Statistical analysis plan:
"Efficacy testing of the VR-game "VR FestLab" - Does a VR based game improve alcohol resistance skills among adolescents?"

Statistical analysis plan (SAP)

"Efficacy testing of the VR-game "VR FestLab" - Does a VR based game improve alcohol resistance skills among adolescents?"

Version 1.0

Intervention group:	VR FestLab simulation game experience of 15 minutes per student followed by structured class-room discussion
Control group:	Oculus Quest – first steps simulation game experience of 15 minutes per student followed by structured class-room discussion
Study population:	Students at Danish public schools
Funding:	Danish Safety Foundation TrygFonden, Denmark (Funding ID 129372)
Clinical trials.gov Identifier:	Current Controlled Trails with study ID ISRCTN11768445, https://www.isrctn.com/search?q=ISRCTN11768445

Approved by:

Christiane Stock, Principal Investigator

Berlin 01/03/2021 C. Hock
Place, Date, Signature

Julie Dalgaard Guldager, Study coordinator

Esbjerg Denmark 27/02/2021
Place, Date, Signature

Ulrike Grittner, Statistician

Ulrike Grittner — Digital unterschrieben von Ulrike Grittner Datum: 2021.02.24 08:09:45 +01'00'

Place, Date, Signature

Content

1. Background

Denmark is among the countries with the highest rate of risky alcohol use in Europe (Kraus and Nociar, 2016 ESPAD Group, 2016). Understanding of what leads children and adolescents to begin using alcohol is important for developing effective programs. In fact, peer pressure has been identified as one of the most common reasons for adolescent uptake and continued consumption of alcohol (Griffin and Botvin, 2010; Herndricks et al., 2015). Virtual Reality (VR) is an emerging educational tool (Durl et al., 2018) that allows for simulating a virtual learning environment that is closely approximating reality (Barko and Sadler, 2013). The convincing nature of characters and environments in VR simulations can improve learning and encourage adolescents to strengthen their own levels of knowledge and skills. This offers unique opportunities to provide learners with similar-to-real-life experiences on how to react positively and in socially accepted ways to avoid social pressure to drink. VR simulations are more engaging to adolescents and have been demonstrated to achieve initial teaching successes (McGrath et al. 2010). We assume that adolescents will benefit from using such tools, because it is engaging and fun to use as well as effective in developing relevant life skills and favourable alcohol expectancies. However, there exist only very few similar virtual reality applications for use in health promotion and preventive interventions worldwide. To the best of our knowledge no virtual reality-based alcohol resistance tool was tested before with respect to its efficacy to enhance alcohol resistance skills in a controlled trial (Durl et al., 2018).

1.1 Study Objective

This study aims to evaluate the efficacy of the VR FestLab game aimed to improve resistance skills of adolescents (15-18 years) who face social pressures to drink via a cluster-randomised controlled trial. The primary outcomes that will be investigated are the drinking refusal self-efficacy (DRSEQ) and alcohol refusal skills constructs.

1.2 Primary Hypothesis

The main hypothesis is that adolescents who experience VR FestLab score higher in *Drug resistance skills* as measured by the social pressure subscale DRSEQ-R at T1 compared to those in an active control group.

1.3 Secondary Hypotheses

Adolescents in the intervention group will be compared to those in the control group and the differences are expected to be:

- Higher score in *Drug resistance skills* at T2 (measured by the social pressures sub-scale of DRSEQ-R)

- No difference in *Gameplay experiences* at T1 and T2 (measured by the short version of the User Experience Questionnaire (Short-EUQ)).

- Higher *Resistance towards peer pressure to drink* at T1 and T2 (measured by the mean score of the 5-item Social Pressure subscale of the DRSEQ-R).

- Higher *Drug refusal skills* at T1 and T2 (measured by the mean score of the 7-item subscale on Drug Refusal Skills from the Brief Assessment Tool of the Life Skills Training).

- Higher *Knowledge/Awareness* at T1 and T2 (measured by a mean score of the items "It is easy for me to estimate my own alcohol tolerance" and "I know how much alcohol I can drink before I get drunk" with response options on a 5 point Likert scale from disagree strongly to agree strongly.)

- Higher *Communication skills* at T1 and T2 (measured using a mean score of the items "If my best friends want me to drink beer with them and I don`t want to, I have ways to say no" and "If someone offers me a drink of alcohol and I say "no", I can make them take "no" for an answer" from the Alcohol Misuse Prevention Knowledge Questionnaire with response options on a 5 point Likert scale from disagree strongly to agree strongly).

- Higher rating of *Positive role-models* at T1 and T2 (measured with the disagreement to the item "Spending time with people who drink alcohol is more fun than spending time with people who do not drink alcohol" from the Regan Attitudes toward Non-Drinkers Scale (RANDS) with response options on a 5 point Likert scale from disagree strongly to agree strongly).

- Higher *Awareness of social pressures* at T1 and T2 (measured using the item "If I am at a party and my friends are drinking alcohol, I would feel left out if I am not drinking alcohol" of the Susceptibility to Peer Pressure Index on a 5 point Likert scale from disagree strongly to agree strongly)

- Higher *Social support willingness* at T1 and T2 (measured with a sum score for the item "If someone is really drunk or sick at a party, the best thing to do is..." with multiple response options as yes answers: "Let him or her recover alone" (0 point), "Help him or her to get better" (1 point), "Ask an adult for help" (1 point), "Call his or her parents" (1 point)).

- Lower level of positive *Outcome expectations* at T1 and T2 (measured using the 4 item outcome expectation scale from the Danish "Ungeprofil" [Youth profile], which asks respondents to rate on a

5 point Likert scale, to what extent they agree to expectations, when they drink alcohol, such as "I become more extroverted" or "I forget my problems").

1.4 Study Design

The efficacy of the simulation game VR FestLab is studied in a cluster randomised controlled trial (cluster RCT) including A) an intervention condition (VR FestLab) and B) an active control condition (Oculus Quest - first steps) with superiority 1:1 parallel design.

1.5 Sample Size Calculation

We used an estimated mean of the Social Pressure subscale of the DRSEQ-R of 19.0 with 6.5 SD based on previous studies (Ehret et al., 2000). No intervention effect measures for DRSEQ-R in similar intervention settings were available. However, we assume to detect an intervention effect of at least 0.44 (Cohen's d), which corresponds to a mean difference for example of 2.85 and a common standard deviation of 6.5 points. Using STATA 15, a power calculation for a clustered randomised design and a two sample means test resulted in a sample size of 135 for the control group (B) and 135 for the intervention group (A) to detect an intervention effect of 0.44 between intervention group and control group assuming a power of 0.80 and an alpha of 0.05. The sample size was calculated based on an ICC for drug resistance skills of 0.01 and 45 students per school. Taking an estimated attrition of 35% into account we will recruit 420 participants for the trial (210 for each group).

2. Analysis Sets

2.1 Definitions

The full analysis set comprises all subjects who were cluster-randomised. Randomisation is performed school-wise with a 1:1 allocation ratio. In case of missing values, multiple imputation methods will be used to estimate missing values. The per protocol analysis set and the safety analysis comprise all subjects who participated in the VR game play session, when conducted in their class. We expect that the per protocol analysis set will comprise more than 98% of the intention-to-treat analysis set.

2.2 Application

The primary efficacy analysis will be done using the full analysis set including estimated values from multiple imputations for missing values (Intention to treat). An analysis of the primary outcome in the per protocol analysis set will not be used as sensitivity analysis, because the datasets will probably differ by less than 2%.

The safety analysis will be done using the safety analysis set and comparing participants according to the actual intervention they received.

3. Trial Centres

3.1 Recruitment

For the trial, approximately 10-13 Danish schools will be recruited including public schools, grammar schools and public boarding schools in the Region of Southern Denmark. The boarding schools will ensure better geographical coverage of study participants, because students typically have their home address and social background in different regions of Denmark. Schools that have students in the age range of students from 15-18 years will be invited to participate via email to be sent to the school principals. In order to increase the willingness of schools to participate in the trial, a session with the prevention programme VR FestLab will be offered to control schools as well after the trial period.

4. Analysis, Variables and Time Points of Assessment

4.1. Time points of assessment

Data collections take place at three time points: before intervention, but after school-based randomisation, immediately after the intervention/control session, and after 6 weeks as 2nd follow-up. The baseline assessment (T0) comprises a questionnaire for the student (see details in the next section). After the intervention/active control session the follow-up assessment (T1) is completed. Six weeks later, the second follow-up (T2) data collection is planned, with the same questionnaire from T1 (see figure 1). Such repeated measurements require linking together data from the same participant at different time points. To link each participant's responses from the three different time points (T0, T1 and T2) an individual anonymous code is generated based on background information such as first letter in mother's name, school grade and month of birth following procedures outlined in Rundle-Thiele et al. (2015).

4.2. Content of the baseline questionnaire (T0)

Demographic information on age and sex is collected. As a proxy for family socioeconomic status, we use the question "How well off do you think is your family?" with 5 answering options from "very well off" to "not at all well off" from the Health Behaviour in School-aged Children study (Currie et al., 2014). Additionally, we ask students, if they had ever been drinking alcohol, been drunk, and been drinking 5 or more drinks at the same occasion based on lifetime measures from ESPAD (Kraus et al., 2016). The last 30 days alcohol use, is measured using three items from the Danish youth survey MULD (Kræftens Bekæmpelse og

6

Sundhedsstyrelsen, 2009), asking the students on how many occasions within the last 30 days they have been drinking alcohol, been drunk, and been drinking 5 or more drinks at the same occasion.

Using items from the HBSC questionnaire students are asked to rate their self-perceived health and their school performance according to their teachers' assessment (Currie et al., 2014). Life satisfaction is measured according to the MULD questionnaire (Kræftens Bekæmpelse og Sundhedsstyrelsen, 2009). We also apply the Strengths and Difficulties Questionnaire (SDQ) is applied in its Danish version (Niclasen et al., 2012). All participants are asked to respond to the Sensation Seeking Scale (SSS) for adolescents with Cronbach's alpha 0.76 (Hoyle et al., 2002).

4.2.1 Primary outcome measure

Resistance towards peer pressure to drink is measured by the mean score of the 5-item Social Pressure subscale of the DRSEQ-R at T1. The Cronbach's alpha of the subscale is 0.87 and split-half reliability of 0.72 (Young et al., 2007). The Cronbach's alpha for the Danish version is 0.92.

4.2.2 Secondary outcome measures

In an explorative analysis the following secondary outcomes are analysed:

Resistance towards peer pressure to drink is measured by the mean score of the 5-item Social Pressure subscale of the DRSEQ-R at T2.

Drug refusal skills is measured using mean score of the 7-item subscale on Drug Refusal Skills from the Brief Assessment Tool of the Life Skills Training (Macaulay et al., 2002). The Drug Refusal Skills scale has a Cronbach's alpha of 0.85 (Macaulay et al., 2002). The Cronbach's alpha for the Danish version is 0.78.

We measure the intermediate effects of the behaviour change functions education, training, modelling and coercion/incentivisation on the following secondary outcomes as potential moderating variables all treated as means of scores or answering categories.

Knowledge/Awareness of blood alcohol concentration is assessed by a mean score of the items "It is easy for me to estimate my own alcohol tolerance" and "I know how much alcohol I can drink before I get drunk" with response options on a 5 point Likert scale from disagree strongly to agree strongly.

Communication skills are measured using a mean score of the items "If my best friends want me to drink beer with them and I don't want to, I have ways to say no" and "If someone offers me a drink of alcohol and I say "no", I can make them take "no" for an answer" from the Alcohol Misuse Prevention Knowledge Questionnaire (Shope et al., 1993) with response options on a 5 point Likert scale from disagree strongly to agree strongly.

Positive role-models are measured with the disagreement to the item "Spending time with people who drink alcohol is more fun than spending time with people who do not drink alcohol" from the Regan Attitudes toward Non-Drinkers Scale (RANDS) with response options on a 5 point Likert scale from disagree strongly to agree strongly (Regan et al., 2017).

Awareness of social pressures is measured using the item "If I am at a party and my friends are drinking alcohol, I would feel left out if I am not drinking alcohol" of the Susceptibility to Peer Pressure index (Dielman et al., 1987) on a 5 point Likert scale from disagree strongly to agree strongly.

Social support willingness is measured with a sum score for the item "If someone is really drunk or sick at a party, the best thing to do is..." with multiple response options as yes answers: "Let him or her recover alone" (0 point), "Help him or her to get better" (1 point), "Ask an adult for help" (1 point), "Call his or her parents" (1 point).

Outcome expectations are measured using the 4 item outcome expectation scale from the Danish "Ungeprofil" [Youth profile], which asks respondents to rate on a 5 point Likert scale, to what extent they agree to expectations, when they drink alcohol, such as "I become more extroverted" or "I forget my problems" (Bendtsen et al., 2015).

4.3. Content of the follow-up questionnaires

The follow-up questionnaires at T1 and T2 contain all questions related to the primary and the secondary outcomes from the baseline questionnaire. In addition, we ask students regards the following aspects at T1:

Potential adverse effects, such as cyber sickness (LaViola, 2000) in an open question.

Gameplay experience is measured using the short version of the User Experience Questionnaire (Short-EUQ) with the eight parameters of a game being; obstructive/supportive, complicated/easy, inefficient/efficient, confusing/clear, boring/exiting, not interesting/interesting, conventional/inventive, and usual/leading edge (Rauschenberger et al., 2013). The Cronbach's alpha of the scale in the Danish version is 0.72. This scale is administered in intervention and control schools at T1.

Items measuring different aspects of *Student satisfaction with VR FestLab* were developed by the research team and consist of seven statements to which degree the students agreed that they: liked to try the game, would like to explore the game further, did not like the VR experience, would recommend the game to his/her friends, thought the characters in the game were unrealistic, thought the game was realistic, and did not like the music. Answers to negatively worded questions will be reversed for analysis (5-point Likert scale from

"disagree a lot" to "agree a lot"). These items assessing satisfaction with VR FestLab are only administered in intervention schools as a measure of process evaluation.
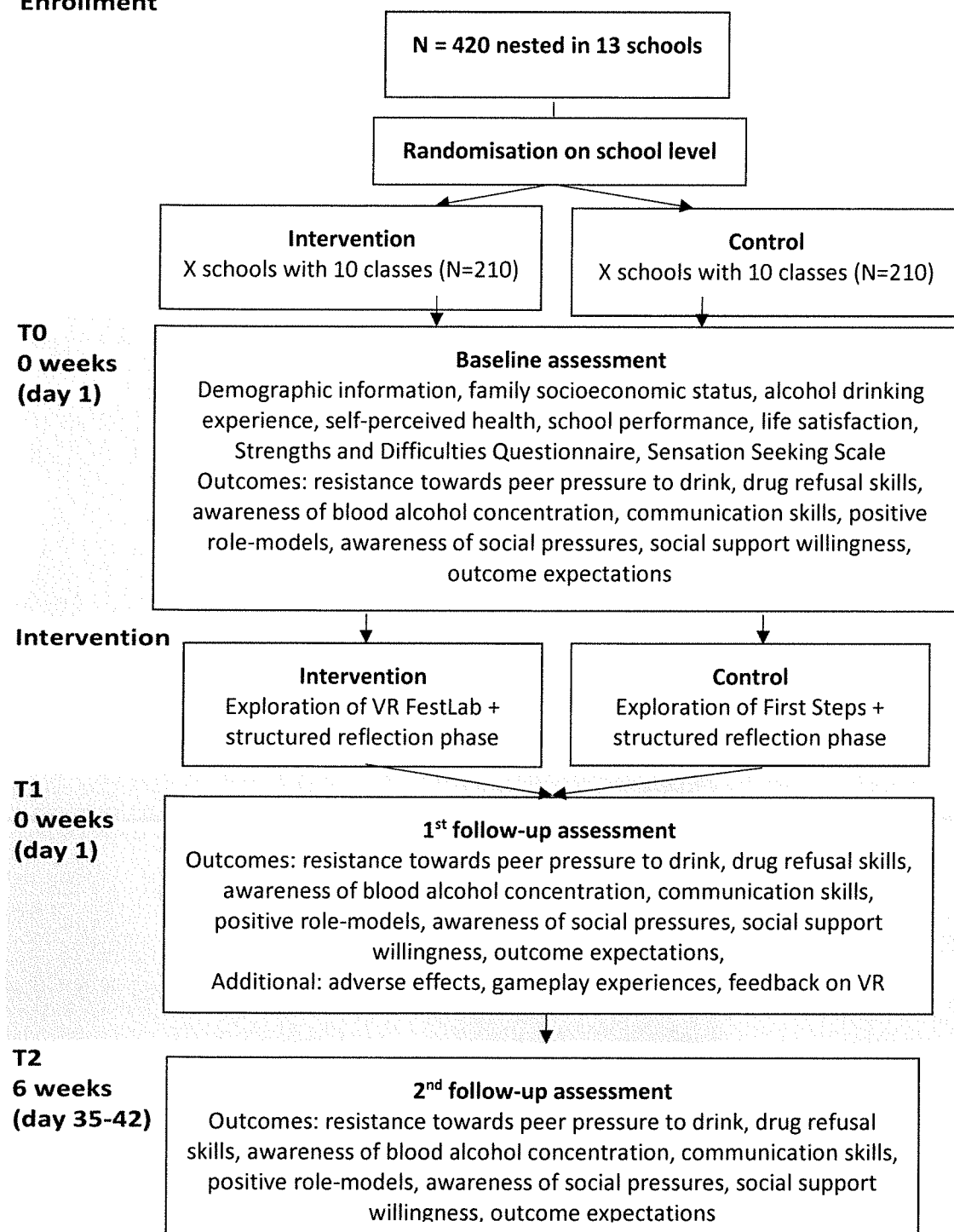
Enrollment

```
                    ┌─────────────────────────────────┐
                    │   N = 420 nested in 13 schools   │
                    └─────────────────────────────────┘
                                   │
                    ┌─────────────────────────────────┐
                    │   Randomisation on school level  │
                    └─────────────────────────────────┘
                        ╱                         ╲
        ┌───────────────────────────┐   ┌───────────────────────────┐
        │        Intervention        │   │          Control           │
        │ X schools with 10 classes  │   │ X schools with 10 classes  │
        │         (N=210)            │   │         (N=210)            │
        └───────────────────────────┘   └───────────────────────────┘
```

T0
0 weeks
(day 1)

Baseline assessment
Demographic information, family socioeconomic status, alcohol drinking experience, self-perceived health, school performance, life satisfaction, Strengths and Difficulties Questionnaire, Sensation Seeking Scale
Outcomes: resistance towards peer pressure to drink, drug refusal skills, awareness of blood alcohol concentration, communication skills, positive role-models, awareness of social pressures, social support willingness, outcome expectations

Intervention

Intervention	Control
Exploration of VR FestLab + structured reflection phase	Exploration of First Steps + structured reflection phase

T1
0 weeks
(day 1)

1st follow-up assessment
Outcomes: resistance towards peer pressure to drink, drug refusal skills, awareness of blood alcohol concentration, communication skills, positive role-models, awareness of social pressures, social support willingness, outcome expectations,
Additional: adverse effects, gameplay experiences, feedback on VR

T2
6 weeks
(day 35-42)

2nd follow-up assessment
Outcomes: resistance towards peer pressure to drink, drug refusal skills, awareness of blood alcohol concentration, communication skills, positive role-models, awareness of social pressures, social support willingness, outcome expectations

Figure 1 VR FestLab study flowchart

4.3. Safety outcomes

The following measures will be used as safety outcomes recorded immediately after game play at Follow-Up T1: symptoms of cyber sickness (dizziness, feeling sick in stomach) or any other complaint. Symptoms will be grouped in classes and summarized including the following parameters: incidence of symptoms.

5. Handling of Missing Values

We expect very limited drop-out (<5%) between baseline (T0) and 1. follow-up (T1), because both data collections take place at the same day. Drop-out between 1. and 2. follow-up is estimated to be 35% based on experiences from a previous trial in Danish schools (Vallentin-Holbech, Rasmussen, & Stock, 2018).

In case of missing values and under the assumption of missing at random (MAR) or missing completely at random (MCAR) as missing data mechanism, data will be estimated using multiple imputation methods with 30 imputed data sets. To estimate values in a realistic range and with values similar as in complete cases, we will use predictive mean matching as this method uses original values of participants without missings for the estimation procedure.

6. Statistical Analyses

For all analyses (including analysis of primary outcome) appropriate descriptive statistics (mean, standard deviation, median, interquartile range, absolute and relative frequencies) depending on the scale and distribution of the outcome variable will be presented.

Statistical analyses will be divided to analyze:

1. immediate intervention effects by including all measures until include T1 (post intervention assessment).

2. long-term intervention effects by focusing on T2 (follow-up assessment 6 weeks after baseline)

6.1. Subject disposition

The study participants do not differ in disposition and no analyses are foreseen.

6.2. Demography and baseline characteristics

Frequencies, measures of location and scale, difference between therapy groups.

First, intervention and control group will be compared with respect to age, gender, and other data in order to assess the comparability between them with Chi Square Tests.

6.3. Exposition to intervention/Compliance

The intervention and control game experimental setting do not differ and no analyses are foreseen to measure compliance.

6.4. Primary Analysis

Intervention efficacy will be established with hierarchical linear regression modelling at follow-up T1 with the primary outcome variable as dependent variable while adjusting for baseline levels with random intercept for school as well as sex and grade as co-variates. The primary analysis will be carried out in the full analysis set including estimated values in case of missings (multiple imputed data sets). All data will be analysed using the intent-to-treat principle.

6.5. Secondary Analysis

Immediate intervention effects

Performance on secondary outcomes will be analyzed in the same manner as the primary outcome, using separate hierarchical regression models for change in secondary outcome from baseline to T1 as dependent variable for each endpoint, intervention group as co-variate, and particular baseline measure as well as grade and sex as covariates. Interaction terms and additional covariates will be included to adjust for possible confounders or to test subgroup differences. Type of regression (linear, ordinal or logistic) will depend on the scaling of the dependent variable.

Long-term intervention effects

Long term intervention effects will be analyzed using separate mixed models at T2 for each outcome. These models will include the baseline measure of the particular outcome, grade and sex as covariates, and a random effect for the participant (random intercept). Type of regression function (logistic, linear, ordinal) will depend on the scaling of the dependent variable. In case of skewed continuous data, variables will be transformed before analysis.

All secondary analyses will be done using the full analysis set with multiple imputed data in case of missing values. All secondary analyses will be done in an exploratory framework.

6.6. Safety/Tolerability

Safety outcomes will be reported separately as incidences (n, incidence rate with 95%CI) in total and by intervention group based on the safety analysis set. Participants will be grouped according to their actually received intervention. Group comparisons will be done using Chi-Square Tests. Results of safety analysis will

be interpreted and discussed thoroughly also for minor group differences, since statistically significance is not of importance here.

6.7. Planned Subgroup Analyses

The possible mediators and moderators will be included as explanatory variables into the hierarchical models after assessing the interactions between the intervention and the potential effect modifying variables.

Potential differences in the efficacy of the intervention on primary and secondary outcomes will be analyzed among the following pre-specified subgroups at time point T1 (subgroups defined by using baseline information):

- based on dichotomization with regard to legal drinking age (<16 and ≥16 years)

- based on dichotomization with regard to sex (male vs. female)

- based on dichotomization with regard to alcohol experience (lifetime binge drinking (5 or more drinks on one occasion) vs. no lifetime binge drinking)

- based on defined categories of sensation seeking personality scale (e.g., median split or defined categories)

- based on family wealth (not at all well off, not so well off, average vs. quite well off, very well off)

- based on defined categories of game play strategy (risk taking vs. risk avoiding)

- based on defined categories of game play experience questionnaire (e.g., median split or defined categories)

All subgroup analyses will be done within an exploratory framework. Moreover, to evaluate the complex relationship between several variables, such as primary/secondary outcomes, intervention group and game play strategies/game exploring, we also plan to conduct structural equation modelling.

7. References

Bendtsen P, Schou Mikkelsen S, Tolstrup JS. Ungdomsprofilen 2014: Sundhedsadfærd, helbred og trivsel blandt elever på ungdomsuddannelser. Copenhagen: Statens Institut for Folkesundhed, 2015.

Currie C, Inchley J, Molcho M, Lenzi M, Veselska Z, Wild F. Health Behaviour in School-aged Children (HBSC) Study Protocol: Background, Methodology and Mandatory items for the 2013/14 Survey: Child and Adolescent Health Research Unit (CAHRU); 2014.

Dielman TE, Campanelli PC, Shope JT, Butchart AT. Susceptibility to peer pressure, self-esteem, and health locus of control as correlates of adolescent substance abuse. Health Education Quarterly. 1987;14(2):207-21.

Durl J, Dietrich T, Pang B, Potter LE, Carter L. Utilising virtual reality in alcohol studies: A systematic review. Health Educ J. 2018; 77(2), 212-225. doi:10.1177/0017896917743534

Ehret PJ, Ghaidarov TM, LaBrie JW. Can you say no? Examining the relationship between drinking refusal self-efficacy and protective behavioral strategy use on alcohol outcomes. Addict Behav. 2013;38(4):1898-904.

ESPAD Group E. ESPAD Report 2015: Results from the European School Survey Project on Alcohol and Other Drugs. The ESPAD Group, 2016.

Griffin KW, Botvin GJ. Evidence-based interventions for preventing substance use disorders in adolescents. Child Adolesc Psychiatr Clin N Am. 2010;19(3):505-26.

Hendricks G, Savahl S, Florence M. Adolescent peer pressure, leisure boredom, and substance use in low-income Cape Town communities. Social Behavior and Personality: An International Journal. 2015;43(1):99-110.

Hoyle RH, Stephenson MT, Palmgreen P, Lorch EP, Donohew RL. Reliability and validity of a brief measure of sensation seeking. Personality Ind Differences. 2002;32(3), 401-414.

Kræftens Bekæmpelse og Sundhedsstyrelsen. Muld. Unges livsstil og dagligdag 2008. Muld-rapport nr. 7. Kræftens Bekæmpelse og Sundhedsstyrelsen.; 2009.

Kraus, L., Leifman, H., and Vicente, J. ESPAD Report 2015: Results from the European School survey Project on Alcohol and other Drugs. Luxembourg: Publications Office of the European Union, 2016.

LaViola JJ. A discussion of cybersickness in virtual environments. ACM Sigchi Bulletin.2000;32(1): 47-56.

Macaulay AP, Griffin KW, Botvin GJ. Initial internal reliability and descriptive statistics for a brief assessment tool for the Life Skills Training drug-abuse prevention program. Psychol Rep. 2002;91(2):459-62.

McGrath D, Wegener M, McIntyre TJ, Savage C, Williamson M. Student experiences of virtual reality: A case study in learning special relativity. Am J Physics. 2010;78(8):862-8.

Niclasen J, Teasdale TW, Andersen A-MN, Skovgaard AM, Elberling H, Obel C. Psychometric properties of the Danish Strength and Difficulties Questionnaire: the SDQ assessed for more than 70,000 raters in four different cohorts. PloS one. 2012;7(2):e32025-e.

Statistical analysis plan:
"Efficacy testing of the VR-game "VR FestLab" - Does a VR based game improve alcohol resistance skills among adolescents?"

Rauschenberger M, Schrepp M, Pérez Cota M, Olschner S, Thomaschewski J. Efficient measurement of the user experience of interactive products. How to use the user experience questionnaire (UEQ). Example: Spanish language version. 2013.

Regan D, Morrison TG. Temporal Stability and Predictive Validity of the Regan Attitudes Toward Non-Drinkers Scale. SAGE Open. 2017;7(2):2158244017698728.

Rundle-Thiele S, Schuster L, Dietrich T, Russell-Bennett R, Drennan J, Leo C, Connor JP. Maintaining or changing a drinking behavior? GOKA's short-term outcomes. J Bus Res. 2015;68(10), 2155-2163.

Shope JT, Copeland LA, Maharg R, Dielman TE, Butchart AT. Assessment of adolescent refusal skills in an Alcohol Misuse Prevention Study. Health Educ Quart. 1993;20(3):373-90.

Vallentin-Holbech L, Rasmussen B, Stock C. Effects of the social norms intervention The GOOD Life on norm perceptions, binge drinking and alcohol-related harms: a cluster-randomised controlled trial. Prev Med Rep 2018; 12: 304-311

Young RM, Hasking PA, Oei TP, Loveday W. Validation of the Drinking Refusal Self-Efficacy Questionnaire--Revised in an adolescent sample (DRSEQ-RA). Addict Behav. 2007;32(4):862-8.

Young RM, Oei TP. The predictive utility of drinking refusal self-efficacy and alcohol expectancy: a diary-based study of tension reduction. Addict Behav. 2000;25(3):415-21