes is addressed in two lessons of 45 minutes each, making the total program consist of eight lessons. The program was evaluated with a semi-randomized controlled trial (RCT) in which schools were matched in pairs. Subsequently, one school was randomized to the intervention or control group, and its matching pair automatically to the other group. A total of 17 schools participated, with 999 students at baseline. Despite theory- and evidence-based program development according to the Intervention Mapping protocol, the program only showed limited effects. For the psychosocial measures, after the intervention, students in the intervention group found it more likely that they might one day get infected with an STI and scored higher on attribution of homosexuality, indicating that they acknowledge more that homosexuality is nature rather than nurture or a personal choice. No effects were found on any of the behavioral outcomes. Therefore, the program was not proven to be effective, despite theory and evidence-based program development, moderate to high self-reported implementation (see Chapter 5), positive program evaluations by teachers and students, and a thorough effect evaluation which matched the program objectives. Given the positive perceptions of the program and the lack of effectiveness, the low percentage of sexually active students in this sample suggests that this target population may benefit more from an intervention at a later stage in life, for instance while at university, when they have become sexually active.

Chapter 5 reports on the process evaluation of LLL+, in which we interviewed 16 teachers about their experiences with implementing LLL+ and held focus group discussions with 60 students to understand how they experienced working with the program. The process evaluation showed rather high levels of implementation (average 75%, ranging from 22 to 100%). Despite the high levels of implementation, time constraints were frequently reported as a barrier. Both teachers and students shared their enthusiasm about the program. The findings showed that teachers appreciated the flexibility offered, including the options to perform exercises online or offline, with a group discussion or in pairs. This may have helped teachers to implement the program

with relatively high levels of completeness. Furthermore, the teachers' positive evaluations of the program seemed partly shaped by their students' positive responses. Students' response was important for teachers' opinions about the program and subsequently their implementation. Both teachers and students enjoyed LLL+. Yet, both seemed to underestimate the importance of acquiring social skills. Although both teachers and students acknowledged the difficulty of certain social skills (e.g. negotiating STI testing), they regarded the LLL+ as easy, seemingly referring to the knowledge components.

Completeness was moderate to high with high fidelity levels, but most teachers added elements to the program. In order to enhance implementation of health education programs in schools, student liking and response, teachers' and students' awareness of the importance and difficulty of social skills, and teachers' awareness of the importance of implementation with high levels of completeness and fidelity need to be considered.

In the discussion section reported in chapter 6 the findings from the studies are summarized and the two interventions are compared. The lack of program effects cannot easily be explained, and this particularly goes for Long Live Love+. The program's systematic development, the rather high levels of implementation, and the positive evaluations of both teachers and students would suggest otherwise. However, both interventions can be regarded as so-called complex interventions. Both programs have multiple interaction components and various outcome goals. The process evaluations of both interventions have shown the large role the students have and how their responses, both positive and negative, influence program implementation. Therefore, a new process model explaining program implementation in the school or institutional context is proposed, which regards '*student response*' as a separate factor rather than simply a part of the interactive context. It is also hypothesized that the so-called fidelity-adaptation tension is partly a result of the student responses and how implementers react upon the responses.

Program implementers should receive training that helps them understand 1) what the critical components of the program, their methods and practical application are; 2) the significant role (power) they have in the delivery, and thus in the translation of evidence-based programs into practice; and 3) their responsibility in maintaining the core components. Ideally, this will prevent implementers from program adaptation in ways that undermine the theoretical methods and their practical applications. Therefore, it is important for program developers to explain implementers why certain methods and applications were (and were not) used. Finally, because of the importance of student responses and how these influence program implementation, student liking must be an important consideration in program development.

Samenvatting

Na het systematisch ontwikkelen van gedragsveranderingsprogramma's is het ook belangrijk dat deze interventies uitgebreid geëvalueerd worden. Een evaluatie geeft inzichten in een programma welke beleidsmakers en andere beslissers ondersteunen om beslissingen te nemen over het gebruik van interventieprogramma's. Een uitgebreide evaluatie bestaat niet alleen uit een effectevaluatie, maar omvat ook een procesevaluatie. Een effectevaluatie evalueert of programmadoelstellingen zijn behaald en de beoogde (gedrags)verandering inderdaad heeft plaatsgevonden. Een procesevaluatie kijkt vooral naar het proces van programmegebruik. Procesevaluaties kunnen meningen en ervaringen van betrokkenen onderzoeken, bijvoorbeeld hoe docenten en leerlingen over een programma denken en hoe ze het gebruik ervan ervoeren. Daarnaast kunnen procesevaluaties ook meten hoe en in hoeverre interventies zijn uitgevoerd. Twee aspecten van implementatie die vaak gebruikt worden zijn volledigheid en waarheidsgetrouwheid. Volledigheid kijkt naar de mate of hoeveelheid van implementatie, terwijl waarheidsgetrouwheid meet of een programma inderdaad werd uitgevoerd zoals bedoeld werd door programma ontwikkelaars en omschreven in de handleiding.

De uitkomsten van procesevaluaties kunnen waardevolle informatie en inzichten toevoegen aan de bevindingen van een effectevaluatie. Ten eerste kunnen inzichten in programma-implementatie helpen om uitkomsten van een effectevaluatie te interpreteren en verklaren; een programma kan bijvoorbeeld niet effectief blijken, maar wellicht was het programma niet of nauwelijks uitgevoerd? Daarnaast is het ook belangrijk wat de meningen over en ervaringen met een interventie zijn. Dit soort bevindingen bevat waardevolle feedback voor programma-ontwikkelaars, en kunnen helpen in het verder optimaliseren van een programma en het ondersteunen van implementatie met hoge mate van volledigheid en waarheidsgetrouwheid.

Het is niet vanzelfsprekend dat interventies worden geïmplementeerd met hoge mate van volledigheid en waarheidsgetrouwheid. De literatuur laat zien dat veel gezondheidsprogramma's op scholen suboptimaal worden uitgevoerd. Implementatie gebeurt niet volledig, niet helemaal waarheidsgetrouw, en er worden vaak ook aanpassingen aan programma's gemaakt. Wiefferink en collega's (2005) hebben een model ontwikkeld dat de implementatie van Lang Leve de Liefde schematisch weergeeft. Dit model hebben we gebruikt voor de procesevaluaties van de twee interventies die we hebben geëvalueerd in deze dissertatie.

De programma's die we geëvalueerd hebben zijn Make a Move (MaM) en Lang Leve de Liefde – Bovenbouw (LLL+). Make a Move is een programma ter voorkoming van seksueel grensoverschrijdend gedrag voor tienerjongens in de residentiële jeugdzorg.

Het is een groepsprogramma dat bestaat uit acht bijeenkomsten van 90 minuten. Het programma is ontwikkeld door Rutgers, het Nederlandse kennis- en expertisecentrum seksualiteit. De doelstellingen van Make a Move zijn door Rutgers geformuleerd als het bevorderen van positieve aspecten van seksualiteit en het voorkomen van seksueel grensoverschrijdend gedrag. Lang Leve de Liefde – Bovenbouw is een gedeeltelijk online seksueel voorlichtingsprogramma voor leerlingen in de HAVO en VWO bovenbouw. Het programma bestaat uit vier modules, elk met twee lessen van 45 minuten. De modules zijn wensen en grenzen, zwangerschapspreventie, SOA preventie en seksuele diversiteit. De doelstellingen van Lang Leve de Liefde – Bovenbouw zijn: 1) adolescenten hebben gezonde, gelukkige seks en relaties, vrij van dwang, 2) adolescenten voorkomen zwangerschap, 3) adolescenten voorkomen SOA en 4) adolescenten zijn niet bevooroordeeld jegens seksuele minderheden.

In hoofdstukken 2 en 3 rapporteren we de effect- en procesevaluaties van Make a Move. Hoofdstuk 2 beschrijft de inhoud van het preventieprogramma, de uitkomstmaten en de bevindingen van de effectevaluatie. De effecten werden gemeten in een semi-gerandomiseerde gecontroleerde studie (Randomized Controlled Trial; RCT) waarin 177 jongens en 20 jeugdzorginstellingen deelnamen. Op geen enkele van de psychosociale determinanten van seksueel grensoverschrijdend gedrag werden effecten gevonden, op de directe nameting noch op de nameting zes maanden later. De effectiviteit van het programma kon derhalve niet aangetoond worden. Dit kan bediscussieerd worden aan de hand van zwakheden in elk van deze drie elementen: het programma zelf, de implementatie ervan en de evaluatie. Zwakke punten in alle drie de elementen kunnen, in verschillende mate, hebben bijgedragen tot het gebrek aan significante programmaeffecten.

Hoofdstuk 3, de procesevaluatie van Make a Move, beschrijft de implementatie van het programma en de ervaringen van trainers met het gebruik van het programma. Implementatie werd gemeten aan de hand van logboeken die de trainers bijhielden. Tevens reflecteerden de trainers in individuele interviews op hun ervaringen. Gemiddeld hebben de trainers 54% van de oefeningen volledig, 28% gedeeltelijk en 18% niet uitgevoerd. Dit zijn echter zelf-gerapporteerde maten en kunnen daarom wat optimistisch zijn. Uit de interviews bleek een scala aan barrières en andere uitdagingen waar trainers tegenaan waren gelopen. Velen van deze waren uniek per jeugdzorginstelling of zelfs per groep. De barrières kunnen worden gecategoriseerd als behorend tot de organisatie van de instelling, het programma zelf, en het gedrag van en reactie op het programma van de deelnemende jongens. De resultaten lieten ook grote verschillen zien in volledigheid en waarheidsgetrouweheid van implementatie tussen trainers en groepen. Gezien de vele uitdagingen en de manier waarop trainers hebben

moeten improviseren, lijkt het erop dat de effectevaluatie het programma niet heeft kunnen evalueren zoals het bedoeld was. Het programma bleek dan ook niet effectief.

Deze uitkomsten onderstrepen het belang van het uitvoeren van procesevaluaties van gedragsveranderingsprogramma's. Naast de organisatorische moeilijkheden en de kenmerken van zowel het programma en als van de individuele trainers, bleken ook het gedrag en de reactie van de jongens op het programma een rol te spelen in de implementatie van het programma. De motivatie en betrokkenheid van de jongens bleek van grote invloed op de implementatie van het programma. Toekomstige implementatie van Make a Move vereist inbedding van het programma in de organisatie, voldoende ondersteuning voor programma-implementeerders en de bevordering van programma-acceptatie onder de doelgroep.

Hoofdstuk 4 beschrijft de effectevaluatie van Lang Leve de Liefde – Bovenbouw, een spin-off van het veelgebruikte Lang Leve de Liefde programma. Terwijl het originele Lang Leve de Liefde programma bedoeld is voor scholieren in de onderbouw van de middelbare school (13-14 jaar oud), is Lang Leve de Liefde – Bovenbouw ontwikkeld voor de HAVO en VWO bovenbouw van middelbare scholen. De doelstellingen van Lang Leve de Liefde – Bovenbouw zijn: 1) adolescenten hebben gezonde, gelukkige seks en relaties, vrij van dwang, 2) adolescenten voorkomen zwangerschap, 3) adolescenten voorkomen SOA en 4) adolescenten zijn niet bevooroordeeld ten opzichte van seksuele minderheden. Elk thema wordt behandeld in een module die bestaat uit twee lessen van 45 minuten. Het programma is geëvalueerd aan de hand van een semi-gerandomiseerde gecontroleerde studie (RCT) waarbij scholen in paren werden gematcht. De ene helft van het paar werd gerandomiseerd toegewezen aan de interventie- of de controlegroep, en de ander van het paar werd vervolgens toegewezen aan de andere groep. In totaal deden 17 scholen mee, met een totaal van 999 leerlingen op de voormeting. Ondanks systematische programma-ontwikkeling aan de hand van het Intervention Mapping protocol bleek het programma nauwelijks effectief. Op de psychosociale uitkomstmaten bleek dat leerlingen in de interventiegroep na de interventie de kans groter achtten dat zij ooit een SOA oplopen en erkenden ze sterker dat homoseksualiteit aangeboren is, en niet het gevolg van opvoeding of een persoonlijke keuze. Op de gedragsmaten werd geen enkel effect gevonden. Dit betekent dat het programma niet effectief is gebleken, ondanks systematische ontwikkeling, matig tot hoge mate van implementatie (zie Hoofdstuk 5), positieve reacties van zowel docenten als leerlingen en een rigoureuze effectevaluatie waarbij uitkomstmaten gekoppeld waren aan programmadoelstellingen. Het gebrek aan effectiviteit, ondanks hoge mate van implementatie (zie Hoofdstuk 5) en positieve ervaringen van het programma, in combinatie met het feit dat slechts een klein deel van de leerlingen al seksueel actief is, suggereert dat de doelgroep misschien meer baat heeft bij een interventie

op een later moment in hun seksuele ontwikkeling, bijvoorbeeld wanneer ze studeren en een groter gedeelte van de doelgroep seksueel actief is.

Hoofdstuk 5 rapporteert de procesevaluatie van Lang Leve de Liefde – Bovenbouw, waarin 16 docenten werden geïnterviewd over hun ervaringen met het gebruik van het programma en 60 scholieren van negen scholen deelnamen aan focusgroep discussies over hun ervaringen met en meningen over Lang Leve de Liefde – Bovenbouw. De procesevaluatie liet matige tot hoge mate van implementatie van het programma zien (gemiddeld 75%, variërend van 22 tot 100%). Tijdsdruk werd vaak gerapporteerd als een barrière voor volledige implementatie. Zowel docenten als leerlingen deelden hun enthousiasme over het programma. De resultaten lieten zien dat docenten de geboden flexibiliteit waardeerden, bijvoorbeeld de opties om oefeningen klassikaal of in kleinere groepjes uit te voeren, en online of offline. Dit kan bijgedragen hebben aan de hoge mate van implementatie. Verder bleek ook dat het enthousiasme van docenten over het programma voor een belangrijk gedeelte gevormd werd door de positieve reacties van de leerlingen. De manier waarop leerlingen reageerden op de oefeningen bleek van invloed te zijn op hoe docenten dachten over bepaalde oefeningen en vervolgens ook op de implementatie en intentie tot toekomstige implementatie. Zowel docenten als leerlingen vonden het leuk om met Lang Leve de Liefde – Bovenbouw te werken. Zowel docenten als leerlingen leken het belang van bepaalde sociale vaardigheden echter te onderschatten. Enerzijds gaven zowel docenten als leerlingen aan dat sommige sociale vaardigheden moeilijk zijn (bijvoorbeeld het bespreken van het doen van een SOA test), anderzijds vonden beide partijen dat Lang Leve de Liefde – Bovenbouw ‘makkelijk’. Het lijkt erop dat zij toen refereerden naar de kennisaspecten van het programma, en niet zozeer naar het opbouwen van vaardigheden en eigen-effectiviteit.

Ook al was de volledigheid van implementatie hoog en werden de oefeningen uitgevoerd met hoge mate van waarheidsgetrouwheid, toch hebben veel docenten aspecten toegevoegd aan het programma. Om implementatie van gedragsveranderingsprogramma's op scholen te verbeteren, moeten daarom de volgende aspecten in acht genomen worden: programma acceptatie onder leerlingen, bewustwording onder zowel docenten als leerlingen van het belang en de moeilijkheid van sociale vaardigheden, en het bewustzijn onder docenten van het belang van implementatie met hoge mate van volledigheid en waarheidsgetrouwheid.

In Hoofdstuk 6, de discussie, worden de bevindingen van de studies samengevat en worden de twee programma's met elkaar vergeleken. Het gebrek aan effectiviteit van de programma's kan niet makkelijk verklaard worden, vooral niet voor Lang Leve de Liefde – Bovenbouw. De systematische ontwikkeling van het programma, de vrij hoge mate van implementatie

en de positieve reacties van zowel docenten als leerlingen laten een potentieel effectief programma zien. Beide programma's kunnen gezien worden als zogeheten 'complexe interventies'. Beide programma's hebben verschillende componenten die op elkaar inwerken, en meerdere uitkomstdoelen. De procesevaluaties van beide programma's laten de grote rol zien die leerlingen hebben, en hoe hun gedrag en reacties op het programma, zowel in positieve als in negatieve zin, invloed hebben op programma-implementatie. Daarom wordt in deze discussiesectie een nieuw procesmodel voorgesteld, die programma-implementatie in scholen of instellingen schematiseert, waarin *student response* (de reactie van leerlingen) een opzichzelfstaande factor is, in plaats van alleen een onderdeel van de interactieve context. We hypothetiseren ook dat de zogeheten *fidelity-adaptation* spanning, dus de balans tussen waarheidsgetrouwde implementatie en aanpassing, gedeeltelijk een gevolg is van de reacties van leerlingen en hoe programma-implementeerders daar vervolgens op ageren.

Het zou goed zijn als programma implementeerders ondersteuning krijgen die hen helpt te begrijpen 1) wat de essentiële componenten van een programma, de theoretische methodes en praktische toepassingen daarvan zijn; 2) de significante rol die zij spelen in en dus controle die zij hebben over hoe een programma uitgevoerd wordt, en daarmee dus in de vertaling van theoretisch-onderbouwde programma's naar de praktijk; en 3) hun verantwoordelijkheid in het behouden van de essentiële componenten. Idealiter weerhoudt dit implementeerders ervan aanpassingen te maken op manieren die de theoretische methodes en praktische toepassingen daarvan ondermijnen. Hierom zou het goed zijn als programma-ontwikkelaars naar implementeerders toe ook onderbouwen waarom bepaalde methodes en toepassingen wel of niet gebruikt worden. Ten slotte, gezien het belang van *student response* en hoe de reactie van leerlingen invloed heeft op implementatie, is programma-acceptatie onder leerlingen ook een belangrijke overweging voor programma-ontwikkelaars die niet onderschat mag worden.

Valorization

Sexual health is an important element of general health and well-being (Douglas & Fenton, 2013). The World Health Organization (WHO) defines sexual health as a state of physical, emotional, mental and social well-being in relation to sexuality, and adds that “*sexual health requires a positive and respectful approach to sexuality and sexual relationships, as well as the possibility of having pleasurable and safe sexual experiences, free of coercion, discrimination and violence*” (World Health Organization, 2006). Although Dutch adolescents have lower rates of unplanned pregnancy and STI’s than their European counterparts (Cook & Cameron, 2015; Part et al., 2013), there is still room for improvement and comprehensive sex education is one way of enhancing the sexual health of youth.

Studies and main findings

Program developers, evaluators, teachers, youth care workers, and policy makers all have different ways in which they try to contribute to the sexual health and well-being of Dutch adolescents. As researchers, we evaluated the two Dutch sex education programs Make a Move and Long Live Love+. We performed effect evaluations to establish whether the program objectives were met and the programs brought about the desired change, and process evaluations to investigate to what extent the programs were implemented and how the program users (Make a Move: trainers; Long Live Love+: teachers and students) experienced working with the program. The results of neither two effect evaluations showed effectiveness, and may thus not be straightforward encouragements for decision-makers in youth care institutions (Make a Move) or secondary schools (Long Live Love +) about the use of the two programs. However, the process evaluations gave many additional insights, such as the importance of student’ liking and response and how that affected program implementation. Time constraints were a barrier for the successful implementation of both programs, which emphasises the importance for future program developers to delimit program content. Similarly, the importance of student liking should not be underestimated by program developers and users.

For both programs, the combination of the effect and process evaluation allowed for gaining a good understanding of the program and what it looked like in practice. This is not only useful information for the program developers but also for decision-makers and potential future program implementers. While some of the chapters in this dissertation have been published in academic peer-reviewed journals, it is also important that these findings are accessible to decision-makers and potential implementers. It is unlikely that teachers or youth care workers would search and find the evaluation studies in academic journals. The majority of journals are not Open Access, and searching and filtering relevant information might be time-consuming

and challenging. Findings from program evaluations like these should therefore be accessible for practitioners, also outside of peer-reviewed academic journals.

Finding suitable programs for your target population as a practitioner

To bridge this gap between research and practice and to allow policy makers and program users to find relevant information about intervention programs, the National Youth Institute (Nederlands Jeugdinstituut) has developed a national database of effective youth interventions (Nationale Databank effectieve jeugdinterventies). This database is a collection of health promotion programs for children and adolescents, addressing topics such as alcohol, bullying, diet, smoking, sex, and physical activity. Programs in this database have been acknowledged by the National Youth Institute (Nederlands Jeugdinstituut). The database can be used as a source of reference for those who are looking for a suitable program to use. The programs that are listed in the database are categorised as ‘well supported’ or ‘effective’. Among effective programs a distinction is made between ‘first indications’, ‘good indications’, and ‘strong indications’. LLL+ (listed under the Dutch name Lang Leve de Liefde – Bovenbouw) is listed as ‘well supported’, indicating that the program has a good description of the problem, target population, program objectives, approach and parameters for use, and that it has a solid theoretical and/or empirical base. The database also gives a summary of the evaluation outcomes. The updated original of Long Live Love for adolescents aged 13-14 (Lang Leve de Liefde Onderbouw) is listed as effective, according to ‘good indications’. Make a Move is not listed in the database, but its counterpart Girls’ Talk is. Similarly to LL+, this program is categorized as ‘well-supported’.

Furthermore, Make a Move and Long Live Love + are owned by Rutgers and SOA Aids Nederland, respectively. Both organisations are well-known among professionals. This collaboration between academia and NGO’s helps to bridge the gap between research and practice. Organisations such as Rutgers and Soa Aids have a good reputation and have strong ties with professionals in the field as well as universities, which can strengthen their work including the performance of evaluations.

Recent developments regarding sexuality in youth care settings

Following revelations about sexual abuse in the Catholic Church and other institutes where children ought to be safe, the Commission-Samson was established. The Commission investigated sexual abuse in youth care (1945-2010). Findings revealed that children and adolescents in care reported twice as much sexual abuse as their peers outside of care, with girls, intellectually disabled, and those in residential care at an even higher risk. Following these findings, the Commission-Rouvoet was established. The task of this Commission was to prevent sexual abuse in youth care. The commission has set up a “*Kwaliteitskader voorkomen seksueel misbruik in de jeugdzorg*” (Quality framework prevention of sexual abuse in youth care), which advises on how to prevent sexual abuse in youth care. The document lists several intervention programs that can be used to foster dialogue regarding sexual harassment in a team of colleagues, in group homes, with foster parents, and/or foster children. Make a Move is one of the interventions suggested, addressing the topic with a group of boys in youth care. Moreover, Rutgers has developed a brochure which advises on youth care institutions’ sexuality policy, called “*Seksualiteitsbeleid of struisvogelpolitiek?*” This brochure also refers to Make a Move. It seems that Make a Move serves a need within the youth care community. Given these developments, it is important that Make a Move is extensively evaluated and the findings have been reported back to Rutgers, and have subsequently been used to optimize the program. Ideally, the updated program, which is currently not implemented by trainers but by youth care staff after a two-day training course with Rutgers, will also be evaluated to see whether the lessons learned and adaptations made have an effect on program implementation and program effectiveness.

Sexual diversity in the school curriculum

Since 2012, Dutch primary and secondary schools have to cover sexual diversity in their curriculum, in order to reduce bullying and discrimination of gay men, lesbians, bisexual and transgender people. However, according to the VO-raad, 5% of schools do not adhere to this legislation (VO-raad, 2017). In April 2017, the Dutch House of Representatives (Tweede Kamer) passed a motion from SP (Socialistische Partij, a Dutch opposition party) to sanction schools that do not comply. As one of the four themes in Long Live Love + is sexual diversity, the curriculum is listed as one of the teaching methods advised to cover sexual diversity and thus comply with the ruling. The findings of the process evaluation of LLL+, including teachers’ appreciation of module 4 (sexual diversity) indicate that the program does not simply serve a need as it helps teachers to fulfil certain requirements, but that teachers found it a great tool which helped lower the threshold to discuss a subject that can be difficult

to address and have meaningful discussions about it. These findings can encourage other teachers to use the program as a whole, or to work with this specific module.

Conclusion

All in all, numerous initiatives have been taken in the past five years to promote the sexual health and well-being of Dutch adolescents. The development of Make a Move took place in a time in which the government and youth care institutes realised that sexuality needs to be addressed in the youth care setting. Similarly, the development of LLL+ was followed by a political agenda to enhance acceptance of LGBT in the school environment. Both programs have played a significant role in these recent developments and seem to serve a need. However, in order to get the best possible program outcomes, ideally a follow-up evaluation would now take place, as youth care workers and teachers are free to work with programs that received positive feedback from users, have been adjusted since by the program developers, but have not proven to be effective yet. Finally, the findings about the time constraints and the importance of student liking are lessons learned that are also relevant to a wider audience of program developers, beyond sex education. These findings are to be taken into account when developing any type of health behaviour change programs targeting adolescents.

References

- Commissie-Rouvoet (2013). *Eindbericht Commissie. Voorkomen seksueel misbruik in de jeugdzorg*. Utrecht: Nederlands Jeugdinstituut
- Commissie-Samson (2012). *Omringd door zorg, toch niet veilig*. Amsterdam: Boom
- Cook, S. M., & Cameron, S. T. (2015). Social issues of teenage pregnancy. *Obstetrics, Gynaecology & Reproductive Medicine*, 25(9), 243–248.
- Douglas, J. M., & Fenton, K. A. (2013). Understanding sexual health and its role in more effective prevention programs. *Public Health Reports*, 128, 1–4.
- Part, K., Moreau, C., Donati, S., Gissler, M., Fronteira, I., & Karro, H. (2013). Teenage pregnancies in the European Union in the context of legislation and youth sexual and reproductive health services. *Acta Obstet Gynecol Scand*, 92, 1395–1406. <http://doi.org/10.1111/aogs.12253>
- Rutgers (2013). *Seksualiteitsbeleid of struisvogelpolitiek? Model voor visie en beleid op seksualiteit, relaties en omgangsvormen 12+ voor de residentiële jeugdzorg*. Utrecht: Rutgers
- VO-raad (2017). Kamer: sancties voor scholen die geen aandacht besteden aan seksuele diversiteit. Retrieved from: <https://www.vo-raad.nl/nieuws/kamer-sancties-voor-scholen-die-geen-aandacht-besteden-aan-seksuele-diversiteit>
- World Health Organization. (2006). *Defining sexual health. Report of a technical consultation on sexual health 28-31 January 2002, Geneva*. Geneva: WHO press

Acknowledgements

Onderwijs geven, weekje congres, hartstikke leuk! Promotiecadeautje? Ik help wel mee. Een werkdag was zo gevuld, het proefschrift bepaald niet zo geschreven. Gelukkig hielden Fraukje en Rob *the big picture* wat beter in het oog dan ik in eerste instantie deed. Ik stond er zeker niet alleen voor!

Ik begin met Fraukje, die heeft na mij de meeste uren in dit proefschrift zitten. We hebben gedurende mijn PhD ups en downs gehad en de laatste maanden een prima lange-afstandsrelatie. Vanuit Michigan kwam je me in februari 2014 opzoeken in Houston, en hebben we een leuke mini roadtrip naar San Antonio en Austin gemaakt, met aansluitend het SPSP congres. Wat soms voelde als eindeloze feedback maakte mijn stukken steeds beter (en feedback van reviewers minimaal!) en ik heb veel van je geleerd. Dankjewel!

Rob: altijd nuchter, rationeel en constructief. Bedankt voor je gedetailleerde feedback en aanmoedelingen wanneer nodig. Bij Gerjo kun je de kamer binnen lopen denkend dat je een lastig probleem hebt, en enkele minuten later opgelucht weer buiten staan. Wordenaantal van papers, sollicitatie perikelen, Intervention Mapping: dank voor al je (geruststellende) hulp! Ook wil ik Gerard graag bedanken voor de privé crashcourse statistiek.

I would also like to thank the ‘leescommissie’ (the external assessors) for taking the time to read and assess this dissertation.

De collega’s van Rutgers die bij Make a Move betrokken waren: Yuri, Annelies, Marianne, Stans, en Willy. Vooral met Marianne heb ik menig telefoontje gepleegd. Dankjulliewel voor de samenwerking. Ook wil ik de trainers van Make a Move bedanken. Arno, Els, Ingrid, Marja, Martine, Odette, en Tineke. Ook dank aan alle jongens die hebben meegedaan en de medewerkers van de instellingen die het belang van een programma als Make a Move erkenden, zich binnen de organisatie hard maakten voor deelname aan het programma en onderzoek, me vragenlijsten lieten afnemen, etc.

De Lang Leve de Liefde mensen. Met de één had ik meer contact dan met de ander, maar toch: iedereen had een inbreng. Een tip hier, een docentencontact daar. Dank! Ook alle docenten en leerlingen die ik gesproken heb en Esri die een deel van de dataverzameling gedaan heeft; bedankt!

I very much enjoyed my stay at the University of Texas, School of Public Health and being surrounded by so many people working on sex education programs. Joining the meetings, talking with different colleagues, as well as the opportunities to come along for observations

in schools have been very valuable. Pat and Lon, thank you for your hospitality. I felt at home right away, learnt so much about the USA, and the evening you took me along to the Prison Show radio broadcast has been a memorable experience. Chris, I'm very grateful for your hospitality (incl. welcoming me at the airport with a hug), time, and thinking along. Thank you for your supervision! En Marieke, ik had me geen gezelligere fellow Dutchie kunnen wensen in Houston

Alicia, Bram, Gerda, Karlijn, Loes, Philippe, Sarah, Stefan, Suzanne, en alle anderen. Ik heb goede herinneringen aan alle lunches op de lelijke zwarte banken en 101 verschillende onderwerpen die daar ter sprake zijn gekomen. Mariëlla, bedankt voor alle praktische ondersteuning en praatjes over vakanties. Mijn ex- kamergenootjes Dilana en Birthe. We hebben veel thee, pepernoten, frustraties en succesjes gedeeld en het was heel fijn met jullie hardop na te kunnen denken. Feziwe, Gilad, and Mehdi, thanks for sharing a wonderful international office! Wishing you all the best, wherever that may be!

Joanne, mijn masterscriptie begeleidster bij Rutgers (toen nog WPF) in 2011. Jij ook bedankt voor je oprechtheid. Je was je proefschrift aan het afronden en had al een nieuwe baan bij WUR toen mijn contract als onderzoeksassistent werd omgezet in een promotieplek. Ik weet nog hoe je me aankeek toen ik dat enthousiast vertelde en je in alle eerlijkheid zei: *"Ik zou op dit moment niemand een PhD aanraden...."* Die fase(s) heb ik ook gekend.

En dan buiten het werk. Marjon en Aïda, we zijn samen een voetbalcarrière gestart en hebben deze ook weer beëindigd. We fietsten 2 à 3 keer per week ouders en liefdesperikelen besprekend en zingend over de dijk naar Huissen maar zijn nu alle drie een andere richting ingeslagen (sja, niet iedereen wil in Huissen wonen, toch Aïda?). We zien elkaar niet meer zo frequent maar onze vriendschap blijft voor mij heel waardevol. Naomi, al meer dan 20 jaar vriendinnen sinds paardrijden. Wat hebben we veel gekletst op de trampoline in Westervoort, op uitjes samen, etc. Ik ben inmiddels bevallen van dit proefschrift, jij van lieve Jules! Ik ben benieuwd wat de volgende 20 jaar ons gaat brengen en hoe we daar over bijkletsen.

Anne, Eline, Ellen, Karlien, Lotte en Marjon, ik heb ontzettend goede herinneringen aan het Scharnerweg tijdperk en het is altijd direct weer gezellig als we elkaar weer zien. Myreen, oud-buuf, Dionne, Anke, Suzan, Esther, Maartje en Kimbria 2. Dank ook voor jullie vriendschap die een decennium Maastricht een heel leuk tijdperk gemaakt heeft. En wat was het leuk toen Karlien weer terug kwam in Maastricht en we na het C1 tijdperk weer even met z'n drietjes waren!

Het Make a Move onderzoek heeft me laten zien hoe bijzonder het is dat ik in een ‘gewoon’ gezin ben opgegroeid, waarin ik mijn verhaal kon doen tijdens het eten. Ellen en ik zijn ook opgegroeid met 3x twee tijdelijke Wit-Russische zusjes en veel vakanties. Dat zijn allemaal heel waardevolle ervaringen geweest. Ook was er thuis een sfeer waarin het vanzelfsprekend dat ik ging studeren, en het niet uitmaakte wat en waar ik dat wilde doen. Bedankt daarvoor, papa en mama. En dan mijn kleine grote zusje Ellen. Weet je nog dat ik vroeger een kop groter was en met m’n kin over jouw hoofd kon schuren? De afgelopen jaren hebben we elkaar bezocht in Mexico, Japan en Engeland, en nu sta je naast me als paranimf, zonder hoge hakken gelukkig. Dankjewel!

Jorinde, mijn andere paranimf. Jaren geleden zei je eens: sja, als ik op een zondagmiddag aan mijn interventie kan werken of een tijdschrift kan lezen, doe ik natuurlijk het eerste! Oei, dacht ik. Dit is het verschil tussen wel en niet succesvol. Wat krijg jij toch veel voor elkaar met je daadkracht, veerkracht, pech, en positiviteit. Vorig jaar was ik jouw paranimf, en jij nu de mijne. Dank voor je support!

CIS colleagues: alternating the PhD days with days in the office kept me sane. I might not be a big fan of Lego Serious Play but I am happy to be part of the Centre!

Thanks also to the Petrou family for having been so warm and welcoming right from the start and making me feel at home in England.

And finally Jacob. Suddenly you were there: good looking, chatty, funny, and passionate about tennis. And did I mention the head band and the soft hands? You finished the coaching with the words *lovely to meet you* - which I only later understood didn’t mean that much... The following week we had our first date and since then we have gotten to know each other inside out, making the above mentioned characteristics seem superficial. Thank you for your love and support (which includes giving my PhD priority when dividing timeslots for using the desk on weekends) and for making me so happy. I promise I will be a bit more relaxed about wedding planning than I was at the end of writing this dissertation.....!

Curriculum Vitae

Curriculum Vitae

Sanne van Lieshout werd geboren op 14 november 1987 in Arnhem. Na het afronden van het tweetalig VWO aan het Lorentz Lyceum, verhuisde ze in 2006 naar Maastricht voor de Bachelor European Public Health, die ze in 2009 afrondde. Daarna deed ze een tussenjaar, waarin ze o.a. een stage bij het Ministerie van VWS deed, voor de GGD Zuid Limburg werkte, en een maand naar Uganda ging. In 2010 startte ze met de Master Global Health, waarvoor ze in 2011 naar Manipal University in India ging voor een symposium en naar Kenia voor haar onderzoeksstage. Deze stage was bij Rutgers in Utrecht en hun partnerorganisatie CSA (Centre for the Study of Adolescence) in Nairobi. In januari 2012 begon ze als parttime onderzoeksassistent bij de sectie Toegepaste Sociale Psychologie, wat in mei 2013 werd omgezet in een fulltime promotietraject. Tijdens haar PhD bracht Sanne twee maanden door aan de University of Texas, Center for Health Promotion and Prevention Research onder begeleiding van Dr. Christine Markham. Momenteel woont Sanne in Engeland en werkt ze bij the Centre for Implementation Science, King's College London, waar ze een programma evalueert dat als doelt heeft om agressie en geweld te verminderen op GGZ opname-afdelingen.

Sanne van Lieshout was born in Arnhem, the Netherlands, on November 14th 1987. After completing her bilingual secondary school education (tweetalig VWO) at the Lorentz Lyceum in Arnhem, she moved to Maastricht in 2006 for the Bachelor European Public Health, which she finished in 2009. She then took a gap year in which she interned for the Dutch Ministry of Health, Welfare and Sports (VWS), worked for the Municipal Health Services (GGD) in Zuid Limburg, and spent a month in Uganda. In 2010 she started her Master's degree in Global Health that included a symposium at Manipal University in India, and a thesis internship at Rutgers (Utrecht) and their partner NGO Centre for the Study of Adolescence (CSA) in Nairobi, Kenya. In January 2012 she started working as a part-time research assistant at the section of Applied Social Psychology on the studies in this dissertation, which transformed into a full-time PhD position in May 2013. During her PhD she spent two months at the Center for Health Promotion and Prevention Research at the University of Texas in Houston, under supervision of Dr. Christine Markham. Sanne is currently living in England and working at the Centre for Implementation Science, King's College London, where she evaluates an intervention to reduce violence and aggression on in-patient mental health wards in two NHS Trusts.