

# Alcohol use among Norwegian 8<sup>th</sup> graders:

*an evaluation of implementation quality in a  
social influence intervention design*

Even Myrtrøen



Masterthesis in Health- and Social Psychology  
Institute of Psychology

UNIVERSITY IN OSLO

November 2013

Alcohol use among Norwegian 8<sup>th</sup> graders:

*An evaluation of implementation quality in a social  
influence intervention design*

© Even Myrtrøen

2013

Alcohol use among Norwegian 8<sup>th</sup> graders: an evaluation of implementation quality in a social influence intervention design

Even Myrtrøen

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

# Abstract

Data collected from the alcohol prevention program «Adolescents and Alcohol» [Ungdom og Alkohol] among Norwegian 8<sup>th</sup> grade students, was analyzed to assess implementation quality. Only data from the intervention group; 586 students, was utilized, to see if aspects of the implementation could predict positive program outcomes reflected in changes in self-reported alcohol consumption. As many students yet have to try alcohol for the first time at age 14, alcohol expectancies are already established, and predict alcohol use. Alcohol expectancies were therefore used as an alternative measure to alcohol use. Durlak and DuPre`s (2008) framework for measuring implementation quality was used, and student self-reports reflect implementation characteristics presented in this framework. Results from analyzing individual scores indicate that a certain amount of teacher control is positive, as well as seeing that the whole video project intervention was delivered in full. These results indicate that fidelity and dosage are aspects of the implementation which have an influence on alcohol use in this study. Results differ when scores are aggregated to represent whole classes instead of individuals; classwise meanscores. More student involvement predicts decrease in both alcohol use and alcohol expectancies, and these results reflect quality of program deliverance and fidelity. No gender differences in alcohol use were reported, but differential alcohol expectancy scores between T1 and T2, showed a larger increase in boys than girls. The conclusions point to several important aspects with implementing the “Adolescents and Alcohol” prevention program, regarding aspects of program delivery and staff characteristics. The conclusions support earlier implementation literature, which states that implementation influences program outcome, and that steps should be taken to ensure that implementation is assessed and verified.

# Preface

A big thank you goes to my supervisor Henrik Natvig, for giving me access to the Adolescents and alcohol [Ungdom og Alkohol] data material. Apart from that, proper thanks are needed for your valuable advice, your flexibility, and your quick responses.

Writing this thesis was interrupted by an internship in the World Health Organization, where I worked with statistics and prevention strategies on alcohol and other substances. I believe this experience might have given me knowledge, ideas and motivation, which is reflected in the thesis. Thanks to my coworkers in the Department of Mental Health and Substance abuse for letting me in, and for your advice and support regarding my work on this thesis.

Thanks also to friends and fellow students at the institute of Psychology, for general support, coffee breaks and motivational discussions. And thanks to my friend Aiden for his technical support.

Oslo, November 2013

Even Myrtrøen

# Contents

Introduction .....	1
The present study: Adolescents and Alcohol [Ungdom og Alkohol].....	2
Implementation.....	3
Young people and alcohol prevention in Norway.....	3
What is implementation (?).....	4
Why measure implementation (?).....	5
The importance of implementation.....	10
Measuring implementation, and factors affecting the implementation.....	11
Alcohol use and alcohol expectancies.....	14
Adolescent alcohol use.....	14
Consequences of alcohol.....	16
Adolescent alcohol expectancies.....	17
The role of social factors.....	20
The intervention.....	22
Methods.....	23
The "Adolescents and alcohol- survey".....	23
Measuring alcohol use and alcohol expectancies .....	24
Measuring implementation.....	25
Use of statistics and analytical procedures.....	27
Results.....	29
Descriptives.....	29
Alcohol use and alcohol expectancies.....	30
Significant component analysis.....	30
Individual implementation scores.....	32
Classwise implementation scores.....	33
Alcohol and gender.....	36
Summary of results.....	37
Discussion .....	39
Implementation and alcohol use.....	39

Implementation and alcohol expectancies.....	41
Gender differences.....	43
What can be learned from the implementation (?). .....	43
Conclusions.....	46
References .....	47





# Introduction

...

The World Health Organization is clear when it comes to harmful use of alcohol: It is a public health problem (World Health Organization, 2013). The Institute of Health Metrics and Evaluation (IHME) calculated the burden of disease attributable to leading risk factors in 2010, published in *The Lancet*; they found alcohol use ranked 5<sup>th</sup>. Slightly higher for men, and slightly lower for women (Lim, Lopez, Murray, & Ezzati, 2012). The same article shows that alcohol's role as a risk factor has increased, compared to numbers from 1990. In 1990 it was ranked 8<sup>th</sup>. The finding suggests that alcohol is a growing problem area in public health. If we look at the global population aged 15-49, alcohol use would be the leading risk factor both in 1990 and 2010, as risk factors such as air pollution, high blood pressure and tobacco smoking are harder on the very young, and elderly population (Lim et al., 2012).

The calculated burden of a variety of diseases are expressed in Disability Adjusted Life Years (DALY's). These numbers have been endorsed by the WHO (World Health Organization, in press). The diagnostic picture is very complex for alcohol. The main risk outcomes of alcohol use are injuries (both intentional, unintentional, and transport injuries), maternal disorders, mental disorders, behavioral disorders, HIV/Aids, tuberculosis, cirrhosis, cardiovascular disease, circulatory diseases, and cancer. There are several other diseases that may be caused by alcohol consumption, but these are the main ones. This myriad of possible outcomes makes it very difficult to measure the real effect of alcohol, in morbidity and mortality figures. The WHO estimates that every year, around 2.5 million people die as a result of harmful use of alcohol, and approximately 320,000 of these people are aged between 15-29. (World Health Organization, 2011). Only diseases where alcohol impact is measurable will turn up accurate mortality numbers. These are diseases such as liver cancer secondary to alcohol use, liver cirrhosis secondary to alcohol use, and alcohol use disorder (Lim et al., 2012). An example is a person who dies from AIDS, who may have been influenced by alcohol when they engaged in unsafe sexual conduct and was infected with HIV in the first place. Alcohol is a powerful risk factor, and it probably causes more morbidity and mortality indirectly than what is known.

## **The present study: Adolescents and Alcohol [Ungdom og Alkohol]**

This study builds on a study by Natvig and Aarø at the University of Oslo, which has been accepted, though still not publicized. They used data material from the questionnaire “Adolescents and Alcohol” [Ungdom og Alkohol], with questions answered at two different times (4 months intermission), by 8<sup>th</sup> graders (14 year olds) on schools throughout Norway. The questionnaire reveals self-reported alcohol use, and a wide range of variables linked with present or future alcohol use. Parts of the respondents received an intervention between T1 and T2, where change scores from T1 to T2 would tell if the intervention had an effect on alcohol use among participants in the intervention group. The intervention was twofold, but both approaches grounded in social influence for behavior change. The intervention group classes from different schools, competed against each other in making a three-minute video, where they advised 7<sup>th</sup> graders against starting to drink alcohol. This exercise was thought to manipulate the 8<sup>th</sup> graders` own attitudes towards alcohol, to become more anti-alcohol. They are “forced” to publicly argue against alcohol, and in turn that may lead them to practice what they preach. The second intervention was that the same group was presented with feedback about actual alcohol use among 9<sup>th</sup> and 10<sup>th</sup> graders at their own and other schools. In that age group it is common to overestimate peer group alcohol use. When the 8<sup>th</sup> graders see that actual use is below what they thought, they will need to correct their beliefs about how “common” alcohol use is, and thereby the normative drinking pressure will be reduced. The study`s overall aim is to make adolescents drink less than what they do, or make them postpone the age in which they start drinking alcohol.

The “Adolescents and Alcohol” questionnaire also comprised several questions about teacher- and student involvement in the video project, as well as perceived engagement and satisfaction with the project. This is valuable information regarding the implementation of the intervention. This feedback makes it possible to link up variations in alcohol use with implementation quality. It shows that there are several concerns that need attention before a prevention program is deemed effective or ineffective. Implementation will be outlined in detail below. The study showed a behavioral effect; the intervention group had a small but significantly lower increase in alcohol use between T1 and T2, compared to the control group (Natvig & Aarø, in press). The control group respondents did not receive any alcohol preventive activities, besides answering the questionnaire. First, implementation will be outlined, followed by an overview on adolescent alcohol use and alcohol expectancies.

# Implementation

Development of effective interventions, are commonly regarded as the first important step towards health improvement. The following will highlight implications that arose with implementing the intervention.

## **Young people and alcohol prevention in Norway**

There are several institutions in Norway, which act towards enlightening adolescents on the dangers involved in drinking alcohol. The Norwegian directorate of health comprises an extensive overview of knowledge, statistics, and advice in matters of alcohol. The webportal; [www.helsenorge.no](http://www.helsenorge.no) is easily available to adolescents, due to the generally high knowledge of using the internet amongst this demographic. The website publicizes articles that they have received from other institutions, and state their sources on these. Here, adolescents can read about how to handle drinking pressure, correcting fallacy beliefs about alcohol, and the association between alcohol and violence, along with other aspects of alcohol as well (helsenorge.no, 2013). There are also suggestions as to how to cope with pressure, and organizations of interest and helplines for further help and are among those embedded on the website (helsenorge.no, 2013).

These institutions are trying to prevent alcohol use at the societal level, by making the information known and available. This method is also an intervention, but it operates on another level than the intervention in this study. A health intervention can be implemented on all levels of the social ecological paradigm (Bartholomew, Parcel, Kok, Gottlieb, & Fernández, 2011). Levels are intertwined with one another, and an intervention implemented at the individual level, for instance, does not emit effects to that level only. An alcohol intervention aimed at one single individual may cause alcohol abstain in his or her closest circle. They may form a group that organize themselves toward working pro abstaining from alcohol among adolescents. The larger and more influential the group grows, the greater is the chance they will be heard in their community. The group may in turn, hypothetically, influence to pass a law on alcohol, which changes individual health behavior. This visualization is merely to show that an intervention aimed at any level may have effect on that

level, and all the levels that are nested within it. Every individual has interpersonal relations, is a part of a community, and a larger society. A society consists of communities, organizations, interpersonal relations, and individuals (Bartholomew et al., 2011). The researcher needs to be aware of an array of factors, which may influence the intervention in one way or other.

### **What is implementation (?)**

Health interventions can be compared to the introduction of any consumer product entering into the market. Durlak and DuPre use a well-known dictum from the business world, “Build a better mouse trap, and the world will beat a path to your door”, to signalize that the new product needs to be better than what we already have, to at least one extent. As with consumer products, this intervention goes through a process called diffusion (Durlak & DuPre, 2008). Diffusion is a series of steps going from dissemination and adoption to implementation, and finally sustainability. The health intervention needs to be introduced and adopted by the target group. Then it needs to be tried out and conducted appropriately, as well as it needs to be effective. Last is the sustainability, to ensure that the product is being maintained over time. Unfortunately, many interventions yield diminishing returns, because one or more of the diffusion phases fail. A program can be correctly adopted in a community, but if it's not adequately implemented it will not be effective and function as intended, and it will not be sustained. But even a successful implementation does not guarantee a sustainable, long lasting program (Durlak & DuPre, 2008).

In this text, only the implementation phase will be thoroughly examined. Based on information retrieved in Natvig & Aarø's study (Natvig & Aarø, in press), implementation is the only phase of the diffusion process that can be adequately investigated. As to dissemination and adoption, schools were first made aware of the intervention program's existence, and it was supplied to each of the volunteering schools. It's also hard to predict intervention outcome sustainability without follow-up studies. Sustainability relies heavily on the other diffusion phases, and especially on implementation. Designs with documented thorough implementation are more likely to be sustained than those which are not (Durlak & DuPre, 2008).

Research on implementation has grown considerably over the last decades. But still, several meta-analyses conclude with a lack of documented implementation in health intervention research. Dane and Schneider found that only 39 of 162 published mental health studies between 1980 and 1994, had documentation of implementation fidelity. Fidelity is one of several important implementation aspects, which will be explained in depth later in this text. Albeit 39 studies showed that steps had been taken to control for and verify program integrity, only 13 of these tried to document implementation effect (Dane & Schneider, 1998). The low number in both documentation and program efficacy, suggest that many programs face the risk of compromising internal validity (1998). Durlak conducted a search among all prevention studies he could find from 1956 to 1995, and found that 5% of 1200 studies provided data on program implementation (Durlak, 1997; Durlak & DuPre, 2008). In 2000, Domitrovich and Greenberg conducted a meta-study on evidence based mental health prevention programs, and found that very few of the programs investigated implementation. Almost all programs paid attention to implementation, but to very varying degrees. Only 21% of the programs assessed more than one implementation dimension (usually fidelity and dosage). One third of the programs linked variability in implementation to differences in program outcome. These results imply that program designers are aware of possible implementation effects, but that either they know too little about assessing them, lack the knowledge of the importance of considering all dimensions of an implementation, or just don't take it seriously enough (Domitrovich & Greenberg, 2000).

### **Why measure implementation (?)**

Implementation research makes it possible to know what actually happens during an intervention. This includes what takes place, quality of the program delivered, and if the target audience was reached (Domitrovich & Greenberg, 2000). A very good reason to conduct a thorough implementation evaluation is to ensure internal validity. It also strengthens the program's role in producing change (Domitrovich & Greenberg, 2000; Durlak, 1997). When we demonstrate a new program, we want to be as sure as possible to the extent of differences caused by the program. There may be outcomes which are not really caused by the design of the program, but which can be related to program delivery, and thus get misinterpreted. This is called a Type III error, and occurs when researchers assume that effects of an intervention are meaningful, when in reality poor program delivery has

invalidated outcome analyses (Dobson & Cook, 1980; Domitrovich & Greenberg, 2000). It may also be that a program is wrongly judged ineffective, when negative outcome analysis is due to shortcomings in delivery.

Implementation research helps the researcher see strengths and weaknesses in the program. An overall picture of how program pieces fit together, and how the users interact (teachers and students) and solve problems, can be of great help understanding program consequences, both anticipated and unanticipated ones. Another rationale for implementation research is that this data gives feedback which is valuable both for the present and future intervention(s). Knowledge generated from implementation research helps advancing the whole field of prevention program developing. It is important that practices on replication, maintaining, and diffusing programs in real world settings are established (Domitrovich & Greenberg, 2000).

As outlined by Durlak & DuPre, there are five main dimensions of implementation, to verify program/intervention integrity. These are *fidelity* (in implementation literature also referred to as *integrity* or *adherence*), *dosage* (also referred to as *exposure*), *quality of program delivery*, *participant responsiveness*, and *program differentiation* (Dane & Schneider, 1998; Domitrovich & Greenberg, 2000; Durlak & DuPre, 2008). It is important to make a few distinctions about the fidelity concept, as implementation literature varies in the conceptual use of it. Proctor and colleagues for instance, have developed a taxonomy including fidelity as a designator to how well the prevention program was delivered as intended, where Durlak & DuPre use *integrity* (Proctor et al., 2011). Dusenbury and colleagues are among the researchers who structure fidelity in the same way as Proctor et al. (2011) (Dusenbury, Brannigan, Falco, & Hansen, 2003). Fidelity according to Proctor et al. is just one of several implementation outcomes in the taxonomy. Alongside are the likes of “acceptability”; stakeholders` satisfaction with design, content, comfort and credibility, and “feasibility”; suitability for everyday use and utility. Also included is “appropriateness, “adoption”, “implementation cost”, “penetration” and “sustainability”. This illustrates that the integrity of the intervention is one of many constructs in implementation literature, but the only outcome to be treated in the current study. The reason for this is twofold: (1) This study, based on Natvig & Aarø`s study, is basically interested in investigating program fit, when moving from the paper to a real world setting for the first time. Traditionally, fidelity (integrity) has been the most measured and the main concern for researchers (Proctor et al., 2011). Also the intervention`s target is to first and foremost provoke an effect among the consumers (the

pupils), to test for program efficacy. Many of Proctor et al.'s suggested outcomes operate on institutional/organizational level or solely on the provider side. Whether the school administration accepts the intervention design or expenses associated with it, is not really an issue here. The study was supported by the Norwegian Directorate of Health, and schools volunteered to take part. Another rationale not to include extra outcomes is based on overlapping constructs. Measures on the provider level for instance are well covered in the questionnaire; ratings on teacher control and engagement, which is integral in integrity, i.e. the fidelity outcome. "Acceptability" and "appropriateness" for the intervention is reflected in attendance, interest and engagement in the intervention, i.e. dosage and participant responsiveness outcome. (2) The other reason to use only fidelity/integrity is down to available constructs. The available constructs measure all the seemingly important parts of intervention integrity. Intervention integrity level is possible to obtain through self-report, whilst other possible implementation outcomes may need observational data (adoption), administrative data (implementation cost), or interview data (acceptability and appropriateness for instance) (Proctor et al., 2011). Implementation outcomes needed for rational and necessary evaluation, as well as available constructs for measurement, is the reason only integrity will be investigated, based on five constructs.

Proctor et al. (2011) divide fidelity in adherence, dosage, quality of delivery, program differentiation and participant responsiveness. Proctor et al. define fidelity on a different level than Durlak & DuPre; as an overall designation on intervention implementation integrity, with underlying implementation variables with different targets, all adding up to the fidelity outcome. It is easy to get confused when different taxonomies are presented, using a different wording or setup level. In this text however I intend to stick with Durlak & DuPre's (2008) descriptions, and use intervention design or integrity as the overall intervention implementation designation. The fidelity concept will be treated as what Proctor et al. (2011) and Dusenbury et al. (2003) call adherence.

*Fidelity* refers to how well the program that was delivered fits with the original intended program. Fidelity data need to be verified, otherwise; it will be hard to attribute changes in dependent variable to manipulations on the independent variable. If you know that there is fidelity between delivered and original program, at least you will be able to control program effects, and assess the effects from program activities (Dane & Schneider, 1998; Moncher & Prinz, 1991). This is especially important in schools, where many prevention programs are

being conducted. Based on this, fidelity may be at risk also in the current study. The implementers are usually teachers or other para-professionals, who may lack the knowledge to play program leaders and follow required procedures (Dane & Schneider, 1998). In addition, school based interventions often compete with regular subjects for class time, in an already busy school schedule. Teachers need to acknowledge that the program needs to take as long as necessary, and not fall for the temptation to rush it, implement it incompletely, or completely discard it. It is important that the program is supported by both school administration and teachers, as this will strengthen the incentive to fit it into the schedule and complete it. The ideal situation would be to have professionals conduct all of the interventions. But, with a typical low cost budget and a broad scope of the prevention program, the researcher(s) cannot possibly conduct or oversee all the interventions. A high number of classes distributed in several parts of the country are often required, to gain sufficient numbers and geographical representation.

Studies reveal that loss of fidelity may reduce program efficacy (G. J. Botvin, Baker, Dusenbury, Tortu, & Botvin, 1990; G. J. Botvin, Dusenbury, Baker, Jamesortiz, & Botvin, 1992; Dane & Schneider, 1998). Moreover, Botvin and colleagues found that inconsistencies in program delivery can hide significant program effects (G. J. Botvin et al., 1990; G. J. Botvin, Dusenbury, Baker, James-Ortiz, & Kerner, 1989; Dane & Schneider, 1998). By verifying fidelity data, the researcher can close down or make adjustments to a program knowing that the programs fidelity was maintained. By not verifying, there will always be the risk of discontinuing a potentially effective intervention (Dane & Schneider, 1998; Felner, Phillips, DuBois, & Lease, 1991). As mentioned earlier, by ignoring or not verifying fidelity data, the whole intervention runs the risk of committing a Type III error. Natvig compares this phenomenon with a problem that has occurred when developed countries experience that well tested antibiotics are ineffective. Often, the medicine in itself works, but the effect does not show when people fail to use them as prescribed (Natvig, 2009). As with antibiotics, to accurately test program effects in the current intervention; we need to verify fidelity.

*Dosage* refers to how much of the original program that was delivered. An intervention that is conducted in full among all respondents will score high on the dosage dimension. It refers to both frequency and duration, and whether these numbers fit with the amount of intervention prescribed in the program design (Carroll et al., 2007). As with fidelity, inconsistencies in dosage threaten internal validity. Dane and Schneider (1998) analyzed several studies, linking



reported dosage level with program outcome. Three of six studies showed clear relationship between high dosage and outcome measures (Allen, Philliber, & Hoggson, 1990; Dane & Schneider, 1998; Felner et al., 1994; Weinman, Schreiber, & Robinson, 1992) . However, dosage effects on implementation seem to be more evident in clinical trials, rather than school based. Overall, research suggests that dosage plays an important role in verifying program integrity. The danger seems to be when a substantial proportion of respondents are absent, and especially in sessions where key concepts of an intervention are being outlined (Dane & Schneider, 1998).

*Quality of program delivery* measures qualitative aspects of the implementation such as implementer enthusiasm and attitudes toward the *intervention* (Dane & Schneider, 1998). The more involved and interested participants are in program activities, the more quality will be reflected in the program delivery. *Participant responsiveness* is a measure that relates to negative or positive response, and interest, attention and enthusiasm to the intervention, among the participants (Dane & Schneider, 1998) *Program differentiation* is a manipulation check, to safeguard that subjects in the intervention group received only planned interventions. In other words; to be sure everything was conducted according to plan, and to protect the program`s uniqueness (Dane & Schneider, 1998).

In addition to the aforementioned dimensions of an implementation, Durlak and DuPre outline another three facets worthy of attention (2008). *Monitoring control- or comparison group(s)*; to observe the tutoring, and the nature of care and services given. Often, it is assumed that control groups receive no alternative services in school based studies. This is almost never the case, as many individuals receive some alternative services (Durlak, 1995). *Program reach* refers to rate of involvement and representativeness among the participants. It differs from participant responsiveness, because *reach* is more concerned with comparing the *types* of participants in one of several interventions, with the original program (Durlak & DuPre, 2008). *Adaptation* is the last dimension, and refers to possible changes in the implementation from the original program. During an implementation, there may occur a need to do modifications to the intervention, depending on the situation (Durlak & DuPre, 2008). Of all the mentioned dimensions, *fidelity*, *dosage*, and also *quality* are the most widely studied. Nevertheless researchers agree that all the five “main-dimensions” are important to gain a comprehensive view of, or a complete picture; of the implementation process (Carroll

et al., 2007; Durlak & DuPre, 2008). Based on this, it is the five first dimensions that will be given attention in this text.

### **The importance of implementation**

As already outlined, we measure implementation to be as sure as possible, that the program/intervention has been implemented according to the program's design. If we find inconsistencies, at least we will be able to locate and control deviations. Otherwise, if no steps are taken to measure implementation, we won't have any facts to back us up- only assuming why or why not the target program seemed to work as intended. Another reason to take implementation seriously is that there is clear evidence that implementation influences program outcomes. Smith et al. (2004) found in a meta-study on school-based anti-bullying programs, that the programs obtained modest effect levels overall. But, programs monitoring implementation obtained twice the effect size than those which did not monitor implementation (Smith, Schneider, Smith, & Ananiadou, 2004). In another meta-study, Wilson et al. found implementation to be the second most important variable overall, and the most important program feature influencing program outcomes (Wilson, Lipsey, & Derzon, 2003). All the 221 programs reviewed here were school based programs targeting aggressive behavior. Another meta study by Derzon and colleagues (2005) found that the three factors with largest impact on outcome, were all related to implementation. Two factors were directly related to implementing the intervention (program consistency and intensity in delivery), and the third factor was that the control group received alternative drug prevention services (Derzon, Sale, Springer, & Brounstein, 2005). Derzon et al. used regression procedures to reestimate program outcome if these three factors were being controlled, and found that programs would be up to 12 times more effective; given perfect control of these three factors (Derzon et al., 2005).

Meta studies on prevention programs outside school, show similar results. DuBois and colleagues found in a review of 59 mentoring studies, that studies monitoring implementation obtained effect sizes three times bigger than studies reporting no implementation monitoring (mean effects 0.18 vs. 0.06) (DuBois, Holloway, Valentine, & Cooper, 2002). When it comes to the impact implementation plays on program outcome, Durlak and DuPre (2008) concluded that implementation plays a significant role. Mean effect sizes are two to three times higher

when the program is carefully implemented and free from serious implementation problems, compared to when these circumstances are absent. They also confirm that higher levels of implementation are often associated with better outcomes, and this is especially the case when both the fidelity and dosage dimensions are assessed (Durlak & DuPre, 2008). Ennett et al. (2011) confirm that fidelity and dosage (adherence and exposure) are the two domains most important in deciding whether a program is properly implemented or not, statistically speaking. Ennett further warns against expecting to achieve complete fidelity. This is unreasonable, and has not even been demonstrated under rigorous research conditions (Durlak & DuPre, 2008; Ennett et al., 2011).

### **Measuring implementation, and factors affecting the implementation**

The “Adolescents and alcohol” questionnaire provides several questions linked to implementation outcome. Students in the intervention group answered several questions related to their work with the movie. Answers to these questions rate level of control of both the implementer- and the respondents` side, as well as how involving, engaging and how interesting it was, and whether the movie project was finished in full or not. Table 1 in the methods section (p. 26), illustrates how implementation indicators represent one or more implementation outcome variable(s). Further details on measurement are presented in the methods section.

By looking at each of the implementation variables, and testing them with alcohol use change scores, it may be possible to discern which part of the implementation was more or less successful. A working taxonomy is useful because it provides clear definitions, and science cannot measure what it does not define. It also provides a similar framework to address implementation quality research to that of earlier frameworks (Durlak & DuPre, 2008), using the same language and conceptual definitions. As the current implementation literature consists of slightly different terminologies; researchers vary in the conceptual use of terms. I find it important to stick with Durlak and DuPre`s (2008) standards to avoid confusion.

Wilhelmsen, Laberg and Klepp (1994) compared two structurally different alcohol prevention programs aimed at Norwegian 7<sup>th</sup> graders. They found a significant interaction between alcohol use and time ( $F(2/838)=6.22, p<.01$ ), where a highly role-specified condition (HRS)

caused a decrease in alcohol use, and a less role-specified condition (LRS) caused an increase in alcohol use. There was no change in the comparison group. In the HRS condition, the teacher and two class selected peers were in charge regarding overseeing and implementing pre-planned activities in detail. This is a marked distinction to the LRS condition, where the teacher and two class selected peers were trained to merely cooperate in implementing pre-planned activities, in less detail than the HRS condition. The same activities were implemented, but implemented in a slightly different manner, regarding degree of controlling roles (Wilhelmsen, Laberg, & Klepp, 1994). The clear difference in the outcome irrespective of the same program being implemented in the end, supports literature reporting that how the implementation takes place, is highly relevant (Dobson & Cook, 1980; Domitrovich & Greenberg, 2000).

Wilhelmsen et al. (1994) further explains that students in the HRS condition felt more involved in the program, and felt more satisfaction with the group work and program participation, based on self-reports, than did the students in the LRS and comparison groups. Degree of involvement is in the current study well covered in implementation variables 3, 4 and 5 (see Table 1, p. 26 for an overview). Implementation variable 2 is well suited to discriminate between roles, i.e. whether students or the teacher control activities associated with the video project. However, Wilhelmsen does not specify whether high role specificity played a more important part on the class selected students or on the teacher. Does it matter most who were assigned which controlling roles, or just that “someone” was in charge of implementing the activities, rather than no one? Wilhelmsen acknowledges that a more specified teacher role (clearer leader) yielded positive alcohol prevention results, but at the same time admits it was an unexpected result. The reason for this is that educational theory advocates student responsibility for filling in details in a preventive program, rather than teachers taking the responsibility away from students (Wilhelmsen et al., 1994). An evaluation of teachers implementing a substance abuse program in American 7<sup>th</sup> grade classes, supports Wilhelmsen’s notion of more teacher control in program implementation leading to more desired results (Pettigrew et al., 2013). Pettigrew et al. (2013) found that authoritative teachers, who gave clear instructions about activities, had most success in delivering the program as intended. Another finding is that the more involved/engaged the teacher is, the more students tended to participate and involve themselves in the program activities. This is also in line with Wilhelmsen et al.’s (1994) finding that students were more

involved with the program when in the HRS condition, than were students in the LRS condition, based on self-reports.

Of other factors affecting the implementation, a challenge may arise as a result of staff turnover, and other organizational issues related to staff. Natvig explains that many questionnaires were never returned at T2, because teachers dropped out of the project (Natvig & Aarø, in press). Staff challenges may also influence attendance in the classes where the intervention was conducted, and it might influence on whether the video project was completed or not, i.e. the dosage outcome. Cross and colleagues (2010) emphasize the importance of staff quality, both in terms of education, level of training, and long-term service, because it seems to affect aspects of implementation. The study investigated implementation quality in after-school programs, where high staff turnover has been problematic (Cross, Gottfredson, Wilson, Rorie, & Connell, 2010). But, staff quality/changes can also influence this study. A change in staff; for instance when a teacher is replaced, may influence interest rates, involvement, quality of delivery, and ultimately if the program continues at all. A replacement teacher may not be briefed on the project, or know how he/she is supposed to play the role. There will also be personality differences between teachers that predict levels of control and involvement, which in turn influences how well the program was delivered and how it fits with the originally intended program (fidelity). Unfortunately, there are neither qualitative measures on teacher behavior, nor self-report data from teachers. Results from the student self-reports will hopefully help to shed light on whether staff quality has been affecting the project.

# Alcohol use and alcohol expectancies

## Adolescent alcohol use

The WHO states that alcohol consumption among people under 18 years of age, based on a five year trend conducted in 2008; has increased substantially (WHO, Global status report on alcohol and health 2011). Out of 73 responding countries 71% reported an increase, 4% reported a decrease, and 24% was stable or inconclusive. The data on young adults aged 18-25 reveal an 80% increase, 11% decrease and 18% stable or inconclusive trends (WHO, 2011). It is important to note that 18 is the most common and also the average national legal minimum age for on- and off premise sale of alcoholic beverages. For instance, the legal minimum age in Portugal is 16, and in the US it is 21. Norway was not one of the responding countries in this report. Alcohol consumption among 15-16 year olds in Norway in 2009, who had tasted alcohol, experienced a decline from the top years 1997-2000 (Vedøy & Skretting, 2009). In total consumption based on litres of pure alcohol, there is also a decline in 2009, compared to the top year 2003 (Vedøy & Skretting, 2009). The evidence of decreased alcohol prevalence among adolescents in Norway, is supported by the most recent ESPAD survey (Hibell et al., 2011). Unfortunately, many other European countries show inconclusive or increasing trends, especially southern and eastern Europe (Hibell et al., 2011). In Norway there have not been any changes in alcohol laws regarding minimum legal age nor distribution in this time frame, which can explain the decrease. A focus on other prevention activities has probably contributed to the alcohol changes we see in recent reports.

Many studies have investigated adolescents' motives for drinking, i.e. why they drink alcohol. Kuntsche et al. (2005) conducted a meta-study consisting of 82 articles from 1989 onwards, where drinking motives among young people 10-25 years old were studied. Motives were derived from both interviews and self-report. Results say most young people drink for social reasons, in the sense of experiencing enjoyment. This includes drinking for social facilitation, to get along well at social gatherings, or to get in a mood for partying; all in all enjoyment related, and not drinking for enhancement reasons or coping. This type of drinking behavior is considered the "healthiest" kind, as the majority of people who drink for social reasons consume a moderate amount of alcohol. An example from the US show that social drinking is negatively associated with intensity ( $b=-.25$ ) in drinking (Kuntsche, Knibbe, Gmel, &

Engels, 2005; Labouvie & Bates, 2002). At the same time, social motives drinkers were less likely to drink excessively ( $r = -.18$ ) compared to those who drank for enhancement or coping reasons (Karwacki & Bradley, 1996; Kuntsche et al., 2005). Based on these results it seems the majority of adolescent drinkers, drink moderately, and for social reasons. Kuntsche et al. (2005) does not provide any information on whether the statistics differ if you divide in age groups, say 14-16, 16-18 etc. As we know, many changes occur in a short time period during the adolescent years. Information on this issue would be interesting. On the other hand, there seems to be some cultural- and gender differences here. Results from Hong Kong indicate that among 15-17 year olds, girls drink for enjoyment, whereas more boys drink for the perceived enhancement effect of alcohol (Kuntsche et al., 2005; Lo & Globetti, 2000).

Alcohol use divided by gender is expected to produce no significant differences, in which would support ESPAD findings from the same year (Hibell et al., 2004). Student alcohol consumption among males in 2003 is higher than females in all European countries reported, but in Norway the gender distribution is reported to be about equal (Hibell et al., 2004). The Norwegian alcohol-picture 8 years later, in 2011, reveals the same pattern; consumption is equally distributed among genders. However, most other European countries see a significant shrinkage in the gender gap. Where males earlier used to uphold a higher consumption than females, the picture is now more unclear (Hibell et al., 2011).

Adolescents, whose drinking motive is enhancement or coping, are more at risk of alcohol related harm compared to social drinkers. When enhancement is the motivation, research reveals that respondents are heavier drinkers (opposed to social/moderate drinkers), who drink to enhance pleasant emotions (Kuntsche et al., 2005). Enhancement motive drinkers also answer "I drink to get drunk" or "I like the feeling of drunkenness". Descriptive scales like these have shown to be good discriminants between heavy and moderate drinking. Heavy drinkers, 14-18 year olds, are three times more likely to state that they "drink to get drunk" than do moderate drinkers (Feldman, Harvey, Holowaty, & Shortt, 1999; Kuntsche et al., 2005). The last motive category involves drinkers who drink for coping reasons. Tension reduction drinking, drinking for escape motives, or drinking to attenuate negative affect, are all drinking strategies associated with heavy drinking (Kuntsche et al., 2005). Kuntsche's study reveal correlations between all three categories, as adolescents tended to give multiple reasons for drinking. For example; drinking more to get in the mood for partying, or drinking to get rid of negative feelings, in order to be sociable and enjoy the party. Even though the

associations between motives are intertwined, social reasons, enhancement and coping stand out as the clearest clusters of factors of motives.

### **Consequences of alcohol**

In a health perspective, drinking for coping is the more dangerous drinking motive. The reason for this is that coping motives are associated with alcohol-related problems (Kuntsche et al., 2005). There are reasons behind negative feelings or stress, which are not being adequately addressed through alcohol intake. Thus, alcohol will only foster new and possibly bigger problems for coping drinkers. Drinking may be effective in short term, but not in long term. Drinking for enhancement is only linked with alcohol-related problems when it is mediated by coping motives. If coping motives are controlled for, the link vanishes (Kuntsche et al., 2005). Enhancement drinkers still tend to be heavy drinkers, and run a higher risk of harming themselves than what moderate drinkers do, due to amounts of alcohol.

In regard to an alcohol prevention strategy, a study by Danielsson and colleagues (2011) show that young high consumers should be the target group. Data were used from the 2007 ESPAD survey (European School Survey Project on Alcohol and Other Drugs), and found that the minority group of heavy episodic drinkers (HED) among all consumers, reported the majority of health problems (Danielsson, Wennberg, Hibell, & Romelsjö, 2011). Heavy episodic drinkers are defined as persons who drink at least 60 grams of alcohol on one occasion, at least once per week (WHO definition). As mentioned earlier, the majorities of young drinkers, drink moderately, and run low health related risks due to their moderate intake. It is the heavy drinkers who run the high risks. A study by Oesterle et al. (2004) supports Danielsson et al's (2011) advice to reach for the young high consumers. There are different kinds of heavy episodic drinkers among adolescents. There are "chronic heavy drinkers", who have an established high consumption pattern already at age 13-14. The group "escalators" usually start around age 15, and increase rapidly from there on, and the group "late onsetters" drink heavier from age 16-17 (Oesterle et al., 2004). Oesterle et al. (2004) compared the pattern-groups at age 24 with a variety of health related outcomes, and found the chronic heavy drinkers worse off. This group is in high risk of long-term negative health consequences, and already at age 24 they were significantly above the other groups in regards such as obesity and blood pressure. The other groups did not show the same adverse health as the chronic



heavy drinkers; but they risk alcohol related health adversities later in life. This evidence points to why Natvig's study is highly relevant. The intervention maps out drinkers at a young age, but even more importantly, it may delay the age people start drinking alcohol and therefore reduce many alcohol related health problems in young people. As this study consists of eighth graders (14 years), alcohol consumption patterns may not yet have established. The data show that a majority of the respondents are still abstainers on time 1, or both on time 1 and time 2. Nevertheless, many pupils report having started experimenting with alcohol. The age 14/15 is usually the phase when many teenagers start experimenting, and is therefore considered to be a crucial age for prevention approaches (Natvig & Aarø, in press).

As a result of investigating 14-year olds, there may not be sufficient numbers on actual alcohol consumption to decide if the implementation of the intervention will yield results on alcohol use. As a consequence, measures on alcohol expectancies will also be tested, as adolescents report alcohol expectancies prior to first personal drinking experience. Alcohol expectancies will be outlined, as it predicts initiation of drinking, and thus can be adequately used after actual alcohol use (H. N. Aas, 1995).

### **Adolescent alcohol expectancies**

It has been reported for several decades, that alcohol outcome expectancies are associated with drinking behavior. Christiansen and colleagues found evidence in 1982, that well established expectancies exist, even before young people try alcohol for the first time (Christiansen, Goldman, & Inn, 1982). Results from the same study suggest that expectancies become clearer with age and drinking experience. The social learning process is not static, and experience with alcohol helps to crystallize the expectations. The social learning theory (Bandura, 1986), suggests that there is a reciprocity between alcohol use and cognitive factors, such as alcohol expectancies. Therefore, alcohol use both influences and is influenced by alcohol expectancies. Wilhelmsen et al. (1997) found in a study of 842 Norwegian 13-14 year olds, that more positive expectancies to the effects of alcohol were associated with lower intentions not to drink. More positive expectancies towards alcohol resulted in lower intentions to abstain from use. This is especially the case with alcohol expectancies which were aimed at social facilitation (social alcohol expectancies), and was found to predict intention not to drink, carrying a significant and moderate weight ( $\beta = -.47$ ) (Wilhelmsen,

Laberg, & Aas, 1997). In regard to gender, there is no evidence that one gender hold significantly more positive alcohol expectancies than the other gender (Schulte, Ramo, & Brown, 2009).

Brown and colleagues developed the first alcohol expectancy measure (AEQ) in 1980 (*Brown, Goldman, Inn, & Anderson, 1980; Christiansen et al., 1982*). A 90-item questionnaire that comprised 7 factor- analytically derived scales, to reveal the positive effects of drinking alcohol. Positive in the sense that more positive expectations resulted in heavier drinking, and more alcohol related problems. The AEQ-A was developed shortly after, which is the version to be used when assessing adolescents. Christiansen and colleagues confirmed much of their earlier work with a one year follow-up study, presenting the power of expectancies in predicting drinking behavior (Christiansen, Roehling, Smith, & Goldman, 1989). Out of the seven AEQ-A factor scales, five expectancy scores discriminated between non-problem drinkers and “potential”/or beginners at problem drinking. The five scales also accounted for a large portion of variance in both a frequency/quantity index, and a problem drinking index. These results lead to conclude that the scales have a strong relation to individual drinking decisions (Christiansen et al., 1989). Of all the scales, the most potent predictor of drinking behavior, was the social scale “alcohol can enhance or impede social behavior”, which is supported by later work (H. Aas, 1993; Wilhelmsen et al., 1997). Young people with high consumption, were most likely to hold strong expectancies to social facilitation for instance. The social scale and the global positive transforming agent- scale (also a potent predictor), are both included in Natvig & Aarø's study, but only the social scale will be included in this study. The social scale explained most variance in alcohol expectancies ( $r = .44$ ), and more than all the scales combined could account for ( $r = .35$ ) (H. Aas, 1993)

The alternative to AEQ-A is the Comprehensive Effects of Alcohol- scale (CEOA). The CEOA is in several respects reported to be the superior expectancy scale (Fromme & D'Amico, 2000), but it depends on what the researcher wants to identify. Overall CEOA was demonstrated to explain more variance in quantity (28%) than did the AEQ-A (20%). The two scales equaled in variance in frequency (15%) (Fromme & D'Amico, 2000). But, social expectancies using the CEOA points more towards subjective evaluations of the effects alcohol has on the individual, whereas AEQ-A points to the expected effects of alcohol in general. For example “drinking would make me friendlier” and “drinking makes people

friendlier”, illustrates scale differences in item specificity. Fromme and D’Amico (2000) support using the AEQ-A when assessing social expectancies, because it shows a stronger association with actual alcohol use. A possible explanation here may be that most adolescents’ lack of alcohol experience makes it more difficult to estimate the specific effects of alcohol. A more general idea of alcohol effects may at this stage be more developed (Fromme & D’Amico, 2000). The use of a well-established questionnaire (AEQ-A) with items belonging to the potent social scale, strengthens the notion that expectancies can predict actual use in this study.

More recent research on alcohol expectancies support Goldman’s results, that positive alcohol expectancies predict both current and prospective alcohol use. Positive expectancies at age 16 predict heavy alcohol use at age 30 (Cable & Sacker, 2007; Patrick, Wray-Lake, Finlay, & Maggs, 2010). Patrick et al. (2010) confirms the link is strong in both genders, but stronger in males. At the same time, it is still unclear whether positive expectancies at 16 can predict heavy use into middle adulthood and throughout life. Adolescence and early adulthood usually consists of several major transitions in life, and the fact expectancies turn up strong throughout this period, points to a very robust predictor. Personality as a whole tends to be well settled at age 30, so we can assume that the link between positive expectancies and heavy use lasts, at least to a certain extent, also later in adulthood. Negative expectancies do not have the same effect on low/moderate use, as positive expectancies do on heavy use. Negative expectancies are thought to depend more on third variables such as gender, age and personality (Patrick et al., 2010).

Results from several studies based on an extensive alcohol expectancy literature, emphasize the need to initiate prevention strategies at an early age (Cable & Sacker, 2007; Patrick et al., 2010; Wiers & Kummeling, 2004; Zamboanga, Ham, Van Tyne, & Pole, 2011), and that even small effect sizes in delaying alcohol insertion can contribute significantly to both individual and society (Foxcroft & Tsertsvadze, 2011). Zamboanga et al. emphasize the need to target both adolescent drinkers and nondrinkers with intentions to drink in the future. Nondrinking adolescents with intentions to drink may wait for two years, but they may also drink tomorrow. Hence should prevention target challenging beliefs and expectancies, to hopefully delay alcohol insertion (Zamboanga et al., 2011). The suggestions from aforementioned studies are therefore in line with what the current study is aiming at (Natvig & Aarø, in press). A few studies have investigated the possibilities of reducing positive alcohol expectancies

through experimentation. Clear tendencies towards reduced positive expectancies, followed by reduced alcohol use, was illustrated in a male college sample (Darkes & Goldman, 1993), and later among female college students (Wiers & Kummeling, 2004). Wiers and Kummeling (2004) point out that the results are from a single study, and that multiple sessions are needed to see if the expectancies last or drop back to baseline. Moreover, the study shows an intuitive way to challenge (and adjust) various social beliefs associated with alcohol use.

The strong current and prospective power of positive expectancies, points to the human beliefs system, and what these adolescents actually believe comes out of alcohol use. When they believe that alcohol use is rewarding, they are more likely to experiment with it, perceive positive consequences, and attribute them to the alcohol. Thus, there will be a continuous feedback route; beliefs→intake→confirmation→beliefs, which reinforces what they “already knew”. At the same time, people who hold positive beliefs about something are especially sensitive to information favoring their beliefs. This effect is called confirmation bias.

Ambiguous information for instance, will more likely be interpreted as supporting evidence for what you believe. So, if you experience both negative and positive consequences with alcohol use, the positive bits are likely to be held in higher regard, if you already held positive beliefs.

### **The role of social factors**

How is it that adolescents who have not even tasted alcohol yet, come to have positive expectancies and positive beliefs about alcohol, resulting in motives to drink? Bandura's social cognitive theory (1986), is probably the best way to describe why this may be so. In short, children and adolescents observe what their models (peers, older children, parents) do, and strive to replicate that behavior, as long as positive opinions are held of the model and his/hers or that groups behavior. Bandura (1986) further explains that there are inner cognitive forces leading to thought and action. The social cognitive thinking is based on connections between human, environmental factors, and personal factors. So, to decide if drinking is a high valued behavior or not, the adolescent will be guided by whether his/her environment is positive towards alcohol, and whether he/she feels drinking would be socially rewarding, and his/her existing schemas on alcohol. The resulting behavior then is influenced by the environment it interacts in, and adapts to it. At the same time, behavior also influences the

environment; the environment is shaped by everyone who interacts in that environment. This process is called reciprocal determinism (Bandura, 1986), and explains why one factor leads to another; the causal relations between these factors.

There is beyond doubt that social factors are important targets in alcohol use prevention interventions. As already described briefly, Natvig and Aarø's study aims at manipulating two cognitive patterns of thinking around alcohol use- both of which are consistent with a social influence approach. The video-contest intervention, which is thought to evoke induced compliance, is grounded in theories on attitude change. Festingers (1957) cognitive dissonance theory explains that people need to act in accordance with their attitudes. Cognitive dissonance (inconsistency) is an unpleasant state of mind, and thus will the individual strive to reduce inconsistency by acting towards achieving consistency. If the person has positive alcohol expectancies, but is set to warn others on the negative sides of alcohol, he/she is likely to change their attitude to anti-drinking. If the person holds no base attitude, then Bem's (1972) self-perception theory predicts that the person will attribute his/her behavior to his or her attitudes. The idea behind is that people have a need to make sense of what they do. Both processes lead to attitude change.

The second intervention; correcting fallacy beliefs about alcohol; stems from a normative education perspective. Normative influence is in short how your thoughts and actions are influenced by what you think the current norm is. Most of us want to follow the norm instead of sticking out, because differing from the norm may ignite negative reactions from peers. Adolescents usually overestimate drinking among other adolescents, and also among adults, which creates inaccurate normative expectations. This way of thinking may result in unfortunate drinking behavior supporting the perceived norm. Normative education is about providing knowledge that corrects misconceptions about alcohol use (G. J. Botvin, 2000). Normative education has been an integral component in a wide range of social influence prevention programs in school, since the 1980s (G. J. Botvin, 2000), providing desired outcomes in alcohol prevention (H. Aas & Klepp, 1992; G. J. Botvin, 2000), and also in adolescent problematic gambling behavior (Donati, Primi, & Chiesi, 2013).

Natvig and Aarø's study showed a difference in change scores on alcohol use, between intervention group and control group. This result alone should not be taken as proof of a successful program design. Evidence is needed to prove the results stem from the intervention, and not confounding variables. The program (any prevention program) must be

verified, it must show sustainability, and it must be replicated. This study covers feedback that makes it possible to measure implementation quality on the video project intervention, which is a precarious part of the verification process. Evidence on implementation adds to scientific strength, and guides the researcher(s) towards adjusting and improving the program design.

### **The intervention**

To summarize, Natvig and Aarø (in press) chose to manipulate the intervention group with two different social influence interventions; attitude change through induced compliance(1), and correcting fallacy beliefs through normative educations(2). The rationale behind the intervention is a constant need to warn off, and prevent adolescent alcohol use, due to both present and prospective health related dangers associated with it. Even though alcohol use among 14 year olds in Norway seem to have stabilized the last years, and actually declined a little compared to the top year 1998 (Vedøy & Skretting, 2009), the need for prevention is still strong. Knowledge on consequences has increased, and it is a growing problem area in public health. Consequently, alcohol use was recognized as a “concern area” at the 63<sup>rd</sup> World Health Assembly (May 2010) (World Health Organization, 2010). Member states were therein urged to take preventive action against alcohol use.

To succeed with all respondents is utopia. But to change the health behaviors of even a small number of the target population, has the potential to save both individuals and society from a range of alcohol related consequences. Based on the theoretical framework on adolescent alcohol use and alcohol expectancies, the following three hypotheses will be tested: (1) Is there an association between implementation quality and alcohol use? , (2) Is there an association between implementation quality and alcohol expectancies? , and (3) Are there any gender differences in alcohol use and alcohol expectancies among the students? If numbers on actual alcohol use turn out insufficient, alcohol expectancies is regarded the best alcohol use predictor, and next best measure after actual use itself. Scores on implementation quality will be based on the students` own evaluations of different aspects with the video project. Answers to the hypotheses can hopefully help decide if adjustments need to be made in the program design.

# Methods

## The "Adolescents and alcohol- survey"

"Adolescents and alcohol" [Ungdom og alkohol] is a survey that was administered to a total of 60 classes from 32 volunteering schools throughout Norway. The survey was administered and financed by the Norwegian Directorate of Health. The total number of 2201 respondents made their own ID-codes to ensure regulations regarding anonymity were followed. The codes were subject to match questionnaires from T1 and T2 only, and could for instance not discern respondents within classes. The survey comprises a range of alcohol related questions, to map out the population's relationship with alcohol. Some questions are dichotomous "yes/no"-questions, others require answers on a continuous likert scale. Here pupils grade their response as to how much they agree or disagree, or how seldom or how often something happens. Relationship with alcohol is revealed from actual alcohol use, perceived behavior control (PBC), descriptive and injunctive norms, alcohol expectancies (AEQ-A), attitudes, experience in drinking situation, anticipated regret, self-image, and risk factors.

Natvig & Aarø explain that 46.1% of the respondents at T2 did not return the questionnaire, because teachers dropped out of the project. 9.4% of respondents could not be matched because of missing or non-corresponding ID-codes. Teachers administered questionnaires to all pupils present in the class at T1, and to the same pupils again four months later. Intervention and comparison classes received identical questionnaires, with only minor changes at T2 (Natvig & Aarø). Punching errors during data collection procedures was about 0.2%, and regarded to be most satisfactory. As a result of this, there were 978 subjects on both T1 and T2. The intervention group consisted of 586 pupils, and was matched by a comparison group consisting of 392 pupils. The sample's mean age was 13.7 years at T1 (+/- 0.4 years). There was no significant distribution skewness among genders. 47.1% of intervention group respondents were boys, 47.4% of comparison group respondents were boys (Natvig & Aarø, in press).

Participation to the intervention started with a letter of invitation to participate, sent from the Norwegian Ministry of Education. The invitation described intervention purpose, and enclosed was a poster explaining rules for the video competition. Also, the poster stated that

some of the videos would be broadcasted on a popular national television show (Tv2`s SONE2). Television viewers were encouraged to submit their vote on Tv2`s webpages, where they could watch the contributions multiple times. One video was chosen by the viewers, and one was chosen by a jury. The two winning classes would be rewarded almost \$2000 each. The competition was to make the best 3-minute video to convince 7<sup>th</sup> graders to postpone the age they start drinking alcohol. One video per class was maximum. The second part of the intervention was a short questionnaire administered to one 9<sup>th</sup> and one 10<sup>th</sup> grade class at the participating 8<sup>th</sup> graders` school. Feedbacks on the results were presented to the 8<sup>th</sup> grade classes, by their teacher, who was encouraged to spend some time presenting the results. The results revealed the percentage of students who had ever consumed an alcoholic beverage, and the percentage who had ever been drunk. The feedback consisted of combined results from both the same and all other participating schools.

### **Measuring alcohol use and alcohol expectancies**

Alcohol use is measured by six items from the questionnaire. Three of the items are dichotomous yes/no questions; “Have you ever had at least one sip of alcohol?”, “Have you ever had at least one glass of alcohol?”, and “Have you ever been drunk?”. In order to create a sum score, the scoring was set to; *Never had a sip* (0); *had a sip* (2); *had a glass* (4); *been drunk* (6). This sum score would discriminate non-drinkers from drinkers, and also reveal how far the respondents have come in “experimenting with alcohol”. To get a perspective on drinking frequency, three partly overlapping constructs were included in the sum score: “Think back on the last 3 months, how often have you been drinking at least one glass of alcohol?” Scoring; *never* (0), *1-2 times in 3 months* (1), *once a month* (2), *2-3 times a month* (3), *once a week* (4), *2-3 times a week* (5), *4-7 times a week* (6). The two last frequency related items were also bound to type of beverage: “For the time being, how often do you drink alcopops?” “For the time being, how often do you drink beer?” Scoring; *never* (0), *seldom* (1), *every month* (2), *every week* (3), *every day* (4). Items on wine and spirits were also included in the original questionnaire, but due to criterion validity only alcopops and beer were included in the sum score (Natvig & Aarø).

As already reported, alcohol expectancies can be accurately measured using the AEQ-A scale. As reported earlier, the two most potent out of seven AEQ-A subscales, are the social



(enhance or impede social behavior), and global (positive transforming agent) scales (H. Aas, 1993; Christiansen et al., 1989; Natvig & Aarø, in press). I have chosen to use only questions answering to the social scale in this study. The reason for this is that this scale is alone the most important predictor of self-reported frequency of alcohol use. Aas found that the social scale explained more variance in alcohol expectancies than any other scale ( $r = .44^{**}$ ), and more than all the scales combined could account for ( $r = .35^{**}$ ) (H. Aas, 1993). This finding is in line with previous work on the AEQ field (Christiansen et al., 1982; Christiansen et al., 1989; Mann, Chassin, & Sher, 1987), and is supported by later studies (Fromme & D'Amico, 2000; Kline, 1995; Wilhelmsen, 1997). In total there are five items on the social scale. The items are as follows: "People become more friendly when they have been drinking and are a little inebriated", "It is ok to drink alcohol because then one can join in with others who are having fun", "Many alcoholic drinks taste good", "Parties become more fun when alcoholic beverages are consumed there", and "People get into better moods when they are inebriated". All items are measured on 7-point likert scales, scoring from *fits poorly* (0) to *fits well* (6). A higher score indicates more positive outcome expectancies, i.e. stronger beliefs to improved social interaction.

### **Measuring implementation**

There are seven indicators from the "Adolescents and alcohol" questionnaire that are judged relevant to assess implementation quality. These are presented in table 1. "(I) Did your class manage to finish the video in full?" is a categorical indicator, with *yes* scoring 1 and *no* scoring 2. In many ways this is the most important indicator of all, because it tells us whether the intervention was completed or not, and points to two precarious implementation outcomes; dosage and fidelity. For example; the intervention was not delivered in full (dosage), and not completing the intervention was not intended in the program design (fidelity). Though the indicator reveals nothing as to what specifically happened from onset to the initial end; it provides a very important "framework" to allow other aspects of the design to happen. A good intervention is designed to fit its environment, and make completion feasible. (II) "To what degree was it the teacher or the pupils who controlled the video project?" indicates fidelity, because the implementer (teacher) is integral in delivering the intervention and see that it is being completed. This indicator is scored on a 7-point likert scale from *mostly the teacher* (1), to *mostly the pupils* (7). (III) "To what degree have you

involved yourself in making a best possible video?”, (IV) “To what degree have the other pupils participating in the video project involved themselves in making a best possible video?” and (V) “To what degree has the teacher involved him/herself in making a best possible video?” indicate fidelity and quality of delivery. Involvement/commitment is clearly important for quality, because it measures qualitative aspects toward the intervention, such as enthusiasm and attitudes. These three indicators were scored on 7-point likert scales, from *not involved at all* (1) to *strongly involved* (7). (VI) “How interesting do you think the video project has been?” indicates quality of delivery and participant responsiveness. Qualitative aspects towards the intervention as well as positive response and interest, should make it more feasible and also more efficient. Scored on a 7-point likert scale from *not interesting at all* (1) to *very interesting* (7). The last indicator; (VII) “Are you happy with the video you made?” measures participant responsiveness. The indicator is scored on a 7-point likert scale from *no, very unhappy* (1) to *yes, very happy* (7).

Table 1: Taxonomy of implementation outcomes

<b><i>Indicator</i></b>	<b><i>Implementation outcome</i></b>
“Did your class manage to finish the video in full?”	Dosage Fidelity
“To what degree was it the teacher or the pupils who controlled the video project?”	Fidelity
“To what degree have <u>you</u> involved yourself in making a best possible video?”	Fidelity Quality of delivery
“To what degree <u>have the other pupils</u> participating in the video project involved themselves in making a best possible video?”	Fidelity Quality of delivery
“To what degree has <u>the teacher</u> involved him/herself in making a best possible video?”	Fidelity Quality of delivery
“How interesting do you think the video project has been ?”	Quality of delivery Participant responsiveness
“Are you happy with the video you made?”	Participant responsiveness

In order to make the 7 implementation variables easier to work with, as well as easier on the reader, they will be referred to as follows, in correspondence with table 1: (1)Completeness, (2)Control, (3)Self-involvement, (4)Involvement of other students, (5)Teacher-involvement, (6)Interest, (7)Satisfaction.

### **Use of statistics and analytical procedures**

IBM SPSS Statistics 20.0 was used for all statistical analysis. T-test was used to assess changes in alcohol use and alcohol expectancies between T1 and T2. An independent t-test was utilized to investigate possible gender differences in alcohol use and alcohol expectancies. Pearson's  $r$  is utilized as measure of effect size on the paired alcohol use, alcohol expectancies, and gender data. A small effect size is  $r$  above .10, a medium effect size is  $r$  above .30, and an  $r$  above .50 is considered a large effect size. Cohen's  $d$  estimates effect size based on the difference between means. A Cohen's below .20 is considered small, between .20 - .50 is a medium effect size, and anything above .50 is a large effect size.

A significant component analysis (factorial analysis) was conducted to see if any of the implementation variables measured the same underlying component(s). The rationale behind this decision is based on correlates, and that some of the variables have a rather similar wording.

To see if any of the implementation variables can predict alcohol use or alcohol expectancies, stepwise multiple regression analyses were conducted. The rationale behind using a stepwise method, rather than a hierarchical regression method, is that the stepwise method allows for a more exploratory approach on the predictive potential of the independent variables (predictors), than does the hierarchical method; which is based on a theoretical hierarchy among the predictors. The stepwise regression method estimates the relative importance of each independent variable, and chooses the variable with strongest predictive power. A new sequence of estimation is then run to see if another predictor is strong enough to contribute to the model. As Field (2009) warns that forward stepwise regression is more likely to exclude predictors involved in suppressor effects; i.e. risking a Type II error, backward stepwise regression was also conducted merely as a safeguard. In addition, results from hierarchical regression are presented in short form in footnotes, to see if results depend on whether

stepwise or hierarchical regression is used. Analyses were conducted on two different levels: Individual level and classwise level. Relevant variables were aggregated to represent a mean score based on class identity. This method eliminates variance within the class, and gives every individual in that class the same score. As each class acted in unity producing the video, this approach is considered highly relevant.

# Results

## Descriptives

First, all the relevant variables were checked to see if they were normally distributed. Exploring the variables by using the Kolmogorov-Smirnov test for assessment, found that none of the relevant variables were normally distributed. Kolmogorov-Smirnov indicates significance if value is below 0.05. To meet demands for normality, this value needs to be above 0.05. Therefore, we can say that the variables in this data material significantly deviate from a normal distribution. The normal distribution is a desired distribution, because it serves as a reality check: In theory it may be difficult to imagine drawing accurate conclusions about reality based on assumptions that clearly differ from what we think reality should be like (Field, 2009). Field also stresses (2009) that there is no pressing need to test large samples for significance, because very small deviates in the data material can turn the test significant from a normal distribution.

Table 2: Descriptive statistics

	N	Range	Mean	Std. Deviation	Skewness	Kurtosis
Alcohol use <sup>a</sup>	586	-9 – 12	,75	2,211	,891	3,650
Alcohol expectancy <sup>b</sup>	586	-6 – 6	,33	1,366	,155	1,777
(1)Completeness	558	1 – 2	1,10	,303	2,634	4,958
(2)Control	548	1 – 7	5,30	1,574	-,982	,696
(3)Self-Involvement	549	1 – 7	4,78	1,712	-,599	-,216
(4)Involvement of other students	549	1 – 7	5,14	1,465	-,873	,750
(5)Teacher-Involvement	548	1 – 7	5,07	1,636	-,801	,147
(6)Interest	549	1 – 7	4,73	1,776	-,560	-,393
(7)Satisfaction	550	1 – 7	4,97	1,805	-,781	-,195
Valid N (listwise)	546					

<sup>a</sup>Differential alcohol use scores between T1 and T2

<sup>b</sup>Differential alcohol expectancies scores (AEQ-A) between T1 and T2

However, the difference from normal distribution comes from a real skewness in scores. The skewness and kurtosis data presented in table 2, clarifies how the data material look. Both

alcohol use and alcohol expectancies show positive values of skewness, which indicate more low scores in the distribution, whereas high scores. The alcohol use value is the more extreme (more low-scores oriented), and a very high positive kurtosis value indicates a heavy-tailed (leptokurtic) distribution. This means that the majority of the scores are distributed on one or a few low values. When considering that a large part of the population has not tasted alcohol, this result comes as no surprise. The extreme skewness and kurtosis of implementation variable 2 is because it is a categorical variable, where the majority of the classes finished their video project; which gave a low value score. Implementation variables 2-7 are all a bit negatively skewed, which means more high scores than low. This means a tendency towards more student control, more involvement from both students and teacher, more interest than disinterest in the project, and more scores on happiness with the video rather than unhappiness.

### **Alcohol use and alcohol expectancies**

The use of a paired samples *t*-test helps to see if there are any changes in alcohol use from T1 to T2. Results show a rise in alcohol use from T1 (M=3.81, SD=3.56) to T2 (M=4.57, SD=3.73),  $t(585) = -8.26$ ,  $p < .001$ ,  $r = .32$ . This is considered a medium size effect. Cohen's *d* was also calculated ( $d = -.21$ ), and indicated a small size effect, bordering on moderate (Becker, 2000). As explained earlier, the age 14 in Norway is an age when young people are experimenting with alcohol, and therefore a common age to start drinking alcohol. Thus, a change in alcohol use with time was to be expected. Results from a paired samples *t*-test on alcohol expectancies show a rise also in expectancies, from T1 (M=2.97, SD=1.48) to T2 (M=3.31, SD=1.65),  $t(586) = -5.86$ ,  $p < .001$ ,  $r = .24$ . This is considered a small effect size. Cohen's *d* ( $d = -.22$ ) indicated a small effect size, bordering on moderate (Becker, 2000). Alcohol use and alcohol expectancies (AEQ-A) correlate significantly with each other both on T1,  $r(586) = .64$ ,  $p < .01$ , and on T2,  $r(586) = .64$ . The result points to a strong positive correlation between the variables. As expected, an increase in alcohol use is associated with increase in positive alcohol expectancies.

### **Significant component analysis**

The different implementation variables are thought to, and which I have argued for, point towards different qualities of the implementation. There are however some of the variables which arguably measure qualities of the same construct. For instance, implementation variable (3), (4) and (5) all measure degree of involvement, albeit involvement among different groups. A bivariate correlation confirms that all implementation variables yield a significant correlation with at least one other variable.

A principal component analysis was conducted on the seven variables with orthogonal rotation (varimax). The rationale for this was to map out underlying components and see if there is a need to add scores on two or more variables together, which may eventually explain more variance in the dependent variable, than what a single predictor does. The Kaiser-Meyer-Olkin measure verified sampling adequacy for the analysis,  $KMO = .77$ , which is considered a good sample size (Field, 2009). KMO values for individual variables were  $< .65$ , which is above the limit of  $.5$  (Field, 2009). Bartlett's test of sphericity  $\chi^2(21) = 881$ ,  $p < .001$ , indicated that correlations were sufficiently large for the analysis. Two components had eigenvalues over Kaiser's criterion of 1, which in combination explained 56.3% of the variance. Component 1 had an eigenvalue of 2.86 and explained 40.81% of the variance; component 2 had an eigenvalue of 1.08 and explained 15.48% of the variance. After rotation component 1 had an eigenvalue of 2.50 and explained 35.73% of the variance, and component 2 had an eigenvalue of 1.44 and explained 20.73% of the variance. Table 3 shows the factor loadings after rotation.

Table 3: Summary of exploratory factor analysis (N=546). Only factor loadings over  $.40$  are present in the table.

Item	Component	
	1	2
(4)Involvement of other students	.73	
(6)Interest	.72	
(5)Teacher-involvement	.71	-.44
(3)Self-involvement	.68	
(7)Satisfaction	.66	
(2)Control		.70
(1)Completeness		-.60
Eigenvalues	2.50	1.44
% of variance	35.73	20.57

The variables that cluster on component 1 suggest that this component represents involvement, interest, and satisfaction with the video project. Component 2 is more difficult to interpret, as variables that have factor loadings on the component represent both involvement, control and whether the video was completed or not. Based on variables with large factor loadings and also based heavily on variable-wording, it was decided to make sumscores. Both individual scores and classwise meanscores were computed of the three involvement variables; (3), (4) and (5), and it was called “General Involvement”. A stepwise regression analysis did not turn up results that “General Involvement” explained more variance in alcohol use nor alcohol expectancies, neither when analyzed individually nor classwise. Each variable tells its own unique history related to implementation qualities, and therefore it was decided to use and analyze all seven variables.

### Individual implementation scores

A stepwise multiple regression analysis was conducted to investigate implementation variables` relationships to alcohol use. Table 4 shows a small but significant correlation between individual alcohol use scores and (2)Control,  $r(548) = .11, p < .05, r^2 = .01$ . The positive b-value (.15) indicates a positive relationship between alcohol use and (2)Control, where more student controlled video projects means higher alcohol use scores, although (2)Control cannot explain more than 1% variance in the change in alcohol use scores. (2)Control ( $t(546) = 2.47, p < .05$ ) is a significant predictor of individual alcohol use. The standardized  $\beta$  (.11) indicates that when (2)Control increases by one standard deviation (1.58) – more student controlled video projects; alcohol use increases with .11 standard deviations.

Table 4: Summary of stepwise regression<sup>1</sup> for variables predicting individual alcohol use (N=546). Dependent variable: Differential alcohol use scores (post-pre).

	Mean	St.dev	R	B	SE B	$\beta$	t
Model 1			.11 <sup>a</sup>				
Constant				-.12	.33		-.04
(2)Control				.15	.06	.11*	2.47

Note:  $R^2 = .01$  ( $p < .05$ ). \*  $p < .05$ ,

<sup>a</sup> predictors: individual alcohol use, control

<sup>1</sup> Hierarchical regression found (2)Control the strongest and only significant contributor ( $\beta = .13, p < .01$ )



The same stepwise regression procedure was conducted with alcohol expectancy scores. Table 5 shows that there is a small but significant ( $r(558) = .10, p < .05, r^2 = .01$ ) positive relationship ( $B = .57$ ) between alcohol expectancies and (1)Completeness. The high score on (1)Completeness means the video project was not completed, so high alcohol expectancies are associated with not completing the video project. The change in (1)Completeness only accounts for 1% variance in the change in alcohol expectancy scores. The b-value differs significantly from zero ( $t(546) = 2.67, p < .05$ ), which means (1)Completeness makes a significant contribution to the model. The standardized  $\beta$  (.11) indicates that when (2)Completeness increases by one standard deviation (.28) – more students not finishing the video project; alcohol expectancy increases with .11 standard deviations..

Table 5: Summary of stepwise regression<sup>2</sup> for variables predicting individual alcohol expectancy (N=546). Dependent variable: Differential alcohol expectancy scores (post-pre).

	R	B	SE B	$\beta$	t
Model 1	.11 <sup>a</sup>				
Constant		-.29	.24		-1.20
(1)Completeness		.57	.21	.11*	2.67

Note:  $R^2 = .01$  ( $p < .05$ ). \*  $p < .05$ ,  
<sup>a</sup> predictors: individual alcohol expectancies, completeness

**Classwise implementation scores**

Instead of looking at individual scores, there are good arguments to group all the respondents' scores in the class they belong in. This way the scores on each of the implementation variables represent a mean of the evaluations by students in every class that took part, 39 classes in all. By aggregating data to classes and calculate means, we obtain data which is less vulnerable to extreme scores. As several of the implementation variables are set to evaluate someone else (the teacher, the other students), a class mean stands out as a relevant measure, objectively. You can argue that a group of people; the entire class, is a better judge of teacher control, than a single student. Another argument favoring the classwise meanscore, is that the making of the video is a class- project. The idea is that each class acts as a unity in this.

<sup>2</sup> Hierarchical regression found (1)Completeness the strongest and only significant contributor ( $\beta = .11, p < .05$ )

The same stepwise regression procedure was conducted on classwise alcohol use and classwise implementation scores. Table 6 shows a summary of the regression output, where (4)Involvement of other students found a moderate significant ( $r(35) = .44, p < .01, r^2 = .19$ ) negative relationship ( $B = -.46$ ) with alcohol use (model 1). The b-value differs significantly from zero ( $t(35) = -2.80, p < .01$ ), which means (4)Involvement of other students makes a significant contribution to the model. When other students get more involved in the video project, alcohol use goes down. (4)Involvement of other students itself accounts for 19% of the variance in changes in alcohol use scores. The standardized  $\beta$  (-.44) indicates that when (2)Involvement of other students increases by one standard deviation (.66) – more involvement of other students in the video project; alcohol use decreases with -.44 standard deviations

However, the contribution of (7)Satisfaction, as presented in Table 6: model 2; improves the predictive power of both (4)Involvement of other students, and the two predictors combined. Combined (4)Involvement of other students and (7)Satisfaction; shows a moderately strong significant ( $r(35) = .53, p < .01, r^2 = .28$ ) relationship with classwise alcohol use. Model 2 improves the amount of change in variance that can be explained in classwise alcohol use scores ( $\Delta R^2 = .09$ ), compared to Model 1. The unique contribution of (4)Involvement of other students improves with a stronger b-value ( $B = -.78$ ) which significantly differs from zero ( $t(35) = -3.51, p < .01$ ). When (4)Involvement of other students increases by one standard deviation (.66); alcohol use decreases with -.75 standard deviations. The unique contribution of (7)Satisfaction on the other hand, is positively associated with classwise alcohol use ( $B = .33$ ), and the b-value significantly differs from zero ( $t(35) = 2.05, p < .05$ ). When (7)Satisfaction increases by one standard deviation (.92) alcohol use increases by .33 standard deviations. This is an unexpected result for two reasons: You would think that high scores on satisfaction with the video project would lead to a smaller increase in alcohol use, rather than a strong increase. And when (7)Satisfaction is correlated with classwise alcohol use, there is no significant relationship ( $r(35) = -.33, p > .05$ ). Influence statistics showed values of Cook's distance well below 1, which is the criterion to decide if any of the cases (classes) have an undue influence on the model (Field, 2009). Casewise numbers on Mahalanobis distance is well below the critical values for an outlier; depending on number of predictors and sample size, recommended by Barnett and Lewis (Barnett & Lewis, 1978), and Field (2009). DFBeta statistics confirmed that no particular case (class) had an especially large influence on regression parameters. These diagnostic indications point to a fairly reliable model, and could

not shed more light to why (7)Satisfaction with the video project somewhat surprisingly increases with alcohol use.

Table 6: Summary of stepwise regression<sup>3</sup> for variables predicting classwise alcohol use (N=35). Dependent variable: Differential alcohol use scores (post-pre).

	R	B	SE B	$\beta$	t
Model 1	.44 <sup>a</sup>				
Constant		3.16	.85		3.72
(4)Involvement of other students		-.46	.17	-.44**	-2.80
Model 2	.53 <sup>b</sup>				
Constant		3.21	.81		3.95
(4)Involvement of other students		-.78	.23	-.75**	-3.51
(7)Satisfaction		.33	.16	.44*	2.05

Note:  $R^2 = .19$  for step 1 ( $p < .01$ ),  $\Delta R^2 = .09$  for step 2 ( $p < .05$ ). \*\*  $p < .01$ , \*  $p < .05$ ,

<sup>a</sup> predictors: classwise alcohol use scores, (4)Involvement of other students.

<sup>b</sup> predictors: classwise alcohol use scores, (4)Involvement of other students, (7)Satisfaction.

(4)Involvement of other students have a moderate significant ( $r(35) = .43$ ,  $p < .01$ ,  $r^2 = .19$ ) negative relationship ( $B = -.29$ ) with classwise alcohol expectancy, as presented in Table 7.

The b-value differs significantly from zero ( $t(35) = -2.74$ ,  $p < .01$ ), which means

(4)Involvement of other students makes a significant contribution to the model. When other students get more involved in the video project, alcohol expectancy goes down.

(4)Involvement of other students itself accounts for 19% of the variance in changes in alcohol use scores. The standardized  $\beta$  (-.43) indicates that when (2)Involvement of other students increases by one standard deviation (.66) – more involvement of other students in the video project; alcohol use decreases with -.43 standard deviations. When (4)Involvement of other students increase by .66, classwise alcohol expectancies decrease by (.43\*.45(alcohol expectancy standard deviation)) .19.

Table 7: Summary of stepwise regression<sup>4</sup> for variables predicting classwise alcohol expectancy (N=35). Dependent variable: Differential alcohol expectancy scores (post-pre).

<sup>3</sup> Hierarchical regression found (4)Involvement of other students the strongest and only significant contributor ( $\beta = -.83$ ,  $p < .05$ )

<sup>4</sup> Hierarchical regression found no predictor to be strong enough to have a significant contribution

	R	B	SE B	$\beta$	t
Model 1	.43 <sup>a</sup>				
Constant		1.83	.55		3.35
(4)Involvement of other students		-.29	.11	-.43**	-2.74

Note:  $R^2 = .19$  ( $p < .01$ ). \*\*  $p < .01$

<sup>a</sup> predictors: classwise alcohol expectancy scores, (4)Involvement of other students.

## Alcohol and gender

To investigate if there are significant differences in alcohol use between boys and girls, an independent samples t-test was used. Independent samples t-test is used when different participants participate in different conditions. Unlike the (dependent) paired samples t-test, which investigated the same participants' changes invoked by alcohol use and alcohol expectancies, the independent t-test investigates alcohol use and alcohol expectancies on different participants. This way, the score will reflect variance in individual differences; gender differences, as well as changes in the experimental condition. See table 8 for an overview of the correlation statistics. On average, girls experienced higher alcohol use on T1 ( $M = 4.00$ ,  $SE = .20$ ), than boys ( $M = 3.61$ ,  $SE = .22$ ). The difference is not significant,  $t(584) = -1.34$ ,  $p > .05$ ,  $r = .06$ . Girls experienced somewhat higher scores also on T2 ( $M = 4.81$ ,  $SE = .22$ ), than boys ( $M = 4.30$ ,  $SE = .21$ ). The difference is not significant,  $t(584) = -1.64$ ,  $p > .05$ ,  $r = .07$ . The result for difference scores between T1 and T2 reveals there are no significant differences between boys and girls. Correlation between girls ( $M = .81$ ,  $SE = .13$ ), and boys ( $M = .70$ ,  $SE = .13$ ) shows no significance,  $t(584) = -.605$ ,  $p > .05$ ,  $r = .03$ .

It is also important to investigate whether there are significant gender differences in alcohol expectancies. The same analytical procedure as with actual alcohol use took place, but swapping alcohol use scores with alcohol expectancy scores. The result showed that on average, girls experienced higher alcohol expectancies on T1 ( $M = 3.07$ ,  $SE = .09$ ), than boys ( $M = 2.87$ ,  $SE = .09$ ). The difference is not significant,  $t(584) = -1.62$ ,  $p > .05$ ,  $r = .07$ . Scores on T2 showed a slightly lower result for girls ( $M = 3.28$ ,  $SE = .09$ ), than boys ( $M = 3.34$ ,  $SE = .10$ ). The difference is not significant,  $t(584) = .45$ ,  $p > .05$ ,  $r = .02$ . The scores in alcohol expectancies between T1 and T2, divided by gender, reveals a lower mean among girls ( $M = .21$ ,  $SE = .08$ ), than boys ( $M = .47$ ,  $SE = .08$ ). This relationship is significant,  $t(584) = 2.31$ ,  $p$

< .05,  $r = .10$ . The difference means that there is variance in the groups' changes in alcohol expectancies that can be attributed to gender. However, the effect size is very small ( $r = .10$ ), which means the gender effect only explains 1% of the total observed variance. Boys experienced more changes in alcohol expectancies between T1 and T2, than did girls. The correlation coefficient had a positive value, and that helps out in deciding the direction of the relationship. In this measure, based on coding, a positive value meant more changes in boys' scores, and a negative value meant more changes in girls' scores.

Table 8: Correlation matrix by gender

	A.use <sup>a</sup> T1	A.use <sup>a</sup> T2	AEQ <sup>b</sup> T1	AEQ <sup>b</sup> T2	A.use-Diff <sup>d</sup> T1-T2	AEQ-Diff <sup>e</sup> T1-T2
Gender	.18	.10	.11	.65	.55	.02*

<sup>a</sup>Alcohol use sum score at T1

<sup>b</sup>Alcohol expectancy (AEQ-A) sum score at T2

<sup>d</sup>Differential alcohol use scores between T1 and T2

<sup>e</sup>Differential alcohol expectancies scores (AEQ-A) between T1 and T2

\*  $p < .05$

It is interesting to see there is a significant difference between genders in alcohol expectancies over 4 months (T1 to T2), and especially so when there is no marked difference between boys and girls in alcohol use change scores. This finding will be discussed later.

## Summary of results

The analysis revealed a difference in changes in variance in alcohol use, depending on whether it was the teacher or the students in control/charge of the video project. Individual alcohol use decreased when the teacher controlled the project. However the same result did not show when alcohol use was compared with each classes' mean score. Here, more involvement from other students, rather than teacher control, seemed the strongest predictor of reduced alcohol consumption. A more surprising finding, albeit weaker- but still significant, was that satisfaction with the video project was associated with more alcohol use.

Positive alcohol expectancies decreased when the video project was completed, rather than not completed. When investigated on each classes' mean score however, this result did not show. But when other students got more involved in the video project, the classes' positive

alcohol expectancies decreased. A gender difference was also found in changes in alcohol expectancies over time, and boys experienced more changes than did girls.

# Discussion

Results from analyzing the data material support alcohol literature reporting that changes in alcohol use are significantly associated with changes in alcohol expectancies (H. N. Aas, 1995; Cable & Sacker, 2007; Christiansen et al., 1982; Christiansen et al., 1989; Patrick et al., 2010). More alcohol use is associated with more positive alcohol expectancies, and vice versa. Alcohol expectancies are well settled long before the first drink of alcohol is consumed, and is therefore viewed as a reliable instrument to assess implementation quality in this study. A frequency analysis on alcohol use reveals a very skewed distribution due to low scores; because a lot of the participants at the time had yet to try alcohol for their first time. Results confirmed a significant change both in alcohol use ( $t(585) = -8.26, p < .001$ ) and alcohol expectancies ( $t(586) = -5.86, p < .001$ ) from T1 to T2. The age around 14 is, in Norway, an age where many adolescents start experimenting with alcohol (Natvig & Aarø, in press). Taking preventive action at this age is therefore in line with several reports, recommending early interventions (Cable & Sacker, 2007; Patrick et al., 2010; Wiers & Kummeling, 2004; Zamboanga et al., 2011).

## Implementation and alcohol use

To sum up the most important findings from self-reported alcohol use, this study showed that a more teacher controlled video project is associated with a lesser increase in alcohol use, than when the video project is controlled by students. As implementation variable (2) Control is measured on a continuous scale, it gives respondents the possibility to rate degree of control. The more extreme scores mean more control, and thereby more distinct roles. Be it the teacher or students, whoever is in control is head of the video project. Recommendations for a more teacher controlled video project is supported by Wilhelmsen et al.'s (1994) research report, suggesting that high role-specificity (HRS) is advantageous to less role-specificity (LRS) when implementing a social cognitive prevention program, at least as a first step. The high role-specified version (HRS) advocated more responsibility on the teacher as well as clearer guidelines for both teacher and the students, which influenced student involvement (Wilhelmsen et al., 1994). Other reports support this finding; Pettigrew et al. (2013) found

teacher control to be integral in prevention program delivery; and student engagement and motivation is associated with a clear and active teacher, rather than a passive teacher (Reeve, 2012). When a prevention program is implemented for the first time, it seems to profit from giving teachers a clear role in being in charge of implementing activities

When it comes to implementing prevention program activities, an influencing factor is implementer characteristics. A teacher who is in control and guiding students is desirable, but a too strict teacher for instance may not be able to get the students as involved as a coordinating teacher might. A classroom environment that grant students freedom within a frame of structured activities, is thought to best promote student engagement (Reeve, 2012). A role specification that encourages teachers to take control, but not to an extent that it diminishes student freedom, is therefore recommended. An outlined role-specification that encourages teachers to take control of implementing program activities should be included in the program design. Based on results from this study, and supported by several reports (Pettigrew et al., 2013; Reeve, 2012; Wilhelmsen et al., 1994), a degree of teacher control may be an indication of implementation fidelity, which in turn strengthens implementation quality.

Involvement is important because it makes the students feel they are part of the project process, a part of the end result, and it strengthens the learning outcome. Ultimately in this project, it strengthens quality of program delivery and fidelity. Results on classwise implementation indicated a moderate negative relationship between (4)Involvement of other students and alcohol use ( $-.44^{**}$ ,  $R^2 = .19$ ). Several reports confirm the importance of involvement in program activities (Domitrovich & Greenberg, 2000; Durlak & DuPre, 2008), and participant involvement has also shown to better program outcomes in long term (Begle & Dumas, 2011).

A high role-specified program that excels teacher control also produces more student involvement, according to reports (Pettigrew et al., 2013; Wilhelmsen et al., 1994). Wilhelmsen et al (1994) also found student involvement to be a mediating variable in influencing adolescents` alcohol use. Literature on teacher vs. student control is not consistent, but several agree that at least some teacher control rather than practically none, produces more student involvement (Hunter, 2012; Pettigrew et al., 2013; Reeve, 2012). Hunter (2012) stresses teachers in student projects at the very least need to be used as consultants, with which comes a naturally leading role, which is the price to pay for



involvement in participation. Hunter (2012) further argues that students should still feel able to make decisions without the teacher's approval, in other words a teacher who guides and encourages without necessarily taking decision-making away from students. Teachers who "push" activities on students and exercise authoritarian class management on the other hand, usually see involvement and engagement rates go down (Pettigrew et al., 2013). There seems to be a thin line between bad control and good control, which favors the teacher who clearly organizes a frame of activities, and thereof acts as active consultants rather than judges, to allow for student decision making.

Classwise (7)Satisfaction with the video project significantly improved the model's predictive power, by adding to amounts of explained variance in self-reported alcohol use. It is the unique contribution of (7)Satisfaction, as it showed the highest semi-partial correlation (.34) with the outcome, second to (4)Involvement of other students, that was high enough to be regarded significant ( $p(.049) < .05$ ). If anything, this outcome was expected to be negatively associated with alcohol use, rather than positive. Also, when all predictors are compared to each other standardized in hierarchical regression, there is no significant contribution from (7)Satisfaction. Diagnostics statistics does not offer any indication that there are outliers or classes that have an undue influence on the model, which can help explain this result from the regression.

### **Implementation and alcohol expectancies**

Results revealed a significant relationship between whether the video project was completed or not, and alcohol expectancies. More of the students who did not complete the video, showed more positive expectancies towards alcohol, than student who completed the video project. This measure refers to dosage, i.e. the amount of intended program delivered. In cases where the video project was not completed, students did not experience the whole intervention. When high levels of dosage are reported, implementation quality is strengthened, and this is often associated with better outcomes (Durlak & DuPre, 2008). Several reports support this association (Allen et al., 1990; Dane & Schneider, 1998; Felner et al., 1994; Weinman et al., 1992). As a high dosage outcome mutually relies on an intervention that is implemented according to program design, this outcome also indirectly influences fidelity. Classes who completed the video project showed better dosage and fidelity, and

thereof; higher degree of implementation quality had an association with better outcomes on students' alcohol expectancies.

As we only had individual data on whether the relevant class managed to complete the video project, much information from the video project *process* is lost. For perfect intervention dosage, you might argue that all students within the class need to be exposed to all the different parts of the project, and not only the end product, and also that exposure is evenly distributed among the students. There may have been instances where some students played more peripheral roles than others, in making the video. If this is true, then there should be variance between subjects on amounts of program delivered. An additional measure on level of exposure/inclusion to program activities, is therefore advisable; to say something about amount of program deliverance in the video project process. The same is true in instances where one or more students are absent from parts of the intervention. Especially being absent when key-concepts of the intervention are being outlined, is critical, and endangers the dosage outcome (Dane & Schneider, 1998). This potential problem can be solved by conducting the whole video project at one time, instead of working with it for shorter sessions through several days for instance. Whether the video project was completed or not is a good and important measure, but including an additional measure on whether each individual attended the whole or only parts of the video project, is advisable. These two extra inclusions would provide more valid measures on dosage, and open up the possibility to measure the amount of the program that is delivered on both process and end product.

(4) Involvement of other students is the most important implementation variable also when alcohol expectancies is the dependent variable. Note that this result was found only when alcohol use, alcohol expectancies, and the different implementation variables were aggregated to represent classes, i.e. classwise mean-scores. When analyzed on individual scores, (4) Involvement of other students did not produce any significant effect neither in self-reported alcohol use nor alcohol expectancies. These two approaches represent two quite different ways of looking at implementation quality. In the first instance we talk about single students, and in the other we talk about whole school-classes, a clear distinction that must not be mixed up.

Alcohol expectancy measures were derived from the AEQ-A social scale (adjusted for use in Norway), which is the expectancy-scale that explains most variance in alcohol use (H. Aas, 1993). Compared to the equivalent scale offered by CEOA, what matters the most is what

kind of expectancies the researcher aspires to identify. There are differences in item specificity between the scales; the AEQ-A reveals general social expectancies to alcohol effects, and CEOA measures subjective evaluations to alcohol effects. These kinds of expectancies border on attitudes to alcohol, and would be advisable to use in instances where the researcher wants to target specific expectancies. CEOA is therefore also adequate in longitudinal studies, to assess changes in specific evaluations and expectancies, throughout adolescence for example. In this study however, AEQ-A is thought to be the more relevant expectancy measure, which has measured social alcohol expectancies reliably.

### **Gender differences**

Results did not show significant differences in gender on self-reported alcohol use, neither on T1, T2, nor on differential alcohol use scores between T1 and T2. This result was expected, and it corresponds to levels of Norwegian adolescents' alcohol use across gender reported in 2003 (Hibell et al., 2004). Statistics from 2011 reveal the same gender pattern in Norway (Hibell et al., 2011).

There was a significant difference in alcohol expectancies, reported on differential scores between T1 and T2. Boys had a higher increase in positive alcohol expectancies than girls. This result means 1% ( $r = .10$ ) of the variance in changes in alcohol expectancies between T1 and T2, can be accounted for by the variance in gender scores. Results did not show gender differences in alcohol expectancies at T1 and T2. A report on gender differences influencing adolescent alcohol use supports the indication that there were no differences in expectancies at baseline; at T1 (Schulte et al., 2009). Whilst Schulte et al. suggest that adult men have more positive expectancies than women, and adolescents report the same degrees of expectancies across gender (Schulte et al., 2009)

### **What can be learned from the implementation (?)**

A very important observation is that results differ depending on whether data are analyzed on individual or classroom level. When analyzed on individual scores, (4)Involvement of other students did not produce any significant effect neither in self-reported alcohol use nor alcohol

expectancies, contrary to what classwise scores did. On the other hand, classwise scores did not produce any significant association with neither (2)Control nor (1)Completeness. These two approaches represent two quite different ways of looking at implementation quality. To most adequately express individual student alcohol use/alcohol expectancy, the individual level results must be used. (2)Control is the outcome variable that best tells us what the most important implementation characteristic is, from the perspective of every single student for instance. If we were to talk about what is best for a single class, or several classes, the classwise mean must be used for interpretation.

Based on results from this implementation analysis, it is clear that both involvement of other students, a certain degree of teacher control, and finishing the video project, are aspects of implementation that improve quality. Groups of participants who experienced more teacher control reported lower increases in alcohol, and participants who finished the video project reported less positive alcohol expectancies- strengthening the fidelity and dosage implementation outcomes. Classes where students (other than the single respondent) involved themselves more than other classes, reported lower increases in alcohol and less positive alcohol expectancies- strengthening quality of delivery and fidelity implementation outcomes. These findings support implementation research reporting dosage, quality of delivery, and especially fidelity; are integral dimensions in producing implementation quality (Carroll et al., 2007; Dane & Schneider, 1998; Durlak & DuPre, 2008), and essential parts in strengthening prevention program designs to produce desired program outcomes (G. J. Botvin et al., 1990; G. J. Botvin et al., 1992; Dane & Schneider, 1998; Domitrovich & Greenberg, 2000; DuBois et al., 2002; Durlak & DuPre, 2008; Ennett et al., 2011).

However, there are also good reasons to believe that self-involvement, teacher-involvement, and interest do contribute, even though the analyses did not find significant contributions - the picture is obviously more complex. Involvement/Engagement as a whole is reported to yield positive social- and academic competence outcomes as well as generally better participation rates (Shernoff, 2010; Shernoff & Vandell, 2007), as well as being central in predicting positive program outcomes (Dane & Schneider, 1998; Giles et al., 2008; Reeve, 2012). The significant component analysis provided information on much the same underlying variables between all involvement variables. They also yield moderate positive correlations between them. Reported high involvement from other students necessarily follows a certain amount of self-involvement and interest in the activities/project. Teacher-involvement is moderately

correlated with student involvement, and it may also come naturally with more teacher control. A classroom with both an involved teacher and involved students, contribute to a dynamical and interactive learning environment (Giles et al., 2008; Payne, Gottfredson, & Gottfredson, 2006). Pettigrew et al. reported that involved teachers often had the most involved students, classes in which also produced the larger prevention program outcomes (Pettigrew et al., 2013). In summary, based on results and implementation quality literature; classes which report less increase in alcohol use and alcohol expectancies, are likely to be characterized by generally high involvement rates and an authoritative teacher in control of implementing program activities.

Taking steps to ensure a degree of staff quality is advised, to strengthen program design and program integrity. Regular teachers may of course quit their jobs or be absent for other reasons, which is hard to predict or do anything about. But teacher self-reports; a few control questions for instance on both T1 and T2, to identify classes which have undergone staff changes, may be a way to handle the problem. This way, it may be possible to control for, and verify effects coming from staff changes, as changes here often affect several aspects of implementation (Cross et al., 2010).

A limitation with this research project is that the data material relies solely on self-report. Ideally, observations made by independent observers should be used to complement the self-reports. Self-reports covering alcohol consumption, as well as rating classmates and teacher, may lead to answers that are socially wanted. This is especially so in instances where the respondents feel that their confidentiality is at risk. Natvig and Aarø (in press) ensure that research norms were followed regarding anonymity and data collection. The nature of the research- many classes located around the country, and limited resources; makes it hard to compliment self-reports with independent observations.

I would recommend to also being careful with the regression results offered from classwise alcohol use and alcohol expectancy scores. (4)Involvement of other students stands out as a strong predictor on both alcohol use and alcohol expectancy scores, but only 35 classes were included in the analysis (N=35). This is enough to satisfy the 5 to 1 minimum ratio of valid cases, but it should still be interpreted with caution. Especially the unexpected significant result with (7)Satisfaction going up with alcohol use should be interpreted with caution, as it also showed no significant contribution when analyzed on hierarchical regression. Influence statistics for this analysis did not reveal cases which had an undue influence on the model, but

pointed to a fairly reliable model. In any case, the results reported from classwise implementation scores, should be read with caution and treated as preliminary. Lewis (2007) suggests that stepwise regression procedures are more vulnerable to small samples and more predictors, than what hierarchical regression is. This may be the case in this study, where stepwise and hierarchical regression produced somewhat different results on classwise implementation scores. Choosing a method of analysis still depends heavily on what the researcher aims at, and in regards of exploratory, predictive research; stepwise regression is the recommended method for use (Lewis, 2007)

## **Conclusions**

The three hypotheses were tested and answered, and results reveal that there is association between implementation quality and alcohol use, and between implementation quality and alcohol expectancies. It also revealed a gender difference in alcohol expectancies, albeit not in alcohol use. Implementation findings support aspects that have been highlighted in earlier research (Cross et al., 2010; Dane & Schneider, 1998; Domitrovich & Greenberg, 2000; Durlak & DuPre, 2008; Dusenbury et al., 2003; Giles et al., 2008; Shernoff, 2010). In this study, measures on fidelity, dosage and quality of program delivery, found associations with positive program outcomes.

For the future, prevention program implementation considerations should be encouraged, so it can be assessed and verified. A school based program`s integrity rests on characterizations both on student/teacher level, organizational level, and it rests on the program itself. The five main dimensions as suggested by Durlak and DuPre (2008), seems to cover and secure an accurate assessment of what happens during an intervention. The dimensions are also useful in their differences; because if assessed adequately, they can point the researcher in the right direction for modifications to the program. The fidelity outcome is especially important, as is also proved in this study. The fidelity dimension checks that the program is implemented according to plan, which should encourage researchers to work carefully with their program designs. Future research should also aim to agree on terms and conceptual definitions, so that a general implementation framework for prevention research can be presented.

# References

- Aas, H. (1993). Adaptation of the Alcohol Expectancy Questionnaire (Aeq-a) - a Short Version for Use among 13-Year-Olds in Norway. *Scand J Psychol*, *34*, 107-118. doi: DOI 10.1111/j.1467-9450.1993.tb01106.x
- Aas, H., & Klepp, K. I. (1992). Adolescents' alcohol use related to perceived norms. *Scand J Psychol*, *33*(4), 315-325. doi: DOI 10.1111/j.1467-9450.1992.tb00920.x
- Aas, H. N. (1995). *Alcohol Expectancies and Socialization: Adolescents learning to drink*. (Doctor Psychologiae), University of Bergen, Norway, Bergen, Norway.
- Allen, J. P., Philliber, S., & Hoggson, N. (1990). School-Based Prevention of Teenage Pregnancy and School Dropout - Process Evaluation of the National Replication of the Teen Outreach Program. *Am J Community Psychol*, *18*, 505-524. doi: Doi 10.1007/Bf00938057
- Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Barnett, V., & Lewis, T. (1978). *Outliers in statistical data*. New York: Wiley.
- Bartholomew, L. K., Parcel, G. S., Kok, G., Gottlieb, N. H., & Fernández, M. E. (2011). *Planning Health Promotion Programs, An Intervention Mapping Approach* (3 ed.). San Francisco: Jossey-Bass.
- Becker, L. A. (Producer). (2000). Effect Size Calculators. Retrieved from <http://www.uccs.edu/~lbecker/>
- Begle, A. M., & Dumas, J. E. (2011). Child and parental outcomes following involvement in a preventive intervention: efficacy of the PACE program. *Journal of Primary Prevention*, *32*, 67-81. doi: 10.1007/s10935-010-0232-6
- Botvin, G. J. (2000). Preventing drug abuse in schools: social and competence enhancement approaches targeting individual-level etiologic factors. *Addict Behav*, *25*, 887-897.

- Botvin, G. J., Baker, E., Dusenbury, L., Tortu, S., & Botvin, E. M. (1990). Preventing Adolescent Drug-Abuse through a Multimodal Cognitive - Behavioral-Approach - Results of a 3-Year Study. *J Consult Clin Psychol*, *58*, 437-446. doi: Doi 10.1037//0022-006x.58.4.437
- Botvin, G. J., Dusenbury, L., Baker, E., James-Ortiz, S., & Kerner, J. (1989). A skills training approach to smoking prevention among hispanic youth. *Journal of behavioral medicine*, *12*(3). doi: 10.1007/BF00844872
- Botvin, G. J., Dusenbury, L., Baker, E., Jamesortiz, S., & Botvin, E. M. (1992). Smoking Prevention among Urban Minority Youth - Assessing Effects on Outcome and Mediating Variables. *Health Psychology*, *11*, 290-299. doi: Doi 10.1037//0278-6133.11.5.290
- Brown, S. A., Goldman, M. S., Inn, A., & Anderson, L. R. (1980). Expectations of Reinforcement from Alcohol - Their Domain and Relation to Drinking Patterns. *J Consult Clin Psychol*, *48*, 419-426. doi: Doi 10.1037//0022-006x.48.4.419
- Cable, N., & Sacker, A. (2007). The role of adolescent social disinhibition expectancies in moderating the relationship between psychological distress and alcohol use and misuse. *Addict Behav*, *32*, 282-295. doi: 10.1016/j.addbeh.2006.04.001
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci*, *2*(40), 40. doi: 10.1186/1748-5908-2-40
- Christiansen, B. A., Goldman, M. S., & Inn, A. (1982). Development of alcohol-related expectancies in adolescents: separating pharmacological from social-learning influences. *J Consult Clin Psychol*, *50*, 336-344.
- Christiansen, B. A., Roehling, P. V., Smith, G. T., & Goldman, M. S. (1989). Using Alcohol Expectancies to Predict Adolescent Drinking Behavior after One Year. *J Consult Clin Psychol*, *57*, 93-99. doi: Doi 10.1037/0022-006x.57.1.93
- Cross, A. B., Gottfredson, D. C., Wilson, D. M., Rorie, M., & Connell, N. (2010). Implementation quality and positive experiences in after-school programs. *Am J Community Psychol*, *45*, 370-380. doi: 10.1007/s10464-010-9295-z



- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin Psychol Rev*, *18*, 23-45.
- Danielsson, A., Wennberg, P., Hibell, B., & Romelsjö, A. (2011). Alcohol use, heavy episodic drinking and subsequent problems among adolescents in 23 European countries: does the prevention paradox apply? *Addiction*, *107*, 71-80. doi: 10.1111/j.1360-0443.2011.03537.x
- Darkes, J., & Goldman, M. S. (1993). Expectancy challenge and drinking reduction: experimental evidence for a mediational process. *J Consult Clin Psychol*, *61*, 344-353. doi: 10.1037/0022-006X.61.2.344
- Derzon, J. H., Sale, E., Springer, J. F., & Brounstein, P. (2005). Estimating intervention effectiveness: synthetic projection of field evaluation results. *Journal of Primary Prevention*, *26*, 321-343. doi: 10.1007/s10935-005-5391-5
- Dobson, D., & Cook, T. J. (1980). Avoiding Type III Error in Program Evaluation. *Evaluation and Program Planning*, *3*, 269-276.
- Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation*, *11*, 193-221. doi: Doi 10.1207/S1532768xjepc1102\_04
- Donati, M. A., Primi, C., & Chiesi, F. (2013). Prevention of Problematic Gambling Behavior Among Adolescents: Testing the Efficacy of an Integrative Intervention. *J Gambl Stud*. doi: 10.1007/s10899-013-9398-1
- DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *Am J Community Psychol*, *30*, 157-197. doi: 0091-0562/02/0400-0157/0
- Durlak, J. A. (1995). *School-based prevention programs for children and adolescents* (D. S. Foster Ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Durlak, J. A. (1997). *Successful Prevention Programs for Children and Adolescents*. New York: Plenum Press, new York.

- 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. *Am J Community Psychol*, *41*, 327-350. doi: 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*, *18*, 237-256. doi: 10.1093/her/18.2.237
- Ennett, S. T., Haws, S., Ringwalt, C. L., Vincus, A. A., Hanley, S., Bowling, J. M., & Rohrbach, L. A. (2011). Evidence-based practice in school substance use prevention: fidelity of implementation under real-world conditions. *Health Education Research*, *26*, 361-371. doi: 10.1093/her/cyr013
- Feldman, L., Harvey, B., Holowaty, P., & Shortt, L. (1999). Alcohol use beliefs and behaviors among high school students. *Journal of Adolescent Health*, *24*, 48-58. doi: 10.1016/S1054-139x(98)00026-3
- Felner, R. D., Brand, S., Mulhall, K. E., Counter, B., Millman, J. B., & Fried, J. (1994). The parenting partnership: The evaluation of a human service/corporate workplace collaboration for the prevention of substance abuse and mental health problems, and the promotion of family and work adjustment. *The journal of primary prevention*, *15*, 123-146. doi: 10.1007/BF02197143
- Felner, R. D., Phillips, R. S., DuBois, D., & Lease, A. M. (1991). Ecological interventions and the process of change for prevention: wedding theory and research to implementation in real world settings. *Am J Community Psychol*, *19*, 379-387; discussion 395-403. doi: 10.1007/BF00938030
- Field, A. (2009). *Discovering statistics using SPSS* (3 ed.). Los Angeles: SAGE Publications Ltd.
- Foxcroft, D. R., & Tsertsvadze, A. (2011). Universal school-based prevention programs for alcohol misuse in young people. *Cochrane Database Syst Rev*(5), CD009113. doi: 10.1002/14651858.CD009113

- Fromme, K., & D'Amico, E. J. (2000). Measuring adolescent alcohol outcome expectancies. *Psychol Addict Behav, 14*, 206-212. doi: 10.1037/0893-164X.14.2.206
- Giles, S., Jackson-Newsom, J., Pankratz, M. M., Hansen, W. B., Ringwalt, C. L., & Dusenbury, L. (2008). Measuring quality of delivery in a substance use prevention program. *Journal of Primary Prevention, 29*, 489-501. doi: 10.1007/s10935-008-0155-7
- helsenorge.no. (2013). Alkohol. from <http://helsenorge.no/Helseogsunnhet/Sider/Alkohol.aspx>
- Hibell, B., Andersson, B., Bjarnason, T., Ahlström, S., Balakireva, O., Kokkevi, A., & Morgan, M. (2004). The ESPAD Report 2003: Alcohol and Other Drug Use Among Students in 35 European Countries. Stockholm, Sweden: The Swedish Council for Information on Alcohol and Other Drugs.
- Hibell, B., Guttormsson, U., Ahlström, S., Balakireva, O., Bjarnason, T., Kokkevi, A., & Kraus, L. (2011). The 2011 ESPAD Report: Substance Use Among Students in 36 European Countries. Stockholm, Sweden: The Swedish Council for Information on Alcohol and Other Drugs.
- Hunter, C. (2012). The politics of participation - with spscific reference to teacher-pupil relationships. In P. Woods (Ed.), *Teacher strategies: Explorations in the sociology of the school* (pp. 213-236). Milton Park, Abingdon: Routledge.
- Karwacki, S. B., & Bradley, J. R. (1996). Coping, drinking motives, goal attainment expectancies and family models in relation to alcohol use among college students. *journal of drug education, 26*, 243-255. doi: Doi 10.2190/A1p0-J36h-Tlmj-0l32
- Kline, R. B. (1995). Eight-month predictive validity and covariance structure of the Alcohol Expectancy Questionnaire for Adolescents (AEQ-A) for junior high school students. *Journal of Studies on Alcohol and Drugs, 57*, 396-405.
- Kuntsche, E., Knibbe, R., Gmel, G., & Engels, R. (2005). Why do young people drink? A review of drinking motives. *Clinical Psychology Review, 25*, 841-861. doi: 10.1016/j.cpr.2005.06.002

- Labouvie, E., & Bates, M. E. (2002). Reasons for alcohol use in young adulthood: Validation of a three-dimensional measure. *Journal of Studies on Alcohol*, *63*, 145-155.
- Lewis, M. (2007, February 7). *Stepwise versus Hierarchal Regression: Pros and Cons*. Academia.edu, Southwest Educational Research Association.
- Lim, S. S., Lopez, A. D., Murray, C. J. L., & Ezzati, M. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *380*, 2224-2260
- Lo, C. C., & Globetti, G. (2000). Gender differences in drinking patterns among Hong Kong Chinese youth: A pilot study. *Substance Use & Misuse*, *35*, 1297-1306. doi: Doi 10.3109/10826080009147483
- Mann, L. M., Chassin, L., & Sher, K. J. (1987). Alcohol expectancies and the risk for alcoholism. *J Consult Clin Psychol*, *55*, 411-417. doi: 10.1037/0022-006X.55.3.411
- Moncher, F. J., & Prinz, R. J. (1991). Treatment Fidelity in Outcome Studies. *Clinical Psychology Review*, *11*, 247-266. doi: Doi 10.1016/0272-7358(91)90103-2
- Natvig, H. (2009). Primary prevention of alcohol and drug use in junior high school. *Nordic Studies on Alcohol and Drugs*, *26*, 536-541.
- Natvig, H., & Aarø, L. E. (in press). *Effects of induced compliance in alcohol use: evaluation of a school-based intervention among Norwegian 8th graders*. *Nordic Psychology*.
- Oesterle, S., Hill, K. G., Hawkins, J. D., Guo, J., Catalano, R. F., & Abbott, R. D. (2004). Adolescent Heavy Episodic Drinking Trajectories and Health in Young Adulthood\*. *J. Stud. Alcohol*, *65*, 204-212.
- Patrick, M. E., Wray-Lake, L., Finlay, A. K., & Maggs, J. L. (2010). The long arm of expectancies: adolescent alcohol expectancies predict adult alcohol use. *Alcohol Alcohol*, *45*(1), 17-24. doi: 10.1093/alcalc/agg066
- Payne, A. A., Gottfredson, D. C., & Gottfredson, G. D. (2006). School predictors of the intensity of implementation of school-based prevention programs: results from a national study. *Prev Sci*, *7*, 225-237. doi: 10.1007/s11121-006-0029-2

- Pettigrew, J., Miller-Day, M., Shin, Y., Hecht, M. L., Krieger, J. L., & Graham, J. W. (2013). Describing teacher-student interactions: a qualitative assessment of teacher implementation of the 7th grade keepin' it REAL substance use intervention. *Am J Community Psychol*, *51*, 43-56. doi: 10.1007/s10464-012-9539-1
- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., . . . Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*, *38*(2), 65-76. doi: 10.1007/s10488-010-0319-7
- Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. L. Christenson (Ed.), *Handbook of Research on Student Engagement* (pp. 149-172). New York: Springer. doi: 10.1007/978-1-4614-2018-7\_7
- Schulte, M. T., Ramo, D., & Brown, S. A. (2009). Gender differences in factors influencing alcohol use and drinking progression among adolescents. *Clin Psychol Rev*, *29*, 535-547. doi: 10.1016/j.cpr.2009.06.003
- Shernoff, D. J. (2010). Engagement in after-school programs as a predictor of social competence and academic performance. *Am J Community Psychol*, *45*, 325-337. doi: 10.1007/s10464-010-9314-0
- Shernoff, D. J., & Vandell, D. L. (2007). Engagement in after-school program activities: quality of experience from the perspective of participants. *Journal of Youth and Adolescence*, *36*, 891-903. doi: DOI 10.1007/s10964-007-9183-5
- Smith, J. D., Schneider, B. H., Smith, P. K., & Ananiadou, K. (2004). The effectiveness of whole-school antibullying programs: A synthesis of evaluation research. *School Psychology Review*, *33*, 547-560.
- Vedøy, T. F., & Skretting, A. (2009). *Ungdom og rusmidler: Resultater fra spørreskjemaundersøkelser 1968-2008*. Oslo, Norway: Statens institutt for rusmiddelforskning (SIRUS).
- Weinman, M. L., Schreiber, N. B., & Robinson, M. (1992). Adolescent mothers: Were there any gains in a parent education program? *Family and Community Health*, *15*(3), 1-10.

- Wiers, R. W., & Kummeling, R. H. (2004). An experimental test of an alcohol expectancy challenge in mixed gender groups of young heavy drinkers. *Addict Behav*, *29*, 215-220. doi: 10.1016/S0306-4603(03)00081-9
- Wilhelmsen, B. U. (1997). *Development and evaluation of two educational programmes designed to prevent alcohol use among adolescents*. (Doctor of Philosophy), University of Bergen, Bergen, Norway.
- Wilhelmsen, B. U., Laberg, J. C., & Aas, H. N. (1997). Alcohol outcome expectancies in adolescence. *Psychology and Health*, *12*, 1037-1044. doi: 10.1080/08870449808407448
- Wilhelmsen, B. U., Laberg, J. C., & Klepp, K. I. (1994). Evaluation of two student and teacher involved alcohol prevention programmes. *Addiction*, *89*, 1157-1165. doi: 10.1111/j.1360-0443.1994.tb02792.x
- Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: a meta-analysis. *J Consult Clin Psychol*, *71*, 136-149. doi: 10.1037/0022-006X.71.1.136
- World Health Organization, W. (2010). Global strategy to reduce the harmful use of alcohol. In WHO (Ed.). Geneva, Switzerland: World Health Organization.
- World Health Organization, W. (2011). Alcohol fact sheet. from <http://www.who.int/mediacentre/factsheets/fs349/en/index.html>
- World Health Organization, W. (2013). Alcohol. from [http://www.who.int/substance\\_abuse/facts/alcohol/en/index.html](http://www.who.int/substance_abuse/facts/alcohol/en/index.html)
- World Health Organization, W. (in press). Global Status Report on Alcohol and Health 2013. In WHO (Ed.). Geneva, Switzerland: World Health Organization.
- Zamboanga, B. L., Ham, L. S., Van Tyne, K., & Pole, N. (2011). Alcohol expectancies among adolescent nondrinkers: they may not be drinking now, but they're "thinkin bout it". *J Adolesc Health*, *49*, 105-107. doi: 10.1016/j.jadohealth.2010.12.004