University of Helsinki, Department of Teacher Education, Research Report

Laura Tuohilampi

Deepening mathematics related affect research into social and cultural

Decline, measurement and the significance of students' multi-level affect in Finland and Chile

Academic Dissertation to be publicly discussed, by due permission of the faculty of Behavioral sciences at the University of Helsinki in Small hall of the main building, Fabianinkatu 33, 19th of March, 2016, at 10 o'clock.

Helsinki 2016

Reviewers
Professor emerita Gilah Leder, La Trobe University, Australia
Professor Magnus Österholm, Umeå University, Sweden

Supervisors
Professor Markku Hannula, University of Helsinki
Docent Anu Laine, University of Helsinki

Opponent
Associate professor Peter Liljedah, Simon Fraser University, Canada

University press Unigrafia, Helsinki

ISBN 978-951-51-1984-1 (paperback)
ISBN 978-951-51-1985-8 (pdf)
University of Helsinki, Faculty of Behavioural Sciences
Department of Teacher of Education
Research Report 384

Laura Tuohilampi

Deepening mathematics related affect research into social and cultural
Decline, measurement and the significance of students' multi-level affect in Finland and Chile

Abstract
The negative development of mathematics-related affect during comprehensive school years is a worldwide phenomenon. This research elaborates that development, the methods and approaches to measure the development, and some possibilities of making a change in such a development.

In the first study, we examined the longitudinal development of Finnish comprehensive school students affect. We were especially interested in the development of the separate affective components, i.e. cognitive component (referring to beliefs and conceptions), motivational component and emotional component. The study confirmed that Finnish students' affect deteriorates rather dramatically during the comprehensive school years. The timing of the deterioration differs depending on the component: the positivity of the enjoyment of mathematics decreased as early as during primary school years, whereas self-efficacy became more negative during the lower secondary school years. Gender differences were also found, as the deterioration was more dramatic among girls than among boys.

Since the use of traditional questionnaires addressing students' individual level of affect has been criticized by a number of researchers, in the second study we analysed the consequences of imposing methods and concepts from the researcher's own culture onto another. We had two different cultures (Western - individual: Finland, and Latin - collectivist: Chile) as a context. The elaboration was done at the individual level of mathematics-related affect. As a result, we identified most of the dilemmas that are claimed to relate to cross-cultural studies. We argued that traditional cross-cultural studies of mathematics-related affect have contained a risk of making unjustified interpretations.

In the third study, we went deeper into the contextual levels of mathematics-related affect. We examined pupils' affect structures simultaneously at the individual, interindividual (social) and cultural levels. We found out that culture made a difference about what level of affect was more positive. This was interpreted as a sign of the level's different significances in different cultures. This study also confirmed that in cross-cultural comparisons more than only one level of context should be covered.

In the fourth and fifth studies we examined longitudinal affective development of Finnish students on both the individual and the interindividual levels. We also investigated the impact of an intervention, aimed to make mathematics classes more active, on those levels. According to the results, the intervention had an impact on the affective learning environment, i.e. the interindividual level of affect. We also noticed that pupils in Finland very seldom experience their mathematics classes as emotionally engaging, positive or inspiring. We argued that without an emotional bond, mathematics is unlikely to become significant to pupils. This was seen as one possible explanation for the negative development of the individual level of affect. Based on the results, it is suggested that more emphasis should be given to engaging social interaction and to students' opportunities to work actively together.

Keywords: mathematics-related affect, longitudinal development, methodology, intervention

Helsingin yliopisto, Käyttäytymistieteellinen tiedekunta,
Opettajankoulutuslaitos, Tutkimuksia

Laura Tuohilampi

Deepening mathematics related affect research into social and cultural
Decline, measurement and the significance of students' multi-level affect in Finland and Chile

Tiivistelmä
Matematiikkaan kohdistuvien affektiivisten tekijöiden on havaittu kehittyvän negatiivisesti ympäri maailman. Tässä väitöstutkimuksessa tarkastellaan tuota kehitystä, sen mittaamiseen käytettyjä menetelmiä sekä mahdollisuuksia saada aikaan aiempaa positiivisempaa kehitystä.

Tutkimuksen ensimmäisessä osassa arvioitiin suomalaisoppilaiden affektiivisten tekijöiden kehitystä peruskouluvuosien aikana. Erityisenä tarkastelun kohteena oli affektiivisten komponenttien erillinen kehitys: tapahtuisiko negatiivinen kehitys samankaltaisesti oppilaiden matematiikkaan kohdistuvien uskomusten ja käsitysten (affektien kognitiivinen komponentti), siihen kohdistuvan motivaation (motivationaalinen komponentti) sekä tunnetilojen (emotionaalinen komponentti) suhteen? Tutkimus vahvisti yleisen negatiivisen kehityksen olevan varsin merkittävää. Tunnetilat heikkenivät jo pian koulun alkamisen jälkeen alakouluvuosina, itseluottamus puolestaan yläkouluvuosina. Tyttöjen kohdalla negatiivinen kehitys oli poikien kehitystä vakavampaa.

Affektitutkimus on nojannut laajasti kyselylomakkeisiin, joiden väitteet kohdistuvat yksilöiden henkilökohtaisiin affektiivisiin kokemuksiin. Tämä perinne on kohdannut viime aikoina paljon kritiikkiä, koska tällöin mitataan vain yhtä kapeaa affektiivista aluetta. Kritiikkiä on kohdistettu lisäksi siihen, että kyselylomakkeita on käytetty samansisältöisinä kulttuurista toiseen. Tutkimuksen toisessa osassa analysoitiin kyseisenlaisen ulkopuolisesta kulttuurista tuotujen menetelmien ja käsitteiden käytön seurauksia. Kontekstina käytettiin Suomea (länsimainen, individualistinen kulttuuri) sekä Chileä (latinalainen, kollektiivinen kulttuuri). Tuloksena todennettiin useimmat kulttuurienvälisiä eroja tarkasteleviin tutkimuksiiin liitetyt ristiriitaisuudet ja haasteet.

Tutkimuksen kolmannessa osassa analysoitiin kontekstia perinteisiä yksilötason kyselylomaketutkimuksia syvemmin. Affektiivisia tekijöitä tarkasteltiin samanaikaisesti yksilötasolla, yksilöiden välisellä tasolla sekä kulttuurisella tasolla. Kulttuurin huomattiin vaikuttavan siihen, minkä tasoiset affektiiviset tekijät koettiin positiivisimmin. Tämä tulkittiin osoitukseksi siitä, että affektien eri tasot koettiin eri tavoin merkityksellisiksi kulttuurista riippuen. Tutkimuksessa todettiin olennaiseksi käsitellä useampaa kuin yhden tason affektiivisia tekijöitä kulttuurien välisiä eroja käsittelevissä tutkimuksissa.

Tutkimuksen neljännessä ja viidennessä osassa suomalaisoppilaiden affektiivisten tekijöiden pitkittäistä kehittymistä tarkasteltiin sekä yksilötasolla että yksilöiden välisellä tasolla. Lisäksi tarkasteltiin oppilaita aktivoivan intervention vaikutusta pitkittäiskehitykseen. Tulokset osoittivat intervention vaikuttaneen positiivisesti erityisesti yksilöiden väliseen affektiiviseen tasoon. Tämä taso kuvastaa suurelta osin oppilaiden oppimisilmapiiriä. Tutkimuksissa havaittiin myös suomalaisoppilaiden kokevan matematiikan oppituntinsa vain harvoin tunnetasolla sitouttavaksi, positiiviseksi tai innostavaksi. Ilman tunnetason kytköstä oppilaan on vaikea kokea matematiikka itselleen merkitykselliseksi. Tämä nähtiin yhtenä selittävänä tekijänä affektiivisten tekijöiden yleiseen vahvaan negatiiviseen kehitykseen suomalaisoppilaiden keskuudessa. Tulosten pohjalta ehdotetaan suurempaa painotusta sosiaaliselle vuorovaikutukselle sekä aktiiviselle yhdessä tekemiselle matematiikan opetuksessa.

Avainsanat: matematiikkaan kohdistuvat affektiiviset tekijät, pitkittäinen kehitys, metodologia, interventio

Acknowledgements

Back in 2009, I had just completed my master's studies, and I was totally exhausted. The idea of not studying anything anymore in my life felt good. My life was a mess, and I had no idea where to go next. Still, only a week after graduating, starting to feel the relief about the completion, I stopped by the office of our professor, Markku Hannula. He happened to be very supportive and encouraged me to go just a bit further.

That is how it all started. Markku became my supervisor, and I started a convoluted examination in the field. I rambled exhaustingly long drafts about all I found interesting, no matter how far it was from my original interest. Markku did not hold me back. Instead, he encouraged me to go deep into wherever I wanted to go. This turned out to be the most productive thing regarding my research. Only after having made examinations based on nothing but my curiosity, I started to realize what I was actually looking for and why. During this journey, Markku has been of extreme importance both as professor Hannula and Markku himself as a person. As a professor so highly appreciated in the field, almost a celebrity with his sharp thinking and the long ponytail, he has given me the most critical feedback and perspective with my research, still allowing me to find my own way of thinking. He has also shown inspiring example of social networking – a skill which cannot be overvalued. As Markku he has supported and helped me in the endless row of obstacles that were on my way during these years.

From the beginning I relied on another person as well. Docent Anu Laine became my mentor, and very soon my second supervisor. With Anu we have had such a number of discussions about how to go over the big challenges I encountered as I tried to combine family, research and personal growth. Anu, a researcher and a mother, but in particular a person with the greatest sense of black humor, has given me courage and hope that will carry me from now on as a researcher and as a human being.

Besides my supervisors, I want to thank all my research fellows in Finland and in all around the world. I have been lucky to meet incredible people, and many of them have become my best friends and collaborators. There are other friends that may not be researchers in my field or researchers at all, but yet have given me support and new ways to see things. I thank you all for being such great people and for giving richness into my life. I also thank my family; my parents, my sisters and my brothers. Kaisa, I wouldn't be here without your enormous amount of babysitting. Besides your dedication, you have taught me determination and acceptance - you are no less than the greatest Kaisa in the whole world. Mum, dad, would you ever have believed this? I know you always believed in me, but I started from nothing, without a single academic nearby, and I guess no one could imagine how far your trust would take me. And this is just the beginning. Mimosa, Frank, Touko and Hilda - my sweet little darlings - thank you for your endless love, and thank you for letting me to be the kind of person I naturally am. Finally, I would like to thank Ilari for being there when I started this challenging journey. It was not easy, and the costs were not small, but it was something which made that person from 2009 to become me.

Helsinki, 24.02.2016
Laura Tuohilampi

Contents

List of original publications

1. Tuohilampi, L., Hannula, M. S., Laine, A. & Metsämuuronen, J. (2014) Examining Mathematics-Related Affect and Its Development During Comprehensive School Years in Finland. In C. Nicol, S. Oesterle, P. Liljedahl, & D. Allan (Eds.), *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and 36th Conference of the North American Chapter of the Psychology of Mathematics Education, Vol (5).* (pp. 281-288). Vancouver, Canada: PME.

2. Tuohilampi, L., Hannula, M. S., Varas, L., Giaconi, V., Laine, A., Näveri, L. M. & Saló i Nevado, L. (2014). Challenging Western Approach to Cultural Comparisons: Young Pupils' Affective Structures Regarding Mathematics in Finland and Chile. *International Journal of Science and Mathematics Education*. Doi: 10.1007/s10763-014-9562-9

3. Tuohilampi, L., Laine, A., Hannula, M., & Varas, L. (2015). A comparative study of Finland and Chile: The culture-dependent significance of the individual and the interindividual level of mathematics-related affect. *International Journal of Science and Mathematics Education.* Doi: 10.1007/s10763-015-9639-0

4. Tuohilampi, L., Näveri, L., & Laine, A. (2015). The restricted yet crucial impact of an intervention on pupil's mathematics-related affect. Accepted to: *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education*. Prague, Czech Republic: ERME.

5. Tuohilampi, L., Laine, A., & Hannula, M. S. (2015). Longitudinal development of pupils' individual affect and the affective learning environment of their mathematics classes. Submitted to: *Educational Studies in Mathematics.*

1 Introduction

When children start going to school, they tend to have very positive affect: they feel good about their skills, enjoy what they are working with, and set high learning goals which they believe they can reach. This is a normative phenomenon, discussed from a developmental psychology perspective by Harter (1999) in her thorough book on self-construction. Overly positive affect in childhood seems to be a specific developmental stage, linked to attachment security, social competence, creative thinking and self-regulation (for a review of the positive affect in childhood, see Davis & Suveg, 2015). During the school years the situation changes and students' affect typically become somewhat less positive in order to contribute to a realistic self-concept.

With the case of mathematics, this development looks different. It seems like during school years, mathematics-related affect becomes unnecessarily negative and even harmful for learning. This is seen in a number of studies showing that students graduate with an undesired affect towards mathematics: in large-scale studies by Lee (2009) and by Sjøberg and Schreiner (2010) the international development has been examined, with the discovery that students all over the world most likely end up with a negative affect towards mathematics. Many students feel anxious about mathematics; they might have experienced mathematics as boring, unpleasant or difficult; otherwise, they might have low confidence regarding their mathematical abilities. In Finland the same problem exists despite the internationally high performance level, and there has not been any significant improvement so far. In a recent study by Hirvonen (2012) Finnish students who were finishing comprehensive school expressed fairly low self-competence in mathematics, and especially students' emotions towards mathematics were negative across the performance level.

There has been extensive research about this phenomenon worldwide. Several instruments to measure mathematics-related affect and its components have been developed, obtaining more and more information about the situation. Recently, there has been a growing interest in using mixed methods and multiple approaches. One of the lately emphasized approaches is the social level of affect, referred to as the interindividual level in this thesis (for a meta-analysis about school climate, see Thapa, Cohen, Guffey and Higgins-D' Alessandro, 2013). Do students collaborate and discuss their ideas? Would that make mathematics learning more meaningful? In what kind of affective learning environment do students usually work with mathematics? If it is difficult to improve student's individual level affect directly, would it help to create learning environments that are more supportive, allowing trial and error, or a wider learning community that constructs mathematical thinking collaboratively?

In this thesis, we present five studies aiming to find out more about the development of mathematics-related affect and about the challenges and the possibilities in measuring it. We started with individual-level research: there was a possibility to analyse a very special longitudinal data covering the whole comprehensive school years that was representative of all Finnish students. From that data the development of the affect and its factors was examined in detail. The four other studies were made within a Finnish-Chilean research collaboration implemented in both countries, where the aim was to develop more activating mathematics classrooms through a three-year intervention. Regarding that data, the first thing was to examine what should be considered when measuring affect in different cultures. We went on to the interindividual level, thus addressing the new social approach in the field – we used students' drawings to analyse what they felt about their mathematics classes. The last two studies reported the results of the intervention in Finland, considering the development of both the individual affect and the interindividual affect of the students. Altogether, two studies out of five operated on the cultural level. Thus, we went deeper into the context, from the social approach into a broader cultural level. The five studies contribute to our understanding about mathematics-related affective levels, the affective level's cultural significance, and the difficulties that can particularly cause the problem of deteriorating mathematics-related affect.

2 Theoretical and practical approaches to mathematics-related affect

The deterioration of mathematics-related affect is a serious problem worldwide. Even though the phenomenon has been studied extensively, the problem still exists, and no philosopher's stone has been found. If we look at the studies made in the field so far, we can see some deficiencies that might explain the resilience of the problem. Over several years studies have almost systematically focused on measuring students' individual-level (personal) affect. In Chamberlin's (2010) study, he provides an overview of the instruments created to assess affect in mathematics. Chamberlin appreciates the evolution of the instruments and concludes that during the development of the questionnaires researchers have become more accurate regarding the concepts as well. However, by taking a closer look at the instruments which Chamberlin presents we can see that the emphasis in the field has been in using paper and pencil questionnaires addressing students' individual level affective factors (traits), such as anxiety, enjoyment and self-efficacy. Many of the instruments are based on the Fennema-Sherman scale (Fennema & Sherman, 1976), which was developed back in the 1970's. In those days, researchers were specifically interested in mathematics anxiety. It would seem that the Fennema-Sherman scale and the interest towards mathematics anxiety are historical starting points or at least cornerstones to the domain, and it is of little surprise that the following studies have followed the Fennema-Sherman line of enquiry most closely.

Has the use of such instruments helped solving the problem of negative mathematics related affect? In 1997, Ma & Kishor made a meta-analysis of studies that have examined the relationship between attitude towards mathematics and achievement in mathematics. In that meta-analysis, they found out that the effect size between the two concepts was relatively small. This indicated that mathematics-related affect is not as deeply connected with performance as was previously thought. Since then developments have taken place in several directions. In this thesis, the following directions are of particular interest: towards more social approach (e.g. Yackel & Cobb, 1996), towards more emotional approach (Pekrun, 2006), towards more culturally sensitive approach (Clarke, 2013) and towards more theoretically accurate approach (Hannula, 2011; 2012).

2.1 Concepts

In a profound work by McLeod (1992), the theoretical approaches regarding mathematics-related affect were widely discussed. In his conclusion, McLeod structured the area into emotions, beliefs, and attitudes. The idea was that the experiences the students face when involved with mathematics learning cause emotional (physiological) reactions. Over time, these reactions develop into a more stable attitude, namely presumptions or expectations about coming mathematics-related experiences. For example, if a student constantly has problems with certain types of mathematics tasks, let us say, with applied problems, the emotional reactions she/he keeps on experiencing reify into a more constant attitude about applied problems. The term attitudes, however, has rarely been adequately defined in studies (Di Martino & Zan, 2015), and it has been used to refer to a variety of concepts. On the other hand, it has been suggested that attitudes are on a higher hierarchy level than other affective factors; Tapia & Marsh (2004; 2005), for example, see attitudes consisting of emotions and feelings that further include value, self-beliefs, enjoyment, and motivation. In addition to this confusion about attitudes, McLeod's view has been criticized for its failure to take into account the motivational factor of affect separately. In 2011, Hannula reconstructed the theoretical model of mathematics-related affect suggesting that affect was constructed of a cognitive component (beliefs, conceptions), an emotional component and a motivational component, all three components having manifestations in both a constancy dimension and a contextual dimension.

2.2 Hannula's model

Multiple constructs have been used in the mathematics-related affect domain as well as in the general affect domain. McLeod (1992) called these multiple concepts "minitheories about parts of the affective domain" (p. 583). In his review, he discussed several concepts, such as confidence, self-concept, self-efficacy and mathematics anxiety. Hannula's (2011; 2012) model serves as a collecting frame to approach the various concepts.

There are three dimensions in Hannula's model (2011, p. 46). The first dimension includes the components: cognitive, emotional and motivational. The different affective concepts used in the field can be placed under these components (see Table 1).

Table 1. A synthesis of Hannula's three-dimensional theoretical model

Level	*Component*	*Cognitive*	*Emotional*	*Motivational*
Psychological	Trait	Beliefs, conceptions	Emotional disposition	Needs, values, desires
	State	Momentarily thoughts	Momentarily feelings	Active goals
	Example of the concept	Self-competence	Anxiety	Mastery goal orientation
Physiological	*Component*	*Cognitive*	*Emotional*	*Motivational*
	Trait	Brain structure, neural connections		
	State	Neural activation, physiological adaptations		
	Example of the concept	-		
Social	*Component*	*Cognitive*	*Emotional*	*Motivational*
	Trait	Norms	Emotional climate	Shared values
	State	Communication	Classroom "weather"	Motivational choices of actions
	Example of the concept	Discourse	Affective learning environment	Competitiveness

The cognitive component includes much of what in earlier mathematics-related affect research has been discussed under the concept of beliefs (see e.g. McLeod, 1992; Pehkonen, 1994). The emotional component is defined by Hannula (2011) as "typical emotional reactions to typical situations in the mathematics classroom" (p. 45), referring to the trait aspect: the reactions that typically arrive. The motivational component refers to conative elements that in general reflect personal preferences and choices. Several theoretical approaches have been used in motivation research (see e.g. Zhu & Leung, 2011), but in this thesis we use the model of Midgley et al. (1998; 2000), who examine goal orientations. Regarding goal orientations, there is a distinction between mastery goal orientation, in which the student wants to learn because of inner, non-profit reasons, and performance goal orientation, in which the student wants to learn because of outside rewards (to gain a degree, to gain social acceptance, etc.). Some studies have also included an avoidance goal orientation, where the student does not want to make a mistake in public.

The trait and state aspects in Hannula's model refer to the constancy of affective components. Constancy is seen in rapidly changing contextual situations: for example, in problem solving the changing emotions and beliefs influence the critical choices that determine whether a problem is solved or not. This would be the affective state. The affective trait, on the other hand, is a more stable pattern affecting how an individual feels and thinks in different contexts and about different situations.

The third dimension in Hannula's model refers to context. The psychological and physiological dimensions seem to refer to individual-level phenomena, while the social dimension refers to interindividual-level phenomena. Even further, there is one more element established in affect research: the intensity of the affective elements, connecting with the degree of stability in response. This has been discussed by Evans (2006) and McLeod (1992), but it is not explicitly present in Hannula's model. In the model, we could consider that the cognitive component is the most stable and least intense, as beliefs are reported to be fairly resilient. The emotional component could be considered the most intense. However, emotions can be very constant and robust, but not necessarily intensive (like boredom). On the other hand, spontaneous emotional reactions can be very intense, but less constant (like anger).

2.3 Combining theory to what has been measured

In Chamberlin's (2010) study about the instruments used in the field, it was seen that mathematics-related affective aspects have largely been measured using questionnaires solely addressing individual affective components. In Chamberlin (ibid.), we can see that the most influential questionnaires cover two elements of affect, both operating on the individual level. The first element is the components (cognitive, emotional, affective, spread into concepts like anxiety, aspirations, attitude, interest, locus of control, self-efficacy, self-esteem and value). The second element is the characteristics, such as target, intensity, and direction. The questionnaires target mathematics, the self in mathematics, or the social context in mathematics (see also Op 't Eynde, de Corte & Verschaffel, 2002). Intensity can be seen as the power the affective components have on the individual (see Evans, 2006). Direction, on the other hand, seems to refer to the positive or negative orientation of affect. For the theory section of study 2 of this thesis, we created a visualization combining Hannula's model with the characteristics that have been emphasized in the measurements (Figure 1).

Psychological, social, physiological levels			
	Cognition	Motivation	Emotion
State	Thoughts in mind	Active goals	Emotional state
Trait	Concepts, facts, scripts, etc.	Needs, values, desires, motivational orientations	Emotional dispositions

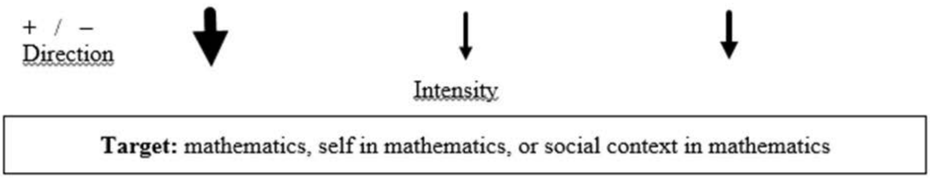

Figure 1. Hannula's (2011; 2012) model integrated with the affective characteristics discussed in Chamberlin (2010)

2.4 Contextual levels

Historically, mathematics-related affect research has focused extensively on individuals. The individual experiences may have been interpreted with the help of contextual variables, such as family expectations, parents' educational levels, or school culture. However, affect as a phenomenon can be studied on other levels as well. Bronfenbrenner (1993) approaches context by dividing it into four levels: (1) the micro level including an individual's interaction with other individuals, (2) the meso level including groups of microsystems interacting with each other (e.g. classes interacting with teachers), (3) the exo level including linkages and processes between social institutions (e.g. the school system, government, family structures), and (4) the macro level including the interaction of all three lower levels. Adapting Bronfenbrenner's division of mathematics education, we can discuss the students individual affect structures, which would represent the micro level. The interaction within students, or between teacher and students, as well as the expectations the students face from their surrounding represent the meso level. When it comes to the exo level, we can talk about the educational systems and politics, such as school hours, teachers' educational requirements and the proportion of public schools. The interactions between the three lower hierarchy levels refer to the macro level, encompassing the overall culture.

The contextual levels were present and visualized in the summary of the affective working group of the 5th Congress of European Research of Mathematics Education (CERME) (Hannula, Op 't Eynde, Schlöglmann & Wedege, 2007, p. 41). In that summary, the structure of the affective domain included three contextual levels (see Figure 2). The three levels resemble Bronfenbrenner's levels: Student / teacher the micro level, Classroom context the meso level, and socio-historical context the exo and macro levels.

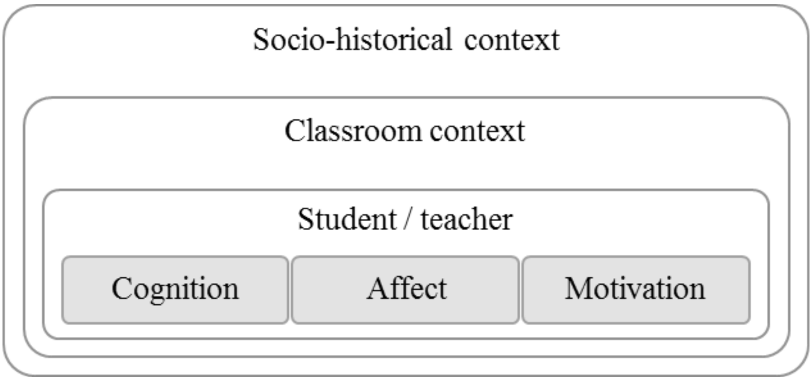

Figure 2. The contextual structure of affect domain

2.5 Social turn

Recently, research on the higher levels of context has been systematically growing, addressing for example school climate (Thapa, Cohen, Guffey and Higgins-D' Alessandro, 2013). In the domain of mathematics education, a decade ago Lerman (2000) discussed the "social turn" in mathematics education. In Thapa and others (2013), the 206 reviewed studies addressed five dimensions of school climate: safety, relationships, teaching and learning, the institutional environment, and the school improvement process.

Evans, Harvey, Buckley and Yan (2009) define the social atmosphere of classrooms through three complementary components: 1) academic, referring to pedagogical and curricular elements of the learning environment, 2) management, referring to discipline styles for maintaining order, and 3) emotional, the affective interactions within the classroom. In this thesis we are interested in the third dimension, as it discusses the affective learning environment of the classes (the interindividual level).

Hannula (2012) discusses the mathematics education research concerning social systems, indicating that the characteristics of learning communities are powerful predictors of students' academic success. He recognizes that the classroom goal structure has been researched within the achievement goal theory, and reports that mastery goal structures have a positive effect on interindividual relations, whereas the performance goal structure has negative or non-significant effects on different types of interindividual relations. Another research area Hannula (ibid.) discusses is the classroom microculture of teacher-student interactions. With respect to mathematical affect, there are studies which analyse social and sociomathematical norms and how these are established in the classroom. Hannula (ibid.) also states that studies on classroom affect and motivation often emphasize the teacher's role in the establishment of classroom discourse and motivational orientation, paying less attention to the students' role in the establishment of the classroom climate. However, students' participation was included in a study by Turner, Midgley, Meyer, Gheen, Anderman, Kang and Patrick (2002), which examined how classroom goal structure and discourse affected students' use of avoidance strategies. They found that there were significantly fewer avoidance strategies in classrooms perceived as emphasizing learning, understanding, effort, and enjoyment. Reyes, Brackett, Rivers, White, and Salovey (2012) also studied the emotional climate of classrooms, finding direct and indirect links with academic achievement. Being impacted by the students, the teacher, or both, the affective learning environment would appear to be a significant element in learning and wellbeing. In sum, Thapa, Cohen, Guffey and Higgins-D' Alessandro (2013) report that school climate has been shown to affect middle school students' self-esteem, mitigate the negative effects of self-criticism, affect a wide range of emotional and mental health outcomes, and have a positive correlation with self-concept.

2.6 Normative affective development

According to developmental psychology, self-construction is influenced by social and contextual factors. The social surroundings play a role in dictating what experiences are most salient and what experiences are encoded (Harter, 1999). During childhood children start to make self-evaluations by comparing themselves with others. At the beginning, children feel very positive about her/himself, having almost an omnipotent view of her/his potential. Little by little this self-construction becomes more accurate, allowing both positive and negative evaluations.

The developmental view presented by Harter (1999) is in line with Op 't Eynde, de Corte, and Verschaffel (2002), who argue that affect comes from what is "first told". This means that if there is nothing that contradicts with some given information (true or false), children tend to take it as true. Only when a contradiction appears do children have a reason to evaluate a former affect, as well as given information in the light of a former affect. Thus, as children get older, it is normal for the affect to become lower over time (i.e. more realistic) because of the child's development. The experiences that children get in schools are part of such a developmental progress, contributing to realistic self-construction. By interacting with peers children start to evaluate their skills and appearance according to the reactions of others. This developmental phase places itself into early school years: from Table 2, you can find Harter's presentation about the normative developmental steps in childhood, and at what age they appear (see Harter, 1999, p. 36).

Table 2. Harter's (1999) developmental stages of self-construction

Age period	Valence / accuracy	Nature of comparisons	Sensitivity to others
Very early childhood (3-4 years old)	Unrealistically positive; inability to distinguish real from ideal selves	No direct comparisons	Anticipation of adult reactions; rudimentary appreciation of whether one is meeting others' external standards
Early to middle childhood (5-7 years old)	Typically positive; inaccuracies persist	Temporal comparisons with self when younger; comparisons with age-mates to determine fairness	Recognition that others are evaluating the self; initial introjection of others' opinions: others' standards becoming self-guides in regulation of behavior
Middle to late childhood (8-11 years old)	Both positive and negative evaluations; greater accuracy	Social comparison for purpose of self-evaluation	Internalization of others' opinions and standards, which come to function as self-guides

2.7 Affective development and mathematics

Regarding mathematics-related affect, it would seem that students do not generate a realistic self, but rather a deconstructive one. In Lee's (2009) study, there were several countries or regions (like Japan, Korea and Hong Kong), where students' mathematics performance level is very high, but they still obtain very low or even negative math self-concept and math self-efficacy: this would not appear to be realistic. In Finland the students' learning achievements in mathematics have also been very good (Mullis, Martin, Foy & Arora, 2012; OECD 2010). Still, in the Trends in International Mathematics and Science Study (TIMMS) from 2011, only one third of Finnish 4th graders' (compared to the international mean of 48%) and only one tenth of 8th graders' (compared to the international mean of 26 %) emotions towards mathematics were positive (Mullis et al, 2012). Many studies have confirmed that teenagers in general tend not to find doing mathematics pleasurable (e.g. McLeod, 1992; Metsämuuronen, 2010).

2.8 Affective development and gender

Studies have produced very consistent results that across age and performance levels, female students tend to have lower self-confidence in mathematics than male students (e.g. Syzmanowics & Furham, 2011; Hannula, Maijala, Pehkonen & Nurmi, 2005; Leder, 1995). In Finland, a similar difference between the genders has been found: even though there are no differences in achievement, girls experience poorer self-efficacy than boys (Hirvonen, 2012). This was also the case back in McLeod's (1992) days. McLeod (ibid.) summarized the main differences between the genders like this: males are more likely to explain their success by ability, and their failure by lack of effort, whereas females are more likely to explain their success by effort, and their failure by lack of ability. In sum, females would be more likely to explain their success by something they can have control over, and failure by things they cannot control. Males, on the other hand, would be more likely to explain their success by their (natural born) ability, which they have a little control over, and failure by lack of effort, something they can control. McLeod (ibid.) also remarked that there are studies which indicate that females are more sensitive to contextual impact, and that especially parents' social norms are often reflected particularly in females.

3 Cross-cultural comparison

Some instruments, such as the Fennema-Sherman scale have been particularly influential in the mathematics-related affect domain. Even large-scale comparative studies, like the Programme for International Student Assessment (PISA) by the Organization for Economic Cooperation and Development (OECD) and TIMSS surveys use a scale that is based on Fennema-Sherman. However, Metsämuuronen (2012) has shown that in PISA and TIMMS studies the Fennema-Sherman scale has led to less reliable scores (especially regarding maths self-concept) in East-Asia, the Middle-East and Europe, compared to North America. According to Metsämuuronen (ibid.), this is due to cultural reasons as well as the instruments' poor adaptation to the range of students' performance levels.

Like the Fennema-Sherman scale, which originates from the States, other dominant instruments have also been developed by Western researchers (Chamberlin, 2010). To use such an instrument with cultures other than Western is an "imposed-etic approach" wherein researchers impose their own perspective on the studied culture (Berry, 1989). According to Berry (1989), there are three approaches to cultural comparisons: the imposed-etic approach, the emic approach and the derived-etic approach. Having an insider's view on the examined topic takes the imposed-etic approach to the next level, i.e. to the emic level. The most desired stage according to Berry (ibid.) is the third stage, the derived-etic stage, which includes comparing features that are common to both (or all) cultures that have been derived from emic research in each culture.

Clarke (2015) points out that cultural differences may invalidate the use of the imposed-etic approach in the field of mathematics education research. In his earlier study, Clarke (2013) discussed the challenges of cross-cultural comparisons, suggesting a validity-comparability compromise. Clarke introduced seven dilemmas: 1) Cultural specificity of cross-cultural codes, 2) Inclusive vs. distinctive, 3) Evaluative criteria, 4) Form vs. function, 5) Linguistic preclusion, 6) Omission, and 7) Disconnection (see Clarke's article for an elaboration of all the dilemmas). We identified four out of seven dilemmas to connect with the imposed-etic approach. First, the dilemma of inclusiveness vs. distinctiveness concerns the presumption that all cultures share similar and universal categories concerning the subjects under investigation and that the content of the categories are easily recognizable and similarly interpreted. Second, the dilemma of evaluation criteria points out that different cultures value things differently. This leads to evaluations that emphasize things differently. In comparative measurements, non-fulfillment of standards is typically interpreted as a sign of failure, but this dilemma points out that "the failure" can be just a sign of those things not have been valued that much (instead of something else that was not addressed in the comparison). Third, Clarke's dilemma of linguistic preclusion reminds us that by using the imposed-etic approach, one might end up comparing variously interpreted or non-existent terminology. Andrews & Diego-Mantecón (2014) have also recently criticized such a presumption about universal congruence between languages and conceptions. Finally, Clarke presented the dilemma of omission, that is, when a researcher is unable to include all that is essential when attempting to research another culture.

In the second study of this thesis, the four dilemmas by Clarke that we connected with the imposed-etic approach were modified into premises that have been used without clear justification. We considered the dilemma of inclusiveness vs. distinctiveness a premise of universalism within the examined components, the dilemma of evaluation criteria a premise of universalism within the components' significances, the dilemma of linguistic preclusion a premise of congruence between languages and conceptions, and the dilemma of omission a premise of components' coverage among the phenomena.

Extensive conclusions are drawn based on tests that may have been biased due to their imposed-etic approach. Consequently mathematics performance has been perceived to be a cultural ability: i.e. that there are high-achieving countries (or regions), such as southern China, Korea, Finland, Hong Kong, and Singapore, and low-achieving countries, such as Kyrgyzstan, Peru, Panama, Chad, and Albania. However, the differences in the measured performance level may also be a result of the dilemmas. This confusion has been referred to by some researchers as a Western bias in cultural comparisons (Hofstede & Hofstede, 2005). In this thesis we consider that as long as the dominant culture in the world is Western-based, the other cultures may become unequally measured, and comparative results may be based on measurements of questionable validity.

3.1 Finland and Chile

In affect research, Finland is perennially placed to fit into the European and North American group of countries. Maths anxiety and maths self-concept seem to have similar roles in Finland as they have inside this Western division, in contrast to, for example, Confucian heritage culture countries (Lee, 2009; Wong, 2008). When it comes to Latin America, Chile, despite in many ways fitting into the mould of a modern country, is reported to have an achievement level very different from most Western (including North European) countries. In this thesis, Finland is perceived as representing individual Western culture and Chile as collectivist Latin culture (see also Hofstede & Hofstede, 2005). We will discuss the two countries' similarities and differences in the four contextual levels presented above.

On the pupils' (micro) level, Finnish students have a high degree of autonomy. Their mastery goal orientation correlates more with their achievement and self-efficacy than their performance goal orientation (Hannula & Laakso, 2011). In Chile, achievement correlates with high expectations for further education (Ramírez, 2005). These expectations seem to refer to performance orientation. When comparing Finnish and Chilean students concerning their engagement with school we can also find a difference. In the latest PISA analysis (OECD, 2014), Chilean students were above OECD average in their drive to learn mathematics, while Finnish students were below average. Further, Chilean students were above average in students reporting they were happy at school, whereas Finnish students were again below the average.

On the meso level, teachers' expectations constitute one structure of context. In Finland, the thing most valued by teachers is students' positive affect regarding mathematics (Niemi, 2010). This does not always happen; although the level of mathematics anxiety in Finland is low (Lee, 2009), many adolescent students have negative overall affect regarding mathematics (Hirvonen, 2012). In Chile, teachers' expectations are addressed to pupils' skills. Ramírez (2005) claims that while pupils in Chile feel they can do mathematics, they are still said to lack adequate skills to deal with the requirements of the national curriculum.

When it comes to the exo level, there are important differences between the cultures, socioeconomic development, and the education systems. Chile has one of the worst Gini index (see studies 2 and 3) in the world (0.52 in 2009 compared to 0.27 in Finland in 2008), and this inequality is mainly reinforced by a highly segregated educational system (Valenzuela, Bellei & De Los Ríos, 2009). In the Chilean educational system, only some schools are public and free of charge, while in Finland almost all schools are public and free. The number of school hours per week in Finland is one of the lowest in the world (23 hours per week is the minimum for the 3rd grade). In Chile, the number of school hours is the highest of all the OECD countries (38 hours per week). In Finland, there are about three mathematics lessons weekly in the 3rd grade, while in Chile, the number of mathematics lessons in the 3rd grade is six per week. In Finland teachers need to have a master's degree in education. The profession is valued in Finland, and teachers' salaries are slightly above the country's median salary. In Chile, teachers have low salaries (approximately half of the OECD average), and a master's degree is not required in order to qualify as a teacher. Generally, Chilean students applying for primary school teacher programs have very low averages in their previous studies. (OECD, 2011). In Finland, almost all primary school teachers are satisfied with the textbooks, materials, remedial instruction facilities, and in-service teachers' supplementary education (Joutsenlahti & Vainionpää, 2010). Teachers wish to have smaller groups than they do, although class sizes are fairly small in Finland (the average is 19 at the primary level; OECD, 2012). However, their wish is reasonable, as teachers report that they need to spend a lot of their time on class management duties. Chilean class sizes are the largest of the OECD countries: the average is almost 40 (OECD, 2012). In addition, teaching is a low-status profession. Teachers are officially assessed, and the results are published nationally.

The macro level consists of the interactions between the three lower hierarchy levels; thus, it encompasses the overall culture. Based on what has been presented above, the two cultures differ in the ways teaching-learning process is controlled and supported. In Finland, support is good, and the process seems fairly little regulated, whereas in Chile there is less support and more control. Independent of that, there are more concerns about the learning environment in Finland than in Chile. If there is a need for strong class management in Finland despite the smaller groups and other kinds of support, we see this as evidence of the social environment of learning being better in Chile and the technical learning environment being better in Finland.

3.2 The centrality of affect

Back in the 1970's, Green (1971) argued that beliefs build up discrete, separate clusters. This makes it possible simultaneously to have beliefs that are in contradiction. Further, this allows some beliefs to become more central than others. McLeod (1992) discussed the same idea, suggesting that the most central beliefs will be those the belief holder has complete consensus about. Those imposed by authorities become the least central. Chapman (2002) has confirmed the centrality of affect structures empirically. Based on her study, a change in the less central part of affect structures does not necessarily have any impact on the overall affect structure that, as the most central components overrule other components. For example, if a student perceives mathematics to be deeply boring, this perception may still remain even if the same student starts to see mathematics as more important. The importance and the boring nature of the subject might be contradictory, but if the boredom is more central to the student, this dominates the student's view on the subject (which is likely, as feelings are typically personally generated, and the idea of the subject's importance is imported by authorities). Chapman (ibid.) showed that to make a clear impact, an explicit contradiction needed to be created which was meaningful to the individual.

Culture determines to some extent how much individuals value norms that are brought from outside. For example, in individual cultures personal experiences may become the most central beliefs. On the other hand, in collectivist cultures social experiences may become more significant. This cultural variation is discussed by Markus and Kitayama (1991), who argued that if the culture is individual, then individual-level phenomena become very significant, structured and easily identifiable, while in collective cultures interindividual phenomena become more significant, more structured and more easily identified. In study 3, we restructured the affective structure to meet the cultural level variation (Table 3).

Table 3. Restructuring affect to meet the cultural variation of affective centrality

	Individual level	Interindividual level
Independent self-organization	Personal beliefs, emotions and values in internal interaction, orientation to maintain internal consistency	Individual differentiation from others
Interdependent (collectivist) self-organization	Personal beliefs, emotions, and values in interaction with surrounding	Others' thoughts and expectations, orientation to fit in

4 The intervention

A specific and persistent problem in classrooms is that they are often emotionally flat, and boredom is one of the most frequently experienced emotions (Nett, Goetz, & Hall, 2011). A change is possible: McLeod (1992) reported studies showing that the specific teaching of social norms impacted students' affective responses at the primary level. He also noticed that the most effective classrooms in terms of problem solving were found to be those wherein the environment was supportive and social norms encouraged students to be enthusiastic and to enjoy mathematical problem solving. Going back to Harter's (1999) view about self-construction, the affective learning environment should allow negative responses that are constructive for the self-concept, but inspire enough positive responses to help the students' self-concept to remain adequately positive. Pekrun, Goetz, Titz and Perry (2002) also call for interventions that "increase opportunities for success by using individual and cooperative reference norms for giving feedback and by inducing a culture of learning from errors" (p. 102).

Such pedagogical changes might help in solving the problem of deteriorating mathematics- related affect. However, most mathematics-related interventions have concentrated on the performance or cognitive problems, such as dyscalculia (see a review of the mathematics-related interventions in Dowker, 2009). Fewer interventions have been done to directly improve affect. When this has been the case, the focus has, for example, been in students' self-control and social interaction. In an intervention by Rimm-Kaufman et al (2014) a Responsive Classroom approach (RC) was used, aimed to foster relationships in the classroom and support students' self-control to enhance student achievement. The intervention discussed here emphasizes the same elements.

To make a change in mathematics classes' affective learning climate, a Finnish-Chilean research group designed a three-year intervention from 3rd (9-year-old pupils) to 5th (11- year-old pupils) grade. The intervention was built around monthly activities with mathematical problems, aiming that classes would become more active and teachers would vary their working methods. The mathematical problems were in most cases open ended, and they were selected or developed by the research group. The teachers were allowed and instructed to execute the problem-solving sessions according to their own preferences, and these choices were discussed in the meetings before and after the implementation. Most commonly, the teachers used collective activities in which pupils were allowed to discuss problems, to move around, and to work collaboratively. The lessons were filmed, and the teachers could use the videos to support their process in developing the activities during the intervention.

In open-ended mathematical problems, more than one solution is possible, and to find a solution, pupils need a linear or a cyclic problem-solving process where they use their resources, heuristics, beliefs, and abilities to monitor and self-regulate (Schoenfeld, 2012). Open-ended problems may enhance pupils' understanding as they allow connections to several or to untypical contexts. To provide an explicit example of intervention, one of the monthly problems will be introduced. It is called "Divide a square". In this problem the prompt was: Make a division to a square so that the two parts of the square are totally equal. How many different solutions can you find?" The problem was given at the 3rd grade, and it was the second problem in the project. The problem, the pupils' solutions and the thinking the pupils indicated have been analysed in a study by Laine, Näveri, Pehkonen, Ahtee, Heinilä & Hannula (2012). According to the study, five levels of thinking were present: level 0 = no solution; level 1 = the two most obvious solutions (two triangles and two rectangles); level 2 = division by a straight line that is not a diagonal, and does not pass through the middle points of the square's sidelines; level 3 = level 2 division plus replacing the straight line with a curve; and level 4 = truly understanding the central symmetry of the task. Because of the five levels of understanding, the problem provided a possibility to work at several levels of difficulty. It was easy to start working with the task, as the pupils could just take a paper and pencil to try out different solutions. The collaboration the pupils were allowed to have during the task made the task active.

The pupils participated in the intervention only through these monthly activities. As mentioned, all these lessons were filmed, but the pupils quickly got used to this. One camera followed the teacher and another followed one group of pupils for the whole lesson. The cameramen did not interfere with what the pupils were doing, and were friendly and supportive if the pupils contacted them. Even though the pupils were only involved in these monthly lessons, the intervention affected their learning climate altogether, as the teachers processed and evaluated their teaching according to the experiences during the intervention. According to the results of Portaankorva-Koivisto (2015) it was seen that the teachers' perceptions about mathematics teaching and learning developed during the intervention, and this affected their teaching in general.

The intervention was implemented in two countries, Finland and Chile between 2010 and 2014. The intervention classes were selected based on the teachers' own wish to participate in the intervention, so it is possible that the teachers in the intervention were above average in their willingness to develop their teaching.

5 Summary of the approaches and terms

In this thesis, one intention is to investigate the limitations and possibilities of commonly used methods to examine mathematics-related affect. Questionnaires that operate particularly on the micro level have been extensively in use (Chamberlin, 2010). By elaborating such a questionnaire this study focuses primarily on the micro level. The social turn mentioned earlier has shifted the focus to the meso level, and in this thesis we follow that shift by examining students' experiences of their mathematics classes. The examinees in the study come from Finland and Chile, countries that differ in their degree of collectivism and individualism, among other differences (see more about the differences in Hofstede & Hofstede, 2005; Harzing, 2006; and OECD, 2014). Because of these cultural differences this study touches on the higher contextual levels, the exo and macro levels, and thus all the four levels of context.

Multiple approaches are used in this thesis. The main approaches are Hannula's (2011; 2012) theoretical model of mathematics-related affect, Harter's (1999) developmental perspective to self-construction, the cross-cultural sensitivity that includes acknowledging Clarke's (2013) dilemmas, plus Berry's (1989) approaches to comparing cultures, and the different contextual levels (Bronfenbrenner, 1993). We use all three of the affective components structured by Hannula: cognitive, emotional and motivational. By operating on several contextual levels, we deepen Hannula's idea of the presence of these components in individual (psychological and physiological) and social levels in mathematics-related affect. On the other hand, in this thesis we discuss only the trait aspect of affect, not the state.

Author's comment. As an author of this dissertation, I would like to remark on my use of the various concepts I have used while progressing with the thesis. During this process, my thinking has been developed, and my understanding and knowledge about the field has increased. This is of course natural and desirable. The problem is that in the later studies of this thesis, I have discarded some of the terms that I have used in previous studies. Some terms have proved to be imprecise, some have referred to concepts I did not mean them to refer to, some terms have turned out to be misunderstood. Moreover, my use of English language has developed – in the first studies some mistakes are due to linguistic difficulties. Below, I have made a brief list of terms that I advise the reader to pay special attention to.

Factor. In the introduction, I use the word factor or concept to refer to an established or specific affective element that can be revealed, for example, by factor analysis (such as: self-competence, enjoyment of mathematics, mastery goal orientation). In the studies, these two words might have been used in various ways.

Dimension. In the introduction, dimension refers solely to Hannula's model's three dimensions, as the model itself is originally visualized in a 3-D structure. In the studies, I have at points used the word dimension to refer to various elements of affect.

Level. In the introduction, I use the word level to refer to the different contextual layers that are present both in Hannula's model (individual and social) and Bronfenbrenner's framework.

Interindividual level of affect. In the studies, I have used several terms to refer to what happens on the meso level. Hannula talks about the social level, and in many studies outside this thesis, terms such as learning climate, learning environment and affective atmosphere have been used, just to name a few. I was aiming to come up with a term that would explain my idea about what I was discussing, and that would also connect with existing research. My idea was that instead of studying the "social level" from the outside, I would use students' drawings. I was investigating students' "social affective landscape" in a way that would include the students' own participation. I have not so far found a similar approach in other research studies. In the last study (study 5), I used the term affective learning environment, which I am satisfied with. However, more constantly I have used the term interindividual level of affect. I find this term quite satisfactory, but it does not go well with the word interpersonal, which has been used by other researches with a slightly different meaning. In this introduction, I have consistently used the term interindividual level of affect. By that usage I have tried to reach the idea of the level at which it operates (the meso level) and to define the meaning of the term (an individual student's affective interaction with others, including the social responses and reflections she/he receives and sends out).

6 Research task

In previous studies on mathematics-related affect domain, the main emphasis has been on individual trait aspects, measured with paper and pencil questionnaires. Many studies have been cross-sectional, and the fewer number of longitudinal studies have typically addressed the individual level of affect. In Table 4, an overview of what has been emphasized and what has been missing in earlier studies in the domain is presented.

Table 4. Overview of the studies in the domain

Contextual level	**Individual (micro)**	**Interindividual (me-so)**	**Cultural (exo, mac-ro)**
Longitudinal de-velopment	General development emphasized, less studies of separate factors development	Less emphasized in the domain	Studied mostly in a different domain: in a domain of education-al policies
How to measure	Long tradition, mostly paper and pencil questionnaires	Less studied so far, growing interest to this exists, mixed methods recom-mended	Studied mostly in the domain of education-al policies
How to make a change (e.g. into more positive)	Interventions mostly addressing the cogni-tive changes, less the affective changes	Less emphasized so far	-

In this study, the five articles address several cells of Table 4. Table 5 shows how the articles distribute to the domain. Numbers refer to articles that address particular topics.

Table 5. How articles of this study relate with the domain

Contextual level	**Individual (micro)**	**Interindividual (me-so)**	**Cultural (exo, mac-ro)**
Longitudinal de-velopment	1, 4	5	
How to measure		3	2, 3
How to make a change	4	5	

Based on what has been presented about mathematics-related affect above and about less emphasized areas in the field, this research aims to answer the following questions:

1. How does Finnish students' mathematics-related affect develop on the individual level during comprehensive school years? (Study 1)
2. How valid and reliable is the commonly used imposed-etic approach in explaining mathematics-related individual affect structure in two very different cultures? (Study 2)
3. What can be revealed when using mixed methods to evaluate mathematics-related affect both on the individual and the interindividual level regarding two different cultures? (Study 3)
4. What is the interaction of individual, interindividual and cultural affective levels? (Study 3)
5. What are the impacts of the intervention aimed to make mathematics learning more active on both the individual and the interindividual levels of mathematics-related affect, and how do the two levels develop during primary school years? (Studies 4 and 5)

7 Overview of the studies

The five studies included in this thesis were done in collaboration with two research projects. The first project was implemented by the Finnish National Board of Education. In that project, the longitudinal development of Finnish comprehensive school students' performance and affect was examined. Study 1 was carried out within that project. The second project was a cross-cultural intervention study implemented in Finland and Chile by a research group from both countries. In that project, the overall aim was to examine the longitudinal impacts of the intervention on students' performance and affect, as well as the impacts on teachers' professional development. Studies 2, 3, 4 and 5 were carried out within this project. Table 6 provides an overview of the five studies of this thesis.

Table 6. Overview of the studies

Article	Design	Participants	Measure	Approach
1	Longitudinal, micro level	Finnish students from 3rd to 9th grade n = 3 502	Questionnaire	Derived etic
2	Cross-cultural, micro level	Finnish and Chilean students n1 = 466 n2 = 901	Questionnaire	Imposed-etic + emic
3	Cross-cultural, micro, meso, exo and macro levels	Finnish and Chilean students n1 = 466 n2 = 901	Questionnaire + drawings	Imposed-etic + derived-etic
4	Longitudinal, micro level	Finnish students from 3rd to 5th grade n1 = 320	Questionnaire	Derived-etic
5	Longitudinal, meso level	Finnish students from 3rd to 5th grade n1 = 320	Drawings	Derived-etic

8 Methods and results

8.1 Study 1 (article 1)

8.1.1 Participants

The data used in this study consisted of 3 502 Finnish students (1 702 girls, 1 800 boys) whose mathematics achievement and mathematics-related affect were followed throughout their comprehensive school years. The measurements were made at the beginning of the third, sixth, and at the end of the ninth grade (in 2005, 2008, and 2012, respectively). All students were selected by using the stratified sampling of the comprehensive schools, with different instruction languages (Finnish/Swedish), provinces and municipal groups (Cities / Population density areas / Rural areas) being represented. The sampling was random based, and its representativeness was very carefully designed. Not all the students could be followed during the whole data collection process, and the students that dropped out from the study were more commonly weak than high achieving. Thus the data, though being representative of all students in Finland, was slightly biased. This means that the results might be little more positive than what they would have been if all the weaker students had been included in the process.

8.1.2 Measure

In the study, the individual level of affect was measured. In all the measurements a modified version of the Fennema-Sherman Mathematics Attitude Scales (Fennema & Sherman, 1976; Metsämuuronen 2012) was used. In the study, two factors of the used instrument, i.e. Self-efficacy regarding mathematics and Enjoyment of mathematics were discussed. With respect to different measurements, the wording of the items was slightly modified to fit the examinees' developmental stage. The examples of the items are as follows: "Mathematics is easy" (Self-efficacy, first measurement), "Mathematics is an easy subject" (Self-efficacy, second and third measurement), "I like to learn Mathematics" (Enjoyment, first measurement), "I like to study Mathematics" (Enjoyment, second and third measurement). A 5-point Likert scale was in use except for the first measurement wherein only a 4-point scale was implemented. For the analyses, all the attitude scores were changed into percentages of the maximum score. Hence, in the most positive case, the student would get 100, which is strictly 100% of the maximum score. In the most negative case, the students would get zero, which corresponds with 0% of the maximum score. The analyses were done by calculating distributions of overall affect (Self-efficacy + Enjoyment) and of the factors separately (Self-efficacy / Enjoyment). This was done with respect to all measurements. Gender differences were investigated by t-tests.

8.1.3 Results

In the study we saw that in the first measurement the pupils' individual level mathematics-related affect was very positive. The mean for overall affect score was 71%, Enjoyment being slightly higher than Self-efficacy (mean for Enjoyment 72%; mean for Self-efficacy 68%). During the following years the affect turned negative. The overall affect score decreased from 71% to 60% by the time of the second measurement, and to 52% with the last measurement.

The decrease did not happen in the same way with both of the affective components. When examining the dimensions separately, we found that the decrease became stronger with respect to Enjoyment. In the first measurement, this component was at a high level (72%), but by the second measurement it had decreased to 54 %. The decrease continued after that, but less dramatically:with the last measurement the degree of the Enjoyment component was 47%. Regarding Self-efficacy, the decrease was very reasonable between the first and the second measurement (from 68% to 66%) despite the decrease in Enjoyment. Instead, by the last measurement Self-efficacy had decreased very clearly, becoming 57 % of the maximum score.

With respect to the different genders, the development of affect was similar, but the decrease was more dramatic regarding girls. In the first measurement, the mean for girls regarding the overall affect score was 68%, whereas the mean for boys was 72%. In the second measurement, the mean for girls regarding the overall affect score was 57%, while for boys it was 64%, and in the third measurement the mean for girls regarding the overall affect score was 50%, whereas the mean for boys was 54%. In all three measurements the difference between genders was statistically significant according to t-tests (the test values were $t = 4.63$, $p < 0.001$; $t = 10.29$, $p < 0.001$; $t = 5.02$, $p < 0.001$ regarding the three measurements respectively).

Examining the gender difference further by making separate analyses regarding the affective components, we saw that the difference was greater concerning the cognitive component (Self-efficacy), and that it was at its highest in the second measurement. According to t-tests, the difference was statistically significant in all three measurements regarding Self-efficacy (the test values were $t = 6.4$, $p < 0.001$; $t = 14.1$, $p < 0.001$; $t = 9.5$, $p < 0.001$; covering the three measurements respectively). The situation seemed less dramatic regarding the emotional component. The difference regarding Enjoyment was largest in the second measurement, and disappeared by the last measurement (the test values were $t = 2.3$, $p < 0.05$; $t = 5.2$, $p < 0.001$; $t = 0.2$, $p > 0.05$ covering the three measurements respectively).

8.2 Study 2 (article 2)

8.2.1 Participants

The data used in this study were collected within an intervention project aiming to develop mathematics learning in Finland and Chile. In Finland, the number of participants was 466 and in Chile 901, making the total number of participants 1 367. The data were collected at the beginning of the 2010–2011 academic year: September–October 2010 in Finland (regions near Helsinki) and March–April 2011 in Chile (Santiago de Chile). The Chilean school year begins 6 months later than it does in Finland for geographical (climatic) reasons, so the research phases were different in the two countries. The questionnaire was administered to 3rd graders. Schools are fairly uniform in Finland (see e.g. OECD, 2010, p. 87), but in Chile, there are private, semi-private, and public schools, and there can be huge variations between schools (OECD, 2010). The data from Chile was collected from all the three types of schools. On the whole, the data were considered to be representative of urban pupils in both countries.

8.2.2 Measure

The questionnaire was a shortened and simplified version of the instrument used by Hannula & Laakso (2011) to measure 4th grade Finnish pupils. The instrument was designed to operate on the individual level of affect, and to measure the three components of affect: the cognitive component, including Self-competence, (an example item: "I have done well in mathematics"), Self-confidence ("I am sure that I can learn math"), and Difficulty of mathematics, referred to as DoM ("Mathematics is difficult"); the emotional component, including Enjoyment of mathematics, referred to as EoM ("I have enjoyed pondering on mathematical exercises"); and the motivational component, including Mastery goal orientation, referred to as MGO ("In every lesson, I try to learn as much as possible"), and intensity, represented by Effort ("I always prepare myself carefully for exams"). The purpose of the instrument was to reveal the trait aspect of affect. All the items were first translated and modified by Roesken, Hannula and Pehkonen (The view of mathematics indicator, 2011), then by Hannula and Laakso, and finally simplified for the purposes of the research in question. The items were originally formulated in either English or Finnish, and then translated into both Finnish and Spanish.

The measurement was done using a three-point Likert scale. Thus the scale could be considered to be only ordinal. Some of the items were presented directly (e.g. "I have done well in mathematics"), while some were indirect claims (e.g. "I am not very good at mathematics"). In the analysis, the items that had an inverse content were re-coded to have the same direction as the directly stated items.

We conducted exploratory factor analyses to find out the structures of pupils' mathematics-related affect. The type of the structure as well as the similarities and differences between the structures in the two countries were inferred using different types of factor solutions. All the factor analyses were carried out separately for both data sets (Finnish / Chilean data). Statistical criteria were used to support decisions for different numbers of factors as well as to justify that the explored solutions were appropriate. We allowed the factor solutions to be rotated, using an orthogonal Varimax rotation. The analyses itself were made using principal component analysis.

The found factors' precision with the theoretical dimensions represented in the questionnaire was evaluated by examining whether the factors were constructed on items from only one or several theoretical dimensions. The number of missing cases varied between 30–41 (6–9%) for Finnish pupils and between 186–216 (21–24%) for Chilean pupils. Though there were more responses missing in the Chilean data, the missing cases did not seem to represent a specific group of pupils according to missing data analysis.

8.2.3 Results

The structure for Finnish 3rd grade students consisted of five factors that complied with instrument construction. The factors were Self-competence, Self-confidence, Enjoyment of mathematics (EoM), Mastery goal orientations (MGO), and Effort. The theory-based factor Difficulty of mathematics (DoM) was merged with the items of Self-competence (meaning that the impression of the difficulty of mathematics in general merged into the impression of one's own skills in mathematics). As much as 56% of variance was accounted for by these five factors. After finding a solution without first deciding on the number of factors, the solution was forced to be three-factorial. The factors were Capability (Self-competence + two items from Self-confidence + DoM), Interest (EoM; MGO had high loadings on this factor), and Investment (two items from Self-confidence + MGO + Effort). In this solution, the motivational component (effort and MGO) was united to Investment, yet all the items of MGO had correlations to Interest, and also connected with EoM. Self-confidence did not construct its own dimension. Its items had correlations with Self-competence and with Investment. As much as 47% of the variance was accounted for by the three factors.

The structure regarding Chilean pupils consisted of six factors. The factors were Inverse self-belief in mathematics (inverse items from Self-competence); Self-confidence; Easiness and fun (one direct item from Self-competence, two direct items from EoM, one direct item from DoM); Displeasure (inverse items from EoM + inverse items from DoM); Mastery goal orientation (MGO); and Effort. Fifty percent of variance was accounted for by these six factors. The scree plot suggested a three-factor solution, which we implemented. The factors were Flow (direct items from Effort + direct item from Self-competence + direct items from EoM + direct item from DoM + 2 items from Self-confidence); Determination (MGO + 2 items from Self-confidence); and Inverse perceptions about mathematics (all the inverse items). The three factors explained 36% of the variance.

When comparing the structures, we noticed that the extent of the congruence between the explanatory factors and the theory-based factors differed between the two countries. Also, the more compressed the structure, the less congruent it was with theory: this was seen with respect to both countries' structures. For the Finnish pupils, the cognitive component of affect was clear and distinct, but the emotional and motivational components were not that separate from each other. More precisely, in the Finnish pupils' structure, Enjoyment of mathematics was connected with Mastery goal orientation. The Finnish pupils linked Effort to Self-confidence and to Mastery goal orientations. Emotions, Self-confidence, and Mastery goal orientation were all connected regarding the Finnish pupils. Further, in the Finnish pupils' structure, Self-confidence was connected with Self-competence. With the Chilean pupils, cognitive and emotional components and intensity were largely combined, whereas the Mastery goal orientation were the clearest distinguishable factor. The Chilean pupils' emotions were connected with Effort and with some items of Self-competence and Self-confidence. Effort was connected with the Enjoyment of mathematics. In both countries Mastery goal orientation was connected with Self-confidence. The most remarkable difference between the countries was the Chilean pupils' inverse factors. All in all, there was no obvious universality between the structures.

8.3 Study 3 (article 3)

8.3.1 Participants

The participants of study 3 were the same as in study 2. The total number of participants was 1 367 (466 Finland, 901 in Chile). The data were collected at the beginning of the 2010–2011 academic year; September–October 2010 in Finland (regions near Helsinki) and March–April 2011 in Chile (Santiago). On the whole, the data were considered to be representative of urban pupils in both countries.

8.3.2 Measure

The goal of the study was to find out what is the degree of theindividual level and interindividual level of mathematics-related affect in Finland and Chile and what is the congruence between these levels. The individual level data were the same as in study 2, i.e. collected through a questionnaire. As the questionnaire items address individual experiences, perceptions and feelings, it was considered that they would reveal the individual level of affect. The interindividual level data were collected through drawings: we asked the examinees to draw their mathematics class. The exact drawing task we gave to our examinees was “Draw your teaching group, the teacher and the pupils in a mathematics lesson. Use speech bubbles and thought bubbles to describe conversation and thinking. Mark the pupil who represents yourself in the drawing by writing ME.” The drawing task regards the drawer’s perception about her/his social environment, not solely inner feelings. Thus, the drawings were considered to represent the interindividual level of affect.

The teachers were instructed to give no more than 45 minutes to the pupils to complete their drawings. Teachers were also instructed to let the pupils know that these drawings would not be looked at by their teachers. The pupils were allowed to draw only one moment in their drawings, but they were free to choose what moment to draw and what to emphasize. The coding was based on the “Classification of drawings” by Tikkanen (2008). In that classification, itinstructions are given to interpret the appearance of the mouths of all the people in the drawing. In addition, written expressions in bubbles were included in the interpretation. Five categories were in use: positive (all the people were positive or neutral, and at least one person had a smiling mouth or something positive was written), ambivalent (there were positive and negative people or writing in the drawing), negative (all the people were negative or neutral, and at least one person had a frowning mouth or something negative was written), neutral (nothing could be inferred from the persons’ facial expressions or bubbles), and not recognizable (there were no people and no bubbles). Only a few drawings were not decipherable. As we started the coding, two researchers interpreted together 100 drawings. As there were no unclearly interpreted drawings, and as the agreement rate between the coders in that subsample was 100%, only one researcher coded the rest of the drawings. Of those drawings, a few turned out to be to some extent unclear; these drawings were interpreted by the same research pair as at the beginning. In this phase, we found three drawings so confusing that we needed one more researcher to make an interpretation. However, in general, the interpretation process was straightforward.

We calculated the distributions of the pupils’ individual (questionnaire-based) and interindividual (drawing-based) affects. The comparison was done separately for both countries: through this, we could find out the effect of the culture on the individual and interindividual levels of affect.

Two types of statistical tests were used in the comparison. First, a t-test was used to compare the questionnaire answers regarding the countries. As the scale was likely to be ordinal, all the results were confirmed with non-parametric tests. When we analysed the drawing-based data, we used a chi-square test of dependence. After statistical analyses, we selected a sample of drawings for further investigation: the ambivalent drawings became of interest. From 19 Finnish classes, five classes having the biggest number of ambivalent drawings were chosen. From 21 Chilean classes, six classes were chosen for the same reason. Consequently, 72 ambivalent drawings were selected: 38 Finnish pupils' drawings, and 34 Chilean pupils' drawings. We analysed the reasons for ambivalence regarding the chosen drawings using thematic analysis. First, all the signs that referred to ambivalence were collected (e.g. someone is sleeping, another one is smiling; someone is succeeding, another one is failing). Second, we grouped the reasons (e.g. throwing erasers refers to making disturbances) and after that we labelled the groups (e.g. cheating/farting/chatting/ teasing/shouting/shooting/homework undone/using a cell phone + at least someone being positive → Disturbance).

8.3.3 Results

According to the questionnaire, the pupils' individual level of affect regarding mathematics was mainly positive in both countries. Still, the proportion of pupils having the most positive affect was considerably greater in Finland than it is in Chile. Almost no pupils had negative affect in either of the countries, but in Chile there were more pupils in the middle category than there were in Finland. With respect to the distributions, there was a statistically significant difference according to the t-test between the countries ($t = -6.613$; $p < 0.001$). Regarding the interindividual level, the pupils seemed to be more negative in Finland. The proportion of positive drawings was clearly smaller in the Finnish responses than it was in Chile, and the proportions of ambivalent and negative drawings were greater in Finland than in Chile. The chi-square test showed a statistically significant difference between the countries ($\chi2$ $(3) = 99.500$, $p < 0.001$). The individual level distributions seemed to be quite similarly independent of the content of the drawing. The chi-square test indicated no statistically significant dependency between the questionnaire and the drawing frequencies among Finnish pupils ($\chi2$ $(3) = 1.372$, $p > 0.05$). In Chile, the situation was similar according to the chi-square test ($\chi2$ $(3) = 3.627$, $p > 0.05$). Thus, the two levels of affect were statistically independent regarding both countries. Summing up these results, in Finland many pupils seemed to think that they personally can do mathematics. They like it, they are confident about it, and they want to learn it. However, they indicated that they were more uncertain about their mathematics classes than were pupils in Chile. In Chile, pupils seemed to enjoy their mathematics classes and yet had a less positive individual affective situation.

In the thematic analysis of the drawings, a number of categories regarding ambivalence was found. There were four categories present in both countries: Easy and difficult (somebody can + somebody cannot; 23 Finnish, 10 Chilean drawings); Disturbance (noise + teacher shouting "silence!" / cell phones / homework undone / throwing erasers / teasing / farting / cheating / unwillingness to do what is asked; 22 Finnish, 7 Chilean drawings); Boredom or sleepiness (pupils saying "this is boring" or somebody was drawn sleeping; 11 Finnish, 4 Chilean drawings); and Positive and negative emotions (typically someone smiling, another one with mouth pointing downwards; 10 Finnish, 4 Chilean drawings). The countries differed somewhat regarding the categories and there were more categories in Chile. Furthermore, the students were not as uniform across the classes in Chile as they were in Finland. In Finland, practically all the categories were present in every class that was examined. In Chile, there were classes where almost all the drawings expressed ambivalence for one or two reasons (categories). None of the Chilean drawings included all the categories at the same time, so the classes varied between each other more in Chile than in Finland. Another difference between the countries concerned the Disturbance category. In Finland, the disturbances mostly targeted the teacher, whereas in Chile this was not always the case: the target could, for example, be another pupil. Finally, in Chile, even in ambivalent pictures, there was a remarkable amount of positive small talk, smiling, and helping; this was not visible in Finnish drawings. All these notions were in line with our hypothesis: in Chile, the interindividual level varied more and the pupils expressed many things with reference to their surroundings, even complicated notions.

8.4 Study 4 (article 4)

8.4.1 Participants

Like in studies 2 and 3, the data used in this study were part of the intervention project. Study 4 focused on the Finnish pupils' individual level data before and after the intervention. The number of pupils that participated either in the pre-test, the post-test or both tests was 320. The pre-test data was collected at the beginning of the academic year 2010-2011 during September-October 2010 in regions near Helsinki. The post-test data was collected at the end of the academic year 2013-2014 during April-May.

In the pre-test, 25 classes were involved; 10 of these classes were intervention groups, while the rest of them were control groups. In the post-test, 6 control groups were not reached and 3 intervention groups had left the project (they had quit doing the tasks, but still participated in the post-test). From the 3 classes that had quit, one had participated in the project for 2 years whereas the other two had participated for only one year. The class that had participated for 2 years (i.e. more than 50% of the intervention tasks) was included, but the classes that had been participating for just one year (i.e. less than 50% of the tasks) were not. There was a teacher change in two of the intervention classes, but no classes were excluded for that reason. Some movement regarding the pupils had occurred, as there were pupils in the pre-test but not in the post-test and vice versa: those pupils' data were excluded from the comparative analysis of the pre-test and the post-test.

8.4.2 Measure

In study 4, we examined the longitudinal development of individual-level affect regarding two groups, the intervention group and the control group. In both tests the same questionnaire was used as in study 2 and 3. In a study by Hannula and Laakso (2011) the same instrument worked well and seemed suitable for measuring the individual level of mathematics- related affect within the Finnish population.

Before starting the analysis, we constructed a sum variable of all the questionnaire items regarding both measurements. The reliabilities (measured by Cronbach alphas) were satisfactory: $\alpha = .895$ in the pre-test, and $\alpha = .858$ in the post-test. To find out the answer to the research problem, we calculated the distributions regarding both measurements. We used a paired sample t-test to compare the means of the distributions regarding the two measurements and an independent sample t-test to compare the means of the distributions regarding intervention and control groups and gender. As the scale was likely to be ordinal, all the results were confirmed using non-parametric tests.

8.4.3 Results

The mean of all items for all pupils in the pre-test was 1.37 (1 = positive, 3 = negative), and the standard deviation was 0.30. In the post-test, the mean of all items for all pupils was 1.64, the standard deviation being 0.29. In a paired samples t-test there was a statistically significant difference between the pre-test and the post-test regarding all the pupils ($t(193) = -11.88$; $p < .000$), the intervention group ($t(108) = -9.72$; $p < .000$), and the control group ($t(84) = -6.98$; $p < .000$). The results indicate that there is a remarkable decline in pupils' affect regarding mathematics from the beginning of the 3rd to the end of the 5th grade in both the intervention group and the control group.

Concerning the differences between the intervention and the control group, no statistically significant difference was found with respect to all items in the post-test ($t(276) = -.67$; $p = .505$). Looking further at the differences between the groups factor by factor in the post-test did not change the picture: $t(287) = -.06$, $p = .954$ regarding Self-competence; $t(294) = -.79$, $p = .433$ regarding Self-confidence; $t(290) = -1.50$, $p = .134$ regarding the Difficulty of mathematics; $t(290) = 0.62$, $p = .533$ regarding the Enjoyment of mathematics; $t(294) = -.57$, $p = .571$ regarding Mastery goal orientation; and $t(290) = .62$, $p = .536$ regarding Effort. Besides the non-significance between the groups, no trend was found regarding minor differences in the different variables, as with respect to one variable the mean could be lower for the control group, but with respect to another the mean could be lower for the intervention group.

With gender differences, we did not find any significant differences in either of the tests (gender difference in the pre-test regarding all items: $t(122) = 1.05$, $p = .295$; gender difference in the post-test: $t(140) = 1.57$, $p = .118$). However, when testing the difference between the control group and the intervention group in the post-test separately for gender, a statistically significant difference was found regarding girls' development (girls: $t(67) = 2.08$, $p < .05$; boys: $t(87) = .42$, $p = .634$). This difference was not present in the pre-test (girls: $t(59) = -,928$ $p = .375$; boys: $t(78) = .30$, $p = .763$). The mean of the control group girls in the post-test was 1.82, and for the intervention group girls 1.65. This means that girls had benefited from the intervention, but not the boys. The significance in the development came through two factors: Self-confidence $t(72) = 2.39$, $p < .05$), and EoM ($t(72) = 2.47$, $p < .05$).

8.5 Study 5 (article 5)

8.5.1 Participants

In study 5, the data gathered within the intervention project was again used. Study 5 focused on the Finnish pupils' interindividual level data before and after the intervention. The number of pupils that participated either in the pre-test, the post-test or both tests was 320. The pre-test data was collected at the beginning of the academic year 2010-2011 during September-October 2010 in regions near Helsinki. The post-test data was collected at the end of the academic year 2013-2014 during April-May.

The pre-test involved 25 classes; 10 of these classes were intervention groups, while the rest of them were control groups. In the post-test, 6 control groups were not reached and 3 intervention groups had left the project (they had quit doing the tasks, but had still participated in the post-test). From the 3 classes that had quit, one had participated in the project for 2 years, whereas the other two had participated for only one year.

8.5.2 Measure

As in study 3, the examinees' drawings of their mathematics classes were used to analyse the interindividual level of affect. A similar drawing task with similar teacher instructions was given to pupils regarding the pre-test and the post-test (Draw your teaching group, the teacher and the pupils in a mathematics lesson. Use speech bubbles and thought bubbles to describe conversation and thinking. Mark the pupil that represents you in the drawing by writing ME). The coding of the drawings was done in a similar way for both datasets, using Tikkanen's (2008) "Classification of drawings".

After the coding, the drawings' differences between the grades were compared with a chi-square test. We also categorized the post-test drawings by considering whether the pupils in the drawing were pictured as activated or de-activated using Pekrun's (2006) division of activating/de-activating emotions. The idea is that emotion can be positive, but de-activating (e.g. satisfaction); even when positive the emotion does not necessarily promote learning process. On the other hand, emotion can be negative but activating (e.g. anger); even being negative the emotion can motivate a pupil to improve learning. For example, the pupils could have been pictured to be very active to find an answer to the teacher's question. In that case, the drawing was coded as activating, and the reason for that was because we considered it as cognitive and performance oriented. Another example is a drawing where some pupils were pictured as enthusiastic ("Yippee, mathematics!"), and other pupils as bored (head down on a table with a thinking bubble showing "zzzzz"). In this case, the drawing was coded "Emotionally mixed: activating and de-activating": as it was both activating because of the positive emotional reason given and de-activating because of the negative emotional and physical reasons given.

8.5.3 Results

Regarding the interindividual level of mathematics-related affect, there was a statistically significant difference between the two groups in the post test ($\chi2$ (4) = 17.292, $p < .01$). The distribution of the intervention group in the post test was almost similar to that of the pre-test, so it seemed as if the intervention helped the intervention group pupils to keep their interindivual affective level similar to what it was at the third grade. On the other hand, the distribution of the control group in the post-test clearly declined compared to the situation in the pre-test. At the 5th grade, the pupils in the intervention group most likely experienced the interindividual level as positive or ambivalent, which typically related to the idea that some of the pupils followed the teaching content, while some of them did not, either because they did not understand the content or because they had no interest in the content. In the control group, the drawings most likely indicated an ambivalent atmosphere. The second likely option was a negative atmosphere, and the proportion of the positive drawings was minor compared to those of the intervention group. Moreover, in both measurements over 10% of the drawings were negative. This indicated that even after the intervention the proportion of pupils who experienced their mathematics classes as negative was remarkable. Only part of the drawings included anything activating or de-activating: this happened in 72 of 119 intervention pupils' drawings (63.0%), and in 77 of 138 control pupils' drawings (55.8%). In the rest of the drawings nothing which referred to being activated or de-activated was present.

When we analysed the activation expressed in the drawings, we found seven categories of activation/de-activation. In the "Cognitive" category there were concerns about how to answer the teachers' questions, how to calculate, how to get the result, or an indication that some pupils did not understand what was going on or needed help with the content. In the "Cognitive and social" category the cognitive concerns were completed by asking for help, offering a helping hand, or in some cases with discussing the learning content. In the "Just social" category there was interaction with the pupils (e.g. chatting with a friend, listening to music together, or even disturbing the lesson), but with no relation to cognition. The category "Emotional: boredom" was noticeably large, but there was also a great difference in the proportions between the two groups (27.8% in the intervention group, 70.1% in the control group). Most drawings in this category expressed nothing else than pupils saying they were bored, or depicting themselves as asleep. There were more intense feelings (e.g. anger, "I'm gonna die" "I'm desperate, I hope this school explodes"), and less intense feelings ("I wish to go home", "Yippee, not much time left before the break"). In the control group, the proportion of this category was extremely worrying. The category "Emotional: enjoying/inspiring" was intuitively the most positive one. In the few drawings included in this category, there were positive expectations about learning mathematics or learning new things, and in some of the drawings there were also aspects of communication, collaboration, excitement, enjoyment and inspiration. There were three drawings which showed both inspiration and boredom; those drawings were coded in the category "Emotional: mixed", and finally, there were two drawings that created the last category, "Other". There was one drawing where envy was present (emotion, but not necessarily negative, nor necessarily positive with respect to learning), and a drawing wherein there were self-expectations: the drawer had not done the homework, and envisioned the day when the homework would be done.

All in all, drawings that were activating and positive (Cognitive, Cognitive and social, Emotional: enjoying/inspiring) characterized more than half of the intervention group drawings (54.3 %), but less than a fifth of the control group drawings (16.9 %). Drawings that were de-activating and negative (Emotional: boredom) characterized almost a third of the intervention group drawings, but more than two thirds of the control group drawings. Drawings that could be interpreted to include illustrations significant to the pupils and to the learning process (Cognitive and social, Emotional: enjoying/inspiring) formed only a small proportion of the drawings, but in the intervention group the proportion was remarkably greater (36.2% in the intervention group; 6.5% in the control group). The chi-square test of independence showed a statistically significant difference between the two groups regarding the distributions in the categories ($\chi2$ (4) = 37.015 $p < .001$). The test was implemented excluding the categories “Emotional: enjoying/inspiring” and “Other” to fulfil the assumptions of the test.

9 Limitations of the study

There are at least four kinds of limitations that might weaken the results of this thesis. The first problem relates to the many concepts that were used in the study. In the field of mathematics- related affect several overlapping concepts and frameworks are in use. While carrying out this study, I learned about new theories and frameworks, and identified new connections with previous ideas. What follows acknowledges the fact that the concepts that were in use at the beginning of the project, i.e. those that were used in the first articles, are slightly maladaptive with the concepts used in later articles.

Another limitation concerns the intervention data, used in studies 2 - 5, where the examinees were not randomly chosen. This is probably one of the most critical limitations. In Finland, schools are fairly equal, and the differences between students are one of the smallest in the world (OECD, 2010). However, in a recent dissertation by Bernelius (2013) it was seen that differences between schools in Finland are becoming more evident, especially in the capital area. The obvious consequence is the decrease of the generalizability of the results of this thesis.

Third limitation addresses the intervention. As the examinees only took part in the problem- solving session monthly, the impact was likely to become small. Thus, there is a possibility of underestimating the effectiveness of the intervention. Moderate results might be a consequence of too few problem-solving lessons rather than their ineffectiveness.

The fourth problem is that as the analysis of the drawings was based on interpretation, the message given by the drawers might be over- or underestimated. Yet, the interpretations were made in a similar vein regarding all the drawings, so the possible misinterpretations were consistent.

10 Discussion of the results

10.1 The development of mathematics-related affect during the comprehensive school years (study 1)

The detrimental development of mathematics-related affect during the comprehensive school years is a widespread problem with undesirable implications (see e.g. Lee, 2009). Students with negative mathematics-related affect are likely to avoid the subject in the future and might suffer from anxious feelings about situations that are mathematics related (Sjøberg and Schreiner, 2010). Earlier studies (Metsämuuronen, 2013) have recognized such an evolution regarding Finnish students as well. In the first study of this thesis, we wanted to confirm what Finnish students' affective development looked like in general, and what would be specific to that development. Confirming previous studies, we found that Finnish students' mathematics-related affect decreases rather dramatically during the comprehensive school years. More specifically, enjoyment of mathematics was most likely to decrease during the primary school years (between 3rd grade and 6th grade), whereas self-efficacy was most likely to decrease during the lower secondary school years (between 6th grade and 9th grade). We could also identify clear gender differences: the decrease was more dramatic among girls than it was among boys.

10.2 Separate factors' development (study 1)

The two affective factors, enjoyment of mathematics and self-efficacy developed independently from each other during the primary school years. In the results of the first study, boys maintained their self-efficacy at a positive level during the primary school years while girls did not. Simultaneously, both genders' enjoyment of mathematics decreased. During the lower comprehensive school years both genders' self-efficacy decreased. The enjoyment of mathematics also kept decreasing with both genders. It is possible that the decrease in self-efficacy can be connected to the move from the primary school to the lower secondary school. The teaching is from there on given by subject teachers, and the change might be too challenging for many students. The content of mathematics becomes more abstract, and students start to notice that some classmates are able to reach the level of abstract thinking more easily.

10.3 Cultural lenses on affect (study 2)

In the second study, we investigated the validity and reliability of an imposed etic approach in measuring affective aspects of mathematics in different cultures.

The questionnaire we elaborated has its roots in Western culture. Using that questionnaire for Finnish and Chilean 3rd grade pupils, we were able to reveal almost all the Finnish but only partially the Chilean pupils' mathematics-related affect structure. Thus, the questionnaire was considered to be valid for Finnish students, allowing us to identify the factors of their affective structure, but less valid for Chilean pupils, as for them the affective factors were not clearly identifiable, nor distinguishable with that instrument. Clarke (2013) has discussed challenges that exist in cross-cultural comparisons, suggesting that one can hardly avoid a clash between validity and comparability. Clarke (ibid.) argues that either one or the other needs to be emphasized, and that the researcher has to make a validity-comparability compromise. We discussed this compromise regarding the affect domain in the second study. If looking for comparability, one could suggest that in international comparisons only the global components should be measured. According to our study, it is not obvious what components would be the universal ones. For example, in Hannula's model (2011; 2012) emotions are distinguished from the cognitive and motivational components of affect. However, in our results the emotional component was connected with the cognitive component among Chilean pupils, whereas among Finnish pupils emotional and cognitive components were distinct whereas the emotional and motivational components were connected. Our results suggest that the structures can be manifested differently in different cultures. Emphasizing comparability by measuring universal factors becomes problematic if we acknowledge that affective structures can be culturally dependent. In Clarke's (2013) words, the dilemma of inclusive vs. distinctive, rephrased in the second study as the premise of universalism regarding the examined components, is present in the affect domain.

The other option would be to emphasize validity. This could be done by evaluating what aspects of affect would be the most beneficial ones. However, Clarke (2013) introduces the dilemma of evaluative criteria, rephrased in the second study as the premise of universalism within the components' significances. This dilemma discusses how an effort to value different aspects of affect might prove to be problematic. In our results, we were not looking for such evaluations, but some discussion is possible based on earlier studies. For example, high self-efficacy regarding mathematics sounds like something that should be acquired. However, in Chile pupils are interpreted as having high self-competence on account of easy tasks (Ramirez, 2005). If this is the case, valuing a high degree of self-efficacy could be a mistake. Another example is a study by Tuohilampi and Hannula (2011), where a remarkably high mastery goal orientation, considered as unrealistic in relation to perceived competence, was connected with poor achievement and displeasure in doing mathematics. This suggests that valuing the high degree of willingness to learn mathematics is not an unambiguous benefit.

One more dilemma relating to the comparability and validity of cross-cultural comparisons was discussed in the second study. People in different cultures may interpret items used in questionnaires in different ways, and very complicated items might become misunderstood or misinterpreted (Harzing, 2006). Metsämuuronen (2012) has shown that especially indirect items are likely to be misunderstood in certain cultures, as well as by low-achieving students. He argues that in international comparisons the items should be as simple as possible, and the number of indirect items should be at a minimum. This approach emphasizes comparability. However, in our study, the indirect items were structured as one affective factor regarding Chilean pupils (alone). In that study we used a derived-etic approach, which in this case means that the results were interpreted by researchers both inside and outside of Chilean culture. This factor, "Inverse perceptions about mathematics", was interpreted as a cultural way of seeing mathematics rather than not understanding the content of the items. We argued in the second study that Chilean people are used to relativizing their statements. Thus, it would be no contradiction in agreeing simultaneously with two opposite statements, as both may depend on the context. With that kind of perspective, it is different to view mathematics through a positive frame than through a negative one. The suggested interpretation goes along with the idea of direction (referring to a positive or negative orientation), presented in Chamberlin's (2010) study. If such cultural factors exist, unpredicted affective factors, or even components might be overlooked when keeping items particularly simple. This would lead to several dilemmas, which Clarke (2013) has noted: the dilemma of omission, the dilemma of linguistic preclusion, and the dilemma of inclusive vs. distinctive. All the dilemmas relate to a decrease in validity. In the second study, these dilemmas were rephrased as the premise of components' coverage among the phenomena (omission), the premise of congruence between languages and conceptions (linguistic preclusion) and the premise of universalism within the examined components (inclusive vs. distinctive). According to the second study, all the three dilemmas exist in the mathematics-related affect domain.

10.4 Combining the individual, the interindividual and the cultural level of affect (study 3)

In the third study, the examined pupils showed mainly positive affect in both tools, i.e. in the questionnaire and in the drawings. The children were also mostly positive about their affective learning climate. These results are in line with their developmental stage (Harter, 1999). They are young enough to see their perceptions in mainly positive terms, but they increasingly make social comparisons in order to justify their self-construct. Though the distributions were mainly positive regarding both levels of the affect, we found incongruence between the levels, indicating a possible lack of similarity between the affective levels. Even more, we saw that the nature of the disparity between the levels seemed to depend on culture, as it was not similar with respect to both countries. In Finland, most pupils expressed the individual level of affect as being very positive, but the interindividual level of affect was clearly less positive. In Chile, the situation was the opposite. In the third study, our hypothesis suggested that one or the other of the levels would appear to be more important in the two countries. We used the idea of Markus and Kitayama (1991), who have argued that if the culture is individual, then individual-level phenomena become very significant, structured and clearly identifiable, while in collective cultures interindividual phenomena become more significant, more structured and more easily identified. In our results, the affect was more positive on the individual level for Finnish pupils, who were considered to represent an individualistic culture. For Chilean pupils, who were considered to represent a collectivist culture, the interindividual level of affect appeared to be more positive. Such a cultural variation in valuing different affective levels suggests that we should use methods that cover more than one affective level when making cross-cultural comparisons. This cultural ontological variation can also make the interpretations of cross-cultural results invalid; a challenge that can be connected with one of Clarke's (2013) dilemmas, namely evaluative criteria.

In the third study, we used thematic analysis for selected drawings, revealing altogether six categories of ambivalence which might exist in mathematics classrooms. These categories were not universal between the countries. We saw that the reasons for ambivalence across the classes were more uniform in Finland than in Chile, as all the categories of ambivalence were present in almost all the Finnish classes. In Chile, the reasons for ambivalence varied across the classes, and in total there were more categories in Chilean drawings. We saw the larger number of categories in Chile and the greater variation between the classes to be an indicator of what Markus and Kitayama (1991) have shown. They argue that individuals in collectivist cultures focus more on their social environment than do independent individuals. Thus, individuals who live in more collectivist cultures might be able to notice different types of variation in their surroundings. Even though there was greater variation between the classes in Chile than in Finland, the classes themselves were more uniform in Chile in terms of reasons for ambivalence. Unlike in Finland, there were no classes in Chile where all the categories were present. In the most likely case, almost all the drawings of any class were positive, and if ambivalence was expressed in some pictures, the origin of this ambivalence was typically identified by others as well.

10.5 The intervention's impact on the individual level of mathematics-related affect (study 4)

In studies 4 and 5, we investigated the longitudinal development of mathematics-related affect at several levels, including the effects of our intervention. Regarding the individual level of mathematics-related affect, the fourth study once again confirmed detrimental development. The individual level as a whole turned from positive to negative from the beginning of the 3rd grade to the end of the 5th grade. This happened both among pupils who participated in the intervention and among pupils who were studying in the control group classes. Even though the intervention could not prevent this detrimental development on the individual affective level regarding an average pupil, the results showed a small-scale impact among girls. For the girls who participated in the intervention the development of self-confidence and the enjoyment of mathematics was less negative than it was on average. This was an important result, bearing in mind that girls in particular suffer from poor and unrealistic mathematical self-confidence worldwide (Syzmanowics & Furham, 2011). Girls' emotions towards mathematics have also been seen as critical (ibid.), so it is a good sign that the intervention could help girls maintain their positive emotions.

10.6 The intervention's impact on the interindividual level of mathematics-related affect (study 5)

The intervention could not make a strong impact on the individual level of mathematics-related affect, but the case was different on the interindividual level. In the fifth study, we reported a clear impact on the pupils' affective learning climate. Pupils expressed their affective learning climate to be very positive at the beginning of the project, as was expected, given the examinees' developmental stage (Harter, 1999). The intervention did not succeed in making the learning climate better than it was at the beginning, but it did prevent it from developing detrimentally. A clear decline was evident in the control group. In the fifth study, we also elaborated on the activation and de-activation illustrated in the drawings. Based on that analysis, we saw that 1) almost three fifths of all the 5th-grade drawings did not relate to any kind of activation or de-activation; 2) the number of drawings expressing clearly positive activation (Enjoyment/inspiration) was minimal in both groups; 3) the intervention group drawings clearly included more positive activation; and 4) the proportion of the emotionally boring mathematics classes was alarming, especially within the control group pupils.

We could identify all the three affective components presented by Hannula (2011) from the drawings. The activation categories "Cognitive", "Cognitive and social" and "Just social" related to social discourse, i.e. how to make certain actions (cognitive). The three emotional categories related to the social atmosphere, i.e. what it is like to be inside a certain microculture (emotional). Those categories also provided background concerning social norms and needs, i.e. why certain actions are made (motivational). This was the case, for instance, when sleepiness resulted from feeling bored.

Similar to the findings of Nett, Goetz, & Hall (2011) and Vogel-Walcutt and others (2012), the fifth study showed that there were quite a few drawings which indicated emotionally flat classrooms. We argued that the drawings suggest that Finnish mathematics classes are fairly non-activating or de-activating. Pupils very seldom saw their mathematics classes as emotionally positive or inspiring, so if there was activation, it was more likely to be related to cognitive actions than emotions.

10.7 Combining the results of the five studies

These five studies show that the detrimental development of affect exists, and it is not limited to individual-level affect. The results show that the problem can be helped through intervention which allows problem-based collaboration. Future elaboration and documentation about this development should include more variation in measurements especially when it comes to cross-cultural studies, as the results indicate that not all the significant elements of affect can be revealed through traditional questionnaires, let alone through imposed etic-based studies. In Table 7, there is a synthesis of the results of the five studies.

Table 7. Combining the main results regarding mathematics-related affect

Contextual level	**Individual (micro)**	**Interindividual (meso)**	**Cultural (exo, macro)**
Longitudinal development	Studies 1 and 4 confirmed detrimental development, and clarified this to begin with a decrease of enjoyment of mathematics and continue with the decrease of mathematical self-efficacy	Study 5 showed that there is a detrimental development on how pupils perceive their learning climate, and their experience about that is fairly unproductive, relating to boredom and deactivation.	-
How to measure	-	Study 3 revealed incongruence between the affective levels by using different types of measurements, suggesting that methods covering more than one affective level are needed in cross-cultural comparisons	Studies 2 and 3 showed that affective structures may not be universal across cultures and raised up that the balance between validity and comparability needs to be carefully considered in cross-cultural comparisons
How to make a change	Study 4 showed that detrimental development can be prevented through a problem based intervention with social collaboration allowed	Study 5 indicated that detrimental development can be prevented through a problem based intervention with social collaboration allowed	-

11 Implications

This thesis discusses several issues concerning the contextual levels of mathematics-related affect. In the first study, the affective development during the comprehensive school years was elaborated on the individual level. In the second study, the emphasis was on how to measure individual-level affect and what appears to be particularly challenging in cross-cultural comparisons. In the third study three levels of affect were connected and the possibility of mixed methods in measuring several affective levels was examined. The fourth and fifth study investigated the development of two affective levels and the intervention's impact on them. The main theoretical bases for the thesis were Bronfenbrenner's (1993) contextual levels, Harter's (1999) developmental perspective on self-construction and Hannula's (2011; 2012) model of mathematics-related affect domain. According to Harter, during the primary school years children modify their self-construct according to the reactions and responses they identify in their surroundings. In Hannula's model, the domain is structured as three-dimensional, namely affective components, trait and state aspects, and different manifestation areas.

11.1 Methodological implications

In the thesis, we identified several cultural aspects that need to be addressed in cross-cultural comparisons regarding mathematics. Based on Clarke's (2013) dilemmas, we confirmed the premise of universalism regarding the examined components, the premise of universalism within the components' significances, the premise of congruence between languages and conceptions and the premise of components' coverage among the phenomena to be problematic in the mathematics-related affect domain. The cultural variation related to these premises needs to be looked at from the perspective of a validity-comparability compromise, as suggested by Clarke (2013). First, in order to acknowledge cultural variety in the evaluation of different affective levels, we should use methods that cover more than one affective level in cross-cultural comparisons. Second, when only focusing on the most global affective components or factors, one has to keep in mind that even the simplest measurements might relate to ontological differences. Third, interpretations should be made using a derived-etic approach; interpretations at least need to be discussed in a dialogue between researches from inside and outside the culture. Another solution was introduced in a study by Bofah & Hannula (2014). By elaborating examinees' factor structure quantitatively, it is possible to identify items that do not match the instruments' purpose, and if several items mismatch in a similar way, the discrepancy can be fixed by using an accordingly designed coefficient. No matter what solution is sought, a culturally sensitive rather than an imposed-etic approach should be used in cross-cultural comparisons.

11.2 Theoretical implications

The categories of ambivalence that were found in the third study deepen the picture of affective experiences students have while learning mathematics. The categories show that during mathematics lessons, pupils have identified or participated in other pupils' cognitive processes, behavioural habits, physically apparent emotions, and emotions expressed by facial expressions. It would be particularly interesting to find out whether all these interactions have (or do not have) a similar significance in terms of pupils' affective development. It is possible that some affective phenomena pupils identify or participate in during their mathematics classes might have a specifically strong impact on self-construction. Future studies should go deeper into finding out possible "cornerstone" experiences, envisioned recently by Hannula (2015).

In the third study, we noticed that the individual and the interindividual level of affect are not always coherent. Finnish pupils' individual level of affect was positive despite the fact that many pupils experienced their interindividual level of affect as ambivalent or even negative. In Chile, the case was different: the individual level of affect was on average lower than it was among Finnish pupils, but the interindividual level was mostly positive. Evans (2006) has argued that individual affect draws from social interaction; however, in our results we could not identify what would indicate such an interaction. The results suggested that it was rather the cultural level of affect that had an impact on the degree of positivity of the affective levels. The level that was hypothesized as the most significant was noticed to be the more positive one regarding both cultures. It is possible that the lower degree of either the individual level of affect (for Chileans) or the interindividual level of affect (for Finns) is simply a sign of the less emphasized construction of the less significant affective level. Based on these results, we created a table to present the interaction and the significances regarding the individual and interindividual levels (Table 8). As this was done for both cultures separately, one can actually find there a combination of all the three affective levels discussed in this thesis. In future studies, the significance of different affective levels should be acknowledged so that the most developed levels would not become omitted.

Table 8. Interaction and significances regarding the two levels of affect in Finland and Chile

	Individual - Interindividual
Finland	**good** - ambivalent The more meaningful level is better: intrinsic harmony, discrepancies in the less significant level
Chile	medium - **good** The more meaningful level is better: harmony with others, discrepancies in the less significant level

In the fifth study, we saw that mathematics-related affective learning climate can be impacted by an activating intervention. The same intervention could make only a moderate impact on pupils' individual-level affect. Such results confirm that by examining more than just one level of affect a wider picture of mathematics-related affect might be achieved. Even though the intervention did not make a clear difference at the individual level, the interindividual level impact might have long-term consequences. In future, it is necessary to find out more information about the long-lasting consequences that the interindividual level of affect might have on individual affect during the school years.

11.3 Practical implications

The thesis as a whole confirms the undesirable development of mathematics-related affect during the comprehensive school years. The degree of negativity in that development seems to exceed the normative progress of the originally very positive affect becoming more realistic. The examinees in this study are predominantly Finnish students, whose performance level is high (OECD, 2010; 2014). It would seem that the high performance level and the affective degree which Finnish students eventually have do not really match. Two studies in this thesis showed the remarkable affective decline during the primary school years, addressing the emotional component in particular. This could indicate that Finnish pupils experience mathematics in too negative a way during the primary school years, and that the reactions and social responses they receive in their surroundings do not serve as a realistic mathematical self-construct. The negative development of the emotional trait (Hannula, 2011; 2012) during primary school years might be a sign of pupils experiencing mathematics as something that causes negative feelings.

A cultural aspect is involved here. In collectivist cultures, a well-performing student may have the same problem of negative development: for example, in Lee's study (2009) it is seen that for students in Korea, the degree of anxiety is remarkably high. However, in collectivist cultures students orientate to community rather than personal aspirations (Markus & Kitayama, 1991). For this reason, the individual level of affect might become less central (Green, 1971) and thus the negative development of that level less harmful. A student might be able to acquire mathematics-related studies even if he or she feels anxious about the subject. However, the individual level of affect is likely to be more significant for Finnish students. For example, Tuohilampi and Hannula (2013) showed that Finnish students who have negative individual affect avoid mathematics in their choices about the future, even if their performance level is good. If such is the case, the achieved performance level becomes irrelevant to the students. It is particularly important to pay attention to the negative development of individual-level affect and to find solutions to prevent this harmful phenomenon.

In the fifth study, some reasons could be found for a negative development. In this study it became clear that Finnish mathematics classes are mostly experienced as non-activating or de-activating. If there is activation, it rarely relates to any emotional aspect. In addition, Finnish pupils very seldom experience their mathematics classes as emotionally positive or inspiring. Without any emotional bond, it is difficult to imagine mathematics becoming significant to pupils. On the other hand, those pupils who develop a negative emotional bond (boredom, despair), tend to feel that mathematics is significant, albeit in a negative way. The proportion of drawings picturing an emotionally boring mathematics class was alarming, especially within the control group pupils. Pekrun (2006) argues that boredom is the result of a lack of control in some cases, but it is possible that the feeling of boredom itself can also reduce the feeling of control. The more physical and emotional a feeling becomes (resulting in sleepiness), the more pupils need to put an effort into controlling it.

One must also remember that it was not necessarily the drawer her/himself who was bored or had more or less intense negative expectations. In some cases, the drawer had put these thoughts into the minds of her/his classmates. This means that the drawer her/himself did not necessarily share these negative thoughts. However, because of her/his affective learning climate, she/he will be subject to experiencing negative social responses, possibly deconstructive ones, about mathematics learning. In the long run, this may result in a measurable negative development regarding the individual level of mathematics-related affect. Luckily, boredom was significantly less present in the intervention pupils' drawings, so it seems like the monthly activities and the teachers' professional development during the intervention succeeded in increasing the subjective experience of control (Pekrun et al 2010), and possibly inspire enough positive responses to help the students' self-concept remain positive (Harter, 1999).

The intervention could not make a strong impact on the individual level directly. In the fourth study, we suggested that the effects of the intervention may have partly become disguised by other features in school that were more significant to the pupils. Chapman (2002) has shown that a significant conflict is needed to allow affect structure to become re-organized. Thus, even if the pupils in the intervention may have received more positive experiences than those in the control group, those experiences might not become significant and effective enough. The analysis we made regarding the activation expressed in the drawings helped to interpret experiences during the intervention. If there were emotionally significant experiences, they were most likely to be negative. The positive experiences were mostly limited to cognitive experiences, which might not be significant enough to create mind-changing conflict, especially if too many emotionally negative experiences have preceded them. On the other hand, there might have been a contradiction between a regular mathematics lesson and the intervention lesson. In that case, it is possible that pupils considered that the two different types of lessons were two totally distinct ways of learning mathematics, and the more traditional idea of mathematics learning dominated in the individual level measurement (in the form of a traditional paper and pencil questionnaire).

The individual-level impact of the intervention addressed the girls, helping them to maintain their enjoyment and self-efficacy on a more positive level compared to the control group girls. Female pupils feel less confident with mathematics in general (Syzmanowicz & Furham, 2011), but on the other hand they estimate that their verbal intelligence is higher than that of males (ibid.). For this reason, girls may find it helpful to work in co-operation with others: Hannula, Kupari, Pehkonen, Räsänen & Soro (2004) have argued that a collaborative atmosphere and collaborative learning methods are connected with increasing self-confidence and improved mathematical performance especially regarding girls. Thus, the benefit the intervention gave to girls might have come through an increase in collaboration. Girls also differ from boys in their interests, as boys tend to be more oriented towards technical aspects of science, whereas girls are likely to show more interest in human issues (Sjøberg and Schreiner, 2010). It is possible that there are different learning styles between genders, and the non-competitive context impacted more on girls' style.

Laura Tuohilampi

12 Concluding words

The study as a whole shows that students' mathematics-related affect develops in an unnecessarily negative way. Even though students' affect might not harm their performance, the negative affective bond with mathematics may result in students giving up mathematics. The centrality of beliefs increases the negative development of mathematics-related effect. One cannot argue that negative feelings are just part of learning; that learning is not supposed to be fun. The question is not about mathematics learning being fun or not, it is about experiencing mathematics in such a way that either makes or does not make the emotional bond easy to carry. If the bond becomes too negative, it becomes necessary for the student to give up this negative bond, and thus leave mathematics behind. In such a case, the achieved performance level becomes irrelevant to the students, and the resources put into educating children in mathematics are wasted. The cultural aspect makes Finnish students follow their individual level of affect more than the interindividual level. Finland has been at the top of mathematics performance assessments (OECD, 2010), but without promoting a positive emotional bond to learn and use mathematics these good results are in risk to become useless. This is why it is particularly important in Finland that students themselves feel the need to engage with mathematics. It is crucial to find ways to stop the negative development to not waste the good outcomes in mathematics.

Based on the results of this thesis, it is suggested to give more emphasis to students' control over their learning, to allow students to engage in social interaction and to increase their opportunities to work with open-ended problems in mathematics. These recommendations have been already given by several researchers, but what have been the implications of such recommendations? Back in 1992 McLeod claimed that when students work with mathematics, they are less likely to work in groups, to develop their research skills, and to work on tasks that are compatible with the development of higher-order thinking skills than they are when working with social studies. McLeod also claimed that in mathematics classrooms students spend a lot of time alone doing "seatwork", and that students' view mathematics as a skill-oriented subject. Since those days, more could have changed. Students still see themselves as unconfident with mathematics and find encountering it emotionally boring. This makes students unlikely to develop an engaging emotional bond with mathematics.

The beliefs students generate reflect their experiences of the typical classroom situations where they encounter mathematics. As McLeod (1992, p. 579) put it: "There is nothing wrong with the students' mechanism for developing beliefs about mathematics; what needs to be changed is the curriculum, (and beyond that, the culture) that encourages such beliefs."

References

Andrews, P. & Diego-Mantecón, J. (2014). Instrument adaptation in cross-cultural studies of students’ mathematics-related beliefs: Learning from healthcare research. *Compare: A Journal of Comparative and International Education*. DOI:10.1080/ 03057925.2014.884346

Bernelius, V. (2013). *Eriytyvät kaupunkikoulut: Helsingin peruskoulujen oppilaspohjan erot, perheiden kouluvalinnat ja oppimistuloksiin liittyvät aluevaikutukset osana kaupungin eriytymiskehitystä*. A Dissertation. University of Helsinki: Finland.

Berry, J. W. (1989). Imposed etics-emics-derived etics: The operationalization of a compelling idea. *International Journal of Psychology, 5*, 165–171.

Bofah, E., & Hannula, M. (2014). Structural Equation Modelling: Testing for the Factorial Validity, Replication and Measurement Invariance of Students' Views on Mathematics. In *SAGE Research Methods Cases*. London, United Kingdom: SAGE Publications, Ltd. DOI: 10.4135/978144627305014529518

Bronfenbrenner, U. (1993). The ecology of cognitive development: Research models and fugitive findings. In. R. H. Wozniak & K. W. Fischer (Eds.), *Development in context: Acting and thinking in specific environments. The Jean Piaget symposium series*. (pp. 3– 44). Hillsdale, NJ, England: Lawrence Erlbaum Associates.

Chamberlin, S. A. (2010). A review of instruments created to assess affect in mathematics. *Journal of Mathematics Education, 3*(1), 167–182.

Chapman, O. (2002). Beliefs structure and inservice high school mathematics teacher growth. In G. C. Leder, E. Pehkonen, & G. Törner (Eds.*), Beliefs: A Hidden Variable in Mathematics education?* (pp. 177-193). Kluwer Academic Publishers, Netherlands.

Clarke, D. (2013). The validity-comparability compromise in crosscultural studies in mathematics education. In B. Ubuz, Ç. Haser & M. A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 1855–1864). Antalya, Turkey: ERME.

Clarke, D. (2015). Faith, hope and charity: Theoretical lenses on affect. In B. Pepin, B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education*. (pp. 119-134). Switzerland: Springer.

Davis, M., & Suveg, C. (2014). Focusing on the positive: a review of the role of child positive affect in developmental psychopathology. *Clinical child and family psychology review, 17*(2), 97-124.

Di Martino, P., & Zan, R. (2015). The construct of attitude in mathematics education. In B. Pepin, B. Roesken-Winter (Eds.*), From beliefs to dynamic affect systems in mathematics education*. (pp. 269-277). Switzerland: Springer.

Dowker, A. (2009). *What Works for Children with Mathematical Difficulties?* University of Oxford. Retriewed from: http://www.nationalnumeracy.org.uk/resources/22/index.html

Evans, J. (2006). Affect and Emotion in Mathematical Thinking. In J. Maasz, & W. Schloeglmann (Eds.), *New mathematics education and practice* (pp. 233-255). Sense Publishers: Netherlands.

Evans, I. M., Harvey, S. T., Bucley, L. & Yan, E. (2009). Differentiating classroom climate concepts: academic, management, and emotional environments. *New Zealand Journal of Social Sciences, 4*(2), 131–146.

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman mathematics attitudes scales. *JSAS Catalog of Selected Documents in Psychology, 6*, 31 (Ms. No. 1225).

Green, T. F. (1971). *The activities of teaching*. Tokyo: McGraw-Hill Kogakusha.

Hannula, M. S. (2011). The structure and dynamics of affect in mathematical thinking and learning. In M. Pytlak, T. Rowland, & E. Swoboda (Eds.), *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education* (pp. 34-60). Rzeszów, Poland: ERME.

Hannula, M. S. (2012). Exploring new dimensions of mathematics-related affect: embodied and social theories. *Research in Mathematics Education 14*(2), 137-161.

Hannula, M. S. (2015). The relevance of affective systems and social factors: a commentary. In B. Pepin, B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education*. (pp. 269-277). Switzerland: Springer.

Hannula, M. S. & Laakso, J. (2011). The structure of mathematics related beliefs, attitudes and motivation among Finnish grade 4 and grade 8 students. In B. Ubuz (Ed.), *Proceedings of the 35th conference of the international group for the psychology of mathematics education, Vol. 3*, (pp. 9–16). Ankara, Turkey: PME.

Hannula, M. S., Kupari, P., Pehkonen, L., Räsänen, P., & Soro, R. (2004). Matematiikka ja sukupuoli [Mathematics and gender]. In P. Räsänen, P. Kupari, T. Ahonen, & P. Malinen (Eds.*), Matematiikka – näkökulmia opettamiseen ja oppimiseen [Mathematics – percpectives to teaching and learning]*. Jyväskylä, Finland: The Niilo Mäki Institute.

Hannula, M. S., Maijala, H., Pehkonen, E., Nurmi, A. (2005). Gender comparisons of pupils' self-confidence in mathematics learning. *Nordisk matematikkdidaktikk, NOMAD10* (3-4), 29-42.

Hannula, M.S., Op ‘t Eynde, P., Schlôglmann, W & Wedege, T. (2007). Affect and mathematical thinking. In D. Pitta – Pantazi & G. Philippou (Eds.) *European Research in Mathematics Education V, Proceedings of the Fifth Congress of the European Society for Research in Mathematics Education, Larnaca, Cyprus 22 - 26 February 2007*, pp. 202 – 208. Department of Education: University of Cyprus.

Harter, S. (1999). *The construction of the self. A developmental perspective*. New York, NY: The Guildford Press.

Harzing, A.-W. (2006). Response styles in cross-national survey research: A 26-country study. *The International Journal of Crosscultural Management, 6*(2), 243–266

Hirvonen, K. (2012*). Onko laskutaito laskussa? Matematiikan oppimistulokset peruskoulun päättövaiheessa 2011. [Are calculating skills decreasing? Mathematics learning outcome at the end of the comprehensive school at 2011.]* Helsinki: The National Board of Education.

Hofstede, G. & Hofstede, G. J. (2005). *Cultures and organizations. Software of the mind*. New York, NY: McGraw-Hill.

Joutsenlahti, J., & Vainionpää, J. (2010). Oppimateriaali matematiikan opetuksessa ja osaamisessa [Learning material in mathematics teaching and learning]. In E. K. Niemi & J. Metsämuuronen (Eds.), *Miten matematiikan taidot kehittyvät? Matematiikan oppimistulokset peruskoulun viidennen vuosiluokan jälkeen vuonna 2008. [How mathematics skills develop? Mathematics learning outcomes after grade 5 in comprehensive school at 2008.]* (pp. 137–148). Helsinki: The National Board of Education.

Laine, A., Näveri, L., Pehkonen, E., Ahtee, M., Heinilä, L. & Hannula, M. S. (2012). Third-graders’ problem solving performance and teachers’ actions. In T. Bergqvist (Ed.), *Proceedings of the ProMath meeting in Umeå* (pp. 69–81). University of Umeå.

Leder G. C. (1995). Equity inside the mathematics classroom: Fact or artifact? In W. G. Secada, E. Fennema & L. B. Adaijan (Eds.), *New directions for equity in mathematics education*. Cambridge: University Press.

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences 19,* 355-365.

Lerman, S. (2000). The social turn in mathematics education research. *Multiple perspectives on mathematics teaching and learning*, 19-44.

Ma, X., & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for research in mathematics education*, 26-47.

Markus, H. & Kitayama, S. (1991). Culture and the self: implications for cognition, emotion and motivation. *Psychological Review, 98*(2), 224–253.

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grows (Ed.), *Handbook of research on mathematics learning and teaching* (pp. 575–596). New York, NY: Macmillan.

Metsämuuronen, J. (2010). Osaamisen ja asenteiden muutos perusopetuksen 3.-5. luokilla. [The changes of attainment and attitudes from grade 3 to grade 5 of comprehensive education. In E. K. Niemi & J. Metsämuuronen (Eds.*), Miten matematiikan taidot kehittyvät? [How do the mathematics skills develop?]* (pp. 93–136). Helsinki: National Board of Education.

Metsämuuronen, J. (2012). Challenges of the Fennema-Sherman Test in the International Comparisons. *International Journal of Psychological Studies, 4*(3), 1-22.

Metsämuuronen, J. (2013). Matemaattisen osaamisen muutos perusopetuksen luokilla 3-9 [The change in mathematics skills on comprehensive grades 3-9]. In: J. Metsämuuronen (Ed.). *Perusopetuksen matematiikan oppimistulosten pitkittäisarviointi vuosina 2005–2012 [The longitudinal assessment of the comprehensive mathematics in years 2005-2012].* National board of education. Finland.

Midgley, C., Kaplan, A., Middleton, M., Maehr, M. L., Urdan, T., Anderman, L. H., ... & Roeser, R. (1998). The development and validation of scales assessing students' achievement goal orientations. *Contemporary educational psychology, 23*(2), 113-131.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L. Freeman, K. E., ... Urdan, T. (2000*). Manual of the patterns of adaptive learning scales*. University of Michigan. Retrieved from http://www.umich.edu/~pals/PALS%202000_V13Word97.pdf.

Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Nett, U. E., Goetz, T., & Hall, N. C. (2011). Coping with boredom in school: An experience sampling perspective. *Contemporary Educational Psychology 36*, 49–59.

Niemi, E. (2010). Matematiikan oppimistulokset 6. vuosiluokan alussa. [Mathematics learning outcome at the beginning of 6th grade]. In E. K. Niemi & J. Metsämuuronen (Eds.), *Miten matematiikan taidot kehittyvät? Matematiikan oppimistulokset peruskoulun viidennen vuosiluokan jälkeen vuonna 2008. [How mathematics skills develop? Mathematics learning outcomes after grade 5 in comprehensive school at 2008.]* (pp. 17-70). Helsinki: The National Board of Education.

OECD (2010). PISA 2009 Results: *What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I).*

OECD (2011*). Education at a Glance 2011: OECD indicators*, OECD Publishing. Retrieved from http://www.oecd.org/education/skills-beyond-school/48631582.pdf.

OECD (2012). *Education at a Glance 2012: OECD Indicators*, OECD Publishing. Retrieved from http://www.uis.unesco.org/Library/Documents/education-at-a-glance-oecd-indicators-2012-en.pdf.

OECD (2014). *PISA 2012 Results in Focus: What 15-year-olds know and 2 what they can do with what they know*, OECD Publishing.Retrievew from: http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

Op' t Eynde, P., de Corte, E., & Verschaffel, L. (2002). Framing students' mathematics-related beliefs. In G. C. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A Hidden Variable in Mathematics education?* (pp. 13-37). Kluwer Academic Publishers: Netherlands.

Pehkonen, E. (1994). On differences in pupils' conceptions about mathematics teaching. *The Mathematics Educator, 5*(1), 3–10.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist, 37*(2), 91–106.

Portaankorva-Koivisto, P. (2015). Chilean and Finnish Teachers' Conceptions on Mathematics Teaching. A conference submission.

Ramírez, M. (2005). Attitudes toward mathematics and academic performance among Chilean 8th graders. *Estudios Pedagógicos, 31*, 97–112.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology, 104*, 700-712. DOI:10.1037/a0027268

Rimm-Kaufman, S. E., Larsen, R. A. A., Baroody, A. E., Curby, T. W., Ko, M., Thomas, J. B., Merritt, E. G., Abry, T., & Jamie DeCoster, J. (2014). Efficacy of the Responsive Classroom Approach: Results From a 3-Year, Longitudinal Randomized Controlled Trial. *American Educational Research Journal 51*(3), 567–603. DOI: 10.3102/0002831214523821

Roesken, B., Hannula, M. S. & Pehkonen, E. (2011). Dimensions of students' views of themselves as learners of mathematics. *ZDM, 43*(4), 497–506.

Schoenfeld, A. H. (2012). *How we think: a theory of human decision-making, with a focus on teaching*. 12th International Congress on Mathematical

Education. Retrieved from: http://www.icme12.org/upload/submission/1900_f.pdf

Sjøberg, S., & Schreiner, C. (2010). *The ROSE project. Overview and key findings.* Retrieved from: http://roseproject.no/network/countries/norway/eng/nor-Sjoberg-Schreiner-overview-2010.pdf

Syzmanowicz, A. & Furnham, A. (2011). Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta analyses. *Learning and Individual Differences 21*(5), 493–504.

Tapia, M. & Marsh, G. E., II. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly, 8*(2), 16–21.

Tapia, M. & Marsh, G. E., II. (2005). Attitudes towards mathematics inventory redux. *Academic Exchange Quarterly, 9*(3), 272–275.

Thapa, A., Cohen, J., Guffey S., and Higgins-D' Alessandro, A. (2013). A Review of School Climate Research. *Review of Educational Research 83*(3), 357-385. DOI: 10.3102/0034654313483907

Tikkanen, P. (2008). *Helpompaa ja hauskempaa kuin luulin. Matematiikka suomalaisten ja unkarilaisten perusopetuksen neljäsluokkalaisten kokemana [Easier and more fun than I thought. Mathematics experienced by fourth-graders in Finnish and Hungarian comprehensive schools].* Jyväskylä Studies in Education, Psychology and Social Research 337. Jyväskylä, Finland: Jyväskylä University Printing House.

Tuohilampi, L. & Hannula, M. S. (2011). High expectations, low confidence – discrepancy in self-image as a reason for displeasure in mathematics. In B. Ubuz (Ed.), *Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education, Vol. 1,* (p. 406). Ankara, Turkey: PME.

Tuohilampi, L. & Hannula, M. (2013). Matematiikkaan liittyvien asenteiden kehitys sekä asenteiden ja osaamisen välinen vuorovaikutus 3., 6. ja 9. luokalla [The development of mathematics-related attitudes and the interaction between attitudes and performance at 3rd, 5th and 9th grade]. In J. Metsämuuronen (Ed.*). Perusopetuksen matematiikan oppimistulosten pitkittäisarviointi vuosina 2005–2012 [The longitudinal assessment of the comprehensive mathematics in years 2005-2012].* National board of education. Finland.

Turner, J. C., Midgley, C. Meyer, D. K., Gheen, M., Anderman, E. M., Kang, Y., & Patrick, H. (2002). The classroom environment and students' reports of avoidance strategies in mathematics: A multimethod study. *Journal of Educational Psychology, 94*(1), 88-106. DOI: http://dx.doi.org/10.1037/0022-0663.94.1.88

Valenzuela, J.P., Bellei, C. & De Los Ríos, D. (2009). Evolución de la segregación socioeconómica de los estudiantes chilenos y su relación con el financiamiento compartido [Evolution of the socioeconomic segregation of Chilean pupils and their relation to the shared funding policy]. In *Evidencias para Políticas Públicas en Educación,* (pp. 231-284). FONIDE, Ministerio de Educación.

Wong, N. (2008). Confucian heritage culture learner's phenomenon: from "exploring the middle zone" to "constructing a bridge". *ZDM—International Journal of Mathematics Education, 40,* 973–981.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for research in mathematics education,* 458-477.

Zhu, Y. & Leung, F. S. (2011). Motivation and achievement: Is there an East Asian model? *International Journal of Science and Mathematics Education, 9*, 1189–1212.