Change in Pupils' and Students' Attitudes toward School as a Function of Age

– *A Finnish Perspective*

Jari Metsämuuronen[1], Renata Svedlin[2] & Jovan Ilic[3]

[1] Faculty of Behavioral Sciences, Helsinki University, Finland

[2] Applied pedagogics ÅboAkademi, Vasa, Finland

[3] British Council, Kathmandu, Nepal

Correspondence: Jari Metsämuuronen, Faculty of Behavioral Sciences, Helsinki University, Siltavuorenpenger 5 A, P.O.B. 9, FIN-00014, Finland. E-mail: Jari.Metsamuuronen@methelp.com

Abstract

The article describes changes in pupil and student contentment as a function of age in the general education system. School satisfaction was measured with a 216-item test with 15 different versions simultaneously in all grades of primary and lower secondary education and in two grades in upper secondary general education. IRT modeling with matrix sampling was used to equate the versions for all the grades. Pupil/student contentment decreased significantly after the 2nd grade with the intensity of the satisfaction declining steadily until the 8th grade. From 9th grade on, contentment recovered as it began to increase once again, albeit moderately. It then stayed constant until the 2nd grade in high school. Pupil/studentcontentment decreased especially in the areas of assessing teachers, classroom activities, and lessons; there was a smaller reduction in the areas of school atmosphere and physical environment. Over the years, the girls were more satisfied with school than were the boys. Pupils with a higher achievement level were more satisfied than those with a low level of achievement.

Keywords: self-assessment, school satisfaction, pupil contentment, attitude toward school, IRT modeling, school attitudes, student attitudes, pupil attitudes, school climate, student-school relationship, teacher-student relationship, primary education, lower secondary education, higher secondary education

1. Introduction

Though an ever greater share of learning nowadays is achieved in non-formal connections or through non-formal pedagogy (Silberman-Keller, 2006; Bekerman & Silberman-Keller, 2004; Frasca, 2003), pupils and students are still attending school and it can be considered as their "working place". One may relevantly ask: should one enjoy being in the working place? So, by extension, one may ask should one enjoy being in school? Some writers (e.g., Olkinuora & Mattila, 2001, 20; Liinamo & Kannas 1995, 110–111) have connected pupils'/students' enjoyment in school to the quality of the school and areas such as "working conditions". Due to administrative development in schools many usually administer some kind of self-assessment in school. One of the first exercises to achieve this is to measure the level of satisfaction of the pupils and the teachers. The layman's judgment is that, if overall satisfaction is not acceptable, then other functions of the school also cannot be done well. This layman's viewpoint does not differ much from research results. The characteristics of well-functioning schools have been connected to the concept of a "positive climate" as a central feature (OECD, 1995, 20). However, Hautamäki and colleagues (Hautamäki *et al.,* 1999) have noticed that school satisfaction and effectiveness do not necessarily go together.

This article approaches school satisfaction from the pupils'/students' contentment viewpoint. It starts with the brief examples of the pupil/student attitude reviewed on the basis of the international research literature. Secondly, relevant Finnish studies are presented as a basis for the comparison of results in the empirical part of the study. Finally, the changes in the pupil/student contentment are described as a function of age. The subjects of the study are pupils from the 1st grade to the 9th grade and students of upper secondary schools.

2. Student Attitude in the Research Literature

In the ERIC database (http://www.eric.ed.gov/), there is a myriad of publications concerning "Pupil/student attitudes" (over 63,000 hits at spring 2012), "teacher student relationship" (over 18,000 hits), "school attitudes" (over 1, 390 hits), or school climate (over 2,000 hits). Obviously, it is not possible to cover all possible aspects of the topic in a brief article. Although these studies have touched on different aspects (e.g., different grades, approaches, background theories, settings, and the research problem) as Ruoho, Koskela and Pihlainen-Bednarik (2006, 110) have noted, they are here classified under the umbrella of "school satisfaction."

2.1 International Literature

On the basis of the ERIC database, school satisfaction research seems to fall into four different sections. The first includes general school satisfaction from the viewpoint of either general attitude towards school (e.g., Holfve-Sabel, 2006; Gray & Mclellan, 2006; Peters *et al*., 2005; Holfve-Sabel & Gustafsson, 2005), positive school climate (e.g., Cohen, 2006; Perkins, 2006; Catalano *et al*., 2004; Wentzel& Watkins, 2002), or belonging to school (e.g., Sanchez, Colon & Esparza, 2005; Ma, 2003; Rostosky *et al*., 2003; OECD, 2000). The second answers the question of how pupils/students feel about different school subjects, i.e., how pupils/students enjoy attending certain lessons? This sector includes the studies of attitudes towards, for example, reading or writing (e.g., McKay, 2006; Garcia-Sanchez & de Caso-Fuertes, 2005; Sainsburg & Shagen, 2004; Wang, 2000), mathematics (e.g., Maraffi, 2007; Schweinle, Meyer & Turner, 2006; Wilkins & Ma, 2003), the sciences (e.g., PISA, 2006; Papanastasiou & Papanastasiou, 2004; Reiss, 2004; Morrell & Lederman, 1998), languages (e.g., Merisuo-Storm, 2007), or art teaching (e.g., Pavloua & Kambourib, 2007). The third study line is bullying in the school either generally (e.g., Hurst, 2005; Siris & Osterman, 2004; Dake *et al.,* 2003; Eisenberg, 2003) or connected to school attendance (e.g., Borup & Holstein, 2007; 2006; Schnohr & Niclasen, 2006). The fourth study line links school satisfaction as a part of pupils'/students' health. This includes general school health surveys (e.g., Stakes, 2007a), including school satisfaction as one of the topics of the inquiry, interest in the school health care system (e.g., Sihvola, 2000), and connection of school attendance to, for example, smoking (e.g., Nutbeam & Aaro, 1991).

It is relevant to note that the educational systems in different countries may produce certain kinds of variations in school satisfaction and pupil/student contentment. Therefore, it makes sense here to focus on relevant Finnish literature when comparing the results of this study with previous ones.

2.2 Finnish Literature

Some of the relevant Finnish literature is seemingly old. For example, Olkinuora (1983) studied 9th graders' motivation and experiences of sensibleness in school attendance. He defined school satisfaction as the affective component of the sense of sensibleness in school attendance. In his definition, he emphasized the feelings pupils may have and the attitudes they have expressed as well as their experiences of sensibleness toward going to school. In the study, the pupils' positive attitude toward school correlated with their feeling of sensible experiences in school. Another study of Uusikylä and Kansanen (1988) covered also on primary education pupils' contentment. They concluded that satisfaction in school was connected to two separate phenomena. First, school satisfaction was linked to a feeling of school as a working place, where working consists of performance-oriented learning, doing homework, and regularly repeating exams. Second, school satisfaction was partly connected with social relations, classmates, and free time between lessons (recess breaks).

In the 1990s, there were published some relevant studies. Scheinin (1990) connected pupil contentment with the self-image and self-esteem of pupils in primary education. He concluded that school achievement (or how the pupil is experiencing it) did not directly explain self-esteem experiences. However, low achievement could be linked to a general feeling of failure and giving up, which on itself was connected with school satisfaction. Sermilä (1991) studied lower secondary education pupils' perceptions at good performing schools. Pupils emphasized that the central factors for a well-functioning school are receiving good marks and a good diploma, healthy friendships, a sense of appreciation, and the possibilities of contributing in class and being motivated to do so. School satisfaction was strongly attached to pupil-teacher relations, pupil-pupil relations, and the sense of a motivating study atmosphere. Linnakylä (1993) studied the quality of school life. The concept was operationalized by the following factors: general school satisfaction, pupil-teacher relations, social status in class, self-esteem in class, perception of possibilities for achievement, and general school negativity. The strongest explanatory variables for satisfaction in school were gender (girls were more satisfied), plans for further academic education (those with a plan to go to high school were more satisfied), and watching television (those who watched less television enjoyed going to school more). Good reading skills were also connected with school satisfaction. Another study of Scheinin (1999) shed light on how well the 6th graders got along with each other

and found that they had a generally positive attitude toward themselves and the school. Only 4% of the pupils were found to have a clearly negative attitude toward school and learning.

More recent studies are published during the new Millennium. Metsämuuronen studied the general attitudes toward the mother tongue in Finnish language (2006a) and in Swedish language (2006b) and compared these (2006c) using a longitudinal design. The pupils were 7th and 9th graders. The attitudes toward the subject declined by approximately 8% during the three years among Finnish-speaking pupils, whereas the data on Swedish-speaking pupils showed a significantly smaller reduction of 0.3%. In surveys of the youth's health and well-being by the National Research and Development Centre for Welfare and Health (Stakes), in 1999, 2001, 2003, 2005, and 2007, school satisfaction has been part of the study (Stakes, 2007a). The important parts of the inquiry have been pupils' feelings and conceptions of being in school, about themselves as learners, the working environment (e.g., possible noise problems, ventilation, temperature problems, and ergonomic factors), experiences of haste, and pupil-teacher relations. The study involved pupils from the 1st, 2nd, 8th, and 9th grades. Stakes (2007a) found that only 2–3% of the 1st and 2nd grade pupils did not like school at all and 11–15% judged school to be a poor working environment. Among the 8th and 9th grade pupils, 5–9% did not like school at all and 25–30% considered their school to not be a good working environment. The discontent or criticism of pupils seemed to increase with age, confirming Metsämuuronen's (2006a) longitudinal data. However, as years went by, the older age groups' attitudes became more positive. Another trace of studies was started by Stakes in 2007 when, together with the Finnish National Board of Education (FNBE), a long-term survey of health and well-being of learning communities was administered (Stakes 2007b). Some surprising facts were reported by Rimpelä (2007, 133): 2% of the Finnish schools have had cases of pupils threatening others with guns, 4% have had cases where pupils have harmed the teacher, and 7% have had cases where pupils have threatened their teacher with violence. At the same time, Finnish newspapers reported an increase in the number of cases of violence against teachers (e.g., *Opettaja* 4/2007; *Helsingin Sanomat* 26th January 2007; *Helsingin Sanomat* 21st November 2007).

International comparisons show that the Finnish pupils' general school satisfaction have been ranked either at the same level as that of other European pupils (Linnakylä, 1993, 44–55; Lie & Kjæensli, 2006, 83; Arinen & Karjalainen, 2007, 63–65) or (at least partly) markedly lower (Linnakylä, 1993, 44–55; Kannas, 1995). The most positive results regarding the attitude of pupils were noted in other Scandinavian countries and in the USA (Linnakylä, 1993, 44–55). The pupil-teacher relation was more negative and school negativity was significantly more common in Finland than in other countries. Also, research conducted by WHO (Kannas, 1995) showed that school negativity was clearly quite common in Finland. In a comparison of 20 countries, Finnish boys were ranked second to last; with only the situation in Israel being deemed worse. Part of the explanation in Finland (and in Israel) can be the appreciation of honesty and speaking frankly (or "ugly straightforwardness" as Israeli columnist Shoher [2007] has called it). Leino (2003, 79) noted that pupils in Finland (as in other Nordic countries) are relatively humble when they describe their knowledge. This "humbleness" may also be reflected in attitude measurements.

In the WHO inquiry (Kannas, 1995), approximately 30% of the pupils in the lower secondary level did not enjoy school. According to the latest school health survey (Stakes, 2007a), approximately 10% of the pupils did not like school at all. Scheinin (1999) noted that less than 4% of the 6th grade pupils had a clearly negative approach to school and to learning. The discrepancy in the figures is noteworthy and it may be wise to consider them critically. Most probably, the difference is caused by confounding factors such as different measurement instruments, grades, or theoretical frameworks.

2.3 Gender and School Satisfaction

Many Finnish studies present similar results with respect to the difference between boys' and girls' satisfaction with school. Girls seemed to be generally more satisfied with school than boys (Uusikylä & Kansanen, 1988; Scheinin, 1990; Sermilä, 1991; Linnakylä, 1993; Kannas, 1995; Linnakylä *et al.,* 1996; Metsämuuronen, 2006c). Scheinin (1990) stated, on the basis of several references, that differences in self-image between girls and boys were clear when connected to school attendance. According to Metsämuuronen (2006c), the negative change in attitude was statistically significantly smaller in the group of girls (decrease of 6%) than in the group of boys (10%).

Linnakylä and colleagues (Linnakylä *et al.,* 1996) observed that the differences between Finnish girls and boys were the widest in the areas of general satisfaction and teacher-pupil relations during lower secondary school. Fifty-seven percent of the girls and 47% of the boys expressed liking the school. The same proportion was found when assessing the pupil-teacher relationship. However, in the group of most unsatisfied pupils, there were an

equal proportion of boys and girls.

Hautamäki and Hautamäki (1999) explained the difference between girls and boys by referring to the differing qualities of boys and girls. Girls who begin school have greater social sensitivity and obedience, which are fitting qualities to the requirements of school learning. Boys, on the other hand, show a need to express physical strength and activity. Working in a group demands the ability to give help when needed and the ability to receive help. Boys face a clear contradiction as the masculine code emphasizes independence, the status of hierarchy, and justification of aggression. The situation is more contradictory for boys and they have to work constantly to create a balance, which may cause them to be, to some extent, more exposed to problems and to more negative attitudes toward school than girls.

2.4 School Achievement and School Satisfaction

Achievement level has been connected with school satisfaction by many Finnish research studies (Olkinuora, 1983; Uusikylä & Kansanen, 1988; Sermilä, 1991; Linnakylä, 1993; Linnakylä *et al.,* 1996; Rimpelä, 1999). According to Hautamäki *et al.* (1999), in most of the groups where the achievement level was higher than average, the pupils were average or slightly positive in their attitude. On the other hand, the groups with the most positive attitudes rarely show better-than-average skill levels. The researchers stated that, in the classes with low achievement and a simultaneous negative attitude, the negative attitude was responsible for the lack of learning progress. It was concluded that rarely can a group be both highly achievers and extremely satisfied.

2.5 Age and Satisfaction in School

Some studies have described change in school satisfaction as a function of age (Fägerlind & Munck, 1981; Uusikylä & Kansanen, 1988; Kannas, 1995; Metsämuuronen, 2006a; 2006b; 2006c; Stakes, 2007a). All these studies show that pupil contentment declined with age. The results of the Fägerlind & Munck (1981) study are interesting from the viewpoint of reducing the trend of contentment. International data were collected from 4th- and 8th-grade pupils and from 12th-year students. The results revealed that dislike and criticism toward school usually increased with age. This was common in all countries, but especially so in Finland. What is interesting is that the decreasing curve is actually the mirror picture of the letter J. For example, 42% of the 4th graders claimed that school was the best time in their life, whereas only 30% of the 8th grader group and 37% of the last graders of the higher secondary education shared the same opinion. The same result was found regarding the question on school time as being the happiest time of their lives.

Previous studies noted school contentment as being connected to many factors–e.g., gender, achievement, and age. Girls are more satisfied with school than boys; those who succeed better in school are more satisfied; increase in age equates to a decrease in contentment, but this grows once again among later classes once the turning point in contentment is reached. However, does contentment among girls decrease in the same way as it does among boys? In which areas are pupils the most contented and where are they the most discontent? This article searches for answers to these questions.

3. Objectives of the Study

The objectives of this research are to examine

1) how school satisfaction and pupil/student contentment differ in different grades from 1st grade of primary education to 2nd grade of upper secondary education,

2) at which age does the decreasing trend of contentment possibly turn to a more positive track, and

3) how the trends differ according to gender and achievement level.

On the basis of the previous studies, it is expected that 1) the pupils'/students' contentment decreases during the years though it starts to increase at some point, 2) the threshold point of the change in trend will be found somewhere between 8th grade and upper secondary education, and 3) the girls and higher achieving pupils/students will be more content than boys and those with lower achievement level.

4. Methods

4.1 Sample and Design

School satisfaction and pupil/student contentment were measured from 4,696 pupils representing grades 1 to 6 in primary education, 7 to 9 in lower secondary, and 1 to 3 in upper secondary education with a cross-sectional design; all the grades were tested at the same time. Five municipalities were selected to represent different types of providers of education (cities/rural areas and southern/central Finland). Several schools in these municipalities ran the attitude tests (see next section). In the final data set, there were 1,060 pupils from grades 1 to 2; 1,876

pupils from grades 3 to 6; 1,209 pupils from grades 7 to 9; and 530 students from the 1st to 3rd grades of higher secondary education (high school or gymnasium). At the time of inquiry (spring – end of the fiscal year), there were not more than a few (n = 15) 3rd graders in the upper secondary level left in the schools, so their results are not reported. There were 2,331 girls and 2,302 boys in the data. Sixty-three pupils (mostly 1st and 2nd graders) did not indicate their sex.

4.2 Preliminary Item Selection and Test Construction

The original purpose of the testing was to create several test batteries for the schools' self-assessment and to collect information about the item parameters for the item bank of these test batteries. The work conducted by the Finnish National Board of Education (FNBE). This viewpoint, however, is too time-consuming to expand upon here. This background explains why the process described here was laborious. The process of item selection and test construction was as follows.

During the first phase, national research reports, doctoral theses, licentiate theses, master's theses, and pre graduation theses concerning – at least partly – school satisfaction were studied. The main studies were referred to in the literature section of this article. When test batteries were published in these studies, the items were extracted from if they covered relevant topics, namely questions related to attitude. Information about the components of school satisfaction was also collected at the same time. Several themes were found: general satisfaction, school negativity/positivity, attitude towards school, appreciation received in class, self-image, attitude towards school (for example, contentment with equipment, teaching, and recesses), relationships with other pupils, relationships with the teacher and to other personnel of the school, motivation, atmosphere, teaching methods, bullying, the fear of/in the school, and personality factors of the pupil or teacher.

During the second phase, the collected items ($k = 1,060$) were classified roughly with qualitative methods to the components of satisfaction found. Thus, the items were taken from their original context for renewed pretests and later use. Some of the items were obviously the same in different publications and tests so these were discarded.

During the third and final phase, after a series of discussions and rounds of selection, 208 items were selected as appropriate to measure the school satisfaction of pupils from the 3rd grade of primary education to the 3rd grade of the higher secondary education. For the pupils in 1st to 2nd grade, it was appropriate to construct a separate test, where different test techniques and different life experiences and reading skills would be taken care of. These items were chosen independently of the other items, a total of 8 altogether, in addition to the original 208 items.

Attitudes were measured with a four-point Likert-type scale, which was anchored to extremes from "totally disagree" to "totally agree". Extremes in the test of grades 1 to 2 were "Happy/Joyful" and "Sad/Difficult." Visual signs were used for the lower level grades to aid response (Figure 1).
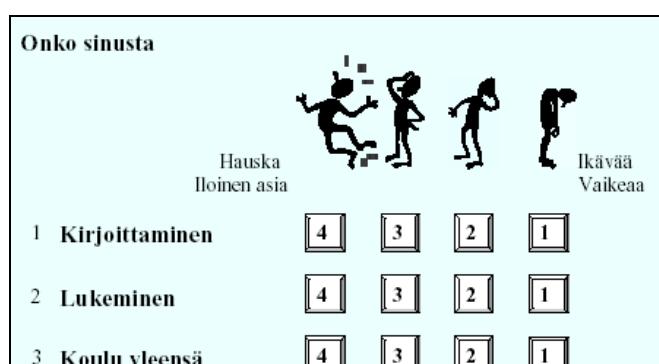

Figure 1. An example of an attitude test for grades 1 and 2

4.3 Matrix Design

The test was run with 15 (= 1 + 7 + 7) versions of 216 items. The matrix sample of items was used (Figure 2) as follows:

1) For grades 1 and 2, separate 13-item tests with five anchoring items to the higher grades was created.

2) For grades 3 to 9, there were seven different test versions. In each version, there were 12 common anchoring items and 28 unique items.

3) For higher secondary education classes 1 to 3, there were also seven different test versions. However, these versions were the same as the versions for grades 3 to 9. The texts were altered to fit the target group (e.g., "pupil" → "student" and "classroom" → "teaching environment"). Otherwise, the versions for the upper secondary students and the lower grades were identical.

1–2		Grades 3–9							Higher secondary education 1–2						
V1-2		V1	V2	V3	V4	V5	V6	V7	V1	V2	V3	V4	V5	V6	V7
■		xx	xx		x				xx	xx		x			
		▢	▢	▢	▢	▢	▢	▢	▢	▢	▢	▢	▢	▢	▢
		■							■						
			■							■					
				■							■				
					■							■			
						■							■		
							■							■	
								■							■

■ = unique items　▢ = anchoring items　　　　x = single anchoring items

Figure 2. The matrix design

4.4 Final Item Selection and Dimensions

In the final item selection, both classical item parameters and one-parametric Item Response Theory (IRT) modeling – that is, Rasch modeling – were used (see Rasch, 1960; Lord & Novick, 1968). Each test version was simultaneously processed with an OPLM program (Verhelst *et al.,* 1995), which produced classical indications for each test version or booklet. Because items reflecting different aspects of school satisfaction were scattered randomly into the tests (excluding the linking items), the classical item- (and test) analysis procedures could not produce the reliabilities of each sub-score or section i.e., ten questions related to bullying could be checked for bullying if all ten were included in the same test but not if spread over many different test versions in groupings of two or three. Instead, the classical reliabilities and item parameters are estimated version-wise. Items were selected for the final test on the basis of classical item discrimination; the item-total correlation was set at *r = 0.40* or higher.

Simultaneously with classical item analysis, the logical shape of the item characteristic curve (ICC) was screened graphically and tested statistically using OPLM. However, selection was made on the basis of graphical judgment and not on the statistical test. When the sample size is large, the statistical test rejects the model too easily because the confidence intervals (CI) of the theoretical ICC become very narrow. Small deviations from the CI are accepted if the item otherwise behaves theoretically decent. Items with illogical behavior (the detected characteristic curve was not ascending or it was even descending) were taken away or if the information function showed them to give less exact information. The item-total correlation of these items is usually low and they would have already been discarded because of low discrimination power.

School satisfaction is examined with five different dimensions. The sub-score of general school satisfaction includes items concerning topics of liking school in the general sense, liking going to school, or feeling nice/comfortable in going to school. The sub-score of the general atmosphere of the school includes items concerning contentment with school atmosphere, pupil evaluation, and class atmosphere. The sub-score of pedagogical learning environment includes items concerning relationships with the teacher, contentment with teaching, and breaks between lessons (recesses). The sub-score of physical learning environment include items concerning contentment with the physical environment of the school, the route/journey to the school, and organizing of school meals. The sub-score of social learning environment includes items concerning contentment with class atmosphere, pupils'/students' self-image, relationships with school mates, and school bullying.

4.5 Equating

The versions of attitude tests were different, first, because of the length (the test for grades 1 and 2 was shorter than the other versions) and second, all the 8 (or 15) test versions contained different items (except for the anchoring items). The logic is somewhat the same as in the PISA studies where more than 10 booklets are administered to pupils and the relevant information is then collected over several versions to find the sub-scores. Prior to making a comparison of test scores meaningful, the scores have to be made comparable. Equating was done by IRT modeling (see Béguin, 2000). The method of equating follows the same procedure as that used in the PISA studies (OECD 2001, 2003, 2007). The advantage of IRT modeling is that it makes it possible to use several test versions and to combine results with different subpopulations into commensurate results.

The scores were transformed into the same (standard score) scale on the basis of characteristics of IRT models that pupils' latent level of attitude (θ) and the "difficulty" (Note 1) level of an item (θ) are identical, when certain preconditions are fulfilled (Wright, 1968). The latent ability level for each pupil can be determined in every version as far as there are so-called linking items connecting the versions. Estimation was made by a one-parametric model (Rasch, 1960) with OPLM program (Verhelst *et al.,* 1995). A brief technical description of the equating process is as follows (see Béguin, 2000, 17–36 for more exhaustive discussion):

1) Define the structure of matrix sampling and fix the "difficulty" level of anchoring items. Each item's parameters of all the 15 versions will be calibrated to the same scale.

2) Use the *conditional maximum likelihood* (CML) procedure to estimate each item's "difficulty" level (parameter).

3) Use the *marginal maximum likelihood* (MML) procedure to estimate each student's latent ability (parameter) over the versions.

4) The latent attitude levels for different grades are combined over the versions.

The results will be described as standardized theta values. In Figures 3 to 7, these particular values of MML-estimation are shown. In some graphs negative values of "school satisfaction" are seen. This is technically possible when there are more negative than positive standardized values in the group.

Differences in the means of the groups were tested with t-tests. They were calculated manually on the basis of the mean, standard deviations, and number of observations in a spreadsheet and adjusted afterward with Bonferroni correction. Cohen's *d* (Cohen 1988) is used as an indicator of effect size.

4.6 Reliability and Validity

The reliabilities of the versions were estimated by Cronbach's alpha (Kuder & Richardson, 1937; Gulliksen, 1950; Cronbach, 1951; see also Hogan *et al*., 2000), which produces the lower bound of the reliability in the case when there are several dimensions built inside all the test versions. However, the reliabilities of final versions 1 to 7 – measured after removal of poor items – were as high as 0.901–0.940. In the version for grades 1 to 2, the reliability remained slightly lower, but it was nonetheless moderate (Table 1). The lower boundary of CI of the alpha (see Feldt, 1965) shows that tests discriminate the pupils/students sufficient enough for reliable conclusions.

Table 1. Reliabilities of the test versions

	Version for grades 1–2	Version 1	Version 2	Version 3	Version 4	Version 5	Version 6	Version 7
Number of pupils/ students	1060	628	591	479	515	526	433	464
Number of items (in the final version)	13	23	28	24	24	23	24	24
Cronbach's	.806	.912	.940	.918	.936	.901	.907	.910
	.788	.901	.933	.907	.928	.889	.894	.898
95% confidential interval for alpha	− .823	− .921	− .946	− .928	− .944	− .913	− .919	− .921

Versions 1–7 were used with grades 3–9; in higher secondary grades, 1–2.

Because items reflecting different dimensions of satisfaction were scattered evenly into different test versions, it is difficult to determine the validity of the separate test versions with traditional computational methods. It is not possible to get coherent covariance or correlation matrix, which could be used to examine the construct validity of the entire test with, for example, structural equation modeling (see e.g., Bollen, 1989). For the same reason, the traditional study of convergence and discriminate validity (see e.g., Trochim, 2002) would, in practice, produce twisted information. Instead, with the method of traditional validity argumentation (see Moss, 1992), it can be stated that the dimensions were found from the literature. According to the classical definition (i.e., Cook & Campbell, 1979), it can be stated with sufficient accuracy that "we are measuring what was supposed to be measured" – school satisfaction.

5. Results

5.1. General School Contentment

General school satisfaction changed radically though logically between different grade levels (Figure 3). Because of the design, it is not known how contentment changes with age when it comes to individual, pupils/students but the results give clear evidence that satisfaction decreases with age (Note 2).

Pupils in grades 1 to 2 are eager and excited about school and going to school. This can be clearly seen in Figure 3. For some reason or another, the excitement of the beginners does not continue for long. Instead, from grade 3 on, general contentment in school decreases year after year, all the way to 8th grade. Possibly, when the turbulent years of teenage life ease and choices for future life start to affect general school contentment and school work, the general satisfaction begins to increase slightly from the 9th grade on. However, the same excitement level as in the lower grades is never reached again. This result, the shape of a turned J, was essentially the same as the result of Fägerlind and Munck (1981). Now, however, it is known more precisely that the actual turning point occurs between grades 8th and 9th. For boys, the turning point seems to come a year later, compared with that for girls. It is surprising how intensively school satisfaction decreased after the first grades. All grades differed significantly from the first and second class contentment average values (with all grades, $t > 8.958$, $p < 0.001$) (Note 3). Effect sizes are large (with all grades, $d > 0.94$).
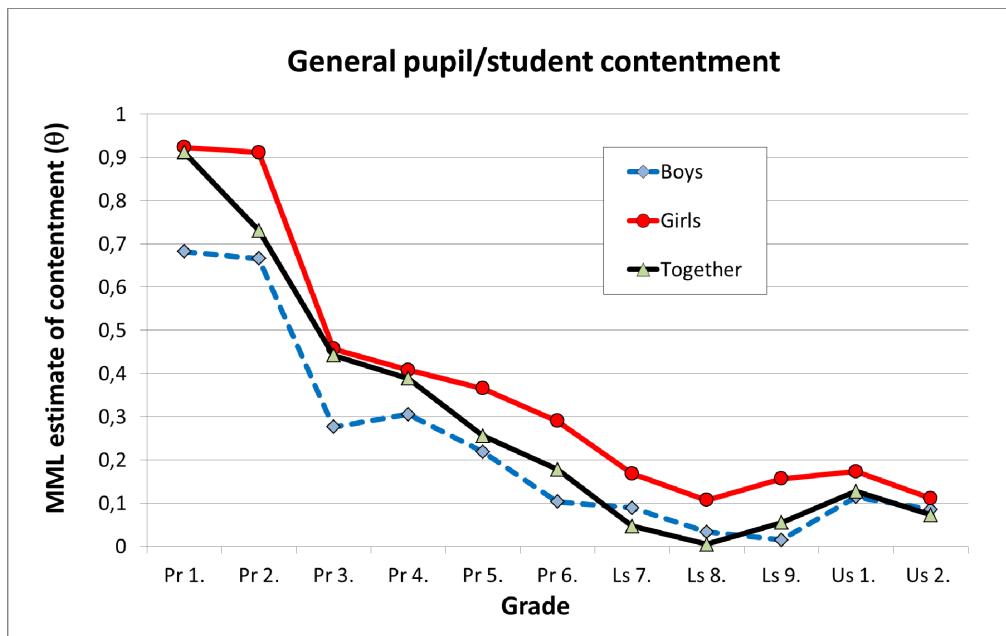

Figure 3. Pupils'/students' general satisfaction with school as a function of age

There are clear differences between genders when it comes to general school satisfaction: girls are thoroughly more content with items of school satisfaction than boys. The difference between girls and boys is statistically significant in grades 1, 2, 3, 5, and 6 ($p < 0.001$). Effect sizes in classes 1 to 3 vary between 0.44 and 0.47, which indicates a moderate effect. The differences between the groups in grades 4 and 9 are also significant ($p = 0.05$

and $p = 0.015$); however, neither of the p values is statistically significant after Bonferroni correction. Thus, differences between girls and boys are significant in all grades other than the 4th, 7th, and 9th grade and in the upper secondary grades. This result is consistent with previous results which show that girls' enjoyment in school is consistently higher than that of boys (Uusikylä & Kansanen, 1988; Scheinin, 1990; Sermilä, 1991; Linnakylä, 1993; Kunnas, 1995; Linnakylä *et al.,* 1996; Metsämuuronen, 2006a; 2006c).

It seems that the more successful one is in school, the higher the general satisfaction toward school is (Table 2). Pupils in grades 1 and 2 were not included in the comparison because they were not asked about their achievement level. Pupils/students, who indicated their average marks to be "poor" (in the Finnish system, teachers' marks of 5 or 6, n = 166) were most unsatisfied about school and about going to school. Those pupils/students (n=1, 849), who had marks indicating average or "good" (marks 7 or 8), were significantly ($p<0.001$) more satisfied with school than those with poorer marks. Most satisfied were pupils (n=865) whose achievement level was "excellent" in terms of school-based assessment (marks of 9 and 10). The difference between "excellently successful" (9 or 10) and "well successful" (7 or 8) is statistically significant ($p < 0.001$). Cohen's d varied between 0.40 and 1.07, indicating high or at least moderate effect sizes. Results are consistent with earlier results of the connection between achievement and school satisfaction (Olkinuora, 1983; Uusitalo & Kansanen, 1988; Sermilä, 1991; Linnankylä, 1993; Linnankylä *et al*., 1996; Stakes, 2007a).

Table 2. General attitude of pupils toward school with different levels of achievement

Achievement level	Mean of MML (Population)	Std. dev. of MML (Population)	Sample size	t value p value Cohen's d	t value p value Cohen's d	t value p value Cohen's d
Marks 5 – 6	-0.065	0.355	166	-9.45	-9.36	-13.82
Marks 7 – 8	0.208	0.374	1849	<0.001	<0.001	<0.001
Marks 9 – 10	0.361	0.407	865	0.733	0.398	1.07

Note 1. Does not include pupils of grades 1 and 2

Note 2. In the Finnish school marking system, 8 is "good" and 4 is "failed". Thus, marks 5 to 6 actually mean "poor," marks 7 to 8 mean "average" or "good," and marks 9 to 10 mean "excellent."

5.2 Contentment with the School Atmosphere

General atmosphere of the school, which included school spirit or -atmosphere, sense of justice in pupil evaluation, and school atmosphere, were factors where all pupils and students felt quite positive compared with the other sub-scores, but it is necessary to say that pupils of the 1st and 2nd grades were substantially more content than pupils/students in the other grades (Figure 4). Generally speaking, school atmosphere seems to be emphasized, especially in the answers of higher secondary education students.

Generally speaking, girls are more content with the school atmosphere, evaluation of achievement, and class atmosphere than boys. After the Bonferroni correction, girls are significantly more content than boys during grades 2 and 6 ($p < 0.001$), 3 ($p = 0.003$), and 5 ($p = 0.009$). In the 9th grade, the difference between girls and boys is not convincing ($p = 0.065$). Differences between other groups are not statistically significant, although with pairwise comparison (without correction), differences show significance in the 4th grade ($p = 0.015$) and 8th grade ($p = 0.037$). Effect sizes varied $0.30 < d < 0.41$, indicating moderate or small effect size.
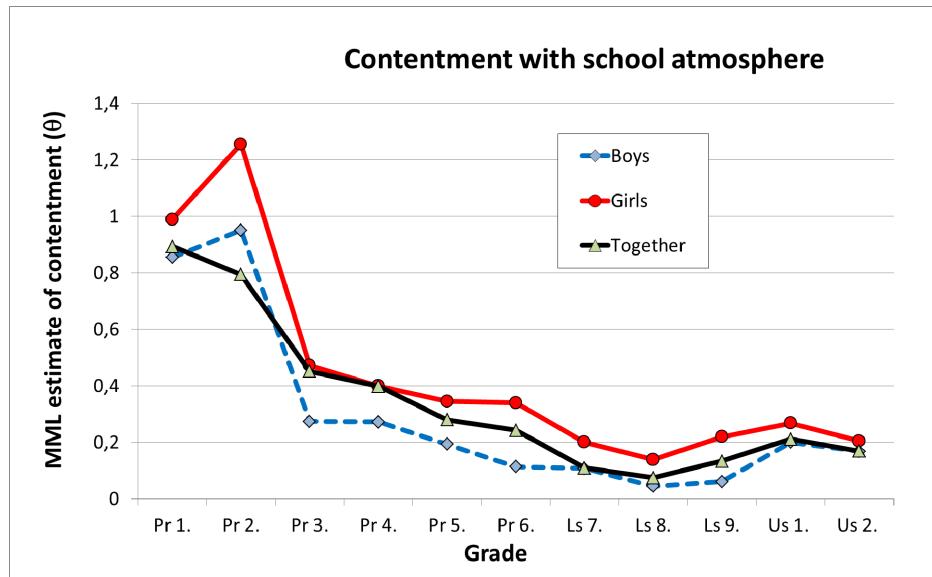
Figure 4. Contentment with school atmosphere, evaluation, and class spirit as a function of age

5.3 Contentment with the Pedagogical Environment

General contentment with school declines with age so that in the 8th grade, it is only 0.5 % (Note 4) and in the higher secondary education 8-14 % from the excitement level of the first grade students. The lowest the contentment is when evaluating the pedagogical learning environment (Figure 4). In grades 7 to 9, the contentment of the boys was even negative. Interestingly, contentment with the teacher was somewhat higher in the second grade than that in the first grade, especially in the group of girls. After the Bonferroni correction, there are statistically significant differences in the contentment of the girls and boys during grades 1 to 6, excluding the group of pupils from 4th grade. Grades 7 to 9 and higher secondary education students did not show statistically significant differences, although girls' attitudes are slightly more positive than boys'. In all the grades, the boys are more discontent with teaching, the teacher, and working in class than girls. The girls are statistically significantly more satisfied with teaching, the teacher, and working in class than boys at 1st ($p = 0.003$), 2nd ($p < 0.001$), 3rd ($p = 0.003$), 5th ($p = 0.015$before correction), and 6th grade ($p = 0.007$). Effect sizes varied $0.29 < d < 0.41$, indicating moderate or low effect size.
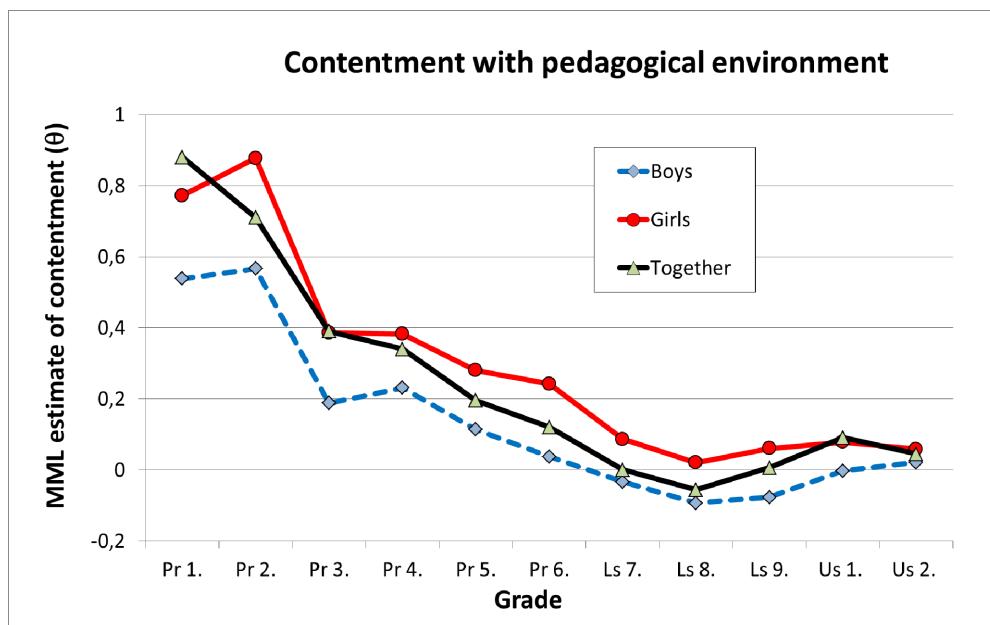

Figure 5. Contentment with pedagogical learning environment at different grades

5.4 Contentment with School's Physical Environment

The physical learning environment of the school (including school buildings, school food, route/journey to school, recess breaks and their length) did not receive a very high mean of contentment compared with general contentment. This was especially true with boys in grades 6–9 and in the 2nd grade of the upper secondary level where dimension was very low (Figure 6).

Girls are generally more content than boys with the physical environment of the school. They are statistically significantly more content than boys in the 1st ($p = 0.002$), 2nd ($p < 0.001$), 3rd ($p = 0.006$), 5th ($p = 0.006$), and 6th grades ($p = 0.0012$). After the Bonferroni correction, the difference between girls and boys in the 7th grade was not convincing ($p = 0.056$). Effect size varied $0.32 < d < 0.37$, indicating moderate or small effect size.
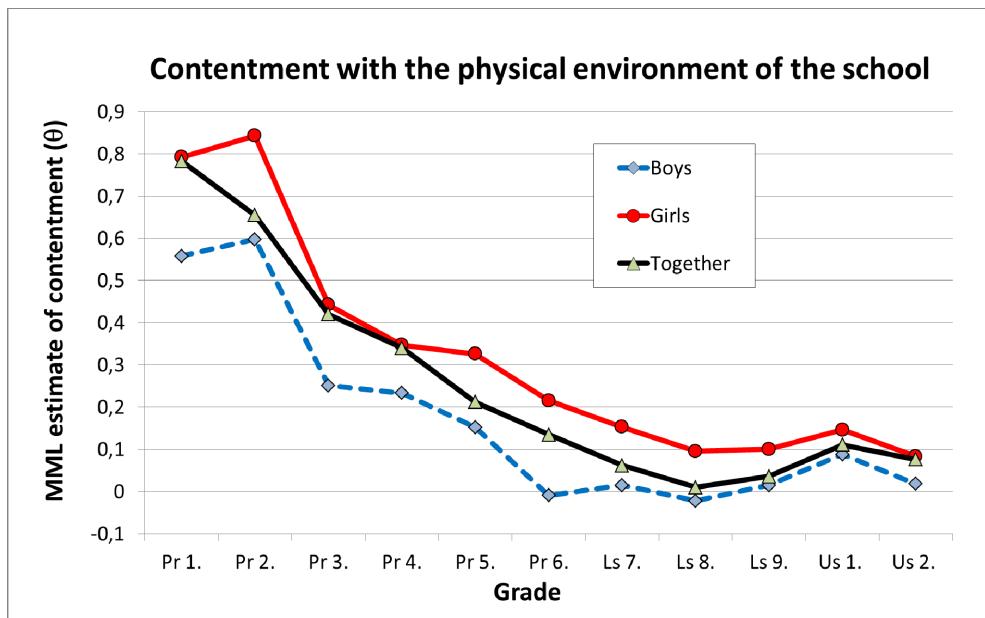

Figure 6. Contentment with the physical environment of the school, maintenance of the school, school distance, and arrangement of dining at different grades

5.5 Contentment with School's Social Learning Environment

What is interesting is that pupils and students are most positive when they evaluate themselves as learners and in relation to other pupils in the dimension of the social learning environment. Self-image, class spirit, and relationships with other pupils were emphasized to be highly satisfying factors in all age groups. This is clearly seen in the answers of higher secondary education students and 9th grade girls. It is still important to notice that the most contented pupils were in the lowest grades and the most discontented were the 7th and 8th graders (Figure 7).

In this area, girls are also more content than boys. After Bonferroni correction, the girls' contentment in relation to this social dimension of school satisfaction is statistically significantly higher than boys' at 1st ($p = 0.004$), 2nd ($p = 0.001$), 3rd ($p = 0.014$), 5th ($p = 0.004$), and 6th grades ($p = 0.001$). There was a difference also in the 9th grade, but the difference was not statistically convincing ($p = 0.055$). Effect size varied $0.31 < d < 0.40$, indicating moderate or small effect size.
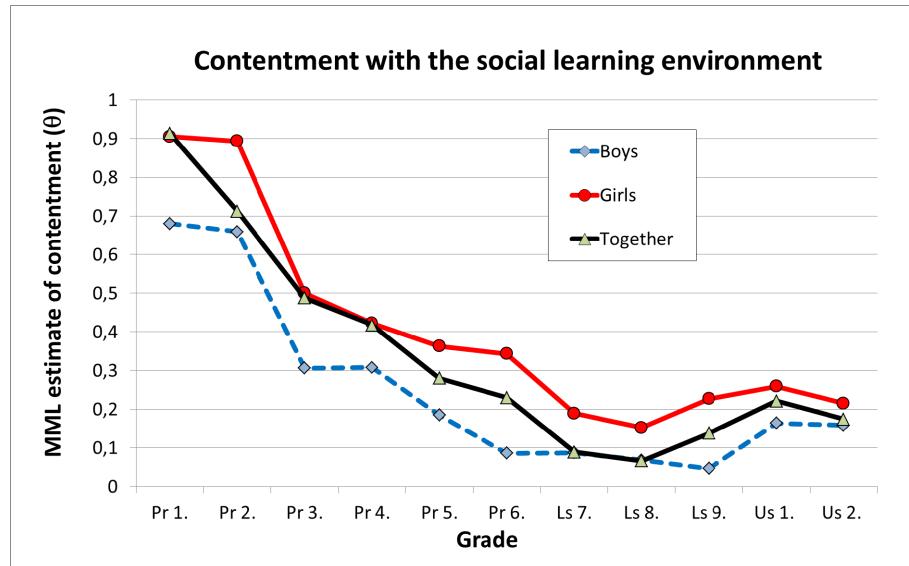
Figure 7. Contentment with the social learning environment at different grades

5.6 Summary of the Results

To summarize, it can be stated that, within all studied dimensions of school satisfaction, contentment declines as a function of age. The decrease in satisfaction after the 1st and 2nd grades is vast and in the 8th grade for girls and in the 9th grade for boys, the trend turns to a more positive track. In some areas (such as contentment with teacher or teaching), the decrease in contentment is severe especially among boys. The higher the pupils'/students' achievement level, the more satisfied they seem to be toward school. There are statistically significant differences in all dimensions between girls and boys, with girls always at a higher level of satisfaction. An interesting deviation from this trend is the pupils of the 4th grade, where there is no statistically significant difference between girls and boys. The differences are moderate in the higher classes. In the higher secondary education, there is not much difference between the girls and the boys, most probably because of selection; only the best pupils gained entry to the upper secondary high schools. However, even after the selection, the girls are more content (though not statistically significantly) than boys.

6. Discussion

School satisfaction and pupil/student contentment have been considered in this article from several perspectives. School satisfaction declines radically after the first grades. This increase of criticism by age is in line with earlier results (Uusikylä & Kansanen, 1988; Kannas, 1995; Metsämuuronen, 2006a; 2006c; Stakes, 2007a). The satisfaction increases slightly after the 9th grade and it follows the shape of a turnaround letter J, which corresponds to the results of Fägerlind and Munck (1981). The turning point of contentment occurs for girls during the 8th grade and for boys during the 9th grade.

Although the differences are not always statistically significant, practically in all grades the girls are more satisfied with school than are boys. This result confirms earlier results (Uusikylä & Kansanen, 1988; Scheinin, 1990; Sermilä, 1991; Linnankylä, 1993; Kannas, 1995; Linnankylä *et al.,* 1996). Generally speaking, pupils are most discontented at the 7th and 8th grades when they evaluate their teacher or lessons. They are most content when they evaluate themselves as learners (self-image) and in relation to other learners, although even then, 8th grade pupils showed the most discontent. The connection between achievement and school satisfaction is also consistent with earlier results: school satisfaction is, generally speaking, higher, the more successful the pupils are (Olkinuora, 1983; Uusikylä & Kansanen, 1988; Sermilä, 1991; Linnankylä, 1993; Linnankylä *et al.,* 1996; Kouluterveys, 2002).

One limitation of the interpretation of the results comes from the fact that though the school contentment evidently declines during the school years, because of the design, nothing much cannot be said of individual students and their development. Another questionable issue is the use of the linking items between the 1st and 2nd graders' test and the test versions. There are not many linking items which may produce some limitation with regard to any extreme inferences on the basis of the data. Technically, however, a couple of linking items are

adequate for equating – even one would have been sufficient.

Pupil/student contentment decreases evidently as a function of age. Contentment with pedagogical environment (with teachers, teaching, teacher-pupil relations, and classroom activities) seems to decline more than the other areas studied. This leads to two kinds of pedagogic implications that are in direct opposition to one another. First, it is possible to think that there is no need to develop teaching, content of classes, and teacher-pupil relationships because it seems that the contentment with school decreases anyhow. It means that whatever is done in the general education, especially at 7th and 8th grades, results will be the same as pupils are in their worst teenage years and they will feel discontent no matter what happens in the classroom.

Second, another way to react to the pedagogical challenges is to think that the decrease in contentment by age brings a special challenge to the development of the pedagogic. It may be relevant to ask whether there is a need to create some kind of new reform pedagogic focusing on teens instead of children. If the answer is yes, then some interesting reference groups would emerge; there are activity groups for teens, such as boy/girl scouts, sport teams, music schools, and churches, where there are many in their early teens who work excitedly in the group. For example, work done by churches with the youth point to the need to keep the early teenagers in the domain of the church by doing activities different from those involving children and older youths. They have arranged many kinds of planned activities aimed for 12- to 14-year-old adolescents. These programmes are usually guided by certain types of active, functional, idea-rich and charismatic male or female (usually a couple) who enjoys and feels it important to work with this age group. In this early teen activity, the main focus is not on intellectual goals, but more on functionalism, participating in activities, and maintaining a safe environment. Of course, this is in contradiction with traditional school teaching, which has as the main focus the intellectual transfer of information, especially when moving from primary education to lower secondary education level. The other contradiction is linked to arranging traditional school teaching in the 7th and 8th grade: psychological development would require something steady in the worst turbulent era of the teenage years. Should there be, at that age, for example, at least a permanent classroom for the pupils?

It may be easy to agree with Metsämuuronen's (1995, 54–59; 1997, 37–38) claim that learning is more effective when it is linked with personal interest and intention. In connection with developing a new kind of pedagogic, the same question could be asked (Metsämuuronen, 1997, 288): what should be done to make school teaching more like a hobby, especially for the pupil groups with the weakest learning results? Would it be reasonable to have more of a club- like style of studying or would it be in the end better if development were viewed from the whole perspective of the pedagogic? Should one change the school set up or alter the entire idea of pedagogy? It is a moot point to state that circumstances have changed, but children in a very real sense are not the same, so is it time to review the pedagogy of learning, to move on from a lecture style to activity based curriculums for example, and incorporate more visual stimuli in delivery using virtual simulations? This new "activity" pedagogic should most probably be based on functionality, play-like action, a hobby-style of teaching, and on being safe together. The most difficult age group would have teachers with specialization on this type of age group. How does the new kind of game-type learning environment challenge the traditional school type of studying (see Manninen, 2004)? How could the non-formal learning environments and pedagogic with its plays and games (Silberman-Keller, 2006; Manninen, 2004; Bekerman & Silberman-Keller, 2004; Frasca, 2003) be used for the benefit of the school? These questions need not be answered in this article – they shall remain for others to ponder about.

References

Arinen, P., & Karjalainen, T. (2007). *PISA 2006 ensituloksia. 15-vuotiaiden koululaisten luonnontieteiden, matematiikan ja lukemisen osaamisesta* [in Finnish]. Opetusministeriö. Retrieved from http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2007/liitteet/opm38.pdf?lang=fi.

Bekerman, Z., & Silberman-Keller, D. (2004). Non-formal pedagogy: Epistemology, rhetoric and practice. *Educational and Society*, *22*(1), 45–63.

Béguin, A. (2000). *Robustness of Equating High-Stake Tests*. Enschede: Febodruk B.V.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: John Wiley & Sons.

Borup, I., & Holstein, B. E. (2007). Schoolchildren who are victims of bullying report benefit from health dialogues with the school health nurse. *Health Education Journal, 66*(1), 58-67. http://dx.doi.org/10.1177/0017896907073787

Borup, I., & Holstein, B. E. (2006). Does poor school satisfaction inhibit positive outcome of health promotion at school? A cross-sectional study of schoolchildren's response to health dialogues with school health nurses.

Journal of Adolescent Health, 38(6), 758-760. http://dx.doi.org/10.1016/j.jadohealth.2005.05.017

Catalano, R. F., Haggerty, K. P., Oesterie, S., Fleming, C. B., & Hawkins, J. D. (2004). The importance of bonding to schools for healthy development: Findings from the social development research group. *The Journal of School Health*, *74*(7), 252-262.

Cohen, J. (2006). Social, emotional, ethical and academic education: Creating a climate for learning, participation in democracy and well-being. *Harvard Educational Review*, *76*(2), 201-237.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cook, T., & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Boston: Houghton-Mifflin Co.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of test. *Psychometrika, 16*, 297-334. http://dx.doi.org/10.1007/BF02310555

Dake, J. A., Price, J. H., Telljohann, S. K., & Funk, J. B. (2003). Teacher perception and practices regarding school bullying prevention. *Journal of School Health, 73*(9), 347-355. http://dx.doi.org/10.1111/j.1746-1561.2003.tb04191.x

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior.* New York: Plenum Press.

Eisenberg, M. E., Nuemark-Sztainer, D., & Perry, C. L. (2003). Peer harassment, school connectedness, and academic achievement. *Journal of School Health, 73*(8), 311-316. http://dx.doi.org/10.1111/j.1746-1561.2003.tb06588.x

Feldt, L. S. (1965). The approximate sampling distribution of Kuder-Richardson reliability coefficient twenty. *Psychometrika, 30*, 357-370. http://dx.doi.org/10.1007/BF02289499

Frasca, G. (2003). Simulation versus narrative: Introduction to ludology. In M. J. P. Wolf & B. Perron (Eds.), *The Video Game Theory Reader* (pp. 221-237). New York: Routledge.

Fägerlind, I., & Munck, I. (1981). *Attitudes towards School and Education in Seven Countries. Comparisons between England, Finland, Hungary, India, Japan, Sweden and the USA Based on IEA Data.* The Yellow Report Series 52. Institute of International Education: Stockholm University.

Garcia-Sanchez, J.-N. & de Caso-Fuertes, A.-M. (2005). Comparison of the Effects on Writing Attitudes and Writing Self-Efficacy of Three Different Training Programs in Students with Learning Disabilities. *International Journal of Educational Research*, *43*(4-5), 272-289. http://dx.doi.org/10.1016/j.ijer.2006.06.006

Gray, J., & Mclellan, R. (2006). A Matter of Attitude? Developing a Profile of Boys' and Girls' Responses to Primary Schooling. *Gender and Education*, *18*(6), 651-672. http://dx.doi.org/10.1080/09540250600980238

Gulliksen, H. (1950). *Theory of Mental Tests* (reprinted 1987). Hillsdale, NJ: Erlbaum.

Hambleton, R. K. (1993). Principles and Selected Applications of Item Response Theory. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed). American Council of Education. Series of Higher Education. Oryx Press.

Hautamäki, A., & Hautamäki, J. (1999). Minä koulussa ja omat koulutavoitteet - sosio-moraalinen minäkäsitys ja usko koulutehtävien suorittamiseen. In J. Hautamäki, P. Arinen, B. Bergholm, A. Hautamäki, S. Kupiainen, J. Kuusela, J. Lehto, M. Niemivirta & P. Scheinin (Eds.), *Oppimaan oppiminen ala-asteilla*. Oppimistulosten arviointi 3/1999. Helsinki: Opetushallitus.

Hautamäki, J., Arinen, P., Bergholm, B., Hautamäki, A., Kupiainen, S., Kuusela, J., Lehto, J., Niemivirta, M., & Scheinin, P. (1999). *Oppimaan oppiminen ala-asteilla.* Helsinki: Opetushallitus [In Finnish].

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability methods: A Note on the Frequency of Use of Various Types. *Educational and Psychological Measurement*, *60*, 523-531. http://dx.doi.org/10.1177/00131640021970691

Holfve-Sabel, M.-A. (2006). A Comparison of Student Attitudes Towards School, Teachers and Peers in Swedish Comprehensive Schools Now and 35 Years Ago. *Educational Research*, *48*(1), 55-75. http://dx.doi.org/10.1080/00131880500498446

Holfve-Sabel, M.-A., & Gustafsson, J.-E. (2005). Attitudes towards School, Teacher, and Classmates at

Classroom and Individual Levels: An Application of Two-Level Confirmatory Factor Analysis. *Scandinavian Journal of Educational Research*, *49*(2), 187-202. http://dx.doi.org/10.1080/00313830500048931

Hurst, M. D. (2005). When it Comes to Bullying, There are No Boundaries. *Education Week, 4*(22), 8.

Kannas, L. (Ed.). (1995). *Koululaisten kokema terveys, hyvinvointi ja kouluviihtyvyys*. Helsinki: Opetushallitus [In Finnish].

Kuder, G. F., & Richardson, M. W. (1937). The theory of Estimation of Test Reliability. *Psychometrika*, *2*, 151-60. http://dx.doi.org/10.1007/BF02288391

Leino, K. (2003). Computer Usage and Reading Literacy. In S. Lie, P. Linnakylä & A. Rue (Eds.), *Northern Lights on PISA* (pp. 71-81). Department of Teacher Education and School Development, University of Oslo, Norway.

Lord, F. M., & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Reading, Mass: Addison-Wesley Publishing Company.

Liinamo, A., & Kannas, L. (1995). Viihdynkö, pärjäänkö, selviänkö turvallisesti: koulunkäynti Oppilaiden kokemana. In L. Kannas (Ed.), *Koululaisten kokema terveys, hyvinvointi ja kouluviihtyvyys* (pp. 109-150). Opetushallitus. Helsinki: HakapainoOy [in Finnish].

Linacre, J. M. (2003). What is Item Response Theory, IRT? A Tentative Taxonomy. *Rasch Measurement Transactions*, *17*(2), 926-927.

Linnakylä, P. (1993). Miten oppilaat viihtyvät peruskoulun yläasteella? Kouluelämän laadun kansallinen ja kansainvälinen arviointi. In V Brunell& P Kupari (Eds.), *Peruskoulu oppimisympäristönä - peruskoulun arviointi 90-tutkimuksen tuloksia.*Jyväskylä: Kasvatustieteiden tutkimuslaitos [in Finnish].

Linnakylä, P., Brunell, V., & Kankaanranta, M. (1996). Miten oppilaat kokevat koulun? Kouluelämän laatu ja sen kehittyminen 90-luvulla. In R. Jakku-Sihvonen, A. Lindström & S. Lipsanen (Eds.), *Toteuttaako peruskoulu tasa-arvoa?*Arviointi 1/1996. Helsinki: Opetushallitus [in Finnish].

Ma, X. (2003). Sense of Belonging to School: Can Schools Make a Difference? *Journal of Educational Research*, *96*, 340-350. http://dx.doi.org/10.1080/00220670309596617

Marraffi, M. (2007). Girls' Attitudes, Self-Expectations, and Performance in Math. An Annotated Bibliography. Retrieved from http://mathforum.org/~sarah/Discussion.Sessions/ biblio.attitudes.html

MacKay, T. (2006). Can hanging children's attitudes to reading bring about long-term gains? A 5 year follow-up study. In T MacKay (Ed.), *The West Dunbartonshire Literacy Initiative: The Design, Implementation and Evaluation of an Intervention Strategy to Raise Achievement and Eradicate Illiteracy. Phase I Research Report*. Dumbarton: West Dunbartonshire Council.

Manninen, T. (2004). *Rich Interaction Model for Game and Virtual Environment Design*. Doctoral thesis, Department of Information Processing Science, University of Oulu. Retrieved from http://herkules.oulu.fi/isbn9514272544/isbn9514272544.pdf

Merisuo-Storm, T. (2007). Pupils' attitudes towards foreign language learning and the development of literacy skills in bilingual education. *Teaching and TeacherEducation, 23*(2), 226-235.

Metsämuuronen, J. (2006a). *Äidinkieli ja kirjallisuus -oppiaineen oppimistulosten ja asenteiden muuttuminen perusopetuksen ylempien luokkien aikana*. Oppimistulostenarviointi 3/2006. Opetushallitus. Helsinki: Yliopistopaino [in Finnish].

Metsämuuronen, J. (2006b). *Förändringar i kunskapnivån i ämnet modersmål och litteratur under de högre årskurserna i den grundläggande utbildningen*. Utvärdering av inlärningsresultat 4/2006. Utbildningsstyrelsen. Helsinki: Yliopistopaino [in Swedish].

Metsämuuronen, J. (2006c). *Oppimistulosten ja asenteiden muuttuminen perusopetuksen ylempien luokkien aikana. Äidinkieli ja kirjallisuus ja modersmåloch litteratur -oppiaineiden näkökulma. Tekninen raportti*. Oppimistulosten arviointi 5/2006. Opetushallitus. Helsinki: Yliopistopaino [in Finnish].

Metsämuuronen, J. (1997). *Omaehtoinen oppiminen ja motiivistruktuurit*. Opetushallitus. Tutkimus 3/1997. Helsinki: Yliopistopaino [in Finnish].

Metsämuuronen, J. (1995). *Harrastukset ja omaehtoinen oppiminen*. Sitoutuminen, motivaatio ja coping. Helsingin yliopiston opettajankoulutuslaitos. Tutkimuksia 146. Vantaa: Tummavuoren Kirjapaino Oy [in

Finnish].

Morrell, P. D., & Lederman, N. G. (1998). Students' attitudes toward school and classroom science: Are they different phenomena? *School Science and Mathematics*, *98*(2), 76-83. http://dx.doi.org/0.1111/j.1949-8594.1998.tb17396.x

Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: implications for performance assessment. *Review of Educational Research,* *62*(3), 229-258. http://dx.doi.org/10.3102/00346543062003229

Nutbeam, B., & Aaro, L. E. (1991). Smoking and pupil attitudes towards school: the implications for health education with young people. *Health Education Research*, *6*(4), 415-421. http://dx.doi.org/10.1093/her/6.4.415

OECD. (2007). PISA 2006 results. Retrieved from http://www.pisa.oecd.org/document/2/ 0,3343,en_32252351_32236191_39718850_1_1_1_1,00.html#ES.

OECD. (2003). *Education at a Glance*. OECD Indicators 2003. OECD: Paris.

OECD. (2001). *Knowledge and Skills for Life*. First results from PISA 2000. OECD: Paris.

OECD. (2000). *Student Engagement at School: A Sense of Belonging and Participation - Publications PISA 2000*. Retrieved from http://www.oecd.org/document/52/0,3343,en_32252351_ 32236159_33688628_1_1_1_1,00.html.

OECD. (1995). *Schools under Scrutiny*. Centre for Educational Research and Innovation. Paris; OECD.

Olkinuora, H. (1983). *Oppimisen ja opiskelun mielekkyydestä peruskoulun päättövaiheessa.Empiirisiä tuloksia.*Turun yliopiston kasvatustieteiden tiedekunta. Julkaisusarja A:91. [in Finnish]

Olkinuora, E., & Mattila, E. (2001). Oppilaan koulutyön mielekkyyden ja koulujen kasvatustoiminnan edellytysten arviointia. In E. Olkinuora & E. Mattila (Eds.), *Miten menee peruskoulussa? Kasvatuksen ja oppimisen edellytysten tarkastelua Turun kouluissa* (pp 17-53). Turku: Turun yliopisto.[in Finnish]

Papanastasiou, C., & Papanastasiou, E .C. (2004). Major Influences on Attitudes toward Science. *Educational Research and Evaluation*, *10*(3), 239-257. http://dx.doi.org/10.1076/edre.10.3.239.30267

Pavloua, V., & Kambourib, M. (2007). Pupils' Attitudes towards Art Teaching in Primary School: An Evaluation Tool. *Studies in Educational Evaluation, 33*(3-4), 282-301. http://dx.doi.org/10.1016/j.stueduc.2007.07.005

Perkins, B. K. (2006). *Where We Learn.The Cube Survey of Urban School Climate*. A project of the Urban Student Achievement Task Force. Council of Urban Boards of Education. National School Boards Association. Retrieved from http://www.nsba.org/site/docs/38100/ 38081.pdf.

Peters, R., De.V., Petrunka, K., Angus, D. E., Arnold, R., Bélanger, J.-M., Boyce, W., Brophy, K., Burke, S. O., Craig, W., Currie, A., Evers, S., Herry, Y., Hoch, J., Irving, L.G., Khan, S., Kocher, P., Mamatis, D., Nelson, G., Pancer, M., Parker, K., Russell, C. C., & Towson, S. (2005). *Do Better Beginnings Lead to Better Futures? Three-Year Follow-up Outcomes for Grade 6 Children and Their Families (FINAL REPORT)*. Kingston, ON: Better Beginnings, Better Futures Research Coordination Unit Technical Report.

PISA (2006). Students Attitudes to Science. Retrieved from http://www.pisa2006.helsinki.fi/ finland_pisa/results/2006/2006_student_attitudes_to_science.htm.

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Danmarks pædagogishe Institut. Studies in MathematicalPsychology I. Nielsen & Lydiche, Copenhagen.

Reiss, M. J. (2004). Students' Attitudes towards Science: a Long-term Perspective. *Canadian Journal of Science, Mathematics and Technology Education*, *4*, 97-109. http://dx.doi.org/10.1080/14926150409556599

Rimpelä, M. (2007). Häiriöitä ja ongelmatilanteita koulussa. In M. Rimpelä, A.-M. Rigoff, J. Kuusela & H. Peltonen (Eds.), *Hyvinvointi ja terveyden edistäminen peruskouluissa - Perusraportti kyselystä 7. - 9. vuosiluokkien kouluille* .Opetushallitus ja Stakes. Vammala: Vammalan kirjapaino Oy [in Finnish].

Rimpelä, M., Rigoff, A.-M., Kuusela, J., & Peltonen, H. (Eds.). (2007). *Hyvinvointi ja terveyden edistäminen peruskouluissa - Perusraportti kyselystä 7.-9. vuosiluokkien kouluille*. Opetushallitus ja Stakes. Vammala: Vammalan kirjapaino Oy [in Finnish].

Rostosky, S. S., Owens, G. P., Zimmerman, R. S., & Riggle, E. D. B. (2003). Associations among Sexual Attraction Status, School Belonging, and Alcohol and Marijuana Use in Rural High School Students. *Journal of Adolescence, 26*(6), 741-751. http://dx.doi.org/10.1016/j.adolescence.2003.09.002

Ruoho, K., Koskela, H., & Pihlainen-Bednarik, K. (2006). Kouluviihtyvyyden käsite ja tutkimus Suomessa. In K. Ruoho (Ed.), *Tehtävänä erityispedagoginen näkökulma*. Tutkimusta erityiskasvatuksen laitoksella. Joensuun yliopisto [in Finnish].

Sainsbury, M., & Schagen, I. (2004). Attitudes to reading at ages nine to eleven. *Journal of Research in Reading*, *27*, 373-386. http://dx.doi.org/10.1111/j.1467-9817.2004.00240.x

Sanchez, B., Colon, Y., & Esparza, P. (2005). The Role of Sense of School Belonging and Gender in the Academic Adjustment of Latino Adolescents. *Journal of Youth and Adolescence*. http://dx.doi.org/10.1007/s10964-005-8950-4

Scheinin, P. (1999). Kouluun liittyvä minäkäsitys ja itsetunto. In J. Hautamäki, P. Arinen, B. Bergholm, A. Hautamäki, S. Kupiainen, J. Kuusela, J. Lehto, M. Niemivirta & P. Scheinin (Eds.), *Oppimaan oppiminen ala-asteilla*. Oppimistulosten arviointi 3/1999. Helsinki: Opetushallitus [in Finnish].

Scheinin, P. (1990). *Oppilaiden minäkäsitys ja itsetunto - vertailututkimus peruskoulussa ja steinerkoulussa.* Helsingin yliopiston opettajankoulutuslaitos, tutkimuksia 77 [in Finnish].

Schnohr, C., & Niclasen, B. (2006). Bullying among Greenlandic school-children: Development since 1994 and Relations to Health and Health Behaviours. *International Journal of Circumpolar Health, 65*(4), 305-312.

Schumacker, R. E. (2005). *Classical Test Analysis*. Retrieved from http://www.appliedmeasurementassociates.com/ama/assets/File/CLASSICAL%2520TEST%2520ANALYSIS.pdf

Schweinle, A., Meyer, D. K., & Turner, J. C. (2006). Striking the Right Balance: Students' Motivation and Affect in Elementary Mathematics. *The Journal of EducationalResearch.99*(5), 271-293.

Sermilä, P. (1991). *Miten koulu toimii - oppilaiden, vanhempien ja opettajien käsityksiä peruskoulusta ja lukiosta.* Helsingin yliopisto, kasvatustieteen tutkimuslaitos. Lisensiaattitutkimus, julkaisematon. [in Finnish]

Shoher, O. (2007). The Iranian Scarecrow. In *Israelenews.com*. Retrieved from http://www.israelenews.com/view.asp?ID=461

Sihvola, S. (2000). *Terveyskeskustelu kouluuntulotarkastuksessa.* Acta Universitatis Tamperensis, 751. Tampere: Tampere University Press. Retrieved from http://acta.uta.fi/pdf/951-44-4826-X.pdf [in Finnish]

Silberman-Keller, D. (2006). The Game of Nonformal Pedagogy. *Educational Practice and Theory*, *28*(1), 37-54.

Siris, K., & Osterman, K. (2004). Interrupting the Cycle of Bullying and Victimization in the Elementary Classroom. *Phi Delta Kappa*, *86*(4), 288-292.

Stakes. (2007a). *Kouluterveyskysely 2007*. Muutokset 1999-2007. Retrieved from http://info.stakes.fi/NR/rdonlyres/0422E3E0-0C92-4971-BFEC-3A93A1F32FFF/0/kokoaineisto2007.ppt#7 [in Finnish]

Stakes. (2007b). Oppimisympäristöjen terveys-ja hyvinvointiseuranta. Retrieved from http://info.stakes.fi/TedBM/FI/kyselyt/oppimisyhteisot.htm [in Finnish].

Trochim, W. M. K. (2002). The Research Methods Knowledge Base, 2nd edition. Retrieved from http://trochim.human.cornell.edu/kb/index.htm (version current as of August 02 2000). Retrieved from especially in http://www.socialresearchmethods.net/kb/convdisc.htm

Uusikylä, K., & Kansanen, P. (1988). *Opetussuunnitelman toteutuminen - oppilaiden tyytyväisyys oppiaineisiin, opetusmuotoihin ja kouluelämään peruskoulun ala-asteella*. Helsingin yliopiston opettajan koulutuslaitos, tutkimuksia 66 [in Finnish].

Wang, Y. (2000). Children's Attitudes toward Reading and their Literacy Development. *Journal of Instructional Psychology*, *27*(2), 120-129.

Wentzel, K. R., & Watkins, D. E. (2002). Peer Relationships and Collaborative Learning as Contexts for Academic Enablers. *School Psychology Review, 31*(3), 366-367.

Verhelst, N. D., Glas, C. A. W., & Verstralen, H. H. F. M. (1995). *One-Parameter Logistic Model OPLM*. Cito, Arnhem.

Wilkins, J. L. M., & Ma, X. (2003). Modeling Change in Student Attitude Toward and Beliefs about Mathematics. *The Journal of Educational Research, 97*(1), 52-63. http://dx.doi.org/10.1080/00220670309596628

Wright, B. D. (1968). Sample-free Test Calibration and Person Measurement. *Proceedings of the 1967 Invitational Conference on testing Problems*. Princeton, NJ: Educational Testing Service.

Note 1. The IRT model has been created for item analysis of achievement items where the parameter is the parameter for "item difficulty" (see, e.g., Hambleton 1993, Linacre 2003, Schumacker 2005). Thus, the parameter tells the "difficulty" level even though in the attitude test, there is not any built-in "difficulty" element. In practice, parameter tells the average attitude level in the item.

Note 2. In Figures 3 to 7, one may note that the line of "together" is not the mean of boys and girls. It refers to a separate analysis where all the respondents are analyzed as a whole. Combining the boys and girls as "students", the item- and person parameters are changed radically.

Note 3. Boys and girls were tested against each other in all the grades. Hence, as outputs, there are 11x5 = 55 t-tests with different degrees of freedoms. To shorten the text, only the main characteristics of tests have been reported here; mainly the p values of the t-tests are shown. In the case, all the values of the t-test statistics are higher than t = 8.958 and thus the notation is unconventionally without the degrees of freedom and equal sign plainly t > 0.958. Also, the corresponding effect sizes (Cohen's d), as many as there are t-test values, lye within a certain boundaries within the family of tests. In the case, all of them are higher than d = 0.94 and thus the notation d > 0.94.

Note 4. From the background of Figure 3, it is known that the latent value for the contentment of the 1st graders was for girls and boys altogether = 0.912 and the parallel value for 8th graders was = 0.005. Therefore 0.005/0.912=0.0055,which equals 0.5%. Correspondingly, at the 1st grade of the upper secondary education, = 0.127 and at the 2nd grade, = 0.073. Thus, 0.127/0.912=0.139, which equals 14 % and 0.073/0.912=0.080, which equals 8%.