

# Challenges in establishing positive mental health effects of a school-based intervention: a cluster-randomized controlled trial of a life skill program in ten public high schools in Norway

Gry Anette Sælid (✉ [grsa@fhi.no](mailto:grsa@fhi.no))

Norwegian Institute of Public Health <https://orcid.org/0000-0003-3347-0749>

Nikolai Olavi Czajkowski

University of Oslo: Universitetet i Oslo

Leif Edvard Aarø

Norwegian Institute of Public Health: Folkehelseinstituttet

Thormod Idsøe

Norwegian Center for Child Behavioral Development: Nasjonalt utviklingscenter for barn og unge AS

Miguel Delgado Hellesester

California State University Channel Islands

Arne Holte

Norwegian Institute of Public Health: Folkehelseinstituttet

---

## Research Article

**Keywords:** mental health promotion, life skills, universal

**Posted Date:** October 21st, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-878704/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

Schools are well positioned for large scale promotion of mental health among young people. We tested the MindPower program delivered universally in the classroom to strengthen self-efficacy, self-esteem, and quality of life in high school students. We used a cluster randomized two-groups delayed intervention design. Participants were 1,673 out of 2,384 students, aged 15–16 years, in all the 110 first year high school classes in a county in Norway, randomized into two intervention groups, IG1 and IG2. IG1 started with MindPower. IG2 waited for four months to participate. Seven repeated measures were conducted with Rosenberg Self-esteem Scale (short version), General Perceived Self-Efficacy Scale (short version), and Warwick Edinburg Mental Wellbeing Scale. Self-efficacy level was compared to scores from a large comparable population study, locally and nationally (UngData). Mixed model analyses adjusted for cluster effects showed no significant increase in the three measures from start to five to nine months follow up. Self-efficacy level was approximately the same as the one found in UngData. The universal implementation in schools was a great challenge and attrition was comprehensive. MindPower did not strengthen students' levels of self-efficacy, self-esteem, and quality of life, partly because the initial levels were already high (ceiling effect). In spite of careful preparations, fidelity was low and attrition high. Results should therefore be interpreted with caution. Learning points in implementing and conducting class based universal mental health programs in high school are discussed comprehensively. ClinicalTrials.gov: NCT03647826. The 27.08.2018. Retrospectively registered.

## Background

Schools are well positioned to promote mental health and prevent mental illness among young people due to their organization and the amount of time all young people spend there (Sturgeon, 2006). Developing personal skills to cope with the challenges of life is one of many aspects of mental health promotion (Lahtinen, Joubert, Raeburn et al., 2005). Personal skills develop in interaction with environmental factors such as parents, peers, schools, and local community. Positive self-perceptions and quality of life are two among many factors that correlate positively with optimal learning, and these are also important in developing good mental health (Taylor, Oberle, Durlak et al., 2017; Chodkiewicz & Boyle, 2017).

There is substantial evidence of increased positive mental health and wellbeing and decreased emotional and behavioral risk factors following implementation of effective school based programs (O'Reilly, Svirydzenka, Adams et al., 2018; Werner-Seidler, Perry, Calear et al., 2017; Taylor et al., 2017). Numerous guidelines and policies on how to integrate health and education in schools have also been published (Atkins, Hoagwood, & Seidman, 2010; Scottish Government 2017; Welsh Government, 2017).

In this study, we assess effects of the MindPower program delivered universally in a classroom setting. The theoretical frame is derived from the Coping With Depression (CWD) course. CWD is one of the most tested programs during the past 30 years (Clarke, Hawkins, Murphy et al., 1993; Cuijpers, Muñoz, Clark et al., 2009; van Zoonen et al., 2014). Adaptions from CWD are ranging from treatment of depression (Lewinsohn (1970; 1984), to promoting mental health in the workplace (Saelid, Czajkowski, Holte et al., 2016a; Saelid, Czajkowski, Holte et al., 2016b). Also, interventions have similarities to CWD without referring to the CWD-concept (Keles & Idsoe, 2020).

Why then conduct still another effect study on CWD/CWS-based interventions?

First, results from Norway's most comprehensive study of high school students' mental health, UngData, show that there has been a continuous increase in self-reported symptoms of anxiety, depression and mental distress among high school students during the past 20 years (Bakken, 2014). Recently, there is an increase of five percentage points among girls (Bakken, 2017), which is consistent with results from another highly representative population study, UngHunt, where 44% of girls 15–17 years of age had a high level of symptoms of depression scores (Hunt, 2019). Taking into account the high degree of impairment associated with mental distress and mental disorder (Knudsen, Tollånes, Haaland et al., 2017), the high costs of illness of common mental disorder (The Norwegian Directorate of Health, 2015), the limited effects of treatment (Cuijpers, Karyotaki, Weitz et al., 2014), and the rate of return to investment from early intervention (Heckman, 2006), action is warranted if not required.

Second, these high prevalence rates can hardly be effectively mitigated at a population level by targeted interventions. According to Rose's (1992) classical and still valid prevention paradigm, underpinned empirically by Mackenbach, Lingsma, van Ravesteyn et al., (2011), Brugha, Morrell, Slade et al., (2011), and others, universal interventions should be preferred rather than targeted ones when a condition is that widespread. In implementing universal interventions among young people, no arena can compete with the school.

Third, since the Foresight report (Jenkins et al., 2008), emphasis on strengthening health and quality of life in the population, i.e. health promotion, rather than trying to prevent specific mental disorders, i.e. disease prevention, has come to the forefront also in the area of public mental health interventions. Several studies have found positive effects on youth mental health after life-skills programs. The effectiveness of such programs in a school context has been meta-analyzed by Taylor et al.'s (2017). However, reviews have also concluded that there are serious risks of biases in a number of these studies. This casts doubt on conclusions from these studies (O'Reilly et al., 2018; Caldwell, Davies, Hetrick et al. 2019).

The CWD/CWS-based interventions, on the other hand, are scientifically solidly evidence based and today the most well documented interventions to prevent common mental disorders. But, common to all these studies is that they only target groups with an elevated risk of mental disorder.

To our knowledge, none of the previous CWD/CWS-based interventions have been designed both to promote positive mental health and quality of life among young people and to be implemented universally, i.e. independently of risk of mental distress or disorder.

Consequently, MindPower was born as a CWD/CWS-based intervention delivered universally in the school, class wise to promote positive mental health and quality of life among high school students. The aim of this controlled trial is to test the hypotheses that MindPower strengthens first year high-school students' levels of self-efficacy, self-esteem, and quality of life.

## Methods

### Study design

We used a cluster randomized parallel design, of a two-group delayed intervention design with mixed model analyses adjusted for cluster effects. Allocation ratio was 1:1. School classes were randomly assigned to one of two intervention groups. The students in both groups completed questionnaires six times at school (T1-T6), and the seventh and final time at follow up at home (T7).

The study design is illustrated in Fig. 1. Intervention group 1 (IG1) started MindPower at T1 and ended their eight weeks course at T2. Intervention group 2 (IG2) waited for four months, started MindPower at T3, and ended their eight weeks course at T4.

There were two booster sessions in each intervention group. IG1 had their first booster at T3 and their second at T4, two and four months after the end of the course at T2. IG2 had their first booster at T5 and their second booster at T6, two and four months after the end of their course at T4. The final data collection in both groups was at T7, which took place nine months after the last booster in IG1 and five months after the last booster in IG2.

## Figure 1 Study design

### Sample size and power

We assumed that 20 out of 30 students from each school class would participate in the study. Then, if there was one teacher per class and 30 classes, there would be 30 clusters (classes) x 20 students = 600 students in each of the two intervention groups. Furthermore, we assumed that the difference between the two intervention groups would have an effect size of 0.30, the statistical significance would be 0.05, power 0.80, size of the cluster 20 and ICC 0.05. We would need 34 clusters with 680 students divided on the two intervention groups (17 clusters with 340 students in each group). However, there were actually 110 classes that wanted to participate in the intervention. The power calculation is based on the calculator at the University of Aberdeen Sample Size Calculator (Campbell, Thomson, Ramsay et al., 2004). The ICC is based on data from a study on children at the HEMIL-senteret at The University of Bergen (Roberts, Freeman, Samdal et al., 2009); a mean score based on an eight-item distress scale showed an ICC of 0.032. We assumed an ICC of 0.05, and our calculation of minimum sample size is thus rather conservative.

### Samples

The MindPower sample consisted of 15–16 years old students in 110 first year classes in all ten public high schools in the county of Østfold in South Eastern Norway. All the students were invited to complete an online questionnaire which contained measures of self-efficacy, self-esteem and quality of life.

The aim was to send invitations by e-mail to all 2,384 students in the 110 clusters. However, a large number of e-mail addresses turned out to be invalid. As a result, the questionnaire was administered to 1,673 students. Among these, 933 (55.8%) students completed the questionnaire at least once: 298 girls in IG1; 242 boys in IG1; 210 girls in IG2; 183 boys in IG2. Attrition increased over time. At the last measurement in total 86 participants responded (Table 1). Data collection took place in the schools, class by class. Students who were unable to respond because they had not received the electronic questionnaire or did not want to participate in the study were allowed to involve in other activities such as reading a book or doing homework.

Self-efficacy scores among the MindPower participants were compared to self-efficacy scores in a large general population sample (UngData) of first year public high school students nationally (n = 18,992 students; 9,423 boys and 9,414 girls) and in the same schools that were included in MindPower program (n = 2,391; 1,227 boys and 1,160 girls) with a response rate of 75%. UngData is a national school-based survey conducted each year in most of the municipalities throughout Norway. It contains questions about physical and mental health, parents, friends, school, community and leisure activities. Like MindPower, students complete the UngData questionnaire in the

Loading [MathJax]/jax/output/CommonHTML/jax.js

University is responsible for the national coordination while the regional Drug and Alcohol Competence Centers are responsible for collecting the data in schools.

## Randomization

The statistical computing platform R was utilized when the random sequence specifying allocation into the two intervention groups was generated. Based on the system clock at the time of allocation (1525868496), an initial random seed was chosen and a sequence of random numbers was generated using the `runif` function. The number of classes in each field ranged from 1 to 7, whilst each school participating in the study were offered a number of different fields of studies (FOS). Randomization was stratified by school ID and FOS. Half of the classes within each FOS at a given school were randomly assigned to each of the two intervention groups. Allocation of classes into the intervention groups was known by the school in advance.

## Measures

*Self-efficacy* refers to belief in being able to deal with difficult demands in life. We used a short version the Norwegian Version of the General Perceived Self-Efficacy Scale (Røysamb, Schwarzer & Jerusalem, 1998) originally developed by Jerusalem and Schwarzer (1992). The short version contains five items scored from 1 to 4, where 1 is “not at all true” and 4 is “exactly true”. The scale is used both in MindPower and in UngData. Previously, the short version has demonstrated high reliability and validity (Tambs & Røysamb, 2014). In MindPower, the short version had a Cronbach’s  $\alpha$  across measurement waves ranging from 0.82 to 0.92. In UngData, the  $\alpha$  value was 0.91.

*Self-esteem* refers to sense of being valuable and important (Blascovich & Tomaka, 1991). We used a short version of Rosenberg Self-esteem Scale (RSES) originally developed by Rosenberg (1965). The short version contains four items scored from 1 to 4, “very agree” to “not agree at all”, two positively and two negatively phrased. The scale has demonstrated high validity and reliability (Tambs & Røysamb, 2014). In MindPower, the short version had Cronbach’s  $\alpha$  across measurement waves ranging from 0.76 to 0.80.

*Quality of life* refers to sense of meaning and satisfaction with life. The Warwick Edinburg Mental Wellbeing Scale (WEMWBS) has fourteen items scored from 1 “not at all” to 5 “all the time”. The Norwegian version has shown high reliability and validity (Ringdal, Bradley-Eilertsen, Bjørnsen et. al., 2018; Smith, Alves, Knapstad et. al., 2017). In MindPower, WEMWBS had a Cronbach’s  $\alpha$  across measurement waves ranging from 0.92 to 0.96.

*Monitoring fidelity.* We asked all students the following question on the last assessment: “Do you receive tuition in MindPower now?” Response alternatives were: “Yes”, “No, MindPower is no longer tuitioned”, and “I don’t know”. Furthermore, the second question was to enter a number from 1 to 10 on the following question: “Now I have course day number XX”.

The questionnaire used in MindPower also included measures on symptoms of depression and anxiety. These measures are reported elsewhere (Sælid, Czajkowski, Aarø et al., 2021). Background variables were gender, age and the ten public high schools. Furthermore, we hypothesized that MindPower increases academic grades and reduces dropouts from school, however, these measures were beyond the scope of this study.

In MindPower, data collections among participants were carried out at all seven time points for self-efficacy, self-esteem and quality of life. From UngData only one measurement time point was included.

## Intervention and procedure

MindPower is a group-based cognitive behavioral intervention. The aim is to give adolescents tools to strengthen their positive mental health and wellbeing by training their capabilities to cope with the strains of daily life. The main topics are: how the brain develops; how feelings, thoughts and behavior are linked together; how relationships with others might help you gain better mental health; common thinking styles; dysfunctional thinking; and coping strategies.

MindPower is rooted scientifically and historically in the well-tested Coping With Depression Course (CWD) (van Zoonen et al., 2014; Dalgard, 2006, Clark et al., 1993). However, MindPower differs from CWD in several ways. CWD courses are usually delivered to participants who are at least mildly depressed or at risk of developing depression. CWD courses are usually delivered in a health care setting, by psychologists, nurses or other health professionals. MindPower is delivered universally to all students irrespective of their mental health challenges. MindPower is administered in a classroom setting during ordinary school hours, delivered by trained public high school teachers of varied professional background. CWD addresses symptoms of depression and anxiety. The MindPower course book does not utilize disorder related terms like “anxiety”, “depression”, “illness” and “disorder”. Instead, it is worded in a language reflecting “everyday challenges”, like “discouraged” and “distressed”. Most of the CWD-related courses, including MindPower, are organized in eight weekly sessions and two booster sessions. However, while in the CWD courses each session lasts for 2.5 hours, each MindPower session is adapted to the school schedule and lasts for 1.5 hours.

Prior to start of the study, approximately 170 public high school teachers attended a five-days course. The course was group-based and included theoretical lectures, guidelines and instructions, homework, role playing and intensive training in how to teach MindPower. The main modules contained information about how the adolescent brain works and develops, and overall models from cognitive behavior theory. The organization, Fagakademiet, planned the training and certified teachers to become MindPower life skills teachers.

At the last day of training, the teachers were explained about the study design and received the study coordinator’s phone number and e-mail address in case they had questions about the study. The teachers provided the information sheet and consent form to the students and posted the consents back to the study coordinator along with students contact information and e-mail addresses in an encrypted file. According to Norwegian law, among the 15 years olds, only students with a valid consent from both of their parents were included in the study. Participants who were 16 years of age responded directly in the online questionnaire to the invitation to consent. In addition to the training course, the teachers were encouraged to gather regularly in meetings to share thoughts and feelings and learn from each other’s experiences of teaching MindPower.

Before start of each data collection, the teachers received a short list of key points in order to help them to remember each task. The list contained information about the next data collection and some questions they might ask themselves: “According to the study design, have I calculated time in class for responding on the questionnaire?”, “Did I remember to administer the questionnaire at the start of the session and not at the end of the class?”, “If there are any changes, have I e-mailed the study coordinator?” and “Did I remember to give myself an applause for all the work I have done with assisting the students in responding to the online questionnaire?”

## Statistical analyses

The data was analyzed with linear mixed models and paired samples t-tests, using R 4.0.2 (R Core Team, 2020) and SPSS 26 (SPSS, I., 2011).

Linear mixed models are statistical models that contain both fixed and random effects and can handle non-balanced data with missing entries and repeated observations (Jackson, 2010; Kwok, Underhill, Berry et al., 2008).

In total, eight linear mixed models were fitted for each of the three outcome variables, Self-efficacy, Self-esteem and Quality of life. Model 1, serving as a baseline reference model, included only fixed and random intercepts. This allowed the dependent variable to vary across participants, but not across time. In subsequent models, the following components were added; a fixed effect of sex (Male, model 2), a fixed effect of time (model 3), a random effect of time (model 4), a random effect of school (model 5), and an indicator variable for intervention group membership (model 6).

In the mixed models, the effect of the MindPower course is tested by the coefficients “Course”, and “Time\*Course”, the interaction between time and the completion of MindPower. A change in scores over the MindPower course, beyond what can be accounted for by the linear effect of time is in model 7 captured by the estimate of the regression coefficient for an indicator variable marking the completion of the course. The variable “course” is coded “0” until T2 and T4 for the IG1 and IG2 respectively, after which point it is coded “1”.

In model 8, a change in the fixed effect of time on scores after the course is completed, is captured by the interaction between time and Course. A consequence of this parametrization is that once the interaction coefficient is included, “Time” quantifies the change “pre course completion”.

## Results

Table 1 shows mean scores and standard deviations on Self-efficacy, Self-esteem and Quality of Life across the seven MindPower measurement points in Intervention group1 (IG1) and Intervention group 2 (IG2). The table also lists Self-efficacy scores in UngData: UngData schools that participated in MindPower; UngData schools that did not participate in MindPower.

(Table 1 Self-efficacy, Self-esteem, Quality of Life: Mean, SD, n)

Table 1 Self-efficacy, Self-esteem and Quality of Life: Mean, standard deviation (SD) and n

Table 1 shows that all the outcome measures were stable over time. Although there was a significant decline in Quality of Life over the eight weeks course (IG1), there were no significant change in mean levels across time for any of the three outcome variables.

Figure 2 shows scores on Self-efficacy (orange), Quality of Life (blue) and Self-esteem (green) across seven measurement waves, standardized with respect to measures at baseline. Results from IG1 are given on the left, and IG2 on the right. The colored region indicates the eight weeks during which the group participated in the active part (eight weeks course) of the MindPower program. The timeline in Fig. 2 shows that baseline starts at zero weeks (T1) and the data collection ends after 60 weeks (T7). Although the lines show changes over time, there were no significant changes in mean levels across time for any of the three outcome variables.

(Fig. 2 Scores on Self-efficacy, Quality of Life and Self-esteem across seven measurement waves)

Loading [MathJax]/jax/output/CommonHTML/jax.js

## Figure 2 Scores on Self-efficacy, Quality of Life and Self-esteem across seven measurement waves

(Table 2 Estimates of the fixed effects coefficients from linear mixed models of self-efficacy scores across the seven measurement points)

Table 2 Estimates of the fixed effects coefficients from linear mixed models of self-efficacy scores across the seven measurement points

Estimates from linear mixed models of self-efficacy is given in table 2. Of the eight models fitted, the lowest AIC values was found for model 4, which included the fixed coefficients for Male and Time, as well as a random effect of Time. In this model, boys were found to have significantly higher scores on self-efficacy ( $p < 0.001$ ), and while we observed no significant fixed effect of time, there was a significant random effect of time, which was also supported by comparison with model 3 ( $\chi^2 = 15.914, df = 2, p < 0.001$ )

Estimates from the same models fitted to the outcome variable self-esteem is given in table 3. Here, the best fitting model (model 5) contained fixed effects of Time and Male, as well as random intercepts over Students and School, and a random effect of Time. In model 5, boys were found to have significantly lower self-esteem ( $p < 0.001$ ), but while a coefficient for time was included in the model, it was not significant. Lastly, estimates for Quality of Life (QOL) is given in table 4. The best fitting model (7) we found a significant positive effect of being male, a significant decline in QOL across the course weeks ( $p < 0.05$ ), and a significant difference between the two intervention groups at time zero. Also here, we observed effects on the random intercept of both Students and School, and a random effect of time.

In summary, no significant change in mean levels was observed across time for any of the three outcome variables, nor did we observe an interaction between the MindPower course and time which would have indicated a change in the developmental trajectory after completing the course.

(Table 3 Estimates of the fixed effects coefficients from linear mixed models of Self-esteem scores across the seven measurement)

Table 3 Estimates of the fixed effects coefficients from linear mixed models of Self-esteem scores across the seven measurement

(Table 4 Estimates of the fixed effects coefficients from linear mixed models of quality of life scores across the seven measurement)

Table 4 Estimates of the fixed effects coefficients from linear mixed models of quality of life scores across the seven measurement

### Drop-out analysis

Participants who dropped out or did not respond on the questionnaire, did not differ from complete responders on their last measurement. This suggests that there was no systematic dropout associated with the outcome variables of interest. Dropout was modest over the first three time points but increased subsequently. At T7, only 86 responded to the final questionnaire. Only 11 out of 1,673 students responded to the questionnaires seven times; 95 out of 1,673 students responded six times.



## Monitoring the implementation

To test fidelity, students answered two questions on their last assessment. Approximately half of the students (n = 194, 49%) in IG1 reported to not have tuition in MindPower at their tenth (final) course day. However, the students also answered to be having “course day 10” (n = 108, 89%) at this assessment.

Nearly half of the students (n = 120, 48%) in IG2 did not have their last and tenth course day. However, the data are not consistent, because 86.3% of the students (n = 82) responded to have had course day number 10 at their final course day. The lack of tuition in IG1 and IG2 at the final course day, however, deviates from the study design and course book manual.

## Discussion

MindPower is an adjusted and upscaled version of the CWD/CWS interventions. CWD/CWS courses have traditionally been delivered within a health service context by health personnel targeting high risk groups to prevent depression. In contrast, MindPower is delivered class wise, in the school, during ordinary school hours, by trained school teachers, universally, i.e. independent of the students’ risk of mental distress or disorder, to strengthen positive mental health and quality of life.

In this study we tested effects on self-efficacy, self-esteem and quality of life by using a two-groups cluster randomized delayed intervention design with linear mixed models statistical analyses. The intervention was offered to all students at all schools in a county in South-East Norway. Level of self-efficacy among the participants was compared to a large scale population survey (UngData), nationwide and from the same schools as were the MindPower-intervention was implemented.

The main finding was that there was no significant effect of the MindPower program delivered universally by trained teachers in a class room situation. Neither did we find effects within any of the two intervention groups, nor did the two groups change consistently as expected from the delayed intervention design if an intervention effect had occurred. These were consistent findings across outcome measures, except a small significant decrease in quality of life.

Another finding was that the level of self-efficacy in both our intervention groups was high and equal both to that of the general population in the same age groups, nationally and in the same schools as were the intervention was implemented. This indicates that the students that participated in the intervention were characterized by high levels of self-efficacy, self-esteem, and quality of life from the beginning to the end of the study. A further increase in these measures through a universal initiative would therefore be a challenging task (ceiling effect).

The high levels of self-efficacy, self-esteem and quality of life uncovered in UngData in combination with very high and steadily increasing levels of self-reported mental distress as well as symptoms of depressive feelings and anxiousness found in Scandinavian studies may seem contradicting. Particularly if the increases in these self-reported negative feelings are interpreted as indications of a deterioration of the mental health among young people.

However, if the increasing reports of negative feelings are regarded as a sign of de-stigmatization and increased openness, such an effect is compatible with the high levels of self-efficacy, self-esteem and quality of life observed

in this study. That would be highly consistent with the very high levels of satisfaction with life reported among close to 90 percent of, at least Norwegian, young people (UngData).

In that case, a displacement of the meanings of values in scales intended to measure mental distress and symptoms of common mental disorders is likely to have occurred. Because the context of openness has changed, items like the Hopkins Symptom Checklist (SCL-8) for symptoms of depression and anxiety, does not mean the same today as it meant a decade or two ago.

This may explain why we with such a carefully and comprehensively prepared intervention, advanced study design and statistical methods, did not find any effects of MindPower on self-efficacy, self-esteem or quality of life in this study. We may already have hit the ceiling before we started.

Another explanation could have been selective drop out as the drop out from this study was progressive and comprehensive. However, the complete responders on the last assessment were not significantly different from those who dropped out. A clinically meaningful effect size of the intervention should therefore have been revealed in spite of the high level of drop out.

A third possible explanation may be lack of fidelity. Our monitoring of fidelity indicates that only half of the students in both intervention groups reported having been tuitioned at their tenth and last Mindpower session. This deviates from the instructions in the course leader manual and the study design. Completion of all the ten sessions are regarded as imperative. Less than complete implementation may attenuate effects of the intervention seriously.

What can we learn from this study?

The planning of this study was careful and comprehensive. The study was launched and conducted in a very positive contextual atmosphere. The government had pledged a change in the curriculum. A new national curriculum on public health and life skills, including mental health, was in the process of being launched in all schools in the country.

Parents and students themselves, teachers, researchers, psychologist and local, regional, and national politicians had for a long time argued for equalization of mental and physical health in the school.

The top administrator of education in the county welcomed the study and was effective in supporting communication and collaboration with the school administration and the schools. There were several motivation meetings with the teachers before and during the project period indicated high motivation to participate among students, school leaders and teachers.

There were comprehensive preparations before the implementation. The organization "Fagakademiet" arranged intensive training of all the involved teachers, e.g theoretical lectures, practical guidelines and instructions, role playing, and homework. According to an internal evaluation from Fagakademiet, teachers reported very high satisfaction regarding the training.

Teachers received a short list of key points prior each data collection, to ease remembering of what to do. All 170 teachers and the ten principals received the study-coordinator's phone number and e-mail address in case they had questions about the study. There was updates and information by e-mail and by phone, and meetings at each

With such a positive context and such a thorough implementation and enthusiastic follow up, how could it go so wrong?

A qualitative study connected to this trial has addressed this question (Cheng, Ihlebæk & Sælid, 2021). Furthermore, we have collected information through a number of informal sources. This has provided us with a number of learning points for others who intend to launch large-scaled universal mental health initiatives among public high school students in a similar context.

The core of a successful evaluation is to know to which extent essential factors have been implemented. In this study, however, a full-scale fidelity on teachers and school administrations study was not possible.

It is crucial that teachers follow the instructions in the course leader manual on how to tuition the intervention program. Also, follow the instructions in the study design regarding which course day are to be tuitioned.

In implementing a study like this, the first point is to monitor the extent to which the teachers start the intervention at the correct time point. In this study, half of the MindPower sample was instructed to start four months prior to the second group. However, without a full-scale implementation study of the teachers, optimal assurance that the expected starting point is realized exactly, is not possible to achieve.

Another implementation challenge is that at some schools not all teachers may have sufficient support from their school administration. Teachers may also lack assistance from their colleagues in adapting the MindPower program into the curriculum. As a consequence, some classes may never have started the intervention, while others may have received a shortened program.

Here, one solution may be, as some schools did, to establish a forum at the school that could include mental health personnel. Their intention was to share positive and negative experiences and to seek advice and support. Such a team might give the teachers the support they need to find time to MindPower in the curriculum. Additionally, this may contribute to feelings of competence in teaching life-skills.

Still another challenge is to what extent the teachers experience to have a clear job description. Some teachers might feel that teaching life-skills is providing therapy, a task which is not a school teacher's responsibility. This may result in lack of motivation and avoidance of teaching mental health-related programs.

One solution to this may be that the program leadership and principals together give counseling on role consciousness and give clear directions in these concrete cases. Life-skills training is not therapy and can be taught in terms of "how life is", implying that challenges in life are normal and that no one is alone in experiencing such feelings.

In our study, some teachers were instructed to participate in the preparatory course without being asked about their motivation. Some school administrators did not allow teachers to choose not to teach MindPower. Especially some teachers who taught practical vocational subjects such as electronics, plumbing, and carpentry were not comfortable carrying out teaching activities implying handling of issues related to emotions.

To address this, one solution is to only admit motivated teachers to teach life-skills programs. If the teachers are not motivated, they will not do a good job. It is also necessary to respect that some teachers feel uncomfortable in teaching life-skills. Consequently, life-skills teaching or training should not be mandatory.

Furthermore, a challenge may be to what extent the MindPower program has been sufficiently well designed and tailored for classroom-based educational purposes. In our study, some teachers experienced difficulties running the MindPower program in the classroom.

A solution to this might be to involve teachers and students in the planning of the implementation. Ideas and opinions of teachers and students are crucial in order to succeed. In this way the program may be better tailored for the particular school-setting. In our study, students and teachers were asked to give feedback prior to the implementation. Based on these feedbacks, the designer of MindPower changed certain points of the program to ease the implementation process, e.g. 90 minutes sessions rather than two and a half hours as in traditional Coping With Depression/Coping with Strain-courses (Cuipers et al., 2009; Sælid et al., 2016).

We experienced that teachers had deviated from the implementation plan on several points. For instance, some teachers did not complete all the ten MindPower sessions. Other teachers shortened the 90 minutes sessions. This emphasizes the necessity do involve all teachers on beforehand and make sure that the school administration expects the program to be implemented as planned and do not accept deviations from the plan. A more flexible approach, however, would be to work on the program to gain an acceptable format for all participants.

The confounding variables and the implementation factors in large studies are not easy to control. The current study has an RCT-design. However, researchers should keep in mind that the findings from smaller, well implemented studies with sufficient statistical power, may be more valid than a large scaled natural study were several confounding factors may be out of control (Dawson, Yeomans & Brown, 2018). In our case, a smaller study might have done it easier to have better control of some of the implementation factors described above.

One of the greater learning points from this study concerns technicalities in collecting data from public high school students. The data collection department at NIPH, a governmental research institution, refused to use the student e-mail addresses provided by the students' teachers. Instead, they insisted on collecting and using the students' officially registered e-mail addresses.

However, this register contained to a large extent the students' parents' e-mail addresses and were not valid contact points to the students. Furthermore, some e-mail addresses received from the teachers were also not valid. Consequently, we ran into difficulties in administering the questionnaires to some students' valid e-mail addresses. These addresses a common problem. Young people may frequently use unofficial e-mail addresses and change e-mail addresses frequently. Washing addresses against official registers may do more harm than good. In addition, at each data collection time point, some students had difficulties in both retrieving and opening the questionnaires.

If a technical data collection partner is involved, which administrates the questionnaire, a contract is advised where the data collection partner fully accepts the plan for what to do if there are technical failures. Such a plan could include setting up additional links with the questionnaires, which can be sent to the participants and teachers if technical problems occur.

Pretesting of the data collection procedures is warranted. This should include making sure that you have the best possible records of e-mail addresses and telephone numbers of all study participants. Finally, if possible, change technical partner if the problems continue, e.g. the questionnaires are not possible to be opened for all students at any time point.

Altogether, these challenges resulted in high drop-out rates and potential sample bias. In our case, fortunately, the complete responders on the last assessment were not significantly different from those who dropped out and such possible bias could be ignored.

### Strengths and limitations

To our knowledge, this is the first study to assess effects of a Coping With Depression/Coping With Strain Course (CWD/CWS) based intervention (Cuijpers et al, 2009; Sælid, 2016) which universally addresses positive mental health and quality of life among public high school students.

Another strength of this study is the adapted randomized control trial design (RCT) with mixed model statistical analyses tested for cluster effects. With the delayed intervention design one group functioned as a control group until the intervention was implemented for this group too. In this way all participants who wanted, could benefit from the intervention, only with a period of delay for some. If the two intervention groups had shown similar positive patterns in the development of the outcome measures, this would have been a strong indication of a positive program effect.

Yet another strength is the inclusion of two comparison samples from the UngData. This made it possible to establish how the outcome measures in the experimental samples related to the relevant general population.

Furthermore, compared to similar studies of CWD/CWS-based interventions, the sample size of this study is relatively large with more than 1,600 in the MindPower sample and more than 18,000 in the national UngData.

The outcome measures used in this study are high quality instruments commonly utilized in studies on adolescents. When measures are administered repeatedly, changes over time can be examined. However, the findings might not have generalizability if the school-sample e.g. background variables, organization and school structures, teacher training, differs in great extend from the current sample.

Limitation are that a full-scale fidelity study could not be obtained. This was to some degree compensated by collecting information from the participants. Unfortunately, there are indications that this information was not 100 percent trustworthy. Therefore, we do not know for sure how exactly the MindPower program manual was followed. On the other hand, we have no indication that this was a widespread problem.

Technical problems with e-mail addresses and with opening of the questionnaires lead to a high level of attrition. Fortunately, the complete responders on the last assessment were not significantly different from those who dropped out. Hence possible sample bias due to selective attrition at this point may be ignored.

## Conclusion

No effect of MindPower delivered universally to public high school students in the classroom was found. One possible explanation is that initial levels of self-efficacy, self-esteem, and quality of life in the intervention groups already were high and almost identical to that of the comparable group in the general population (ceiling effect). Hence, there was little to gain. This large universal school trial suffered from a high level of attrition, and full mapping of intervention fidelity was not possible. Caution in interpreting the data is therefore warranted. This study yielded a long list of practical learning points for researchers who wish to evaluate effects of universal

mental health promoting interventions implemented class wise in public high schools where the mental health intervention is not yet an integrated part of the ordinary curriculum.

## Declarations

## Acknowledgement

Special thanks to Ingri Myklestad at The Norwegian Institute of Public Health for valuable reflections and support, and to Anders Bakken at The Norwegian Social Research who made their Ungdata materials available. The Norwegian foundation Stiftelsen Dam for funding the research project, and Charlotte Elvedal at the The Norwegian Council for Mental Health for assistance through the funding process. Additionally, Erik Bråthen and all the students and teachers at the public high schools, school administrations, municipalities and county for their enormous effort and positive attitudes.

## Authors` contributions

Leif Edvard Aarø and Gry Anette Sælid contributed to the study conception and design. Material preparation for the data collection was performed by Gry Anette Sælid. Nikolai Czajkowski generated the random allocation sequence. Gry Anette Sælid contacted the teachers about the enrollment and assignment of participants to interventions. Nikolai Czajkowski and Gry Anette Sælid analyzed the data. All authors contributed to interpreting the data. The first draft of the manuscript was written by Gry Anette Sælid. All authors commented several times on previous versions of the manuscript, and all authors read and approved the final manuscript.

## Founding

The role of the Norwegian foundation Stiftelsen Dam was to fund the postdoctoral salary of the first author and the data collection. Grant number 2017/FO143251.

## Conflicts of Interest/Competing Interests

The authors declare that they have no competing interest or conflict of interests.

## Availability of data and material

The datasets analyzed during this study is not publicly available because of ethical restrictions. Contact the corresponding author.

## Code availability

Not applicable.

We confirm that the study was approved by the Regional Committee for Medical and Health Research Ethics (REC) for South-East Norway, registration number 2017/1908. Furthermore, we confirm that the study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki.

MindPower is not therapy. The teachers were instructed to contact health professionals at the school if they suspected mental health problems or risk of suicide among any of the students. For students, speaking in the class was voluntary. The teachers spoke to the whole class and did not ask individual students about their experiences.

## Consent to participate

Only students (15-years of age) with a valid and written informed consent from both of their parents or legal guardians were included in the study. We confirm that informed written consent was obtained from all parents/legal guardians. According to Norwegian law and REC, participants who were 16 years of age and older can respond directly on the consent in the online questionnaire without consent from parents or legal guardians. We confirm that informed written consent was obtained from all participants in the study.

## Consent for publication

Not applicable.

## References

1. Atkins, M.S., Hoagwood, K.E., Kutash, K., & Seidman, E. (2010). Toward the integration of education and mental health in schools. *Administration and Policy in Mental Health and Mental Health Services Research*, 37(1):40–47.
2. Bakken, A. (2019). *Ungdata 2019. Nasjonale resultater 2019, NOVA Rapport 9/19*. Oslo: NOVA, OsloMet. Retrieved April 14 <https://oda.oslomet.no/oda-xmlui/bitstream/handle/20.500.12199/2252/Ungdata-2019-Nettversjon.pdf?sequence=3&isAllowed=y>
3. Bakken, A. (2017). *Ungdata 2017. Nasjonale resultater*. Oslo: NOVA, OsloMet. Retrieved May 15 <https://oda.oslomet.no/oda-xmlui/bitstream/handle/20.500.12199/5117/Opprettet-Ungdata-rapport-2017-4-august-2017-web-utg-med-omslag.pdf?sequence=1&isAllowed=y>
4. Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 6. –160). San Diego, CA: Academic Press.
5. Brugha, T. S., Morrell, C. J., Slade, P., & Walters, S. J. (2011). Universal prevention of depression in women postnatally: cluster randomized trial evidence in primary care. *Psychological medicine*, 41(4), 739–748.
6. Caldwell, D. M., Davies, S. R., Hetrick, S. E., Palmer, J. C., Caro, P., López-López, J. A., Gunnell, D., Kidger, J., Thomas, J., French, C., Stockings, E., Campbell, R., & Welton, N. J., (2019). School-based interventions to prevent anxiety and depression in children and young people: a systematic review and network meta-analysis. *Lancet Psychiatry*, 6(12), 1011–20.

9. Campbell, M. K., Thomson, S., Ramsay, C. R., MacLennan, G. S., & Grimshaw, J. M. (2004). Sample size calculator for cluster randomized trials. *Computers in biology and medicine*, *34*(2), 113–125.
10. Cheng, S., Ihlebæk, C., & Sælid, G. A. (2021). Implementation of a life-skill and health promoting program in High school - A qualitative study of teacher's experiences. [Unpublished Manuscript].
11. Chodkiewicz, A. R., & Boyle, C. (2017). Positive psychology school-based interventions: A reflection on current success and future directions. *Review of Education*, *5*(1), 60–86.
12. Clarke, G. N., Hawkins, W., Murphy, M., & Sheeber, L. (1993). School-based primary prevention of depressive symptomatology in adolescents: Findings from two studies. *Journal of Adolescent Research*, *8*(2), 183–204.
13. Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. *Journal of affective disorders*, *159*, 118–126.
14. Cuijpers, P., Muñoz, R. F., Clarke, G. N., & Lewinsohn, P. M. (2009). Psychoeducational treatment and prevention of depression: the “Coping with Depression” course thirty years later. *Clinical psychology review*, *29*(5), 449–458.
15. Dalgard, O. S. (2006). A randomized controlled trial of a psychoeducational group program for unipolar depression in adults in Norway (NCT00319540). *Clinical Practice and Epidemiology in Mental Health*, *2*(1), 1–8.
16. Dawson, A., Yeomans, E., & Brown, E. R. (2018). Methodological challenges in education RCTs: reflections from England's Education Endowment Foundation. *Educational Research*, *60*(3), 292–310.
17. HUNT. (2019). *Folkehelseutfordringer i Trøndelag. Folkehelsepolitisk rapport med helsestatistikk fra HUNT inkludert tall fra HUNT4 (2017-19)*. Trondheim: Hunt Research center.
18. Jackson, D. L. (2010). Reporting results of latent growth modeling and multilevel modeling analyses: Some recommendations for rehabilitation psychology. *Rehabilitation Psychology*, *55*(3), 272.
19. Jenkins, R., Meltzer, H., Jones, P. B., Brugha, T., Bebbington, P., Farrell, M., Crepaz-Kay, D., & Knapp, M. (2008). *Foresight mental capital and wellbeing project. Mental health: Future challenges*. London: The Government Office for Science.
20. Jerusalem, M., & Schwarzer, R. (1992). Self-efficacy as a resource factor in stress appraisal processes. In R. Schwarzer (Ed.), *Self-efficacy: thought control of action* (pp.195–213). Hemisphere, Washington, DC.
21. Keles, S., & Idsoe, T. (2021). Six-and Twelve-Month Follow-up Results of a Cluster Randomized Controlled Trial of a CBT-Based Group Course. *Prevention Science*, *22*(4), 409–418.
22. Knudsen, A. K., Tollånes, M. C., Haaland, Ø. A., Kinge, J. M., Skirbekk, V., & Vollset, S. E. (2017). Sykdomsbyrde i Norge 2015. Resultater fra Global Burden of Diseases, Injuries, and Risk Factors Study 2015 (GBD 2015). [Disease Burden in Norway 2015. Results from the Global Burden of Diseases, Injuries, and Risk Factors Study 2015 (GBD 2015)] Rapport 2017. Bergen/Oslo: Folkehelseinstituttet.
23. Kwok, O. M., Underhill, A. T., Berry, J. W., Luo, W., Elliott, T. R., & Yoon, M. (2008). Analyzing longitudinal data with multilevel models: an example with individuals living with lower extremity intra-articular fractures. *Rehabilitation psychology*, *53*(3), 370.
24. Lahtinen, E., Joubert, N., Raeburn, J., & Jenkins, R. (2005). Strategies for promoting the mental health of populations. *Promoting Mental Health*.
25. Lewinsohn, P. M., Antonuccio, D. O., Steinmetz, J. L. & Teri, L. (1984). *The coping with depression course: A*



26. Lewinsohn, P. M., Weinstein, M. & Alper, T. (1970). A behavioural approach to the group treatment of depressed persons: A methodological contribution. *Journal of Clinical Psychology*, 26(4), 525–532.
27. Mackenbach, J. P., Lingsma, H. F., van Ravesteyn, N. T., & Kamphuis, C. B. (2013). The population and high-risk approaches to prevention: quantitative estimates of their contribution to population health in the Netherlands, 1970–2010. *The European Journal of Public Health*, 23(6), 909–915.
28. Norwegian Directorate of Health. (2015). *Samfunnskostnader ved sykdom og ulykker 2013*. Oslo: Helsedirektoratet.
29. O'Reilly, M., Svirydzenka, N., Adams, S., & Dogra, N. (2018). Review of mental health promotion interventions in schools. *Social psychiatry and psychiatric epidemiology*, 53(7), 647–662.
30. R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
31. Ringdal, R., Bradley Eilertsen, M. E., Bjørnsen, H. N., Espnes, G. A., & Moksnes, U. K. (2018). Validation of two versions of the Warwick-Edinburgh mental well-being scale among Norwegian adolescents. *Scandinavian journal of public health*, 46(7), 718–725.
32. Roberts, C., Freeman, J., Samdal, O., Schnohr, C. W., De Looze, M. E., Gabhainn, S. N., Iannotti, R., & Rasmussen, M. The Health Behaviour in School-aged Children (HBSC) study: methodological developments and current tensions. *International journal of public health*, 54(2), 140–150.
33. Rose, G. (1992). *The strategy of preventive medicine* (Oxford Medical Publications). Oxford: Oxford University Press.
34. Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton university Press.
35. Røysamb, E., Schwarzer, R., & Jerusalem, M. (1998). *The Norwegian version of the General Perceived Self-efficacy Scale*. Retrieved January 30 <https://userpage.fu-berlin.de/~health/norway.htm>
36. Saelid, G. A., Czajkowski, N. O., Holte, A., Tambs, K., & Aarø, L. E. (2016a). Coping With Strain (CWS) course—its effects on depressive symptoms: A four-year longitudinal randomized controlled trial. *Scandinavian journal of psychology*, 57(4), 321–327.
37. Scottish Government. (2017). *Mental health strategy 2017–2027*. Retrieved 21 June 20 <http://www.gov.scot/Publications/2017/03/1750/5>
38. Smith, O. R., Alves, D. E., Knapstad, M., Haug, E., & Aarø, L. E. (2017). Measuring mental well-being in Norway: validation of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS). *Bmc Psychiatry*, 17(1), 182.
39. SPSS, I. (2011). *IBM SPSS Statistics for Windows, version 20.0*. New York: IBM Corp.
40. Sturgeon, S. (2006). Promoting mental health as an essential aspect of health promotion. *Health promotion international*, 21(suppl\_1), 36–41.
41. Saelid, G. A., Czajkowski, N. O., Aarø, L. E., Idsøe, T., Hellesester, M. D., & Holte, A. (2021). Effects of a school-based intervention on levels of anxiety and depression: a cluster-randomized controlled trial of the MindPower program in ten high schools in Norway. *BMC psychology*. [Manuscript submitted for publication].
42. Saelid, G. A., Czajkowski, N. O., Holte, A., Tambs, K., & Aarø, L. E. (2016b). Positive mental health effects of the Coping With Strain (CWS) course on employees: a four-year longitudinal randomized controlled trial. *International Journal of Mental Health Promotion*, 18(3), 158–175.
43. Tambs, K., & Røysamb, E. (2014). Selection of questions to short-form versions of original psychometric instruments in MoBa. *Norsk Epidemiologi*, 24(1): 195–201.

44. Taylor, R. D., Oberle, E., Durlak, J.A., & Weissberg, R.P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child development*, 88(4):1156–7111.
45. van Zoonen, K., Buntrock, C., Ebert, D. D., Smit, F., Reynolds III, C. F., Beekman, A. T., & Cuijpers, P. (2014). Preventing the onset of major depressive disorder: a meta-analytic review of psychological interventions. *International journal of epidemiology*, 43(2), 318–329.
46. Werner-Seidler, A., Perry, Y., Calear, A.L., Newby, J.M., & Christensen, H. (2017). School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clinical psychology review*, 51, 30–47.
47. Welsh Government (2017) *Written statement—providing for the emotional and mental health needs of young people in school*. Retrieved May 10  
<http://gov.wales/about/cabinet/cabinetstatements/2017/mentalhealthneeds/?lang=en>

## Tables

Table 1 Self-efficacy, Self-esteem, Quality of Life: Mean, SD, n

UngData schools that participated in MindPower

*Self-efficacy*

Mean = 3.03

N = 2176

SD = 0.641

UngData schools that did not participate in MindPower

Mean = 3.04

N = 14627

SD = 0.617

	<i>Self-efficacy</i>	T1	T2	T3	T4	T5	T6	T7
MindPower: IG1	Mean	3.05	3.03	3.03	3.04	3.18	3.02	2.98
	N	549	567	460	389	181	189	53
	SD	0.53	0.55	0.61	0.64	0.57	0.64	0.57
MindPower: IG2	Mean	3.00	2.99	3.02	2.97	3.01	3.08	3.03
	N	400	389	544	403	331	250	33
	SD	0.55	0.57	0.54	0.62	0.62	0.64	0.56
	<i>Self-esteem</i>	T1	T2	T3	T4	T5	T6	T7
MindPower:IG1	Mean	2.03	2.07	2.08	2.03	2.03	2.11	2.00
	N	549	568	459	390	180	188	53
	SD	0.65	0.66	0.68	0.69	0.65	0.72	0.65
MindPower: IG2	Mean	2.14	2.06	2.10	2.10	2.05	2.05	2.31
	N	400	390	547	403	331	249	33
	SD	0.70	0.69	0.66	0.64	0.66	0.67	0.75
	<i>Quality of life</i>	T1	T2	T3	T4	T5	T6	T7
MindPower:IG1	Mean	3.73	3.62	3.59	3.60	3.59	3.52	3.48
	N	546	570	458	388	179	189	53
	SD	0.67	0.74	0.83	0.83	0.87	0.85	0.77
MindPower: IG2	Mean	3.64	3.60	3.58	3.51	3.56	3.66	3.40
	N	401	388	544	400	332	247	33

Loading [MathJax]/jax/output/CommonHTML/jax.js

SD 0.71 0.76 0.78 0.83 0.85 0.82 0.69

Table 2 Estimates of the fixed effects coefficients from linear mixed models of self-efficacy scores across the seven measurement points

<b>Self-efficacy</b>								
<i>Predictors</i>	<i>Model 1</i>	<i>Model 2</i>	<i>Model 3</i>	<i>Model 4</i>	<i>Model 5</i>	<i>Model 6</i>	<i>Model 7</i>	<i>Model 8</i>
(Intercept)	3.02 ***	2.95 ***	2.95 ***	2.95 ***	2.95 ***	2.91 ***	2.97 ***	2.97 ***
Male		0.16 ***	0.16 ***	0.15 ***	0.15 ***	0.15 ***	0.15 ***	0.15 ***
Time			0.00	0.00	0.00	-0.00	0.00	0.00
Group						-0.03	-0.03	-0.03
Course							-0.01	-0.01
Time x Course								0.00
<b>Random Effects</b>								
$\sigma^2$	0.20	0.20	0.20	0.19	0.19	0.19	0.19	0.19
Pupil Intercept	0.15	0.14	0.14	0.14	0.14	0.14	0.14	0.14
School Intercept					0.00	0.15	0.00	0.00
Time				0.00	0.00	0.00	0.00	0.00
$\rho$ Pupil intercept,slope				0.61	0.60	0.60	0.61	0.61
Observations	4602	4602	4602	4602	4602	4602	4602	4602
AIC	7312.76	7272.40	7273.95	7262.04	7262.57	7286.23	7264.47	7266.49
log-Likelihood	-3653.38	-3632.20	-3631.98	-3624.02	-3623.28	-3634.11	-3622.24	-3622.24

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

Table 3 Estimates of the fixed effects coefficients from linear mixed models of Self-esteem scores across the seven measurement

### Self-esteem

Predictors	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
(Intercept)	2.08 ***	2.22 ***	2.22 ***	2.22 ***	2.23 ***	2.22 ***	2.22 ***	2.22 ***
Male		-0.30 ***	-0.30 ***	-0.30 ***	-0.29 ***	-0.29 ***	-0.29 ***	-0.29 ***
Time			-0.00	0.00	0.00	0.00	0.00	0.00
Group						0.03	0.03	0.03
Course							-0.00	-0.00
Time x Course								0.00
<b>Random Effects</b>								
$\sigma^2$	0.14	0.14	0.14	0.12	0.12	0.12	0.12	0.12
Pupil Intercept	0.31	0.29	0.29	0.29	0.29	0.29	0.29	0.29
School Intercept					0.01	0.01	0.01	0.01
Time				0.00	0.00	0.00	0.00	0.00
$\rho_{01}$				-0.01	-0.01	-0.00	-0.00	-0.00
Observations	4610	4610	4610	4610	4610	4610	4610	4610
AIC	7053.30	6956.00	6957.99	6895.09	6891.83	6892.66	6894.61	6896.61
log-Likelihood	-3523.65	-3474.00	-3474.00	-3440.54	-3437.91	-3437.33	-3437.31	-3437.31

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

Table 4 Estimates of the fixed effects coefficients from linear mixed models of quality of life scores across the seven measurement

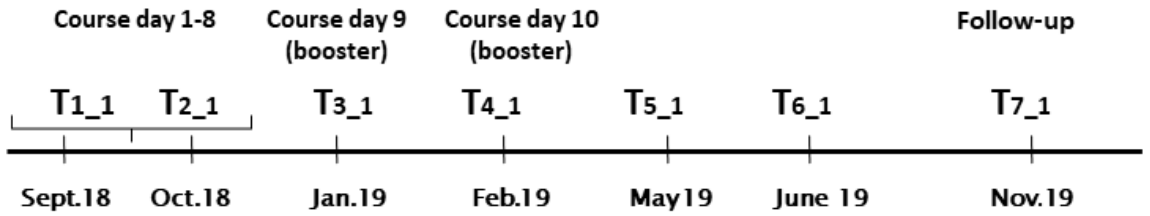
## Quality of life

Predictors	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
(Intercept)	3.60 ***	3.47 ***	3.47 ***	3.47 ***	3.45 ***	3.50 ***	3.54 ***	3.52 ***
Male		0.27 ***	0.27 ***	0.27 ***	0.27 ***	0.27 ***	0.27 ***	0.27 ***
Time			-0.00 ***	-0.00 ***	-0.00 ***	-0.00 ***	-0.00	-0.00
Group						-0.09 **	-0.10 **	-0.10 **
Course							-0.06 *	-0.04
Time × Course								0.00
<b>Random Effects</b>								
$\sigma^2$	0.25	0.25	0.25	0.24	0.24	0.23	0.23	0.23
Pupil Intercept	0.37	0.35	0.35	0.35	0.35	0.35	0.35	0.35
School Intercept					0.01	0.01	0.00	0.01
Time				0.00	0.00	0.00	0.00	0.00
$\rho$ Pupil intercept,slope				0.51	0.50	0.51	0.51	0.51
Observations	4411	4411	4411	4411	4411	4411	4411	4411
AIC	8839.35	8779.08	8766.45	8729.21	8728.67	8723.82	8721.27	8722.12
log- Likelihood	-4416.68	-4385.54	-4378.23	-4357.60	-4356.33	-4352.91	-4350.63	-4350.06

\*  $p < 0.05$  \*\*  $p < 0.01$  \*\*\*  $p < 0.001$

## Figures

# Intervention group 1



# Intervention group 2

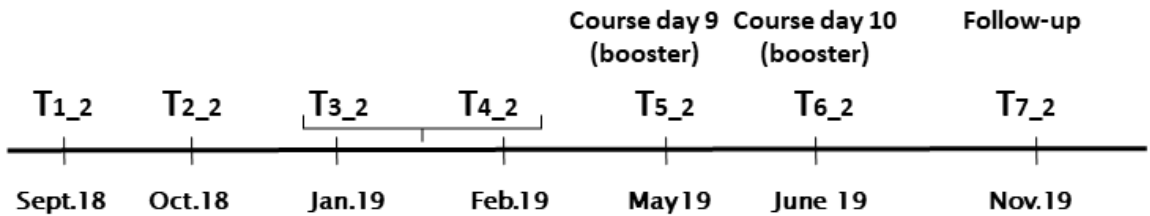


Figure 1

## Study design

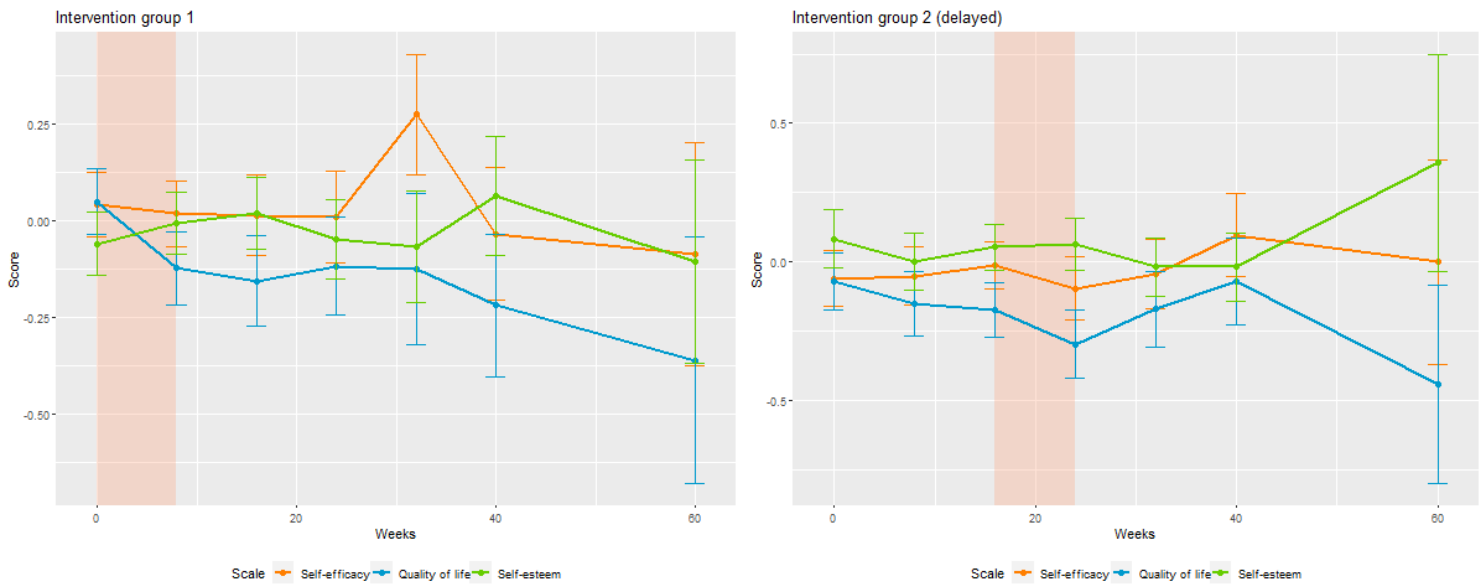


Figure 2

Scores on Self-efficacy, Quality of Life and Self-esteem across seven measurement waves