UNIVERSITY OF GOTHENBURG
DEPARTMENT OF SOCIOLOGY AND WORK SCIENCE

Master thesis in Sociology, 30 higher education credits

Impact of educational system on pupils' well-being: Measurement construction of psychological, cognitive and social well-being

Thea Klapp

Supervisor: Hans Ekbrand
Spring Term, 2020

Abstract: The aim of the study was to examine how the psychological, cognitive and social dimensions of well-being were constructed regarding Swedish Grade 6 pupils' school experiences, and whether the three constructs differed between pupils. In 2011, major educational reforms were implemented in the Swedish school system, which resulted in earlier introduction of grades (in 6th Grade), which overall created a stricter high-stakes assessment system. In the study, the educational reforms were used to investigate implications on pupils' school-related psychological, cognitive and social well-being, by comparing two cohorts of pupils born 1998 (did not experience reform) and pupils born 2004 (experienced full reform). The additional variables were sex, educational background and cognitive ability. Data were retrieved from the longitudinal project Evaluation Through Follow-Up. Confirmatory factor analysis and structural equation modeling were used. The results supported the theoretical construction of measurement instruments for psychological, cognitive and social well-being. Further, the results indicate that the reforms affected pupils' psychological and social well-being negatively. Girls reported lower psychological well-being compared to boys, while pupils' cognitive ability had particularly strong effects on psychological and cognitive well-being. Generally, the effects were stronger for the 2004 cohort.

Keywords: pupils; psychological well-being; cognitive well-being; social well-being; measurement construction.

Introduction

Children and youths' mental health and well-being have lately become a more and more pressing issue in Sweden. In 2017, the Swedish Government assigned the Public Health Agency of Sweden (PHAS) the mission of investigating well-being amongst children and youth in Sweden (Government decision S2017/01227/FS), based on the worrying results reported from the survey study Health Behaviour in School-aged Children (Public Health Agency of Sweden, 2014). PHAS (2014:6; 2018:10) report that the mental illness of adolescents in Sweden has increased since the mid 1980's, with psychosomatic symptoms such as sleeping problems, depression, headaches and stomach-ache. Moreover, this negative development is more extensive in Sweden than it is in the neighbouring countries (ibid.), indicating that the development isn't necessarily a Nordic trend. The development of child and youth well-being can, according to PHAS (2018:64), partly be connected to school related factors, such as school stress. Seeing that children and youth spend a considerable amount of time in school during their compulsory education makes it reasonable to discuss not only child well-being, but specifically school-related child well-being. This is supported by research findings, for example by Lewis and Frydenberg (2002), whose results indicate that the family home and school are the two factors that tend to have most influence on children's well-being.

In the beginning of the 90's, several major reforms were implemented that changed the Swedish educational system profoundly. These reforms included changes such as decentralization, criterion-referenced grading with a fail step in upper secondary school (National Agency for Education, 2014), independent schools became government funded with vouchers for each student (which led to an increased number of independent schools as well as increased competition between schools) (Proposition 1991/92:95). In the years around 2011,

the liberal government reformed the Swedish educational system in a conservative direction, with focus on results rather than learning (National Agency for Education, 2015:244). They argued, among other things, that the reforms were implemented in order to increase Sweden's competitiveness in international measurements of student achievement, such as the Program for International Student Assessment (PISA), in which the result of Swedish pupils had dropped compared to earlier years and to other countries (National agency for Education, 2010:128). The reform implementations took effect between the years 2009-2012 and included several changes in the Swedish educational system.

The problem of pupils' well-being and mental health is investigated in academic fields such as sociology, psychology and social work. However, the term well-being is neither defined nor measured consistently between projects, which have created confusion about what is actually measured. Academic measurements of well-being can therefore be argued to have a problem of validity. Pollard and Lee (2003) argue, for instance, that well-being should be understood as a construct that consists of several facets, which means that well-being should not be generalized as "well-being" unless all facets are represented in the measurement. An incomplete construction and presentation of a measurement instrument could thus mean that important nuances are overlooked. Psychological, cognitive and social well-being are examples of facets that are a part of the general well-being construct, and these facets will be the main focus of the present study.

Purpose and research questions

The overall purpose of the study is to investigate the dimensionality of the construct well-being, with specific regards to the psychological, cognitive and social facets of the construct, by conducting confirmatory factor analysis (CFA) with data from Evaluation through Follow-Up (Härnqvist, 2000). Furthermore, comparisons of the construct dimensionality between pupils unaffected by the educational reforms and pupils affected by the reforms will be made with tests of measurement invariance, with data of Swedish pupils born 1998 and 2004. Structural equation modeling (SEM) will be conducted as a last step, to investigate the structure of well-being in relation to pre-reform/post-reform schooling, sex, educational background and cognitive ability. Specifically, the study intends to answer the following research questions;

- How can school-related psychological, cognitive and social well-being for Swedish pupils in Grade 6 be constructed, and to what extent are the measurements invariant over time?
- What was the overall effect of the 2011 educational reforms on pupils' school-related psychological, cognitive and social well-being?
  - How did the effect of the reforms vary with sex, educational background and cognitive ability?

Background of the educational system

To understand whether or how the school environment might affect school children's well-being, a brief review of the Swedish compulsory school system is in order. Here, four of the changes regarding compulsory school will be presented. First, and perhaps most importantly, prior to the reforms, pupils received grades for the first time during their 8th school year. After the reform implementations (2011-12), pupils receive grades for the first time in 6th Grade, which is followed by grading assessments every term throughout compulsory school (National Agency for Education, 2015:46). This specific reform implementation took effect autumn 2012, for Grade 6 pupils (ibid.:50). Second, in Sweden the national tests are today introduced in Grade 3 instead of in Grade 5 and 9. The national tests today also include testing in more subjects than Swedish, English and Mathematics (ibid.:49). Third, prior to the reforms, pupils who did not meet the grade requirements for a passing grade, received no grade in the specific subject. After the reforms, the grade Fail was introduced in Swedish compulsory school, which means that pupils from the age of 11 face a risk of failing at school (ibid.:47). Fourth, the requirements for entering upper secondary school became stricter. Prior to the reforms, passing grades in the subjects Swedish, English and Mathematics were required to be able to apply to upper secondary school (ibid.:31). After the reforms, pupils need passing grades in at least 12 subjects if they are interested in college preparatory programs, and 8 passing grades if they are interested in vocational programs (National Agency for Education, 2019). In 2018, 15.6 % of Swedish Grade 9 pupils were not qualified to enter upper secondary school, according to statistics presented by the Swedish National Agency for Education (2018). Overall, the educational reforms have contributed to a school system with a stronger focus on tests and results in comparison to the pre-reform school system. Grades have been given an even more important role in the Swedish educational system and have thus become high-stakes for pupils in the sense that the consequences of grades highly affect the possibility of continuing within the educational system in Sweden.

Previous research

The Public Health Agency of Sweden is not alone in highlighting the connection between school-related factors and child and youth well-being, several researchers have contributed to the subject of pupils' well-being. School factors such as high demands, tests and grading have been shown to give negative effects on stress and anxiety amongst children and youth (von der Embse et al., 2013; Pollard & Lee, 2013; Högberg et al., 2019; Konu et al., 2002; Chamberlin et al., 2018).

One of the reasons to study psychological, cognitive and social well-being is because they don't seem to be distributed randomly across populations. Bradburn (1969:5) notes, for example, that psychological well-being is unequally distributed across economic and social classes, as well as between men and women. Later studies have also shown that girls tend to report lower well-being in comparison to boys (Högberg et al., 2019; Nordlander & Olofsdotter Stensöta, 2014; Haugland et al., 2001; Sonmark et al., 2016; Giota & Gustafsson, 2016). However, seeing that well-being in some cases has been inconsistently defined and measured (Pollard & Lee, 2003), the opposite result regarding gender effects can also be

found (Miller et al., 2013; Pietarinen et al., 2014). Furthermore, studies have shown that grades in different ways can have negative impact on pupils' mental health and self-rating. Giota and Gustafsson (2017), as well as Schraml et al. (2011), found that perceived failure decreased compulsory school pupils' self-esteem, which is highly relevant considering the current grading system in the Swedish compulsory school, in which it is possible for pupils to formally fail in school subjects. Grades are also related to pupils' academic self-concept (Marsh et al., 2007), which has been shown to relate to factors such as anxiety, social comparison and perceived self-worth (Marsh, 1986; Marsh, Walker & Debus, 1991; Bong & Clark 1999).

Chamberlin et al. (2018) found that grades didn't function as an intrinsic academic motivation for students, but rather engendered extrinsic academic motivation. Extrinsic motivation can be described as an externally controlled form of motivation with lack of personal autonomy, while intrinsic motivation is autonomously driven (Deci & Ryan, 2000). Intrinsic motivation has been shown to positively affect psychological well-being (Bailey & Phillips, 2016), conceptual learning and academic persistence (Vansteenkiste et al., 2006) and engagement and academic performance (Taylor et al., 2014). In comparison, extrinsically oriented motivation has been shown by some researchers to relate to decreased academic persistence, achievement and well-being (Ryan & Deci, 2017). Further, research findings show that grades contribute to social comparison and competition between peers, which fosters school environments that oppose peer solidarity and instead encourage more individualistic approaches (Chamberlin et al., 2018; Tannock, 2015). Other studies have examined the effect of high-stakes testing and exams on pupils' mental health. In the Swedish educational system, grades are high-stakes, but assessed during a school term instead of during one or a couple of exams. Banks and Smyth (2014) found that although high-stakes exams had a negative impact on Irish pupils' school-related stress, well-functioning social relations with peers and teachers mitigated the negative effect. West and Sweeting (2003), on the other hand, found that pupils' family occupational backgrounds had significant effects on whether pupils worried about matters such as school performance and high-stake exams.

Högberg et al. (2019) suggest that the 2011 educational reforms have affected the health of Swedish pupils negatively, as post-reform pupils reported an increase in psychosomatic symptoms compared to pre-reform pupils. They also included school-related stress and academic self-esteem in the analysis. Only one indicator was used to measure each of the two phenomena, however, which could be argued to be a limitation. Nordlander and Olofsdotter Stensöta (2014) found a positive association between school grades and general subjective well-being for Swedish pupils. However, the measurement instrument didn't represent the multidimensionality of general well-being, and can therefore be argued to not be an adequate measure of general well-being. This means that there could be a discrepancy in what is presented and what is actually measured. The present study will hopefully contribute to further knowledge of the multidimensionality of the construct, in a school context.

Pupils' school-related well-being and health has been studied, both nationally and internationally and often in regard to gender differences. However, the dimensionality of the

construct well-being, as well as differences between subgroups of pupils, needs further investigation. Little or no research has been made that investigates how school-related dimensions of well-being are structured and can be measured. The same goes for whether Swedish pupils' have been negatively affected, regarding psychological, cognitive and social well-being, by the 2011 educational reforms. Grades existed and functioned as the selection system for higher education before the educational reforms as well, and likely inflicted stress and anxiety for 8th and 9th Grade pupils (during the ages of 14-16). However, considering that the introduction of testing and grading occurs earlier (during the ages of 12-13) and more frequent, it is reasonable to think that pupils' school-related well-being has been affected for that age group. Hopefully, the present study can contribute to definitions and measures of psychological, cognitive and social well-being and show the effect of the educational reforms.

Sex, educational background and cognitive ability

In this section, three subgroups of pupils are presented. These subgroups will later be presented as variables and used to study whether they affect school related psychological, cognitive and social well-being. To investigate differences between pupils, sex or gender is often used as a background variable. Sex is commonly defined as biological sex; while gender is understood as socially constructed roles of masculinity or femininity (see e.g. Carl, 2012). This variable was collected by Statistics Sweden when the sampled pupils were in Grade 3, based on their registered biological sex, meaning that in the present study, sex is the correct term to use. As mentioned previously, well-being has in several studies been shown to differ between girls and boys (Högberg et al., 2019; Nordlander & Olofsdotter Stensöta, 2014; Haugland et al., 2001; Sonmark et al., 2016; Giota & Gustafsson, 2016), which is why sex is included as a categorization in the current study. Further, an important category to consider is pupils' educational background. Parents' educational level functions as an indicator of pupils' social origin, and has been shown to be an important aspect to consider regarding pupils' educational attainment and other academic outcomes (Bukodi & Goldthorpe, 2013). However, when studying parents' educational level together with parents' income and social status, the measurement would give a fuller picture of pupils' social origin (Bukodi & Goldthorpe, 2013; Bukodi et al., 2014). Even so, educational level alone is an important component to include when investigating school-related issues, because it says something about the educational resources available at home, which can influence educational support and guidance, as well as knowledgeable guidance through the educational system (Bukodi & Goldthorpe, 2013).

A third factor that has documented effects on school outcomes is pupils' cognitive ability. Cognitive ability refers to pupils' ability to perform in academic settings with respect to problem solving and knowledge acquisition, but should be distinguished from academic achievements, such as school grades or test scores, because they measure different aspects of pupils' academic capability (Chen et al., 2012). Verbal, inductive and spatial tests are components that often are included in the cognitive ability term (Svensson, 1964:6). These tests are not a part of the Swedish school curricula, and are thus not something pupils have trained specifically for. Research (e.g. Marsh, 1990) shows that cognitive ability has direct and positive effects on academic achievement and academic self-concept, while Stoeger and Ziegler (2010) found that cognitive ability affects pupils' academic self-efficacy and

motivation positively. Academic self-concept and self-efficacy are terms of academic self-perception, that is, pupils' beliefs and thoughts of their own academic ability (Marsh, 1990). Further, Svensson (1964:36) found evidence of a relation between social groups and cognitive ability, indicating that socioeconomic status influence pupils' cognitive ability positively. Next follows a theoretical presentation of psychological, cognitive and social well-being.

Theory and concepts

Pollard and Lee (2003) discusses the dimensionality of the well-being construct. According to their compilation of the literature, general child well-being is constructed of five main facets: the psychological, cognitive, social, economic and physical facets. In the current study, the psychological, cognitive and social facets of well-being are accounted for, leaving the economic and physical facets out due to data limitations and project demarcation. In the present study, well-being refers to pupils' coping mechanisms of everyday school events. That is, how pupils are affected by everyday difficulties in school. The constructed well-being measurements are therefore not a measure of severe mental illnesses, such as depression, in the present study. Furthermore, it is the subjective form of well-being that is of interest in the present study, namely pupils' own reports of their school experiences (Angner, 2010; Pollard & Lee, 2003).

Psychological, cognitive and social well-being should, according to Pollard and Lee (2003), be measured by indicators that say something about children's emotions and mental health or mental illness. In the school-related versions of the three facets, the indicators should reasonably be defined as school-related emotions and school-related mental health or illness.

Three facets of well-being

Psychological well-being

The psychological well-being facet should, according to Pollard and Lee (2003), be measured by indicators that say something about children's emotions and mental health or mental illness. Anxiety, distress, nervousness, stress and self-esteem are examples of psychological well-being indicators (ibid.). For these indicators to say something about school-related psychological well-being, they should reasonably be related to and studied in a school context. The emotions and mental health or mental illness, expressed through the mentioned examples of indicators, should therefore be directly connected to school. For example, school-related stress could be caused by events at school or expectations on certain performance from surrounding people.

Cognitive well-being

Cognitive well-being refers to a well-being that is related to individuals' intellectual capability and the opportunities or limitations it entails (Pollard & Lee, 2003). That is, cognitive well-being isn't a reference to actual intellectual capacity, but rather how individuals' perceive that their own capability is functioning in an intellectual context, such as school. Clearly, cognitive well-being is already in its natural state a construct that captures well-being in a school context. Cognitive well-being is often measured by indicators such as

self-concept of academic ability, perceived academic incompetence, ability to concentrate or emotions connected to academic achievement (ibid.). In short, school-related cognitive well-being can be understood as pupils' emotions concerning their cognitive functioning in the school context.

Social well-being

Social well-being can be understood as having supporting and well-functioning social relationships (Pollard & Lee, 2003). Individuals who feel that they are able to develop and maintain satisfying and healthy relationships with the surrounding people tend to have high social well-being, while individuals who feel the opposite are more likely to experience social isolation and because of that are more exposed to depression and lack of well-being (Bandura et al., 1996). Relationships with peers and family are examples of social relationships that tend to be important for social well-being (Pollard & Lee, 2003). According to Pollard and Lee (2003), this facet of well-being is often measured by indicators such as peer problems, anti-social behaviour, social support and social acceptance. Translated into a school context, the indicators should reasonably say something about how pupils perceive their peer and student-teacher relationships.

In all, the literature shows some inconsistent results and lacking of a "standard" definition of general well-being, as well as of the different dimensions, which leads to an inconsistent discussion of child well-being. Next, the methodology of the study will be presented.

Methodology

The study used a cross-sectional secondary questionnaire data design, to investigate school-related psychological, cognitive and social well-being with two samples of Swedish pupils born 1998 and 2004. The quantitative methods confirmatory factor analysis (CFA) and structural equation modeling (SEM) are employed in the present study. To conduct the analyses the following statistical software programs were used; IBM SPSS Statistics 26.0 was used for data management, descriptive analysis and graph making, while Mplus 7 (Muthén & Muthén, 1998-2012) was used for modeling and data analyses. First, the participants, procedure of data collection, background variables and measurement instrument will be presented, followed by methods of analysis and the modeling of measurement- and structural models.

Participants

Data were retrieved from the Swedish longitudinal database Evaluation through Follow-up (ETF) (Utvärdering genom Uppföljning) (Härnqvist, 2000). The sampling was a two-step stratified procedure, where municipalities were selected in the first step and catchment areas in the second step. In the present paper, questionnaire data collected from the two cohorts with Swedish pupils born 1998 (N = 9180) and 2004 (N = 9775) were used, when the pupils in both cohorts were of the age 12-13 in Grade 6. The samples are 10 % of the populations, and are nationally representative of their respective population. The two cohorts were chosen because the pupils in the 1998 cohort were in the school system before the educational reforms were implemented and did thus not experience the stricter assessment system. The

2004 cohort, on the other hand, was in the school system after the reforms were implemented and did therefore experience the stricter school system. Examinations of the self-reported well-being from the two cohorts thus enable comparisons of pupils' school-related well-being prior and post the 2011 reform implementations. The sampled pupils answered identical questions over cohorts, which further enables comparisons across cohorts.

Procedures

The database consists of register and questionnaire data, compiled by Statistics Sweden for nationally representative samples of individuals born between 1948 and 2004, approximately every fifth year (Härnqvist, 2000). Background variables, such as sex, age and parents' level of education, were collected when the pupils were in the 3th Grade by Statistics Sweden. Questionnaire data for the ETF project were collected for the first time when the pupils had started 6th Grade. The questionnaire data for the 1998 cohort were collected spring 2011 (prior to the educational reform implementations), while the questionnaire for the 2004 cohort were collected spring 2017, when the reform concerning earlier grades had been active for about 4-5 years.

Cohort, sex and educational background

The background variables used as independent variables in the study are cohort, sex and educational background. The data were retrieved by Statistics Sweden when the pupils had entered Grade 3. Cohort is coded as a dichotomous dummy variable and is operationalized as year of birth with the values 0 = pupils born 1998 and 1 = pupils born 2004. Similarly, sex is coded into a dichotomous dummy variable, with the values 0 = boy and 1 = girl. Family educational level is operationalized as the educational level of the highest educated parent (it doesn't have to be a biological parent). It has 5 categories, 1 = pre-secondary education, 2 = Vocational preparatory two- or three-year upper secondary education and two-year study preparatory upper secondary education, 3 = Three-year study preparatory upper secondary education and post-secondary education less than two years, 4 = Two or three years of college education, 5 = College education covering four years or more. Originally, the variable has an additional category, "Data not available", which has been coded as missing observations.

Cognitive ability

The fourth independent variable used in the study is cognitive ability (COGN). It is an interval scale variable operationalized by three added variables that each represent a test of cognitive ability (Svensson, 1964:6). The tests are the following: vocabulary antonyms, a spatial- and an inductive test. The test of vocabulary antonyms is a measure of verbal ability. The pupils are presented with a word and are asked to choose between four alternative words to find the word that is an antonym to the first word (ibid.). The spatial test is designed as a paper folding test, where the pupils are presented with a picture of an unfolded paper figure and thereafter asked to choose among four alternatives of folded figures, to find the matching one (ibid.). The last test is the inductive test, which is designed as a test of number series. The pupils are asked to find the pattern in six presented numbers, to continue the number pattern with two more numbers (ibid.). These three tests are conducted at the same time as the

questionnaire in Grade 6. The cognitive ability variable is standardized into z-scores, in order to make the results more comprehensible.

Instrument

The measurement instruments were first composed based on theory and then further investigated with confirmatory factor analysis. They consist of three latent factor variables, psychological, cognitive and social well-being. Every factor consists of a number of questionnaire items (see Table 1), on 5-point Likert scales. The items function as measurable indicators of theoretical latent factors. Psychological well-being is operationalized by four indicators that are connected to school-related stress and worries, which is in line with previous research (Pollard & Lee, 2003). Cognitive well-being is operationalized by six indicators that are connected to cognitive functioning and behaviour in school, as well as academic self-concept in Swedish, English and Mathematics. Social well-being is operationalized by four indicators that says something about pupils' social relations within school, both with peers and teachers, as well as in different contexts, such as in the classroom and during breaks. The variables are coded so that the higher the value, the greater well-being.

Table 1. Questionnaire items

Items	Answers (5-point scales)
Psychological well-being	
How true are the following? I worry about things that happen in school	Always – never
How true are the following? I worry about tests on homework	
How true are the following? I worry about how I am going to pass exams	
How true are the following? Experience stress	
Cognitive well-being	
How true are the following? Find it difficult to concentrate in lessons	Always – never
How true are the following? I normally manage to do the tasks that I am given	
How true are the following? I find it difficult to keep up in lessons	
How good do you think you are in the following subjects? Swedish	Very poor – very good
How good do you think you are in the following subjects? English	
How good do you think you are in the following subjects? Mathematics	
Social well-being	
How true are the following? I have friends who I can be with	Always – never
How do you like your current class?	Very poor – very good
How do you like the teachers?	
How do you like it during breaks?	

Note: The questionnaire items were answered on 5-point Likert scales.

Methods of analysis

In this section, the methods of analysis utilized in the study are presented, as well as the procedures of implementing the methods.

Descriptive statistics

As a first procedure, descriptive analyses were conducted in IBM SPSS Statistics 26.0, to investigate trends in the data. In order to see how the observations were distributed, frequency distributions were scrutinized. Additionally, the internal consistency of the items was estimated with Cronbach's alpha.

Introduction to confirmatory factor analysis

To conduct the proposed study, confirmatory factor analysis (CFA) was used as the first method of analysis. CFA is a theory driven statistical method, that is useful for investigations of the relation between indicators (observed variables) and latent factors (unobserved variables) (Brown, 2015:10; Harrington, 2009:3; Muthén & Muthén, 1998-2017:55). Factors are theoretically constructed variables, often built by questionnaire items, and are suitable to use as measurement instruments in analytical methods such as structural equation modeling (SEM). The purpose of using CFA in the current study was to test if the data fitted the theoretically assumed measurement models of the three constructs of well-being (psychological, cognitive and social well-being). That is, to test the construct validity of the measurement models. Construct validity refers to whether a measurement model measures what is theoretically intended and can be divided into two parts; convergent validity and discriminant validity (Brown, 2015:2). Convergent validity refers to the relation between the indicators that are set to a certain factor. The factor loadings should therefore indicate appropriately strong interrelations, for the convergent validity to be adequate. Regarding discriminant validity, neither factors nor indicators, which theoretically belong to different factors, should have a strong interrelation. For instance, if two factors have a correlation of around or above .80 to .85, the reason could be that they aren't two separate factors but rather measure the same phenomenon (ibid.:28).

The relation between factor and indicator can be understood as the relation between independent and dependent variables, respectively, because latent factors are assumed to be the cause of the variance in the observed indicator (Harrington, 2009:17). The observed variables are influenced by the latent construct in the factor, meaning that they are interrelated because they are caused by the same latent construct (Brown, 2015:10). The estimate of the relation between factor and indicator, factor loading, can be understood as a regression coefficient (ibid.). Factor loadings can range from -1 to 1 and the further it is from 0, the stronger the relation between indicator and factor. As a minimum, factor loadings should be over .30, but preferably around at least .50-.70 (Harrington, 2009:17). From the factor loadings, common and unique variance of the factors was calculated.

Multiple-groups CFA

To test whether the factor structures were invariant across subgroups of pupils, measurement invariance was tested with multiple-groups CFA. With multiple-groups CFA models, it is possible to investigate if there are group similarities or differences regarding the measurement structure (Vandenberg & Lance, 2000). Testing measurement invariance with multiple-groups CFA have advantages compared to more traditional methods, such as ANOVA. With the CFA approach, for example, measurement error is adjusted for when estimating factor means,

variances and covariances, which contributes to the statistical strength of multiple-groups CFA (Brown, 2015:267).

Configural, metric and scalar invariance have been investigated. Configural invariance refers to equal factor structures across groups (Brown, 2015:242), meaning that the number and pattern of factors and factor loadings, respectively, are constructed to be identical across groups. At this stage, two models were run. One in which the factor loadings were restricted to have equal patterns of fixed and free components across groups and one that varied freely. This to control that the restricted model was not significantly worse than the freely estimated model, as well as controlling whether the restricted model fitted the data (Marsh et al., 2018). Metric invariance refers to equal factor loadings across groups, which means that the factor loadings should have approximately the same value across groups in order to be invariant (Brown, 2015:243). Similar to the configural model testing, two models were run to test metric invariance, one in which the factor loadings were constrained to be equal across groups and one that was estimated freely, to control that the constrained model was not significantly worse compared to the freely estimated model and that it fitted the data (Marsh et al., 2018). If the constrained model fits the data, the metric model is invariant and the indicators can thus be said to relate equally to the factors across groups, which further enables comparisons of factor variances (Brown, 2015:243). The third test, scalar invariance, refers to equal indicator intercepts across groups. It means that the indicator intercepts should be of the same value across groups in order for the measurement to be scalar invariant (Brown, 2015:243). Two models were compared to test scalar invariance, one in which the indicator intercepts were constrained to be equal across groups and one which was estimated freely, to control that the constrained model didn't have a significantly worse fit than the freely estimated model (Marsh et al., 2018). In order to compare factor means across groups, the scalar model needs to be measurement invariant across groups (Brown, 2015:243). However, scalar invariance is many times difficult to achieve with real data (Muthén & Asparouhov, 2013; Asparouhov & Muthén, 2014; Marsh et al., 2018). When groups answer significantly different on questionnaire items, such as between cohort differences, there will be scalar non-invariance.

Introduction to structural equation modeling

The measurement model constructed with CFA laid the groundwork for the next step, structural equation modeling (SEM), in which additional structural models were constructed. SEM is thus built by two types of models, a measurement model and a structural model (Jöreskog, 1993; Byrne, 2012:14). The measurement model is constructed with CFA, as previously described, and consists of relations between manifest and latent variables. The structural model, on the other hand, defines the relationships between the latent factors with structural regression coefficients. The structural regression coefficients in SEM have the same function as ordinary regression coefficients, namely identifying the effect of one variable on another (Byrne, 2012:11). SEM can be understood to consist of several statistical techniques, such as path analysis and multiple regression, which contributes to the benefits of the method. Other strengths of SEM is the ability to conduct analyses with both manifest and latent variables, either as predictor or outcome variables (Kline, 2015:13), as well as the possibility

of testing both direct and indirect effects, and testing for reciprocal relations between the model variables (Gustafsson, 2009:269).

Model specification

The measurement models are theoretically constructed, meaning that theoretical assumptions of the latent factors guided the use of indicators. In a CFA model, the parameters need to be considered when specifying a model. A parameter can be understood as population characteristics and can be specified in three different ways; free, fixed and constrained (Brown & Moore, 2012:365). With a free parameter, the analysis finds the ideal values that have limited discrepancy between observed and predicted covariance, whereas a fixed parameter is manually specified to a certain value. The constrained parameter is neither set to be of a certain value, nor free to be of any value. Instead, a restriction is set on the predicted values of the solution, which needs to be fulfilled (ibid.). When conducting measurement invariance evaluation, constrained parameters are used in relation to unconstrained parameters.

The metrics (or units of measurements) of the latent variables are fixed by one marker indicator per factor, meaning that the metric of the marker indicator is transferred to the factor. The three marker indicators were selected by default in Mplus (Muthén & Muthén, 1998-2017). Further, the construction of the structural model is similarly based on theoretical assumptions. The measurement model consists of the three latent factor variables; psychological, cognitive and social well-being, and 14 indicators. The first structural model consists of the three latent factors and the three background variables sex, educational background and cognitive ability (see Figure 1), while cohort is added in the second structural model. Latent variables are represented by circles and ellipses in the models, while manifest variables are represented by squares.

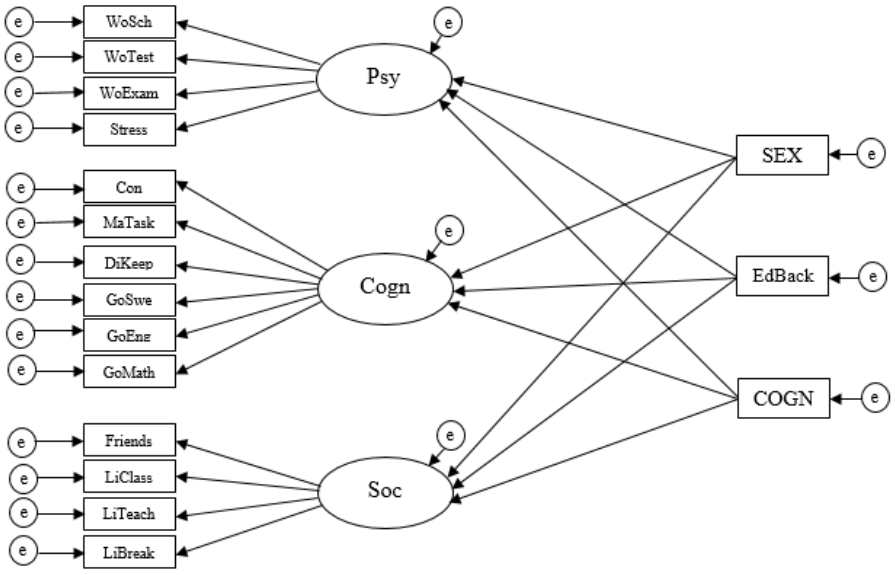

Figure 1. Model specification of the structural equation model with a 3-factor measurement model

Model estimation

To estimate the parameters in the identified models, the estimator MLR has been used. MLR is a robust maximum likelihood estimator, which means that it finds the set of parameters which differs the least between the predicted and the observed parameter estimates in the variance- covariance matrix (Brown, 2015:62). A robust estimator is suitable to use for managing non-normal distribution, because it corrects the model estimates and standard errors for non-normality in the sample (Muthén & Muthén, 1998-2017:668; Brown, 2015:65). Further, the type = complex option is used in the input analysis command, together with the cluster option in the input variable command, enabling corrections of the standard errors and tests of model fit that, for example, accounts for observations that aren't independent (Muthén & Muthén, 1998-2017:57).

Model evaluation

To test how well the models fits the data, model fit information have been estimated. The following goodness-of-fit indices are used: root mean square error of approximation (RMSEA), confidence intervals for RMSEA, comparative fit index (CFI), Tucker-Lewis index (TLI) and standardized root mean square residual (SRMR). RMSEA evaluates how well a model fits in the population and should be equal or below .06 to be considered acceptable (Brown, 2015:74), while the confidence intervals suggest the accuracy of the RMSEA estimate. CFI is a comparative index that measures the specified model in relation to a restricted baseline model, while the TLI is a measure of unnecessary model complexity (ibid.:73). Both CFI and TLI should be around .90-.95 to indicate a good model fit. Lastly, SRMR measures the average difference between the predicted correlations and the observed correlations and should be equal or below .08 to be considered acceptable (ibid.:74). Additionally, statistical significance of model parameter estimates will be evaluated (Brown & Moore, 2012:369).

Ethics

Participants in the ETF project are protected by secrecy and the data have been collected and handled in accordance with the Data Protection Directive (Central bureau of Statistics, 2011; Berndtsson & Svensson, 2012), which is in line with the ethical principles of the Swedish Research Council (2017:72). It is therefore not possible to find out about the identity of the participants in the ETF project. Since the participants weren't of age when participating in the project, both the pupils and the parents were informed about the purpose of the project, as well as that participation was voluntary and anonymous. Additionally, the current thesis is carried out transparently regarding design, method and results. Consequently, the four principles stated by the Swedish Research Council (2017:13) and ALLEA (2017); information, consent, confidentiality and utilization, are met.

Results

In this section, the results from the analysis will be presented in four steps: descriptive statistics; measurement modeling (CFA); measurement invariance and lastly; structural equation modeling (SEM). The results presented here will then be theoretically analysed in the discussion.

Descriptive statistics

Internal missing is the primary type of missing observations, meaning that participants have skipped either certain or all questions in the questionnaire (Berndtsson & Svensson, 2012). In the 1998 cohort, 18.5% of the observations are missing in the cognitive ability test, while 46% of the observations are missing in the 2004 cohort. Regarding the questionnaire, in the 1998 cohort, 17.3-18.6% of the observations are missing, while 46.6-48.1% of the observations are missing in the 2004 cohort (see Table 2). The number of internal missing observations is thus greater in the 2004 cohort than it is in the 1998 cohort. The number of observations for the 1998 cohort are N = 7476 in the cognitive ability test and N = 7516-7591 in the questionnaire items. For the 2004 cohort, the number of observations are N = 5097 in the cognitive ability test and N = 4895-5040 in the questionnaire items. Sex and educational background are background variables collected by Statistics Sweden, which means that they have no internal missing. The missing of 0.5% (1998) and 0.7% (2004) in the educational background variable represent the response option "Data not available" which has been coded as missing.

Table 2. Descriptive statistics for the manifest variables; sex, educational background, cognitive ability and questionnaire data, for two cohorts of pupils born 1998 and 2004.

Variables	N				% missing		M		SD	
	98		04		98	04	98	04	98	04
	Girls	Boys	Girls	Boys						
SEX	4454	4721	4651	4786	0.0	0.0				
Educational Back.	9144		9379		0.5	0.7				
Cognitive Ability	7476		5097		18.5	46.0	64.90	67.13	16.75	17.71
Questionnaire data										
14 items	7516-7591		4895-5040		17.3-18.6	46.6-48.1				

Note: The questionnaire data contain missing on every item to approximately the same amount.

Further, the 2004 cohort has a somewhat higher mean in the cognitive ability variable than the mean in the 1998 cohort, suggesting that low-performing pupils in the 2004 cohort did not respond in the same extent as high-performing pupils. To control if the missing data in the 2004 cohort were skewed, a calibration weight was used. This revealed that there was a minor skewness but it did not affect the estimates considerably. The internal consistency of the factor indicators was measured with Cronbach's alpha, and the estimates are as follows, for the 1998 and 2004 cohorts respectively; psychological well-being = .74/.79, cognitive well-being = .68/.68 and social well-being = .73/.75. The internal consistency of the items in the cognitive well-being factor is just below the value of .7, which often is acknowledged to be the minimum of an acceptable value (Field, 2013:709). However, seeing that the measure of internal consistency in this case is a way to describe the data rather than a part of the main analysis, the somewhat low estimate doesn't have to be considered a problem. Consequently, the scale reliability is overall acceptable.

Confirmatory factor analysis

Presented in Table 3 are the standardized factor loadings and goodness-of-fit estimates for the confirmatory factor analyses conducted separately for the two cohorts. The factor loadings for

the psychological well-being factor indicate that all four indicators have a strong influence on the factor in both cohorts and specifically in the 2004 cohort. The factor loadings for the cognitive well-being factor indicate that the cognitive factor has a somewhat lower influence on the indicators than for the psychological factor. The factor loading for the item self-concept in English is .27 for the 2004 cohort, which is below .30 and could thus be regarded as having a too weak influence on the cognitive factor to be included. However, seeing that the nature of the current study is comparative, the factors must be identically structured between the two cohorts, for differences in well-being to be properly compared. The other factor loadings indicate an acceptable to good fit for both cohorts, with factor loadings ranging from .32 to .70. The factor loadings for the social well-being factor indicate that the indicators have a stable relation with the factor. Overall, the factor loadings imply an adequate convergent validity of the measurement models.

Table 3. Standardized factor loadings for the 3-factor measurements models, for each cohort separately. Indicator intercepts in parentheses.

Indicators	Psychological w-b		Cognitive w-b		Social w-b	
	98	04	98	04	98	04
Worry school	.61 (3.46)	.66 (3.11)				
Worry tests on homework	.70 (3.47)	.70 (3.22)				
Worry pass exams	.68 (2.69)	.74 (2.55)				
Stress	.60 (3.57)	.67 (2.71)				
Hard to concentrate			.63 (3.58)	.66 (2.89)		
Manage tasks			.48 (7.04)	.50 (5.50)		
Difficult to keep up			.68 (4.22)	.70 (3.61)		
Good SWE			.35 (6.03)	.32 (4.95)		
Good ENG			.34 (4.73)	.27 (3.92)		
Good MATH			.44 (4.57)	.44 (3.72)		
Have friends					.59 (8.01)	.58 (6.22)
Like current class					.79 (5.66)	.76 (4.69)
Like the teachers					.48 (5.10)	.52 (4.52)
Like it during breaks					.75 (6.88)	.80 (5.52)

Note: All estimates are significant on the .001 level. Goodness-of-fit for 1998: RMSEA = .057 [C.I .055-.060] CFI = .905 TLI = .878 SRMR = .045. Goodness-of-fit for 2004: RMSEA = .054 [C.I. .052-.057] CFI = .922 TLI = .900 SRMR = .043

To further evaluate the construct validity and specifically the discriminant validity, factor correlations are presented. For the cohorts 1998 and 2004 respectively, the psychological factor has a correlation of .66/.65 with the cognitive factor and .40/.38 with the social factor; while the cognitive factor and the social factor have correlations of .42/.42. The psychological and cognitive factors thus have more to do with each other than with the social factor. They are not, however, on the level of correlation that indicates poor discriminant validity (Brown, 2015:28). The acceptable levels of convergent validity and discriminant validity indicate that the constructs are valid, meaning that the measurement models measures what is theoretically intended. The two measurement models both have good model fit, with RMSEA below .60 and CFI above .90. To compare, a model in which all three factors were included as one was

estimated for each cohort. The goodness-of-fit estimates for these models were unacceptable, which gives further evidence of the multidimensionality of the well-being construct.

Moreover, the tests of measurement invariance with multiple-groups (see Table 4) show that the configural (equal factor structure) and metric (equal factor loadings) invariance models have good model fit, with a somewhat low TLI, while the model fit for the scalar model (equal indicator intercepts) slightly decreases. The $\chi 2$ difference testing between the models further suggests that the difference in model fit between the configural and the metric model is appropriately low, while the differences of 837 and 794 $\chi 2$ between the scalar model and the configural- and metric models, respectively, are high. However, RMSEA indicates acceptable model fit for the scalar model, which implies that the scalar model is acceptable but not perfect. That is, the indicator intercepts does not differ significantly between the two cohorts. As stated previously, however, perfect scalar invariance is difficult to achieve (Muthén & Asparouhov, 2013; Marsh et al., 2018)

Table 4. Tests of measurement invariance of psychological, cognitive and social well-being of pupils born 1998 and 2004, with unstandardized parameter estimates.

Model	$\chi 2$	df	RMSEA (90% CI)	SRMR	CFI	TLI
Single-group solutions						
1998 (n = 7670)	1855***	71	.057 [.055-.060]	.045	.905	.878
2004 (n = 5051)	1130***	71	.054 [052-.057]	.043	.922	.900
Measurement invariance						
Configural	2977***	142	.056 [.054-.058]	.045	.912	.888
Metric	3001***	153	.054 [.052-.056]	.047	.912	.895
Scalar	3876***	164	.060 [.058-.061]	.051	.885	.873
Metric vs Configural	59***	11				
Scalar vs Configural	837***	22				
Scalar vs Metric	794***	11				

Note: N = 12 721.

Structural equation models

In table 5, the first structural equation model is presented with the two cohorts separated in order to compare the effect of the background variables on the latent factors (psychological, cognitive and social well-being) for pupils born 1998 and 2004. The relation between the factors was accounted for in the models. Girls in both cohorts have reported lower psychological well-being compared to boys in their respective cohort, but the sex differences are greater in the 2004 cohort. While sex does not have a significant effect on cognitive well-being for the 1998 cohort, the effect on the same factor in the 2004 cohort is significant and negative, meaning that 2004 girls have reported lower cognitive well-being compared to the boys in their cohort. Further, sex has negative effect on social well-being in both cohorts, meaning that girls in both the 1998 and the 2004 cohort have reported lower social well-being compared to boys in their respective cohort. The difference is, however, greater in the 2004 cohort, meaning that the effect of sex is more distinguishable in the 2004 cohort than in the 1998 cohort. Regarding pupils' educational background (EdBack), the only significant effects

are found on the cognitive well-being factor. The positive estimates indicate that higher education relate positively to pupils' cognitive well-being in both cohorts. Educational background does not have any significant effects on psychological and social well-being.

Cognitive ability (COGN) has positive and significant effects on psychological well-being in both cohorts, meaning that pupils who performed higher on the cognitive ability tests tended to report higher on the psychological well-being items. The effect of cognitive ability is, however, greater in the 2004 cohort, meaning that cognitive ability became more important for pupils' psychological well-being. The same goes for the effect of cognitive ability on cognitive well-being. The effects are positive, significant and strong for both cohorts. The difference in effects across cohorts is, however, extensive. This indicates that cognitive ability became even more important for cognitive well-being in the 2004 cohort. See Figure 2 for a visual presentation of the estimates. The two SEM models both fit the data, with a RMSEA of .058 [C.I. 0.056-.060] for the 1998 model and .062 [C.I. 0.060-.065] for the 2004 model.

Table 5. Structural equation models with standardized factor coefficients. The cohorts were estimated in separate models. Latent factors are psychological, cognitive and social well-being.

	Psy w-b		Cogn w-b		Soc w-b		RMSEA/SRMR	
	98	04	98	04	98	04	98	04
SEX[1]	-.23	-.29	-.02^{ns}	-.10	-.03*	-.10		
EdBack	.01^{ns}	-.00^{ns}	.07	.07	.00^{ns}	.02^{ns}		
COGN	.18	.20	.37	.50	-.01^{ns}	.10		
Model fit							.058/.046	.062/.049

Note: [1]Boys are the reference. Educational background (EdBack), cognitive ability (COGN). All estimates are significant on the .001 level when no marking is presented. * p≤.05.

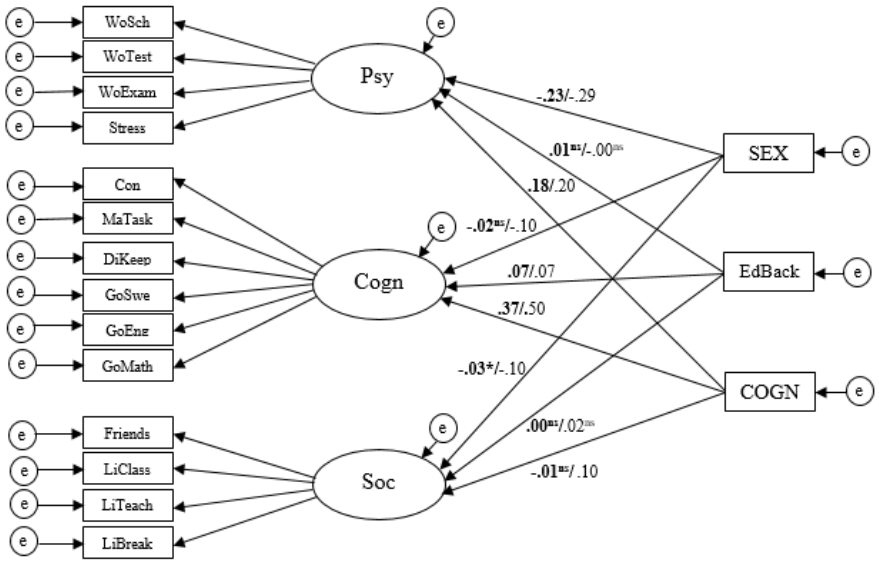

Figure 2. Structural equation models, with the cohorts 1998 (bold font) and 2004 (normal font) in two separate models. Background variables are sex, educational background (EdBack) and cognitive ability (COGN).

In the second structural equation model (see Table 6), the two cohorts were included in the same models as a dummy coded variable (0 = 1998, 1 = 2004) as an additional background variable, in order to investigate the direct effect of cohort on the latent factors. Further, a fully saturated model with interaction terms was estimated to control for possible moderation effects, whereof the two significant interaction terms are presented. In model 1, cohort has a significant and negative effect on both psychological and social well-being, meaning that pupils belonging to the 2004 cohort have reported lower well-being, compared to pupils belonging to the 1998 cohort. The effect of cohort is strongest on psychological well-being, with a coefficient estimate of -.20. The effect of cohort on cognitive well-being, on the other hand, is significant and positive, indicating that 2004 pupils have reported higher cognitive well-being, compared to the 1998 pupils. The remaining coefficient estimates in model 1 are in line with the previously presented estimates, when considering that the effects of two distinct groups have been collapsed. By collapsing two separate groups who had different contexts to relate to, goodness-of-fit is affected negatively. With RMSEA on .082 [C.I. 0.081-.084], CFI on .736, TLI on .660 and SRMR on .057, the model can't be said to fit the data well. Even so, the model fills an important role in evaluating the effect of birth cohort on the latent factors.

Table 6. Structural equation model with standardized factor coefficients. Both cohorts were estimated in the same model. Latent factors are psychological, cognitive and social well-being.

Variables	Model 1			Model 2			Model 3		
	Psy w-b	Cogn w-b	Soc w-b	Psy w-b	Cogn w-b	Soc w-b	Psy w-b	Cogn w-b	Soc w-b
Cohort[1]	-.20***	.16***	-.12***	-.12**	.12***	-.09***	-.21***	.16***	-.12***
SEX[2]	-.23***	.05***	-.06***	-.14***	.01[ns]	-.03*	-.23***	.05***	-.06***
EdBack	.01[ns]	-.06***	.00[ns]	.01[ns]	-.06***	.01[ns]	.01[ns]	-.06***	.01[ns]
COGN	.22***	-.39***	.04***	.22***	-.39***	.04***	.23***	-.32***	-.01[ns]
Cohort*SEX				-.16***	.07***	-.06**			
Cohort*COGN							-.02[ns]	-.11***	.07***
RMSEA/SRMR		.082/.057			.084/.057			.081/.056	

Note: [1] Pupils born 1998 are the reference, [2]Boys are the reference. Educational background (EdBack), cognitive ability (COGN). *** p≤.001, **p≤.01, *p≤.05.

When interacted together in model 2, high values in the interaction term Cohort*SEX (girls born 2004) are negatively related to psychological- and social well-being, in comparison to the reference groups. The interaction effect mitigates the strengths of the original and separate relations, from model 1. This indicates that the negative effect of cohort, and by implication the reform of 2011, is larger for girls than for boys. Further, high values in the interaction term Cohort*SEX have a positive relation to cognitive well-being, in comparison to the reference groups. The relation between Cohort and cognitive well-being is somewhat mitigated by this, and it fully rules out SEX as a single explaining variable to cognitive well-being. Girls in the 2004 cohort have thus reported somewhat higher on the cognitive well-

being scale, compared to boys within their cohort, as well as to boys and girls in the 1998 cohort.

In model 3, the interaction term Cohort*COGN has significant effects on cognitive- and social well-being only. The estimate for the single term COGN on social well-being is non-significant, meaning that for the 1998 cohort, the variable does not have a significant effect on social well-being. Otherwise, the original and separate effects remain approximately of the same strength. The negative estimate of the interaction term on cognitive well-being (see Figure 3), suggests that pupils in the 2004 cohort, with lower cognitive ability, reported lower cognitive well-being compared to pupils in the 1998 cohort with the same cognitive ability, as well as compared to pupils with higher cognitive well-being, within both cohorts. The positive effect of the interaction term on social well-being (see Figure 4), suggests that for pupils with higher cognitive ability, the effects on social well-being were similar between the cohorts. For pupils with lower cognitive ability, however, the difference between the cohorts was larger, meaning that the 2004 cohort reported significantly lower social well-being compared to the 1998 cohort.

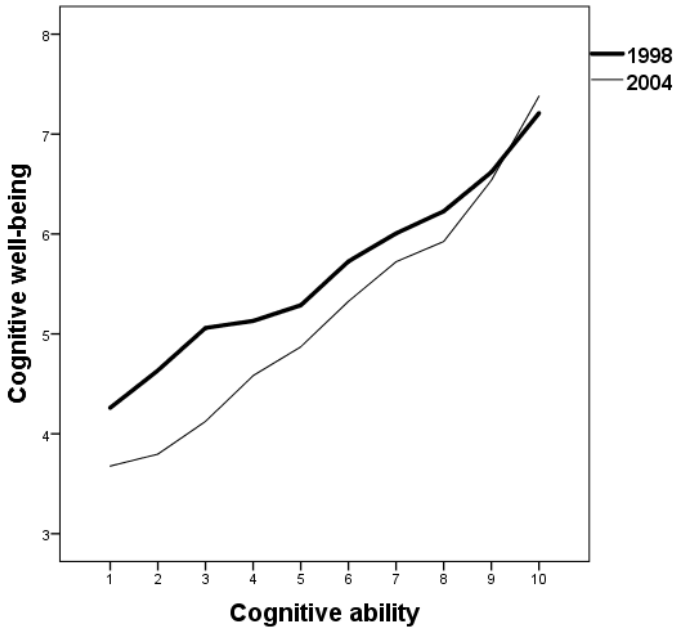

Figure 3. The interaction effect of Cohort*COGN on cognitive well-being. All values are standardized; cognitive ability is grouped in 10 deciles while cognitive well-being is presented as average grouped/decile.

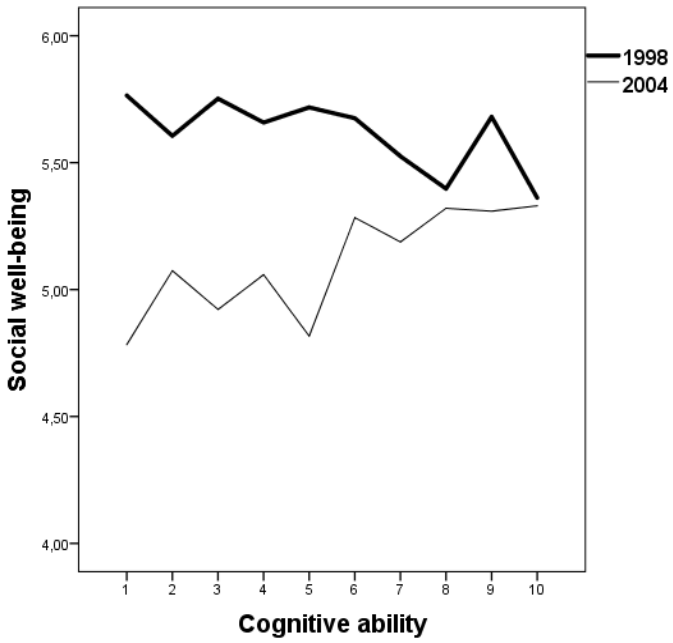

Figure 4. The interaction effect of Cohort*COGN on social well-being. All values are standardized; cognitive ability is grouped in 10 deciles while social well-being is presented as average grouped/decile.

In all, the results suggest that there are differences in psychological, cognitive and social well-being between subgroups of pupils, and that pupils belonging to certain subgroups have stronger relations to certain factors. Additionally, pupils born 1998 and 2004 differ in their

reports of well-being; 2004 pupils seem to have lower psychological- and social well-being compared to the 1998 pupils, but higher cognitive well-being.

Discussion

The overall purpose of the study was to investigate the psychological, cognitive and social dimensions of the construct well-being, and compare between two cohorts of pupils born 1998 and 2004, to measure the reform effects. Pupils born 1998 were unaffected by the 2011 educational reforms in Sweden, while pupils born 2004 were affected in full by the reforms. An additional purpose was to investigate the psychological, cognitive and social well-being constructs in relation to pupils' sex, educational background and cognitive ability.

The multidimensionality of well-being

One important aspect in the results is that significant nuances can be overlooked when well-being is investigated as a unidimensional construct. In previous studies, well-being is commonly presented as a unidimensional construct which is problematic in relation to the results of the present study. It is, on the contrary, of importance to regard well-being as a multidimensional phenomenon and should thus be empirically investigated as such. In the present study, it was found that the well-being construct has at least three separate facets (the psychological, cognitive and social facets), in relation to a school context. This suggests that a generally constructed measure of well-being, which lacks full representation of the separate facets, does not necessarily measure what is stated or intended. The results suggest, for example, that the measures of school-related psychological, cognitive and social well-being have good construct validity, meaning both that the indicators within the factors have strong interrelations, and that the factors are appropriately independent from each other (Brown, 2015:2). This means that the questionnaire items in the respective facets of well-being have enough in common with each other to be good measures of well-being, while the facets are different enough from each other to be regarded as good measures of school-related psychological, cognitive and social well-being. This is in line with arguments presented by Pollard and Lee (2003), who claim that unless all facets (psychological, cognitive, social, economic and physical) are represented in the measure, it is not a full measurement of well-being and should thus not be referred to as simply "well-being".

The multidimensionality of well-being in relation to subgroups

Yet another interesting result, which further emphasises the importance of measuring well-being as a multidimensional construct, is that the dimensions of well-being seem to relate differently to different subgroups of pupils. That is, there seem to be significant patterns regarding the subgroups, which indicate that pupils have different levels of psychological, cognitive and social well-being, conditioned by their sex, educational background and cognitive ability.

The results regarding sex show that both 1998 and 2004 girls reported lower overall well-being, compared to the boys in their respective cohort, which support the previous findings of sex differences in well-being (Högberg et al., 2019; Nordlander & Olofsdotter Stensöta, 2014;

Haugland et al., 2001; Sonmark et al., 2016). However, the results indicate that the sex differences were greater in psychological well-being in both cohorts, compared to the differences in the cognitive- and social well-being dimensions. Thus, within the same cohort, differences between the sexes is primarily associated with psychological well-being, such as feeling worried and stressed over school-related tasks such as tests and exams. This is in line with several studies (e.g. Högberg et al., 2019; Haugland et al., 2001; Sonmark et al., 2016) and indicates that girls seem to be more affected by pressure to perform well in school, compared to boys. Additionally, it is reasonable to think that these outcomes are the results of gender expectations, rather than biological sex.

Further, previous research indicates that educational background has a positive impact on pupils' educational attainment and other academic outcomes (Bukodi & Goldthorpe, 2013), which is supported by the results of the present study. One possible cause for the influence of educational background on pupils' cognitive well-being could be the benefits of having parents with educational resources, who to a higher degree than others can provide educational support and guidance to their children (Bukodi & Goldthorpe, 2013; Bukodi et al., 2014). It is reasonable to believe that educational resources influences pupils' cognitive well-being positively, seeing that educational resources tend to influence pupils' cognitive ability (Svensson, 1964), which in turn is an important aspect for cognitive well-being. When it comes to previous research on the influence of family background on psychological well-being, however, the results are mixed. West and Sweeting (2003) found a relation between parental occupation and pupils' psychological well-being, while Högberg et al. (2019) found no significant impact of parents' employment on pupils' school-related stress, which is in line with the present study. That is, pupils' school-related worries and stress were not dependent on their family's educational background. It is reasonable to think that the cause for the non-significant effect of educational background on psychological well-being was that the source of pupils' worries and stress came from processes in school, which weren't affected differently by low- or highly educated parents. That is, it is likely that parents had similar short-term expectations regardless of educational level, which could be argued to differ from the effect of highly educated parents on academic outcomes such as educational attainment, that Bukodi and Goldthorpe (2013) note, which thus would be long-term expectations.

Cognitive ability has in previous research been found to positively affect pupils' academic achievements and self-concept (Marsh, 1990), as well as their self-efficacy and motivation (Stoeger & Ziegler, 2010), which is confirmed by the results of the current study. Pupils' cognitive ability seems to have strong effects on their psychological- and cognitive well-being in both cohorts. Academic self-concept items were a part of the measure of cognitive well-being, and there are reasons to believe that pupils with high cognitive ability also perceived themselves as academically capable. It is reasonable to believe that this, in turn, affects pupils' cognitive well-being positively.

Well-being measured over time

A third important aspect in the results is that there seems to be differences in school-related psychological, cognitive and social well-being over time, between the two cohorts with pupils

born 1998 and 2004, where each cohort went through their 6th school year under different circumstances.

The comparison of the cohorts show that pupils who were exposed to the 2011 reforms reported lower psychological- and social well-being, compared to pupils who were not exposed to the reforms. Pupils' school-related worries and stress thus increased after the reform implementations 2011, which supports the results presented by Högberg et al. (2019). This may be caused by increased pressure as a consequence from earlier grading and testing in school. Regarding social well-being, previous research findings suggest that grades foster a school climate of social comparison and peer competition (Chamberlin et al., 2018; Tannock, 2015), which could explain the decrease in social well-being from cohort 1998 to cohort 2004. It is possible that the 2011 educational reforms influenced the 2004 pupils' social relations in school negatively by introducing testing and grading at an early age, fostering pupils to compare results instead of creating an open learning environment. Overall, the differences across groups in psychological- and social well-being indicate that pupils were negatively affected by the 2011 educational reforms.

The relation between cohort and cognitive well-being, however, differs from the overall pattern. Pupils who were exposed to the 2011 educational reforms seem to have higher cognitive well-being, compared to the pupils who were not exposed, which does not support previous research in the subject (Schraml et al., 2011). However, the educational reforms were only beneficial for pupils' cognitive well-being if they had high cognitive ability, while pupils who did not score the highest at the cognitive ability test were negatively affected by the reforms. Cognitive ability seems to be an important aspect for both psychological- and cognitive well-being in the cohorts, but particularly so in the 2004 cohort. This indicates that the relation between cognitive ability and psychological- and cognitive well-being has increased in strength from the 1998 to the 2004 cohort. There is thus reason to believe that cognitive ability was more important for pupils in the 2004 cohort, in order to develop functioning psychological- and cognitive well-being, than it was for pupils in the 1998 cohort. The group difference is particularly large when it comes to cognitive well-being, indicating that 2004 pupils' beliefs and thoughts of their ability to perform in school were highly influenced by their cognitive ability, which could be a consequence from the pressure from the result driven system introduced by the 2011 educational reforms. One possible interpretation is presented by Schraml et al. (2011), who argue that high levels of stress tend to increase cognitive difficulties, which could explain why the difference in cognitive well-being between high- and low-performing pupils increases substantially from the 1998 to the 2004 cohort. That is, seeing that the 2004 pupils are put under greater school-related stress than their older counterparts, it becomes even more important for the 2004 pupils to have a high cognitive ability, in order to develop a positive cognitive well-being.

Further, the results indicate that girls in the 2004 cohort experienced more worries and stress in school in relation to boys with the same educational background and cognitive ability, compared to the same relation in the 1998 cohort. In previous research, sex differences in perceived school-related stress have been suggested to be caused by the often higher

expectations from parents' and teachers' on girls, regarding academic performance and behaviour, which thus result in a higher exposure to stress for girls compared to boys (e.g. Schraml et al., 2011; Högberg et al., 2019). Considering that the year when grades were introduced had been lowered to 6th Grade, it is reasonable to believe that the 2004 pupils, with girls in particular, were more exposed to expectations on their academic performance from an early age, than their older counterparts. Thus, the increase in difference between the sexes from the 1998 to the 2004 cohort may have been a consequence of the 2011 educational reforms (Public Health Agency of Sweden, 2018), which created a school climate with more emphasis on tests, results and performance than learning processes (National Agency for Education, 2015). Educational background seems to have an equal influence on cognitive well-being for both cohorts. This indicates that the 2011 educational reforms did not particularly affect the relation between pupils' educational background and their cognitive well-being. One reason for this could be that the educational resources that parental education can provide to pupils (Bukodi & Goldthorpe, 2013), are not dependent on the school climate.

Conclusion

The Swedish school has not been disconnected from the neo-liberal development that has taken place in recent years. As a consequence of this, conservative ideals are now prevalent in the educational system, with disciplining practices such as early introduction of grades (National Agency for Education, 2015:46). These implementations have been criticized, for example by Högberg et al. (2019), who investigated the neo-liberal reforms from a health perspective. The current study has had a similar aim, but with a focus on measurements of well-being, as well as potential differences between subgroups of pupils, since the literature is somewhat lacking of consistent definitions and measures. Previous studies regarding well-being have had a tendency to fail discussing the dimensionality of the construct well-being, which in actuality is a multifaceted construct. By overlooking the complexity of the construct well-being, statistical results may be portrayed incorrectly and lack important nuances. It could also be argued to be misleading to describe and measure well-being as a "general" unidimensional construct, if not all dimensions are represented in the measure. The present study contributes with knowledge of how the psychological, cognitive and social well-being dimensions can be measured and how the effects of exposure to the reforms, sex, educational background and cognitive ability differ between Swedish Grade 6 pupils.

The results of the study indicate that the current result-driven-educational-system in Sweden seem to affect compulsory pupils' well-being negatively, as well as increasing the gaps between pupils' prerequisites to succeed in school. This calls for a re-evaluation of the values the Swedish school currently is built on, from result based values to learning based values, to create school environments that foster well-being and health (Public Health Agency of Sweden, 2018:88). In conclusion, to deal with the current downward trend of Swedish pupils' health and well-being, more research on school-related well-being and mental illness is necessary.

Limitations and future research

It could be so that additional interesting nuances were overlooked because all facets were not included in the measurements of the present study. In future research, therefore, the economic and physical facets of the construct should be included in the measures, to further investigate school-related well-being as a full construct. It would allow for a more general discussion about school-related well-being. Additionally, the present study lacks academic achievement as an outcome. In following research, the scope of the current study can thus be developed by including final grades as an outcome, which would make the results more relatable to actual academic achievements and making it possible to further investigate and compare how school-related well-being affected pupils in the 1998 and 2004 cohorts. The final grades for the 2004 cohort will be available in autumn 2020 and further analyses will be carried out in a project funded by the Swedish Research Council (dnr. 2019-04531).

References

ALLEA – All European Academies. (2017). *The European Code of Conduct for Research Integrity*. Berlin: ALLEA – All European Academies.

Angner, E. (2010). Subjective well-being. *Journal of Socio-Economics*, *39*(3), 361-368.

Asparouhov, T., & Muthén, B. O. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling*, *21*(4), 495–508.

Bailey, T. H., & Phillips, L. J. (2016). The influence of motivation and adaptation on students' subjective wellbeing, meaning in life and academic performance. *Higher Education Research & Development*, *35*(2), 201–16.

Bandura, A., Barbaranelli, C., Caprara, G.V., and Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development*, *67*(3), 1206–1222.

Banks, J., & Smyth, E. (2015). 'Your whole life depends on it': academic stress and high-stakes testing in Ireland. *Journal of Youth Studies*, *18*(5), 598-616.

Berndtsson, Å., & Svensson, A. (2012). *Insamling av enkätuppgifter från elever i grundskolans årskurs 6 våren 2011 för UGU-projektets nionde kohort (födda 1998)*. Gothenburg: University of Gothenburg.

Blohm, L. (2018, feb 08). Ny modell ingen succé. *Tidningen Gymnasiet.* Retrieved from https://tidningengymnasiet.se/ny-modell-ingen-succe/

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*(3), 139–154.

Bradburn, N. M. (1969). *The structure of psychological well-being.* Aldine Press: Chicago, IL.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), Handbook of structural equation modeling (pp. 359-379). New York: Guilford Publications.

Bukodi, E., Erikson, R., & Goldthorpe, J. (2014). The effects of social origins and cognitive ability on educational attainment: Evidence from Britain and Sweden. *Acta Sociologica*, *57*(4), 293-310.

Bukodi, E., & Goldthorpe, J. (2013). Decomposing 'Social Origins': The Effects of Parents' Class, Status, and Education on the Educational Attainment of Their Children. *European Sociological Review*, *29*(5), 1024-1039.

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge.

Carl, J. (2012). Gender vs. sex: What's the difference? *Montessori Life*, *24*(1), 26-30.

Central Bureau of Statistics. (2011). Uppföljningsundersökning Elever - Teknisk rapport. Stockholm: Central Bureau of Statistics.

Chamberlin, K., Yasué, M., & Chiang, I.-C. A. (2018). The impact of grades on student motivation. *Active Learning in Higher Education.* https://doi.org/10.1177/1469787418819728

Chen, S.-K., Hwang, F-M., Yeh, Y-C., & Lin, S. S. J. (2012). Cognitive ability, academic achievement and academic self-concept: Extending the internal/external frame of reference model. *British Journal of Educational Psychology*, *82*(2), 308-326.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–68.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics: and sex and drugs and rock 'n' roll*. (4. ed.) Los Angeles: Sage.

Giota, J., & Gustafsson, J.-E. (2017). Perceived Demands of Schooling, Stress and Mental Health: Changes from Grade 6 to Grade 9 as a Function of Gender and Cognitive Ability. *Stress and Health, 33*(3), 253-266.

Government decision. S2017/01227/FS. *Uppdrag avseende skolbarns hälsovanor*. Retrieved from https://www.regeringen.se/494aba/contentassets/ff98176eae454498a26752e03149b0e3/uppdrag-till-fohm-avseende-skolbarns-halsovanor.pdf

Gustafsson, J.-E. (2009). Strukturell Ekvationsmodellering. [Structural Equation Modeling]. In G. Djurfeldt & M. Barmark (Eds.), *Statistisk verktygslåda 2 Multivariat analys*. (pp. 269-321). Lund: Studentlitteratur.

Harrington, D. (2009). *Confirmatory factor analysis*. Oxford: Oxford University Press.

Haugland, S., Wold, B., Stevenson, J., Aaroe, L. E., & Wyonarowska, B. (2001). Subjective health complaints in adolescence. *European Journal of Public Health, 11*(1), 4-10.

Härnqvist, K. (2000). Evaluation through follow-up. A longitudinal program for studying education and career development. In C.-G. Janson (Ed.), Seven Swedish longitudinal studies in behavioural science. Forskningsrådsnämnden: Stockholm.

Högberg, B., Lindgren, J., Johansson, K., Strandh, M., & Petersen, S. (2019). Consequences of school grading systems on adolescent health: evidence from a Swedish school reform. *Journal of Education Policy, 34*, 1-23.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. Scott Long (Eds.), *Testing Structural Equation Models* (pp. 294-316). California: Sage Publications.

Kline, R. B. (2015). Principles and Practice of Structural Equation Modeling. New York: The Guilford Press.

Konu, A. I., Lintonen, T. P., & Rimpelä, M. K. (2002). Factors associated with schoolchildren's general subjective well-being. *Health Education Research, 17*(2), 155-165.

Lewis, R., & Frydenberg, E. (2002). Concomitants of failure to cope: What we should teach adolescents about coping. *British Journal of Educational Psychology, 72*, 419-431.

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*(4), 646-656.

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*(1), 129–149.

Marsh, H. W., Guo, J., Parker, P. D., Nagengast, B., Asparouhov, T., Muthén, B., & Dicke, T. (2018). What to do When Scalar Invariance Fails: The Extended Alignment Method for Multi-Group Factor Analysis Comparison of Latent Means ccross Many Groups. *Psychological Methods, 23*(3), 524–545.

Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The Big-Fish-Little-Pond Effect: Persistent Negative Effects of Selective High Schools on Self-Concept after Graduation. *American Educational Research Journal, 44*(3), 631-669.

Marsh, H. W., Walker, R., & Debus, R. (1991). Subject-specific components of academic self-concept and self-efficacy. *Contemporary Educational Psychology, 16*(4), 331–345.

Miller, S., Connolly, P., & Maguire, L. K. (2013). Wellbeing, academic buoyancy and educational achievement in primary school students. *International Journal of Educational Research, 62*, 239-248.

Muthén, B., & Asparouhov, T. (2013). *New methods for the study of measurement invariance with many groups*. Retrieved from http://statmodel2.com/download/PolAn.pdf

Muthén, L. K., & Muthén, B.O. (1998-2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B.O. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

National Agency for Education. (2019). *Antagning till gymnasieskolan*. Retrieved 2020-02-19 from https://www.skolverket.se/regler-och-ansvar/ansvar-i-skolfragor/antagning-till-gymnasieskolan#h-Behorighetochurval

National Agency for Education. (2014). *Betygshistorik.* Retrieved 2020-04-16 from https://web.archive.org/web/20141015220948/http://www.skolverket.se/bedomning/betyg/betygshistorik-1.46885

National Agency for Education. (2018). *Fler elever behöriga till gymnasiet*. Retrieved 2020-02-19 from https://www.skolverket.se/om-oss/press/pressmeddelanden/pressmeddelanden/2018-09-27-fler-elever-behoriga-till-gymnasiet

National Agency for Education. (2010). Rustad att möta framtiden? PISA 2009 om 15-åringars läsförståelse och kunskaper i matematik och naturvetenskap. Stockholm: National Agency for Education.
National Agency for Education. (2015). *Skolreformer i praktiken - Hur reformerna landade i grundskolans vardag 2011–2014.* Stockholm: National Agency for Education.
Nordlander, E., & Olofsdotter Stensöta, H. (2014). Grades – for better or worse? The interplay of school performance and subjective well-being among boys and girls. *Child Indicators Research, 7*(4), 861-879.
Pietarinen, J., Soini, T., & Pyhältö, K. (2014). Students' emotional and cognitive engagement as the determinants of well-being and achievement in school. *International Journal of Educational Research, 67*, 40-51.
Pollard, E.L., & Lee, P.D. (2003). Child well-being: A systematic review of the literature. *Social Indicators Research, 61*(1), 59-78.
Proposition 1991/92:95. Proposition om valfrihet och fristående skolor. Retrieved from https://www.riksdagen.se/sv/dokument-lagar/dokument/proposition/valfrihet-och-fristaende-skolor_GF0395
Public Health Agency of Sweden. (2014). *Health behavior in school-aged children*. Solna: The public Health Agency of Sweden.
Public Health Agency of Sweden. (2018). *Varför har den psykiska ohälsan ökat bland barn och unga i Sverige? Utvecklingen under perioden 1985–2014*. Stockholm: The Public Health Agency of Sweden.
Ryan, R. M., & Deci, E. L. (2017). *Self-determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York: The Guilford Press.
Sanmartín, R., Inglés, C. J., Vincent, M., Gonzálvez, C., Díaz-Herrero, À., & Garcia-Fernández, J. M. (2018). Positive and negative affect as predictors of social functioning in Spanish children. *PLoS ONE, 13*(8), 1-10.
Schraml, K., Perski, A., Grossi, G., & Simonsson-Sarnecki, M. (2011). Stress symptoms among adolescents: The role of subjective psychosocial conditions, lifestyle, and self-esteem. *Journal of Adolescence, 34*(5), pp.987-996
Sonmark, K., Godeau, E., Augustine, L., Bygren, M., & Modin, B. (2016). Individual and Contextual Expressions of School Demands and Their Relation to Psychosomatic Health a Comparative Study of Students in France and Sweden. *Child Indicators Research 9*(1), 93–109.
Stoeger, H., & Ziegler, A. (2010). Do Pupils with Differing Cognitive Abilities Benefit Similarly from a Self-Regulated Learning Training Program? *Gifted Education International*, *26*(1), 110–123.
Svensson, A. (1964). *Sociala och regionala faktorers samband med över- och underprestation i arbetet: pedagogisk-sociologiska studier jämte en beskrivning av skolstatistikens individualuppgifter* (Report 1964:13). Gothenburg: University of Gothenburg.
Swedish Research Council. (2017). *Good research practice*. Stockholm: Swedish Research Council.
Tannock, S. (2015). No grades in higher education now! Revisiting the place of graded assessment in the reimagination of the public university. *Studies in Higher Education, 42*(8), 1–13.
Taylor, G., Jungert, T., Mageau, G. A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: The unique role of intrinsic motivation. *Contemporary Educational Psychology, 39*(4), 342–58.
Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods,3*(), 4–70.
Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic Versus Extrinsic Goal Contents in Self-Determination Theory: Another Look at the Quality of Academic Motivation. Educational Psychologist, 41(1), 19-31.
von der Embse, N., Barterian, J., & Segool, N. (2013). Test Anxiety Interventions for Children and Adolescents: A Systematic Review of Treatment Studies from 2000–2010. *Psychology in the Schools, 50*(1), 57-71.
West, P., & Sweeting, H. (2003). Fifteen, female and stressed: changing patterns of psychological distress over time. *Journal of Child Psychology and Psychiatry, 44*(3), 399-411.
Wróbel, M., Finogenow, M., Szymańska, P., & Laurent, J. (2019). Measuring positive and negative affect in a school-based sample: A Polish version of the PANAS-C. *Journal of Psychopathology and Behavioral Assessment, 41*(4), 598-611.