University of Helsinki, Department of Teacher
Education, Research Reports series 390

Emmanuel Adu-tutu Bofah

A cross-cultural analysis of the dimensions of mathematics-related affect

Assessing the psychometric properties and the relationship with achievement

To be presented, with the permission of the Faculty of Behavioural Sciences of the University of Helsinki, for public discussion in the Auditorium XII, the university main building (Fabianinkatu 33), on Saturday May 21, 2016 at 10 o'clock

Helsinki 2016

Reviewed by
Professor Dr. Stanislaw Schukajlow-Wasjutinski, University of Münster, Münster Germany
Professor Petri Nokelainen, Tampere University of Technology, Tampere, Finland

Custos
Professor Markku Hannula, University of Helsinki

Supervised by
Professor Markku Hannula, University of Helsinki
Docent Anu Laine, University of Helsinki

Official Opponent
Professor Paul Andrews, University of Stockholm

Unigrafia, Helsinki

ISBN 978-951-51-2055-7 (paperback)
ISBN 978-951-51-2056-4 (PDF)
University of Helsinki, Faculty of Behavioural Sciences

Department of Teacher of Education
Research Report 390

Emmnauel Adu-tutu Bofah

A cross-cultural analysis of the dimensions of mathematics-related affect
Assessing the psychometric properties and the relationship with achievement

Abstract
The aim of the present set of studies in this dissertation was to examine the psychometric properties of measures of mathematics-related affect developed and normalized in one culture for use in another, how these properties transcend cross-culturally in an African context, and the methodological challenges associated with the process. Further aims were to examine the relationships between these constructs on a cross-cultural level, and to explore any associations between students' background variables and mathematics achievement. With these aims in mind, we conducted four original empirical studies based on different types of structural equation modeling.

Studies I and II explored the problems of importing an instrument from one culture into another, and the associated methodological challenges. More specifically, Study I gives a detailed account of the processes involved in applying structural equation modeling to validate mathematics-related affective measures developed in one culture (Finland) for use in another (Ghana). Reliability estimates and confirmatory factor analyses indicated that the Ghanaian data set did not fit the original hypothesized model (seven-factor structure). A series of factor and confirmatory factor analyses indicated a four-factor structure for the Ghanaian sample. Study II examined the possible causes of the differences in the factor structures from a cross-cultural perspective. The results indicate that measurement artifacts, cultural differences, and construct validity and adaptability were possible causes of the observed differences in factor structure between the Ghanaian and the theoretical model. In conclusion, it is suggested that researchers should be aware of construct importation and adaptation, and of the fact that measurement errors, question order, negatively worded item, translation, and content overlap may influence the reliability and validity of survey measures. Moreover, it is necessary to consider cultural variation and the methodological approaches involved in the theoretical settings in order to make any meaningful comparative assessment. Researchers focusing on cross-cultural mathematics-related affect are recommended to acquire the theoretical and practical knowledge necessary to address these issues using appropriate tools such as structural equation modeling.

Study III investigated the psychometric properties (factor structure, reliabilities, method effect, and measurement invariance—country and gender) of the mathematics-related affective constructs used in the 2011 Trends in International

Mathematics and Science Study (TIMSS 2011) across the five participating African countries. It also examined the relationship between these mathematically related affective constructs, as well as the associations amongst the constructs, and between the students' background variables and mathematics achievement cross-culturally. The results empirically support the multidimensionality of the construct, and the measures were largely invariant across the five educational systems and gender. There was also some evidence to suggest that negatively worded items in a construct can attenuate the reliability and validity of the measures: a series of confirmatory factor analyses (CFAs) revealed the need to control for the method effects associated with such items. It seems that responses to negatively worded items differ systematically across countries and are systematically linked to students' achievement. The analyses also covered the relationships between the student affect, achievement, and background variables such as parental education, gender and students' educational aspirations. Lower mathematics achievement was associated with students from countries reporting higher motivational belief. On the association between students' mathematics-related affect, achievement, parental education, educational aspirations and gender, it seems that there are culture-specific as well as cross-cultural universal outcomes. For instance, the relationship linking parental involvement, teacher responsiveness and mathematics achievement turned out to be culture-specific whereas, boys rated their mathematics competence more highly than girls did. Parental education, gender and long-term educational aspirations also influenced student achievement and motivation, but to different extents in different countries.

Study IV, based on non-recursive structural equation models, tested theoretical and methodological models of the reciprocal relationship between mathematics-related affect (e.g., self-concept) and achievement using the TIMSS 2011 cross-sectional data set. The results in different countries support the existence of a unidirectional influence of affect on achievement and of achievement on affect, and a direct feedback-loop relationship between affect and achievement. According to the evidence, the reciprocal determinism between affect and achievement is dependent on the national context. Moreover, there was a cross-cultural universal pattern among males: reporting a stronger mathematics self-concept and higher long-term educational aspirations significantly predicted higher mathematics achievement. The effects of socioeconomic status on mathematics achievement and self-concept among the students were dependent on the national context.

Keywords: Africa, cross-cultural research, instrument adaptation, mathematics-related affect, mathematics achievement, Trends in International Mathematics and Science Study, TIMSS, structural equation modeling, method effect, measurement invariance

University of Helsinki, Faculty of Behavioural Sciences
Department of Teacher of Education
Research Report 360

Emmanuel Adu-tutu Bofah

Kulttuurien välinen tutkimus matematiikka-asenteen ulottuvuuksista
Asennemittarien psykometriset ominaisuudet sekä asenteen ja osaamisen välinen suhde

Tiivistelmä
Tämän väitöskirjan tavoitteena oli selvittää matematiikan asennemittareiden psykometrisiä ominaisuuksia, kun niitä käytettiin toisessa kulttuurisessa kontekstissa kuin missä ne alun perin oli kehitetty, miten nämä ominaisuudet eroavat afrikkalaisten maiden välillä, sekä prosessiin liittyviä menetelmällisiä haasteita. Tavoitteena oli myös vertailla miten asennetekijöiden väliset suhteet erosivat eri maiden välillä sekä selvittää oppilaan taustamuuttujien yhteyksiä hänen matematiikan osaamisensa. Näiden tavoitteiden saavuttamiseksi teimme neljä empiiristä tutkimusta käyttäen erilaisia rakenneyhtälömallintamisen menetelmiä.

Tutkimuksissa I ja II selvitimme mittarin siirtämistä kulttuurista toiseen sekä siirtoon liittyviä menetelmällisiä haasteita. Tutkimuksessa I selvitettiin yksityiskohtaisesti miten rakenneyhtälömallintamisella voidaan validoida Suomessa kehitetyn matematiikka-asennemittarin käyttöä Ghanassa. Reliabiliteetin estimaatit ja konfirmatorinen faktorianalyysi osoittivat, että Suomessa löydetty seitsemän faktorin malli ei sopinut ghanalaiseen aineistoon. Sarja exploratiivisia ja konfirmatorisia faktorianalyysejä osoitti, että neljän faktorin malli sopii ghanalaiseen aineistoon. Tutkimuksessa II selvitimme kulttuurien välisenä vertailuna faktorirakenteiden erojen mahdollisia syitä. Tulokset osoittavat, että mittarivirheet, kulttuurierot sekä mittarin heikko käsitevaliditeetti sekä heikko käännös ja sovittaminen olivat mahdollisia syitä alkuperäisen mallin ja ghanalaiseen aineistoon pohjautuvan mallin välillä havaittuihin eroihin. Johtopäätöksenä ehdotamme, että tutkijoiden tulee huomioida käsitteiden alkuperä ja mittarin kääntäminen ja sovittaminen uuteen kontekstiin. Kun mittari siirretään kulttuurista toiseen, on huomattava mittausvirheiden, väittämien järjestyksen, väittämien negatiivisen muotoilun, käännöksen ja sisällöllisten päällekkäisyyksien mahdolliset vaikutukset kyselymittareiden luotettavuuteen ja pätevyyteen. Tämän lisäksi on otettava huomioon tutkimuksen teoreettisen viitekehyksen tunnustama kulttuurien välinen vaihtelu, mikä pitää huomioida valituissa menetelmissä, jotta voidaan tehdä kulttuurien välistä vertailua. Suosittelemme, että tutkijat, jotka paneutuvat matematiikka-asenteiden kulttuurien välisiin eroihin, hankkivat teoreettisen ja

käytännöllisen tiedon käsitellä näitä ongelmakohtia rakenneyhtälömallintamisella tai muilla soveltuvilla työkaluilla.

Tutkimuksessa III selvitettiin vuoden 2011 *Trends in International Mathematics and Science Study* (TIMSS 2011) –tutkimuksessa käytettyjen matematiikan asennetekijöiden psykometrisiä ominaisuuksia (faktorirakenne, luotettavuus, metodiefekti ja mittarin invarianssi maan ja vastaajan sukupuolen suhteen) viidessä tutkimukseen osallistuneessa Afrikan maassa. Tutkimuksessa selvitettiin myös matematiikkaan liittyvien eri asennetekijöiden välisiä yhteyksiä sekä oppilaan matematiikan osaamisen ja taustamuuttujien vaikutusta asennetekijöihin ja niiden välisiin yhteyksiin. Kulttuurien välinen vertailu tukee olettamusta asenteiden moniulotteisuudesta. Mittarit olivat pääsääntöisesti invariantteja viiden tutkitun koulutusjärjestelmän sekä vastaajien sukupuolen suhteen. Havainnot viittaavat siihen, että negatiivisesti muotoillut väittämät saattavat heikentää käsitteen mittaamisen luotettavuutta ja pätevyyttä. Tehdyt konfirmatoriset faktorianalyysit vahvistivat että negatiivisiin väittämiin liittyvä metodiefekti tulee huomioida. Maiden välillä oli systemaattisia eroja negatiivisiin väittämiin vastaamisessa ja negatiivisilla väittämillä havaittiin myös systemaattinen yhteys oppiaan osaamiseen. Analysoimme myös oppilaiden asenteiden, osaamisen sekä sukupuolen, koulutustavoitteiden ja vanhempien koulutuksen kaltaisten taustamuuttujien väliset yhteydet. Mitä parempi matematiikan osaaminen kussakin maassa oli, sen heikompi oli oppilaiden keskimääräinen motivaatio. Oppilaiden matematiikka-asenteiden, matematiikan osaamisen, vanhempien koulutuksen, koulutustavoitteiden sekä sukupuolen välillä havaitsimme sekä kulttuurispesifejä että kulttuurisesti universaaleja yhteyksiä. Esimerkiksi vanhempien tuen, opettajakontaktin ja matematiikan osaamisen väliset yhteydet osoittautuivat kulttuurispesifeiksi. Sen sijaan poikien käsitys omista matemaattisista kyvyistään vaikuttaa olevan universaalisti myönteisempi kuin tytöillä. Vanhempien koulutustaso sekä oppilaan sukupuoli ja koulutustavoitteet vaikuttivat oppilaan osaamiseen ja motivaatioon, mutta vaikutuksen voimakkuus vaihteli tutkittujen maiden välillä.

Tutkimuksessa IV testasimme matematiikka-asenteiden (mm. minäkuva) ja osaamisen välisiä teoreettisia ja menetelmällisiä vastavuoroisen riippuvuuden malleja käyttäen TIMSS 2011 –tutkimuksen poikittaisaineistoa. Eri maissa saadut tulokset antavat tukea kaikille mahdollisille eri malleille: yksisuuntaiselle vaikutukselle asenteista osaamiseen, vastakkaissuuntaiselle yksisuuntaiselle vaikutukselle, sekä myös suoralle takaisinkytkennälle asenteiden ja osaamisen välillä. Tulosten valossa asenteen ja osaamisen välisen suhde on luonteeltaan kulttuurispesifi. Miesten osalta havaitsimme kulttuurien rajat ylittävänä universaalina piirteenä, että myönteinen matemaattinen minäkuva ja korkeammat koulutustavoitteet ennustivat merkittävästi parempaa matematiikan osaamista. Sosioekonomisen taustan vaikutukset matematiikan osaamiseen ja matemaattiseen minäkuvaan puolestaan olivat erilaiset eri maissa.

Avainsanat: Afrikka, kulttuurien välinen vertailu, matematiikka-asenteet, matematiikan osaaminen, metodiefekti, mittarin invarianssi, mittarin kääntäminen, rakenneyhtälömalli, TIMSS-tutkimus

Acknowledgements

I am immensely proud of myself, but it is very true that I absolutely could not have done this alone. It all started during my Master's degree studies in 2009 when it was time to write my Master's thesis. I was introduced to Markku, and our friendship is the benefit I am reaping today.

Markku became my supervisor when I obtained my Master's degree, and he is now my mentor. He has not held me back. On the contrary, he has encouraged me to proceed with whatever I wanted to do, for which I am very grateful. I thank him for his expertise, systematic thinking, never-ending ideas and his most perceptive, unyielding, persevering and precise feedback that I could not escape. I am grateful for the many meetings in which we went repeatedly through my texts and reviews, even during the summer breaks. I have been extremely lucky to have a supervisor who cared so much about my work, and who responded to my questions and queries so promptly. You always found the time to propose consistently excellent improvements. Through your guidance and your belief in me and in my skills I embarked on an intellectual journey to places I would never have dreamed of exploring. You have turned me into a writer, tirelessly correcting my cumbersome and complex writing and thinking. I am also very grateful for all your trust and support in my academic pursuits. I consider it a great honor to have worked with you and to have been part of your team. Not only do you stand for everything I hope to be as a scientist and contributor to society, the fantastic and excellent response from everyone who comes into contact with you is indicative of a true leader and educator. I thank you for the time you have invested in my work, and for all the lessons you have taught me about being an educationist and a scientist.

I would also like to thank my co-supervisor Docent Anu Laine for her encouragement, insightful comments, and tough questions. Anu taught me how to present my research in very simple language. Thank you for your patience, encouragement, and immense knowledge, which motivated me throughout the PhD process. You gave me invaluable perspectives on my research when I needed them most. I also followed your example in terms of combining a family life and children with PhD education.

Completing this work would have been all the more difficult had it not been for the support and friendship of Markku and Anu thesis seminar group on mathematics education at the Department of Teacher Education. I would especially like to thank Garcia Moreno-Esteva for helping me with the proofreading and good suggestion. Grateful thanks are also due to the many mentors I met at various conferences in recent years, whose views, recommendations, and practices I have incorporated into the dissertation.

Conducting research in the classroom is not possible without the effort and help of the teachers and participating students. I am therefore very grateful to all the students who took part in the studies and the teachers who facilitated that. My deep gratitude goes to Kofi Ayebi-Arthur for helping me with the data collection in Ghana, and to Maame Yirenkyiwaa Adjei for helping in entering the data set into the statistical software. I would also like to thank Joan Nordlund for helping me with the language revision. I thank my friends (too many to list here but you know who you are!) for providing the support and friendship that I needed.

As we all know, life and education co-exist, and the immense support from my family and friends outside academia have helped me to advance on the inside. Moreover, I would not be the researcher I am today without my wife and children. I especially thank my mom, dad, and siblings, I am grateful to my beloved father, John Adututu Bofah, who has passed on: I wish he could have lived for another few months to see my graduation. My hard-working parents sacrificed their lives for me, and I would not have made it this far without them. I also wish to thank my family for supporting me spiritually throughout my life and throughout this journey. This PhD is a testament to your faith in me: I hope I have made you proud. I also gratefully acknowledge all the support and love received from my sister Christiana Bofah Agyeman and her husband throughout this process. My sister and her children have been my rocks, and I thank them for all their encouragement. I know I always had my family to count on when times were rough. I am also grateful to my wonderful "mother-in-law", Anna Yeboah for her strong support and for providing help with the children during this academic process, and to my in-laws Harry, Kwame, and Dennis for their immense logistical support. Your prayer for me was what sustained me thus far.

My strongest source of happiness during these years and the one for which I am most grateful is my best friend, partner, and wife Josephine. Josephine is the only person who appreciates my quirkiness and sense of humor. Not only do you make me a better person, your contributions to this work on a personal, social, and scientific level are incalculable. Josephine has unconditionally supported, encouraged and loved me during my good and bad times. She was instrumental in instilling hope and confidence in me. These past years have not been easy, academically or personally. I truly thank Josephine for sticking by my side even when I was irritable, confused and depressed. I had the support of your mind, your heart, and your hands whenever I needed it, and you never asked for anything in return. For that I can say that I am truly blessed. You shouldered far more than your fair share of the parenting and household duties while I pursued this final degree. My children, Emmanuella, Emil, Emphine and Elspeth, are a constant reminder both of my aspirations and of my limitations as a person.

This list of acknowledgements can only capture a small fraction of the people who supported my work. I send my deep thanks to all. Your contributions to this dissertation were vital, but the inevitable mistakes in it are very much my own.

Finally, I thank God for letting me through all the difficulties. You deserve the gratitude for this work. I will keep on trusting and believing in you. Thank you, Lord.

I dedicate this thesis to
my family, my wife, children and to the memory of my beloved father.
I love you all dearly.

Espoo, April 2016
Emmanuel Adu-tutu Bofah

Contents

List of empirical studies

Study I

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

Study II

Bofah, E. A., & Hannula, M. S. (2015). Studying the factorial structure of Ghanaian twelfth-grade students' views on mathematics. In B. Pepin & B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education: Exploring a mosaic of relationships and interactions* (pp. 355–381). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-06808-4_18

Study III

Bofah, E. A., & Hannula, M. S. (2015). TIMSS Data in an African Comparative Perspective: Investigating the Factors Influencing Achievement in Mathematics and their Psychometric. *Large-scale Assessments in Education, 3*(1), 4 doi: 10.1186/s40536-015-0014-y.

Study IV

Bofah, E. A. (2015). Reciprocal determinism between students' mathematics self-concept and achievement in an African context. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (CERME 9) (pp. 1688-1694). Prague, Czech Republic. https://hal.archives-ouvertes.fr/hal-01287995

1 Introduction

It is widely acknowledged that students' affective dispositions influence their learning of mathematics, for better or for worse (DeBellis & Goldin, 2006; Hannula, 2006, 2015a; Lim & Chapman, 2012; Xin Ma, Kishor, & Kisor, 2014; Xin Ma & Kishor, 1997; Xin Ma, 2011; McLeod, 1992, 1994; Tighezza, Tighezza, & M'hamed, 2014; Valentine, DuBois, & Cooper, 2004). In fact, mathematics is seen as a yardstick for Science, Technology, Engineering, and Mathematics (STEM) programs (Robinson, 2003; Sahin, Morgan, & Erdogan, 2012; Sells, 1973; Semela, 2010), and as a "powerful career filter" to higher pay and prestigious jobs (X. Ma & Johnson, 2008; Sells, 1973). The world is gradually becoming "quantified", and these students need to keep up with mathematical and technological thinking to live productive lives (Mullis, Martin, Foy, & Arora, 2012). Given the important role of mathematics-related affect in mathematics learning and instruction, most educators and researchers in the field would agree that, "if research on learning and instruction is to maximize its impact on students and teachers, affective issues need to occupy a more central position in the minds of researchers" (McLeod, 1992, p. 575). Various researchers (e.g., Chaman, Beswick, & Callingham, 2014; Hattie, 2009; McLeod, 1992; Stankov, Lee, Luo, & Hogan, 2012) further emphasize the need to maximize research in mathematics education through the integration of affective issues into the study of cognition and instruction.

How is mathematics-related affect in education measured? More often than not, it is via questionnaires. From where do these mathematics-related affective questionnaires derive? They appear to have their roots in "Western-mathematics-related affect research". The pioneering work and the first widely used construct on the assessment of mathematics-related affect involved the construction of the 98-item Mathematics Anxiety Rating Scale (MARS) (Richardson & Suinn, 1972), followed by several abbreviated versions (Fennema & Sherman, 1976; Sandman, 1980; Tapia & Marsh, 2004). Over the years, the Mathematics Attitude Scales (Fennema & Sherman, 1976) have become among the most influential in the field. The instrument comprises a total of nine scales, namely: (1) Mathematics Anxiety, (2) Confidence in Learning Mathematics, (3) Attitude Toward Success in Mathematics, (4) Effectance Motivation in Mathematics, (5) Mathematics as a Male Domain, (6) The Usefulness of Mathematics, (7) The teacher scale, and (8) and (9) Parental (mother/father) interest in mathematics.

The "Indiana Mathematics Belief Scale," which was designed for secondary school and college-level students (Kloosterman & Stage, 1992), has also been influential in the field. The instruments presented in the aforementioned studies and others inspired by them (Op't Eynde, De Corte, & Verschaffel, 2006; Roesken, Hannula, & Pehkonen, 2011) have been found useful in exploring the structural properties of mathematics-related beliefs.

Survey questionnaires are commonly used for studying students' mathematics-related beliefs. The Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) are examples of such surveys, and both have been used to measure student's mathematics-related belief systems worldwide (Bos & Kuiper, 1999; Frempong, 2010; House, 2006; Liu & Meng, 2010; Marsh et al., 2013; Zhu & Leung, 2011). The samplings associated with these large-scale surveys are robust. For instance, the International Association for the Evaluation of Educational Achievement (IEA), which organizes the TIMSS survey, has long and extensive experience of organizing large surveys. These surveys are 'heavenly data sets' for developing countries such as Ghana that lack the resources to organize any meaningful large-scale educational surveys of their own. One problem with these instruments is that they were developed in Western countries with relatively ethnically homogenous populations. Moreover, the approach associated with them relies on the unwarranted assumption that the structure of affect is cross-culturally invariant (van de Vijver & Leung, 2000).

Given the above concerns, the first aim of this thesis is to determine the psychometric properties (factor structure, reliabilities, method effect, and measurement invariance—country and gender) of measures of mathematics-related affect, as well as how these properties transcend cross-culturally and the methodological challenges associated with the process. Further aims are to examine the relationship between these mathematics-related affective dimensions from a cross-cultural perspective, and to determine how students' mathematics-related affect relates to their background variables and mathematics achievement.

There is a strong need for cross-cultural comparisons of mathematics-related affect among students to challenge the foundations of current theories and provide a heuristic basis on which to test the external validity and generalizability of measures, theories, and models (Chiu & Klassen, 2010; Herbert W. Marsh & Köller, 2004). More often than not the most challenging problem in existing studies concerns instruments that have been developed and normalized in one cultural setting (often very homogenous societies) and used in another, either in their original linguistic form or in a translated full or abridged version of the original instrument (Byrne, 2000). The new instrument tends to be adapted under the vague and unrealistic assumption of measurement equivalence across cultures. Cross-cultural researchers should understand that any differences in the outcomes of such surveys may be attributable to particular biasing effects in the data, or may directly relate to methodological problems and/or item bias (Herbert W. Marsh & Köller, 2004). Within this context, van de Vijver & Leung, (2000) and Byrne (2000) argued strongly for the use of multiple-group Structural Equation Modeling (SEM) in cross-cultural comparisons.

Furthermore, theories and studies associated with these constructs tend to be products of Western cultures and societies. Issues related to adaptation, construct validity, and outcome interpretation become problematic when the constructs are used in cultures with very heterogeneous populations (cf. Clarke, 2013). In fact, empirical studies have shown that the reliabilities of these adapted scales are, more often than not, highest in Western countries and lower in non-Western countries (e.g., Rutkowski & Rutkowski, 2010). F. F. Chen (2008), for instance, conducted a search in the PsycINFO database for cross-cultural studies carried out between 1993 and 2006, using "cross-cultural invariance" and other similar expressions as key words. According to the results, 74.2 percent of the average factor loadings were higher in the reference group (North America), where the instrument originated.

1.1 The affective domain and mathematics education

One problematic issue in the research on mathematics-related affect is the ambiguity in the terminology and definitions associated with the concept (Furinghetti & Pehkonen, 2002; Grootenboer & Hemmings, 2007; Hannula, 2012). McLeod (1992) proposed in his influential paper that mathematics-related affect could be conceptualized as a continuum: "… we can think of beliefs, attitudes, and emotions as representing increasing levels of affection involvement, decreasing levels of cognitive involvement, increasing levels of intensity of response, and decreasing levels of response stability" (p. 579). McLeod's (1992) framework identifies the following three constructs in the research: beliefs, attitudes, and emotions, further suggesting that students' mathematics belief systems comprise four structural qualities: (a) beliefs about mathematics, (b) beliefs about the self, (c) beliefs about mathematics teaching, and (d) beliefs about the social context. Later, DeBellis and Goldin (1997) proposed 'values' as the fourth construct.

Other authors have identified other mathematics-related affective domains such as motivation (Hannula, 2006) and identity (Frade, Roesken, & Hannula, 2010). However, Op't Eynde and colleagues (2002; 2006) further developed a framework of students' mathematics-related belief systems, clustering them as implicitly or explicitly held subjective conceptions students hold to be true, as follows: "Beliefs about mathematics education: (a) beliefs about mathematics, (b) beliefs about mathematical learning and problem solving, (c) beliefs about mathematics teaching; Beliefs about the self as a mathematician: (a) intrinsic goal orientation beliefs, (b) extrinsic goal orientation, (c) task-value beliefs, (d) control beliefs, (e) self- efficacy beliefs; Beliefs about the mathematics class context: (a) beliefs about the role and the functioning of the teacher, (b) beliefs about the role and the functioning of the students in their classes, (c) beliefs about socio-mathematical norms in their own class" (Op 't Eynde et al., 2006).

Although, McLeod's (1992) framework has been very influential in the research, terminological ambiguity is still a major challenge. Experts in the field of mathematics education have not been able to come up with unambiguous definitions of concepts such as attitudes and beliefs, as acknowledged in the literature (e.g., Furinghetti & Pehkonen, 2002). According to McLeod (1992, p. 590): "a major difficulty is that research on affect has not usually been grounded in a strong theoretical foundation" , and more recently Cretchley (2008, p. 147) supported this assertion: "this neglect is easy to understand, given the difficulties researchers face in this new field: theories not yet well-developed, terminology used differently and ambiguously, and varying research instruments, some untested, make the literature difficult to interpret, and leave researchers open to criticism."

Hannula (2011, 2015a, 2015b), in a critique of McLeod's framework highlighting the missing concepts, proposed a model of mathematics-related affect comprising three relatively stable overlapping constructs: motivation, cognition and emotion. Cognition deals with the self and the environment; motivation directs behavior; and emotions reflect goal-directed behavior, thereby providing a feedback system related to cognitive and motivational processes (2015a, p. 144). Motivation, for example, integrates goals, needs, and values, at the intersection of these three overarching constructs. A good example of this is the TIMSS 2011 motivational framework (Martin & Mullis, 2012) for large-scale assessment, in which the term motivational construct is used to describe students' views of their own mathematics skills: it encompasses *self-confidence*, a specific form of *self-concept* (often referred to as *competence belief*), *positive affect* (i.e. intrinsic motivation; students like learning mathematics), and *task value, including importance* (i.e. extrinsic motivation; students' value mathematics) (see also Marsh et al., 2013).

Hannula (2015b) also developed a meta-theory of affect in an attempt to link and contextualize theories of mathematics-related affect, proposing the following three levels of theorizing affect: as a social phenomenon, as part of individual psychology, and as a biological or physiological phenomenon. He distinguishes the *cognitive*, *emotional* and *motivational* affective traits from the more evolving changing state of the three mathematics-related dimensions. The *cognitive traits* include beliefs and other mental representations to which it makes sense to attribute a truth value" (p. 270), the *emotional traits* include emotional dispositions such as enjoyment and anxiety in doing mathematics, and the *motivational traits* include personal preferences and choices (e.g., needs, values, desires, and motivational orientations) (ibid).

What is missing from this affective construct in Hannula's framework is a link between mathematics-related affect and achievement-related performance and choice. With a view to narrowing this gap, it was decided in the present study to explore Eccles and her colleagues' modern Expectancy-value theory (EVT) (Eccles & Wigfield, 2002; Eccles, 1994, 2011a, 2011b; Eccles [Parsons] et al., 1983;

Eccles [Parsons], Adler, & Kaczala, 1982; Wigfield & Eccles, 2000). Self-belief is a critical component of this theory of motivation (Wigfield & Eccles, 2000), according to which motivational beliefs include anticipated interests that are likely to be experienced while carrying out various tasks or making activity choices, their attainment and utility value, and the anticipated psychological, economic and social costs. The model further takes into account the role of social structures/cultures/social class in determining the range of options available to individuals when they make educational and occupational decisions (Eccles, 2011b). More generally, Eccles' EVT takes into consideration the social and psychological factors that affect an individual's expectations of success and task value (Nagy, Trautwein, Baumert, Köller, & Garrett, 2006, p. 325).

The modern EVT further posits that individuals excel in subjects in which they expect to succeed or which they value (Leaper, Farkas, & Brown, 2012), and that students' competency beliefs and task value are positively related, directly influencing academic performance and goals (Jacobs & Eccles, 2000). Finally, parents' and teachers' expectations, and parental involvement, values, and beliefs are assumed to influence students' motivation, their goals (e.g., educational aspirations), their competence beliefs (e.g., self-confidence), and task values (e.g., value to mathematics) (Eccles, Jacobs, & Harold, 1990; Jacobs & Bleeker, 2004; Jacobs & Eccles, 2000; Wigfield & Eccles, 2000).

In a way, these two frameworks depict motivational belief as the outcome of social-structural influence and individual experience. Adopting Hannula and colleagues' model (2011, 2015a, 2015b) together with the EVT, this thesis examines the relationships between motivation, the emotional and cognitive trait of affect, and other background variables such as parental education, gender and students' long-term educational aspirations.

1.2 Cross-cultural differences in affect and achievement

Little is known at present about cross-cultural differences in affect and achievement in the African context. Cross-cultural comparison of belief systems is a problematic enterprise (Andrews, Diego-Mantecón, Eynde Op't, & Sayers, 2007; Andrews & Diego-mantecón, 2014; Clarke, 2013: see in particular Clarke's insightful paper (2013) on the "validity-comparability compromise"). The problems of conceptual measurement and linguistic invariance in survey instruments across cultures has been acknowledged and discussed in comparative research, for example (Andrews & Diego-mantecón, 2014; Osborn, 2004; van de Vijver & Leung, 2000).

However, cross-cultural invariance in affect has been demonstrated in most Western, Asian and some Arabic Countries (Andrews et al., 2007; Lee, 2009; Marsh & Hau, 2004; Marsh et al., 2013; D. P. Schmitt & Allik, 2005). J. Lee, (2009), for instance, showed that the factorial structure of self-beliefs (i.e., self-efficacy, self-concept, and anxiety) is the same cross-culturally, in other words the pattern of factor loadings can be generalized as measuring the same dimensions in different countries. Marsh, Abduljabbar and colleagues reported a similar outcome in Arabic-speaking and English-speaking Anglo-Saxon countries (Marsh et al., 2013). Lee (2009), in turn, found a cross-cultural variability in the relative strength of self-belief in predicting mathematics performance: the self-belief components in respect of self-efficacy and self-concepts were very low whereas a high level of anxiety was reported within the Confucian Asian countries under investigation. A stronger association between self-concept and achievement was found in Western European regions, whereas the strongest associations between self-efficacy and achievement were found in countries from Confucian Asia and Eastern Europe. With regard to anxiety, a stronger association with mathematics achievement was found among European countries, and a weaker association within Asian countries.

Marsh, Abduljabbar, et al. (2013), having compared latent mean differences in achievement and motivation, describe the results as "perplexing", students from Anglo-Saxon countries reporting substantially higher achievement than students from the Arab countries, but substantially lower motivation (e.g., the value of mathematics, liking mathematics, and self-concept) across all the measured constructs. However, "paradoxical" findings have been recorded whereby top nations in terms of mathematics performance (e.g., Chinese Taipei, Korea, Japan, Hong Kong, and the Netherlands) show relatively low levels of mathematics enjoyment and self-concept, as opposed to the relatively high levels reported in low-performing countries (e.g., South Africa, Ghana, Botswana, Morocco, Egypt, and Saudi Arabia) (Shen & Tam, 2008).

In conclusion, the findings revealed a negative relationship between self-belief and achievement with the country as a unit of analysis. In other words, students in countries with relatively high achievement levels tend to report lower mathematics self-belief, as opposed to those in countries with relatively low achievement who usually report higher self-belief (Shen & Tam, 2008; Shen & Pedulla, 2000). These findings were consistent in Grades 3, 4, 7 and 8 (Shen & Pedulla, 2000).

According to many comparative studies focusing on East Asian countries and the West, East Asian countries characterized as the Confucian Heritage Culture (Biggs, 1996, p. 46), perform better, and report lower self-belief (e.g., self-concept and self-confidence) and higher mathematics anxiety than their Western counterparts (Leung, 2002; Morony, Kleitman, Lee, & Stankov, 2013). Whereas some authors argue that the common values they share contribute to their similar higher performance (e.g., in TIMSS) (Zhu & Leung, 2011), others disagree on the grounds that most students in other countries also share the same values (Leung, 2002). In the case of mathematics-related affect, Leung (2002, p. 106) attributes the reported low mathematics confidence, for instance, "… to the stress in the cultures of these countries on the virtue of humility or modesty."

In-country studies on the relationship between mathematics achievement and affect have established differences across ethnic groups (Else-Quest, Mineo, & Higgins, 2013; Hyde, Lindberg, Linn, Ellis, & Williams, 2008) and gender (Hannula, Maijala, Pehkonen, & Nurmi, 2005; Meelissen & Luyten, 2008). Indeed, as Schoenfeld points out (1983, p. 330), in any classroom setting students' social-cognitive and metacognitive actions "…are often the result of consciously or unconsciously held beliefs about (a) the task at hand, (b) the social environment within which the task takes place, and (c) the individual problem-solver's perception of self and his or her relation to the task and the environment."

1.3 Mathematics related affect and achievement

Mathematics-related affective measures (e.g., self-concept) appears to be the best non-cognitive predictor of individual achievement in mathematics (Chaman et al., 2014; Hansford & Hattie, 1982; Hattie, 2009; Lee & Stankov, 2013; Lee, 2009; Stankov et al., 2012). It consistently stand out among the student variables in meta-analyses of predictors of academic achievement listed by Hattie (2009).

Other studies acknowledge the important role that other factors such as the school environment, school type, the role of the teacher, parents, day-to-day classroom interactions, and cultural differences play in students' mathematics-related affect (Andrews et al., 2007;Andrews & Diego-mantecón, 2014; Davis, 2003; Guo, Marsh, Morin, Parker, & Kaur, 2015; Ice & Hoover-Dempsey, 2011; Marsh et al., 2013; Op 't Eynde et al., 2002; Vukovic, Roberts, & Green Wright, 2013; Williams & Williams, 2010).

Student achievement in mathematics is also known to be influenced by a range of mathematics-related factors including self-confidence (Brown, Brown, & Bibby, 2008; M.-S. Chiu, 2009; Grootenboer & Hemmings, 2007; Reyes, 1984), self-concept (Marsh & Yeung, 1997; Nagengast & Marsh, 2012; Skaalvik & Skaalvik, 2008; Jianjun Wang, 2007; Wilkins, 2004), anxiety (Ashcraft, 2002; Cates & Rhymer, 2003; Frenzel, Pekrun, & Goetz, 2007; Hembree, 1990; Xin Ma & Xu, 2004a; Miller & Bichsel, 2004; Reyes, 1984), attitude (Hannula, 2002), self-efficacy (Hannula, Bofah, Tuohilampi, & Metsämuuronen, 2014; Lee, 2009; Pietsch, Walker, & Chapman, 2003; Skaalvik & Skaalvik, 2008; Williams & Williams, 2010), and motivational beliefs (e.g., how students value mathematics, and like learning mathematics) (Guo et al., 2015; Marsh et al., 2013; Reyes, 1984; Seaton, Parker, Marsh, Craven, & Yeung, 2014; Trautwein et al., 2012; Wigfield, 1994). Other factors include parental involvement (Gunderson, Ramirez, Levine, & Beilock, 2012; Jeynes, 2007; Tocci & Engelhard Jr, 1991; Wilkins, Zembylas, & Travers, 2002), parental education (C. Chen & Stevenson, 1995; P. E. Davis-Kean, 2005), the student's socio-economic background (Schulz, 2005; Sirin, 2005), teachers (Nye, Konstantopoulos, & Hedges, 2004), educational aspirations (Guo et al., 2015; Philip D Parker et al., 2012; Philip David Parker, Marsh, Ciarrochi, Marshall, & Abduljabbar, 2014), and gender (Kenney-Benson, Pomerantz, Ryan, & Patrick, 2006; Miller & Bichsel, 2004).

It is a known fact that students with positive views of themselves and their academic competences will engage in achievement-related activities to confirm their self-perceptions (Marsh, 1993; Pajares, Britner, & Valiante, 2000; Reyes, 1984; Rosenberg, Schoenbach, Carmi, & Rosenberg, 1995; Swann Jr., 1997; Valentine et al., 2004).

As indicated above, there is a strong relationship between students' affective disposition and performance, although the direction of causality between them is unclear. Students' beliefs or views on mathematics influence the choices they make and, significantly, constitute a determining factor in their problem-solving skills (Schoenfeld, 1985). Indeed, many affect-enhancement programs and educational-policy statements throughout the world are based on the fact that improvements in affect will lead to gains in achievement. There is also evidence of a reciprocal relationship between students' success and their mathematics-related affect (e.g., mathematics self-concept) (Calsyn & Kenny, 1977; Guay, Marsh, & Boivin, 2003; Hannula, Bofah, Tuohilampi, & Metsämuuronen, 2014; Marsh & Yeung, 1997; Parker et al., 2014; Seaton et al., 2014;Valentine et al., 2004).

The reciprocal determinism of affect and mathematics achievement has been the focus of a considerable amount of affect-related research given the practical and theoretical implications (e.g., Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005). The three commonly postulated axioms concerning the relationship between affect and achievement are: (a) there exists a unidirectional influence of affect on achievement; (b) there exists a unidirectional influence of achievement on affect; and (c) there is a reciprocal relationship between affect and achievement (Marsh et al., 2005).

Theories of affect (e.g., self-concept) extend across several branches (for a review, see Wang & Lin, 2008). One such branch concerns the "self", given that self-beliefs (e.g., self-concept) operate differently across cultures (M. M. Chiu & Klassen, 2010; Markus & Kitayama, 1991), and that the self is highly influenced by the frame-of-reference effect—social comparison, causal attribution, and reflected appraisals from significant others (Bong & Skaalvik, 2003; H W Marsh, 2007; Herbert W Marsh, Hau, & Wen, 2009), as well as standards against which to judge one's own traits and accomplishments (Bong & Skaalvik, 2003). Marsh and colleagues further argue that there is a strong frame-of-reference effect associated with students' self-belief (Marsh 2007; Marsh, et al., 2013). In most cases comparisons are made with reference to other students in the same school or class, hence country-level differences, for instance, may not be reflected in individual students' mathematics self-beliefs. However, a cross-cultural frame-of-reference effect has also been validated (e.g., Marsh & Hau, 2004), and additional evidence has been reported in Germany (Jansen, Schroeders, & Lüdtke, 2014; Moller, Streblow, Pohlmann, & Köller, 2006) and in some Arab- and English-speaking Anglo-Saxon countries (Marsh et al., 2013).

Another branch concerns the differing relationships between self-belief (e.g., self-concept) and achievement (Calsyn & Kenny, 1977; M.-S. Chiu, 2012a; Marsh et al., 2005), incorporating the contrasting self-enhancement and skill-development models (Calsyn & Kenny, 1977). The former posits that affect is a determinant of academic achievement (M. M. Chiu & Klassen, 2010; Marsh & O'Mara, 2008; Singh, Granville, & Dika, 2002; Valentine et al., 2004), whereas in the latter affect is simply seen as a reflection of performance (S.-K. Chen, Yeh, Hwang, & Lin, 2013; M. M. Chiu & Klassen, 2010; Xin Ma & Xu, 2004a, 2004b; Skaalvik & Valås, 1999; Jian Wang & Lin, 2008). However, a more realistic and logical compromise between the self-enhancement and the skill-development models is the reciprocal-effects model (Guay, Marsh, & Boivin, 2003; Hannula, Bofah, Tuohilampi, & Metsämuuronen, 2014; Marsh & Yeung, 1997; Parker et al., 2014; Seaton et al., 2014), which posits that prior affect influences subsequent achievement and prior achievement influences subsequent affect (Guay et al., 2003).

Valentine and colleagues found strong support for a reciprocal-effects model in their meta-analysis of self-belief measures (Valentine et al., 2004). They also showed that the effects of self-beliefs on achievement were strongest when the assessment focused on self-belief and achievement within the same domain, such as a particular subject area in school (e.g., mathematics achievement and mathematics self-concept). In other words, reciprocal determinism is stronger when it is based on domain-specific measures rather than global measures such as self-esteem (Jansen et al., 2014; Marsh et al., 2005).

Empirical tests have generally supported reciprocal-effects models in diverse populations including Germany (Marsh & Köller, 2004), Finland (Hannula et al., 2014), Canada (Guay et al., 2003) and Australia (Marks, McMillan, Hillman, & Australian Council for Educational Research [ACER], 2001), and many Organisation for Economic Co-operation and Development (OECD) countries (Marsh & O'Mara, 2008; Williams & Williams, 2010), but there appear to have been no studies on such models in the African context. The present work will thus add to the literature in narrowing this gap. However, some studies examining the relationship between motivational belief and achievement identify affect as the strongest predictor of achievement (Hattie, 2009; Marsh & O'Mara, 2008; Marsh et al., 2005; Stankov & Lee, 2014), whereas others imply that achievement is the strongest predictor of affect (Hannula et al., 2014; Xin Ma & Xu, 2004b), or that there is no causal relationship (e.g., Williams & Williams, 2010). All in all, no firm conclusions can be drawn about the causal ordering of mathematics-related affect (e.g., self-concept) and achievement on a cross-cultural level, possibly because the relationship between the two seems to be culturally specific.

1.4 Parental and teacher involvement, parental education and achievement

Studies focusing on various dimensions of parental and teacher behavior indicate that higher expectations, support, and encouragement from parents and teachers both at home and at school influence students' motivational beliefs and school performance (Desimone, 1999; W. Fan & Williams, 2010; X. Fan, 2001; Gonzalez-DeHass, Willems, & Holbein, 2005; J. Hughes & Kwok, 2007; Ice & Hoover-Dempsey, 2011; Jacobs & Eccles, 2000; Marchant, Paulson, & Rothlisberg, 2001). W. Fan and Williams (2010), for instance, found that eight percent of the variance in mathematics self-efficacy was attributable to parental involvement, whereas Campbell and Mandel (1990) found that the parental-involvement variables accounted for 22 percent of the variance in mathematics achievement. However, some children succeed in school without much parental involvement if other social actors such as the teacher can provide the necessary guidance and encouragement (Epstein, 2010). Moreover, parents' and teachers' affective dimensions and expectations are known to shape students' academic performance and expectations (Eccles [Parsons] et al., 1983; Jacobs & Eccles, 2000).

Parents and teachers are known to be strong "socio-actors" with various "social networks" that shape children's education as well as significantly influencing their educational performance (Astone & McLanahan, 1991; Eccles [Parsons] et al., 1983; Sheldon, 2002). The family social context has also been shown to influence parental involvement (Astone & McLanahan, 1991; Desimone, 1999; Sui-Chu & Willms, 1996a) and school performance. For instance, there appears to be an established relationship between the socioeconomic background and achievement of students and their self-beliefs (M. M. Chiu & Klassen, 2010; Howie, 2013; Lareau, 1987; Sirin, 2005; Sui-Chu & Willms, 1996a; Williams & Williams, 2010). Likewise, it has been found that parental education (C. Chen & Stevenson, 1995; P. E. Davis-Kean, 2005; Haveman & Wolfe, 1995) and the level of educational resources at home are related to students' achievements even when parental education and other factors are controlled for (e.g., Teachman, 1987). Empirical evidence also indicates that parental education is positively associated with educational aspirations (Teachman & Paasch, 1998). Investigating the link between parental education and achievement, C. Chen and Stevenson (1995) found that students whose fathers had a postgraduate education scored almost 10 points higher on the achievement test than students whose fathers had a junior-high-school or lower level of education.

Different research outcomes are associated with parental involvement in their child's education. Both positive (Abu-Hilal, 2001; Bhanot & Jovanovic, 2009; Epstein, 2010; X. Fan & Chen, 2001; Jeynes, 2003; Keith et al., 1993; Marchant et al., 2001; Sharp, 1995; Sheldon & Epstein, 2005; Vukovic et al., 2013) and negative (Desimone, 1999; Levpuscek & Zupancic, 2008) impacts have been found, or no impact at all (Jeynes, 2007; Sui-Chu & Willms, 1996b; Topor, Keane, Shelton, & Calkins, 2010). However, other studies report cultural specificity in terms of the impact of parental involvement and teacher responsiveness on the child's education (Desimone, 1999; Huntsinger & Jose, 2009).

According to the meta-analysis conducted by, X. Fan and Chen (2001), the relationship between parental involvement and academic achievement is stronger when academic achievement is described in terms of a global indicator (e.g., school GPA) rather than a subject-specific academic indicator (e.g., mathematics grade). Some studies report a significantly strong relationship between parental involvement and mathematics achievement in the 8th grade (Bhanot & Jovanovic, 2009; Desimone, 1999; Keith et al., 1993), whereas others identify children's self-perception about their mathematics competence as functions of parental and teacher perceptions of their competence (Bhanot & Jovanovic, 2009). Investigating the relationship between parental involvement and mathematics anxiety, Vukovic et al. (2013) found that parents influenced their children's mathematics achievement by reducing anxiety, particularly in the case of more challenging mathematics involving word problems and algebraic reasoning. Evidence accumulated during the past decade indicates that the quality of students' relationships with their teachers is associated with current and future adjustment to schools, classroom responsiveness and negativity (Herring & Wahler, 2003; S. P. Wright, Horn, & Sanders, 1997), and attitudes toward mathematics (Middleton & Spanias, 1999). Children who experience supportive, positive relationships with their teachers are more academically engaged and perform better (Midgley, Feldlaufer, & Eccles, 1989).

Teachers have been regarded as surrogate parents with a very similar socializing influence as birth parents (Herring & Wahler, 2003). Indeed, a positive student-teacher relationship has been found to be significantly associated with students' academic achievement (J. Hughes & Kwok, 2007; J. N. Hughes, Gleason, & Zhang, 2005; Topor et al., 2010). When parents and teachers were considered jointly in predicting students' academic motivational beliefs and academic performance, teacher involvement turned out to be a stronger predictor of academic outcomes than parental involvement (Levpuscek & Zupancic, 2008; Marchant et al., 2001; Stiller & Ryan, 1992). Moreover, teacher and parental responsiveness appear to be positively associated with student motivation and academic performance (Levpuscek & Zupancic, 2008; Marchant et al., 2001). For instance, there is evidence that the quality of the student-teacher relationship fully mediates the relationship between parental involvement and classroom academic performance (Topor et al., 2010): children who experience positive relationships with their teachers have positive attitudes toward schooling (Ryan, Stiller, & Lynch, 1994) and perform better academically (Midgley et al., 1989).

The potential explanations for these positive findings involving the complex nature of the relationships between parental involvement, motivation and achievement are many and various. When parents constitute a home resource in being involved in their child's education, they bridge the gap between the school and home environments, creating a school within the home. Consequently, the child feels more capable of mastering academic demands at school. Parents who get involved in their children's day-to-day schoolwork can facilitate their children's learning of new ideas by helping them scaffold new concepts (Gonzalez-DeHass et al., 2005). Students are more likely to partake in a learning activity and to be motivated if they have the necessary support and encouragement both at home and at school. Similarly, children who see their parents as role models may develop the ability to mimic them, and thereby learn to assess their own competence and performance.

The mixed findings reported above may reflect the fact that student achievement and motivation may increase because of parental involvement, and that parental involvement may increase as a result of students' poor performance, or that parents may become more involved when they believe their child is motivated, or that the motivated child may seek more parental involvement (Gonzalez-DeHass et al., 2005). There may well be a feedback loop (i.e. reciprocity) in the relationship between parental involvement, motivation and performance. However, these parental and teacher practices occur in a somewhat specific cultural or educational context. Furthermore, parental involvement in their children's education is a multidimensional construct defined by parental behavior in this respect (see Epstein, 2010; X. Fan & Chen, 2001; Gonzalez-DeHass et al., 2005 for a comprehensive review). The dimensions of parental involvement include academic pressure, academic support, academic help, the encouragement of academic success, assistance with children's schoolwork (e.g., homework), parenting styles, participation in school activities and parent-teacher interactions, and control and support of the child's academic progress (Epstein, 2010; X. Fan & Chen, 2001; Gonzalez-DeHass et al., 2005; Jeynes, 2007).

1.5 Educational aspirations, socioeconomic background and achievement

Research consistently shows a positive relationship between students' socioeconomic status (SES), aspirations and educational outcomes (Gil-Flores, Padilla-Carmona, & Suárez-Ortega, 2011; e.g., Sirin, 2005). In particular SES has been associated with children's reading and mathematics-related cognitive development (Paxson & Schady, 2007; Yeung & Conley, 2008), performance (M. M. Chiu & Xihua, 2008; Erberber, Stephens, Mamedova, Ferguson, & Kroeger, 2015; M. Jurdak, 2014; Kupari & Nissinen, 2013; Marjoribanks, 2003; Michaelowa, Filmer, & Pritchett, 2001; Mullis et al., 2012; Pfautz et al., 1967; Sirin, 2005; L. Wang, Li, & Li, 2014; Williams & Williams, 2010), attainment (Filmer & Pritchett, 1999; Teachman, 1987; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2013), and tasks involving addition, subtraction, ordinal sequencing, numeracy, and mathematics word problems (Coley, 2002; Siegler, 2009). Socioeconomic status has also been found to influence parental involvement (McNeal Jr., 1999; Park, 2008), school absenteeism, enrollment and dropout (Langhout, Drake, & Rosselli, 2009; McKenzie, 2005; National Center for Education Statistics, 2008; Zhang, 2003), and teacher quality or responsiveness (Akiba, LeTendre, & Scribner, 2007).

With regard to educational aspirations, there is documented evidence indicating that as the level of education rises individuals with higher qualifications are at an advantage in terms of earnings and employability (Sanders, Field, & Diego, 2001). It has also been found that higher-level educational aspirations are significantly associated with higher academic achievement (Gil-Flores et al., 2011; Marjoribanks, 2003; Saha, 1994; Sanders et al., 2001). Other findings indicate that students' expectations and goals are directly related to their mathematics self-concept, performance, and the perceptions of their parents as well as their teachers' attitudes towards their mathematics competence (Eccles [Parsons] et al., 1983; Gunderson et al., 2012). In general, studies have historically shown that parents and teachers' educational expectations are higher for males than for females (Eccles [Parsons] et al., 1983).

Numerous empirical studies have tested and support the assertion that motivational belief (e.g., self-concept) is a significant predictor of educational and career choices (Guo et al., 2015; Philip D Parker et al., 2012; Philip David Parker et al., 2014). Guo et al. (2015) found, for example, that motivational belief predicted educational aspirations after controlling for prior achievement and aspirations. Other studies have produced empirical evidence of a strong relationship between educational attainment and performance (Sewell, Hauser, & Wolf, 2011; Sewell & Shah, 1967; Teachman & Paasch, 1998).

Several explanations have been put forward to explain the complex relationship between these variables. The modern expectancy model, which posits that motivational beliefs influence educational aspirations and educational attainment, can be used to explain the relationship between these three factors (Eccles, 2009; Eccles [Parsons] et al., 1983; Guo et al., 2015; Wigfield & Eccles, 2000). The encouragement and support that parents give to their children with regard to schooling affect the children's desire for education (Eccles [Parsons] et al., 1982). Children whose parents are involved are more likely to take personal responsibility for their learning, and to become motivated when they see that their parents are taking an active interest in their schooling. In other words, parental involvement helps children to internalize educational values: it has been found that when parents place strong emphasis on academic effort and performance at home, students report higher academic self-belief than their peers whose parents put less value on these attributes (Marchant et al., 2001).

Students from a low-SES background experience lower levels of involvement because of limited financial resources, which constrain the ability of their parents to provide them with learning materials (e.g., books) (Orr, 2003). This financial burden also limits the ability of parents to provide a cognitively stimulating learning environment (Klebanov, Brooks-Gunn, & Duncan, 1994; Orr, 2003; Yeung, Linver, & Brooks-Gunn, 2002), which in turn affects their children's educational performance and aspirations (Teachman, 1987).

1.6 Gendered beliefs about mathematics

Gender differences in mathematics-related affects (e.g., mathematics self-concept) and mathematics achievement have received a lot of attention in the literature on mathematics education. Many social, cultural, political and economic barriers influence gendered motivational beliefs and achievement. These barriers vary across and within countries, but generally, the influences on girls are significant. For instance, when the choice to study mathematics is optional it becomes highly gendered, stereotypically favoring males (Mendick, 2006).

Others have argued that gender differences in affect and mathematics achievement are culture-specific (Boaler & Sengupta-Irving, 2006; P. P. Chen, 2002; H. Forgasz, Leder, Mittelberg, Tan, & Murimo, 2015; Hyde & Mertz, 2009). This assertion stems from studies indicating that there are no gender differences in mathematics achievement in gender-equal cultures (Else-Quest et al., 2013; Fryer & Levitt, 2010; Guiso, Monte, Sapienza, & Zingales, 2008; Hyde & Mertz, 2009): females have reported lower self-belief (Andrews et al., 2007; Bharadwaj, De Giorgi, Hansen, & Neilson, 2012; Frenzel et al., 2007; Hannula et al., 2005; Meelissen & Luyten, 2008; Organisation for Economic Co-operation and Development [OECD], 2015); there is evidence of competing cultural roles in students' beliefs about themselves as learners of mathematics (Barkatsas, Forgasz, & Leder, 2001; H. J. Forgasz & Mittelberg, 2008; H. Forgasz et al., 2015); females are reported to perform better and to have higher mathematics confidence in single-sex schools (Fryer & Levitt, 2010; Marsh et al., 2014); and males are more likely to bring technology into mathematics pedagogy (H. Forgasz et al., 2015; Leder & Forgasz, 2008; Ursini & Sánchez, 2008). There are also studies indicating that educational performance (e.g., in mathematics and science) is higher among females, or that there are no gender differences, in Middle Eastern Arabic countries in which the system of education is predominantly single-sex (Fryer & Levitt, 2010; Hassan & Khalifa, 1999; Marsh et al., 2013); that females respond favorably to intervention measures focusing on achievement-related motivational issues (e.g., Wiest, 2010); and that gender-science stereotyping is related to national gender differences in science and mathematics achievement (Nosek et al., 2009).

According to other studies, the mathematics-achievement gap "is due, in a large part, to sociocultural and other environmental factors" (Hyde & Mertz, 2009), and "… the fact that girls' progress in mathematics has been improving over time, even though boys still perform better, suggests that mathematics ability is not innate but susceptible to social influences and instruction" (Stromquist, 2007, p. 37). There is also evidence to suggest that there is no performance gap when boys and girls with similar levels of mathematics self-belief (e.g., self-efficacy, self-concept and mathematics anxiety) are compared (OECD, 2015), but that self-efficacies differ among boys and girls with similar levels of performance (Hannula et al., 2005).

Parents and teachers play a significant role in a child's education. Several studies indicate that parents and teachers have higher educational aspirations for males than for females (Eccles et al., 1990). It has also been suggested that even without any significant evidence of achievement difference between males and females in mathematics, parental perceptions of their children's mathematics achievement is gender-stereotyped (Eccles et al., 1990; Eccles[Parsons] et al., 1982), parents of males having a stronger belief that their children have higher mathematics abilities than parents of females (Eccles et al., 1990). There is a considerable amount of evidence in the literature, indicating that parents' and teachers' gender stereotypes, self-beliefs, and expectations with regard to mathematics performance influence the children's subsequent mathematics achievement in a way that indicates gender-stereotypical roles (Bharadwaj, De Giorgi, Hansen, & Neilson, 2012; Eccles et al., 1990; Eccles & Jacobs, 1986; Eccles [Parsons], Adler, & Meece, 1984).

This gender stereotyping of mathematics abilities in children may influence their self-concept because they internalize the stereotypes directly or indirectly (Eccles et al., 1990; Gunderson et al., 2012; Nagy et al., 2010). P. Davis-Kean and colleagues (2007) found a negative correlation between mathematics interest and parental stereotyping: girls' interest in mathematics decreased as their fathers' gender stereotyping increased, and vice versa for boys. There were further indications that parents provided a more mathematics-supportive environment for boys than for girls, including the purchase of home educational support materials (P. Davis-Kean et al., 2007). On the other hand, teachers' responsiveness to mathematics anxiety appears to be related to female students' mathematics achievements via their motivational beliefs (Beilock, Gunderson, Ramirez, & Levine, 2010). In other words, parents and teachers' expectations of their children's success in mathematics are biased by their own gender stereotypes (Eccles et al., 1990; Gunderson et al., 2012).

The above findings are indicative of a significant influence of gender-role stereotyping and societal influence (e.g., parents and teacher) on students' motivational beliefs and performance. In other words, cross-cultural gender differences in mathematics-related affect and motivational belief stem from sociocultural, ethnic, and other environmental factors rather than gender per se. There are several theoretical explanations for these gender disparities. Modern expectancy-value theory is a useful framework for conceptualizing the relationship between achievement, motivational beliefs and gender. One aspect of the theory relates to gender stereotypes (Eccles et al., 1990; Spencer, Steele, & Quinn, 1999; Steele, 1997). Indeed, gender influences consistent with gender stereotypes (i.e., higher mathematics self-concepts and values in males) have been reported in many cross-sectional and longitudinal studies (Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2006; Nagy et al., 2010; Watt, 2004). First, consistent with the gender-role stereotypes theory (Eccles et al., 1990; Eccles [Parsons] et al., 1984; Spencer et al., 1999; Steele, 1997), there are opposing views on the effect of gender on affect and mathematics achievement (Eccles et al., 1990; Eccles & Jacobs, 1986). For instance, the stereotype threat theory (Steele, 1997) posits that stereotyping strongly affects the stereotyped individual's performance on a task in a stereotypical environment (Ambady, Paik, Steele, Owen-Smith, & Mitchell, 2004; Plante, de la Sablonnière, Aronson, & Théorêt, 2013; Steele & Aronson, 1995), and it is possible to remove the stereotype if the stereotyped individuals are made aware that there are no such stereotypes anymore (Spencer et al., 1999).

There is also evidence indicating that women who endorse gender stereotypes about women's mathematics abilities (Schmader, Johns, & Barquissau, 2004), and believe that mathematical ability is a stable and fixed trait (Aronson & Good, 2003) show stronger stereotype-threat effects (Eccles, 1994). For instance, the stereotypical assertion that males are naturally more talented in mathematics and science as well as related subjects is known to influence the choice of STEM programs, as well as the mathematics achievements and aspirations of both boys and girls (Kiefer & Sekaquaptewa, 2007; Li, 1999; Meece, Parsons, Kaczala, & Goff, 1982; Nosek, Banaji, & Greenwald, 2002; Ramsey & Sekaquaptewa, 2011; Schmader et al., 2004; Steinberg, Okun, & Aiken, 2012). This results in negative self-perceptions of mathematics competence, lower levels of mathematics confidence and achievement, and less interest in continuing to study mathematics-related subjects (Bharadwaj, De Giorgi, Hansen, & Neilson, 2012; Chaman et al., 2014; Eccles [Parsons] et al., 1982).

Nosek et al., (2009) and colleagues, in turn, using the TIMSS data set and National indicators of implicit gender–science stereotypes found that a national indicator of implicit gender–science stereotyping was related to national gender differences in science and mathematics achievement regardless of age or gender. However, Flore and Wicherts (2014) argue in their meta-analysis that, given the small average effect size, there is no strong evidence to suggest "...that stereotype threat manipulations will harm the mathematical performance of girls in a systematic way or lead women to stay clear from occupations in the STEM domain" (Flore & Wicherts, 2014, p. 17).

Another theoretical approach is Deaux and Major's interactive model of gender differences, which acknowledges the importance of cognitive and cultural influences on gender roles: it is suggested that more immediate factors, such as expectations and situational pressures, as well as social and cultural patterns of discrimination, shape gender-related behavior (see also Eccles, 1994). The model

> "assumes that men and women are relatively equal in their potentialities for most social behaviors and that behaviors may differ widely as a function of personal choice, the behavior of others, and the situational context; and that these behaviors take place in the context of social interaction, either explicitly or implicitly [emphasis added]." (Deaux & Major, 1987, p. 371)

At the center of Deaux and Major's interactive approach is (a) a perceiver (expectancy holder, i.e. the teacher or a parent), who enters the interaction with a set of beliefs about gender and with personal interaction goals; and (b) a target individual (receiver, the student), who enters the interaction with his or her own gender-related self-conceptions and interaction goals (p. 371). Once a gender schema is activated in a perceiver it will, together with the perceiver's goals for the interaction, channel the perceiver's behavior toward the target (student). This situation may vary from a very strong gender-stereotypical manner in terms of the degree to which it makes gender-related issues salient.

Summary studies on gender, mathematics-related affect, and performance indicate that girls tend to hold more negative mathematics attitudes relative to boys: they tend to report more mathematics-related emotions (e.g., anxiety) and lower intrinsic and extrinsic motivation, lower confidence and self-efficacy about mathematics, and a lower mathematics self-concept (Andrews et al., 2007; Else-Quest, Hyde, & Linn, 2010; Else-Quest et al., 2013; Fredricks & Eccles, 2002; Frenzel et al., 2007; Ganley & Vasilyeva, 2011; Hassan & Khalifa, 1999; Herbert & Stipek, 2005; McGraw, Lubienski, & Strutchens, 2006; Meece, Glienke, & Burg, 2006; Stipek & Gralinski, 1991; Tella, 2007), and boys perform better than girls, although the performance gap has reduced over the years (Else-Quest et al., 2010, 2013). Parental and teacher involvement are culture-specific, but in general, girls enjoy better support than boys (cf. Andrews et al., 2007).

As far as students' long-term educational aspirations are concerned, the literature gives mixed outcomes in terms of gender. Some studies report no gender difference (e.g., Garg et al., 2002; Watt et al., 2012), some report higher educational aspirations among females (e.g., Mau & Bikos, 2000), and others indicate that males have higher aspirations (e.g., Mendez and Crawford 2002; Wilson and Wilson 1992).

1.7 Mathematics education in the African context

There are educational disparities between the rich, the poor, and individuals of different genders in many countries. These disparities are huge in sub-Saharan and Arabic African countries. Education is free and compulsory, as well as a human right enjoyed by every child in most parts of the world, however educational opportunities in Africa are deeply marked by inequalities that are linked to wealth, gender, ethnicity, and other social divisions (Van Fleet, & Watkins, 2012). For instance, in the case of Ghana the 2000-2012 UNESCO Deprivation and Marginalization in Education (DME) index indicates that primary-education attainment rates cover 69 percent of females and 71 percent of males. Of those who finish primary school, 35 percent are from poor homes whereas 86 percent are from rich homes. The respective figures for those who complete lower-secondary school are 48 percent of females and 66 percent males, 20 percent coming from poor homes and 80 percent from rich homes. Being poor and female carries a double disadvantage in many Arabic African countries, such as Egypt (UNESCO, 2010, p. 114). According to the DME report, gender disparity was a significant factor explaining the relatively high level of educational poverty—the proportion of the population aged 17-22 with less than four years in school (the minimum required to gain basic literacy and numeracy skills), and extreme education poverty—fewer than two years in school—in most sub-Saharan and Arabic African countries. For instance, young women in Egypt are twice as likely as young men to have fewer than four years of education- and four times as likely if they are poor, and under half of the poor females aged 17 -22 in rural Egypt have fewer than four years of education: the respective rate in Morocco is 88 percent (UNESCO, 2010).

Mathematics education in Africa is generally poor, although mathematics has been used as a gateway to social progress. For instance, entry to some universities is based on a mathematics qualification even if the subject to be studied has no resemblance or connection to mathematics (Gates & Vistro-Yu, 2003). In most developing countries, the need for school mathematics is broadly tied to its relevance to the country’s sociocultural and socioeconomic needs (Gates & Vistro-Yu, 2003). However, most children in most Africa countries who stay in school do not learn mathematics (Africa Learning Barometer; Van Fleet & Watkins, 2012), according to minimum competency thresholds based on data from regional examinations such as Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC) and the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), as well as national assessments of fourth or fifth graders from 28 countries in sub-Saharan Africa (Van Fleet, 2012; Van Fleet & Watkins, 2012). The findings from the Africa Learning Barometer studies led

Van Fleet (2012, para. 3) to conclude that "under the current educational model, half of sub-Saharan Africa's total primary school population – 61 million children – will reach adolescence without the basic skills needed to lead successful and productive lives [emphasis added]".

According to the Africa Learning Barometer, for instance, 32.1 percent of schoolchildren in Ghana are not learning (21.1% not learning to read, 43.1% not learning mathematics), 33.7 percent in South Africa are not learning (27.2% not learning to read, 40.2% not learning mathematics), 58.3 percent in Nigeria are not learning (65.7% not learning to read, 51.0% not learning mathematics), and 16.5 percent in Botswana are not learning (10.6% not learning to read, 22.4% not learning mathematics)(Van Fleet & Watkins, 2012). These figures indicate that schooling does not always translate into learning for children in Sub-Saharan African countries who are able to access education. Although Africa has made progress over the past decade, the barometer indicates that there is a learning crisis at the compulsory-education level. To remedy this situation African governments and the international community should work together and implement policies that will help raise standards and improve learning outcomes, otherwise the future of a whole generation of African youth will be "wasted" and Africa's economy will stagnate (Van Fleet, 2012).

Most African countries do not participate in international large-scale assessments such as TIMSS and PISA, but the few that partake in TIMSS perform abysmally. According to the TIMSS 2007 report, 17 percent of students in Ghana scored above the low international benchmark, of which only nine percent had some knowledge of whole numbers and decimals, operations, and basic graphs (Mullis, Martin, & Foy, 2008). Similar results are reported in TIMSS 2011 (Mullis et al., 2012). Like many others, African countries face the problem of low achievement in mathematics and related subjects, mathematics anxiety, and low university enrolment in programs that require high mathematical skills (Gates & Vistro-Yu, 2003). For instance, in TIMSS 2011, the validity of the average achievement scores in grades four and eight for all participating African countries (except Tunisia) was too low as such for any comparisons to be made without a disclaimer because the percentage of students with achievement at too low a level for estimation exceeded 25 (Mullis et al., 2012). One possible explanation for the low achievement is that the curricula of some of the participating African countries are not aligned with the three components of the TIMSS assessment (Ndlovu & Mji, 2012). The greater the difference between standardized international survey items and the national curriculum, the greater the likelihood that the assessment of an education system will be poor. This misalignment calls into question the value-neutrality of TIMSS, which currently appears to be constantly shifting targets with which only well-resourced, developed countries can cope and improve performance (Ndlovu & Mji, 2012). It is also likely that differential mathematics achievement in African countries and the rest of the world may be associated with differences in culture.

There is some evidence to suggest that the quality of learning and achievement of pupils in a country varies in accordance with its economic and general educational level (Altinok, 2008), and that there are cultural (country) differences in what constitutes school mathematics (Cogan & Schmidt, 2002). Moreover, teachers and schools in Sub-Saharan African countries are on different levels, and often on lower levels in the fields of mathematics and science. It has been suggested that there is a need for more sensitivity to the possibilities and limitations associated with developments in mathematics and science education in Sub-Saharan countries to avoid disappointment with the lack of impact of educational reforms (Ottevanger, van den Akker, & de Feiter, 2007). Many reports advocate the application of the *start where the learner is* credo in schools and national education systems (Ottevanger et al., 2007, p. 66).

According to a report by the developing countries' strategies group of the International Mathematical Union (IMU), African countries differ in many respects, yet are still very similar with regard to their institutions and national issues that hinder mathematical development. In most of them, for instance, mathematical development is limited by the low numbers of qualified mathematics teachers from pre-school to the secondary level, as well as of mathematicians at the Master's and PhD levels, and the curricula and instruments for mathematics and science courses date back to colonial days (International Mathematical Union [IMU], 2009). Other contributing factors include automatic promotion, the poor grounding at the primary and lower-secondary levels, ambitious curriculum content and problems with mixed-ability teaching, poorly resourced schools, large classes, curricula that are not relevant to the daily lives of pupils, and inadequate teacher education (Ottevanger et al., 2007). Furthermore, teacher absenteeism, ethnic and cultural differences in school, in the attitudes and beliefs of parents, and in parental involvement all influence mathematics achievement in Africa.

Culture, societal and vocational, and ICT play a crucial role in shaping national educational goals and policies, which in turn influence the conceptual guidelines in curriculum design and implementation. For instance, even if all different approaches to mathematics schooling yielded identical performance in a given grade, it could still happen that the particular curricular pattern observed in one country provided a firmer foundation for advanced mathematics than other alternatives (Cogan & Schmidt, 2002, p. 38).

Mathematics education in Africa is riddled with cultural issues such as religion and the language of instruction. According to the education-for-all global monitoring report based on DME data, "gender, poverty, language and culture often combine to produce an extremely heightened risk of being left far behind in education [emphasis added]" (UNESCO 2010, p. 9). The vital role that language plays in mathematics achievement has long been recognized (Aiken, 1971). The official language policy in most African nations is complex, reflecting the historical influences, socioeconomic aspirations, cultural identity and power relations that shape national and local society rather than being objective and ideologically neutral (May, 2008; Trudell, 2011). The language of instruction in the case of mathematics is thus a second language. This phenomenon is very common in countries where the official language of instruction in schools is an adopted foreign language such as English or French, or a national language that is not the mother tongue of the learner but a post-independence adopted colonial language (Jones, 1982; S. Wright, 2004). Although as some scholars argue, "… the language of mathematics is culture-fair because it is culture-free" (M. E. Jurdak, 2014, p. 52), according to Jurdak:

> As soon as mathematics is applied in problems and situations, the language of instruction and learning becomes a cultural carrier in terms of behaviors, social relations, habits, and values. (M. E. Jurdak, 2014, p. 52)

The linguistic structures of their native language influence the mathematical deductive reasoning of students (M. E. Jurdak, 1977). For instance, the linguistic structures of foreign languages are significant predictors of item difficulty, which is not the case for native languages. Indeed, the values of mathematics education have been described as incompatible with the value systems of the native culture in most countries. As M. E. Jurdak comments:

> "If values of mathematics education are incompatible with the value system of the mother culture, then mathematics will be "appended" to the culture as a "technology" rather than assimilated as a "mode of thinking". (M. E. Jurdak, 2014, p. 50)

The effect of using a foreign language as a medium of instruction in mathematics as well as the cultural role of mathematics education have been on the agenda in most mathematics education conferences such as PME (the research group reporting on culture, language and multilingualism), CERME (the working group focusing on mathematics and language), and ICME (the Topic Study Group on language and communication in mathematics education): for a discussion on the psychological and pedagogical implications of using a foreign language as a medium of instruction, see the discussion in Berry (1985). Various studies discuss the effect of language on mathematics achievement (Abedi & Lord, 2001; Aiken, 1971; Martiniello, 2008, 2009; Vukovic et al., 2013), especially in Africa with its national language for schooling and the native language spoken at home (e.g., Howie, 2013). For instance, among the African countries participating in TIMSS 2011, proficiency in the medium of instruction at school was a strong predictor of student success in mathematics (Howie, 2013; Mullis et al., 2012). The TIMSS

2011 results for those in the sixth and ninth grades show that only about 26 and 12 percent, respectively, of the students in Botswana spoke the language of the test at home or before starting school, and that these students performed better. A similar pattern was observed in South Africa and for the ninth and eighth graders in Ghana.

Given the medium of instruction in most African countries and how language influences mathematics performance, it may be that assessing mathematics competence on items with high linguistic demands may measure construct-irrelevant language competence (Haag, Heppt, Stanat, Kuhl, & Pant, 2013; Martiniello, 2009; Wolf et al., 2008). M. K Wolf et al., (2008) found in their analysis of the academic language used in mathematics and science items, for instance, that academic vocabulary was the most frequent feature.

Religion plays a significant role in mathematics education in most Arab states, including those in Africa. The problem is that the values of mathematics education are incompatible with the value systems of the native cultures (M. E. Jurdak, 2014, p. 55). The conflict is not so visible in the kind of mathematics taught or in the specialized mathematical techniques and procedures used in the classroom, but it becomes evident during the production and use of mathematical knowledge to cope with reality (p. 51).

The question is what needs to be done? There is no single solution to the problem of African mathematics education. The model at hand is to offer direct support of mathematical expertise at all levels especially, in primary and secondary education.

2 Research questions

There is a need for cross-cultural studies on mathematics-related affect that challenge the foundations of current thinking and contribute to establishing more useful and universal theories. Such studies should focus on enhancing understanding not only of the cross-cultural differences, but also of the psychometric properties of the measured constructs. This dissertation is based on four original publications, which are indicated throughout the text by the Roman numerals I, II, III and IV. Within the conceptual and theoretical framework discussed above, the general objective of the present study is to:

- Determine the psychometric properties (factor structure, reliabilities, method effect, and measurement invariance) of measures of mathematics-related affect and how these properties transcend cultures, and to identify the methodological challenges associated with the process;
- Examine the relationship between these mathematical-related affective constructs on a cross-cultural level;
- Explore the relationship between mathematics-related affect, students' background variables, and mathematics achievement on a cross-cultural level.

The present dissertation is based on two separate cross-sectional data sets involving four studies. The subjects of Studies I and II were Ghanaian twelfth-grade students. Those participating in Studies III and IV came from the five African countries (Ghana, Botswana, South Africa, Morocco, and Tunisia) participating in TIMSS 2011: students in Ghana, Morocco and Tunisia were in the eighth grade and those in South Africa and Tunisia were in the ninth. A more detailed breakdown of the samples and aims of the individual studies is given in Tables 1 and 2, and in the overviews (Chapter 4). The aim in Studies I and II was to examine the structure of students' views on mathematics using both confirmatory and exploratory structural equation modeling. Specifically, the study illustrates the problems associated with the importation of survey instruments developed in one cultural setting into a new one by examining the psychometric properties of the constructs assembled in Finland and replicated in Ghana. Thus, the discussion concerns the process of construct validation using structural equation modeling to propose and statistically test an alternative factorial structure in the event of model misfit.

Studies III and IV are based on the TIMSS 2011 data set, the aims being to investigate the psychometric properties of the mathematics-related affect constructs, the relationship between affect and achievement, and the effect on achievement of background and instrumental variables such as gender, long-term educational aspirations, parental education, and students' socio-economic background.

Along with these more general goals, each study had more specific aims. Study I describes in detail the processes of construct validation in mathematics-related affect from a cross-cultural perspective. Study II extends the discussion with a view to shedding light on the possible causes and explanations of the construct differences between the two nations. The first aim in Study III was to examine the psychometric properties (factor structure, reliabilities, method effect, and measurement invariance by country and gender) of the TIMSS 2011 mathematics-related measures of affect (conceptualized as motivational construct in the TIMSS 2011 framework), and how these properties transcend the cross-cultural and methodological challenges associated with the process. The second aim was to investigate the relationship between these mathematics-related affective dimensions cross-culturally, and the third was to examine the relationship between mathematics-related affect, students' background variables, and mathematics achievement. Finally, Study IV investigated the reciprocal relationship between mathematics-related affect (e.g., self-concept) and achievement using the TIMSS 2011 cross-sectional data associated with the five different educational systems/nations.

3 Methods

3.1 Participants

Studies I and II

Data were collected from 2,034 12th-grade students (mean age = 18.49, Median age =18, and 58.2% female). Education in Ghana follows a two-tier model with three stages of compulsory schooling (Kindergarten: two years for 4-year-olds; Primary school: six years for 7 to 12-year-olds; Junior High School: three years from age 13 to 15; and Senior Secondary School (upper-secondary): four[1] years from age 15 to 18). In more simple terms, Ghana operates a 6-3-3 system of education. Children enter primary school in the calendar year of their sixth birthday. Promotion is automatic in Grades 1–6, and is dependent on academic progress in Grades 7–9: in general, it is automatic in public schools, whereas in private schools it tends to be based on assessment.

The sole language of instruction is English except in kindergartens where the children are taught in the native language spoken in the vicinity. All prescribed textbooks and materials used in the various schools are in English. The most prominent school type is co-educational (90%) in the secondary sector, single-sex schooling for boys accounting for four percent and for girls six percent of the school-age population. As of 2009, enrolment for girls in senior secondary school stood at 44.3 percent.

Nine senior high schools were selected from urban and rural schools according to the rankings of the Ghana Education Service, which are based on past performance in matriculation examinations. Schools were selected for the present study from each category of the ranking. Permission was sought from the institutional heads, and the questionnaire was administered to the students during their normal class hours in the summer of 2011. The questionnaires were voluntary and students had the right to withdraw or skip any question they did not wish to answer. The schools included private, religious and public schools, some falling into more than one category. The students were enrolled in one of two mathematics classes: core mathematics (49.3%) and elective mathematics (50.7%). All the respondents were pursuing General Arts (33%), Business (19.2%), Science (29.1%), or Vocational Science (18.7%) as their elective major-subject areas. The study sample comprised 63 different classes with an average size of 32 students.

[1] At the time of the data collection, senior secondary schooling in Ghana took four years: it now takes three years. This present sample was one of the four-year groups.

Studies III and IV

Data were obtained from 38,806 students who participated in the TIMSS 2011 in the five African countries (see Table 1). The samples used in Studies III and IV comprised eighth-grade students from Ghana, Morocco and Tunisia, and ninth-graders from Botswana and South Africa. The participating countries were given the option of testing a higher grade if the questions were too difficult for the eighth-graders: Botswana and South Africa tested ninth-graders. The TIMSS 2011 sampling design is a two-stage cluster design involving the sampling of schools, and of intact classrooms from the target grade in the schools. There were 989 intact classrooms (clusters) with an average of 39 students per class.

TIMSS is an international assessment tool developed by the IEA to measure trends in mathematics and science achievement in the fourth and eighth grades in participating countries. First administered in 1995, it has been conducted in a regular four-year cycle, making TIMSS 2011 the fifth assessment involving 45 countries or school systems participating in the eighth-grade assessment. Participating countries use the assessment in various ways to explore their educational issues: monitoring system-level achievement trends in a global context, establishing achievement goals and standards for educational improvement, stimulating curriculum reform, improving teaching and learning through research and analysis of the data, conducting related studies, and training researchers and teachers in assessment and evaluation (the International Association for the Evaluation of Educational Achievement [IEA], 2014, para. 2).

Table 1. Sample Breakdown and Associated Characteristics for Studies III and IV.

Country	N	Girls (%)	Schools	Clusters (intact classes)	Average Class size	Grade
Ghana	7,323	48	161	173	41	8
Morocco	8,986	48	279	279	32	8
Tunisia	5,128	51	207	207	25	8
Botswana	5,400	49	150	151	35	9
South Africa	11,969	49	285	317	37	9

Note: N= number of students in each country.

A brief description of the educational systems of the five African countries follows. Information on the educational system in each country was retrieved from the World Data on Education (UNESCO International Bureau of Education, 2011). The duration of compulsory education until the upper-secondary level varies in the countries from 12 to 13 years, lasting until the end of Junior High School or until the students reach the age of 16. The upper-secondary phase is not compulsory in all the countries. As described above, Ghana follows the 6-3-3 (six years primary, three years junior- and three years senior-secondary) system of education, whereas Botswana follows a 7-3-2 structure (seven years primary, three years junior, and two years upper-secondary). Education systems in South Africa, Morocco and Tunisia generally follow the 6-3-3 structure, but with variations in basic and secondary education (UNESCO International Bureau of Education, 2011). For instance, primary education in South Africa is divided into junior (grades 1-3) and senior primary (grades 4-6). High-school education is divided into lower secondary (grades 7-9), which is the last stage of compulsory education, and senior-secondary (grades 10-12) levels.

3.2 Measures

3.2.1 Affective measures

Studies I, and II

The Views on Mathematics (VOM) scale developed in Finland by Pehkonen's research team was used as a measurement instrument in the studies (Hannula et al., 2005; Rösken et al., 2011). The scale comprises 55 items, most of which originated from a qualitative study on student teachers' views of mathematics (Pietilä, 2002). Four additional items were taken from a previous study on Finnish comprehensive schools (Nurmi et al., 2003), 10 items originated from the self-confidence scale in the Fennema–Sherman mathematics-attitude measurement instrument (Fennema and Sherman, 1976), and some novel items were developed by the team to measure student-perceived success in mathematics. All except the 10 Fennema–Sherman items were originally in Finnish and had been translated into English. The items were assessed on a 5-point Likert scale. The statements in the questionnaire were grouped around the following topics: (1) Experiences as a mathematics learner (A1–A29); (2) Image of oneself as a mathematics learner (B1–B15); and (3) Views on mathematics and its teaching and learning (C1–C11).

Studies III and IV

The TIMSS2011 student-motivation questionnaire was designed to measure students' affects related to mathematics. The *students like learning mathematics* (SLM) scale is based on the extent to which respondents agree with five statements; the *students value mathematics* (SVM) scale measures the extent of agreement with six statements; and the *student confidence in mathematics* (SCM) scale similarly measures agreement with nine statements. Both the *parental* involvement (PIV) and the *teacher responsiveness* (TRES) scale consist of four items. For instance, the latter scale measures the extent to which students perceive their teachers as caring, helpful and responsive to their learning needs. The scales were reversed in the present study so that higher values correspond to higher response values. The mathematics self-concept (MSC) construct used in Study IV is an abbreviated five-item version of the mathematics self-confidence construct in TIMSS 2011.

3.2.2 Background variables

The student-background variables used in Study III were gender, long-term educational aspirations, and parental education. Socioeconomic background, gender, and long-term educational aspirations were used as instrumental variables in Study IV. The socioeconomic-background scale was derived from students' reported home educational resources based on their responses to three questions concerning the number of books in the home, the highest level of education of either parent, and the number of home-study supporting features such as having their own room and access to an Internet connection. The long-term educational aspirations (LEA) scale is a self-report measure of education: students are asked to indicate the highest level of education they expected to attain on a scale ranging from (1) "lower-secondary education" to (6) "university program–Master's/Doctorate". In terms of parental education the respondents were asked to record the highest level of education completed by their male and female guardians on a seven-point scale ranging from (1) "Some primary, or [some] lower-secondary, or did not go to school" to (7) "university program- Master's/Doctorate". The male-and-female-guardian variable was combined and recoded into five categories: if both parents (5) "Finished university or higher", (4) "Finished post-secondary education but not university", (3) "Finished upper-secondary", (2)"Finished lower-secondary", (1) "Finished some primary or lower-secondary schooling or did not go to school". Gender was coded "1" for girls and "2" for boys.

3.2.3 Mathematics achievement

Studies III and IV

To ensure maximum content coverage and to reduce the testing burden on students, and due to time constraints, the participants were given a survey comprising different subsets of items from a large pool of items, a design known in the literature on large-scale assessment as multiple matrix sampling (Shoemaker, 1971). Currently, TIMSS and most large-scale surveys report students' achievements in terms of so-called plausible values. Plausible values are estimates of the scores individual students would have obtained if they had taken the test that includes all the items in the pool (Wilkins et al., 2002). One major reason for using plausible values is that they provide an accurate representation of the underlying relationships in all the available student data, including responses to the items they were given as well as all the background data (Foy, Brossman, & Galia, 2013). Including all the available student-background data in the imputation process ensures that the relationships between the background data and the plausible values are "appropriately accounted for" (Foy et al., 2013). Another advantage is that the resulting analyses give more precise estimates of the population parameters than single performance scores (Rutkowski & Rutkowski, 2010; von Davier, Gonzalez, & Mislevy, 2009).

The TIMSS 2011 assessment divides mathematical domains into two dimensions: four content and three cognitive domains. The cognitive domains cover Knowing, Applying, and Reasoning, whereas the content domains are Numbers, Algebra, Geometry, and Data and Chance. Within the content domain, students are expected to draw on the cognitive domains. With regard to the four mathematics-content domains the splits were: Numbers 30%, Algebra 30%, Geometry 20%, and Data and Chance 20%; and for the cognitive domain they were: Knowing 35%, Applying 40%, and Reasoning 25% (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009). In the case of mathematics achievement, TIMSS 2011 computed five sets of plausible values for each student performance in the four mathematics-content and three cognitive domains. It is highly recommended to use all the five plausible values in all analyses to reduce the measurement error associated with estimation.

3.3 Statistical analysis

This section describes the main methods of statistical analysis applied in the present dissertation. Mplus software was used to test the hypotheses. The models in question provide a robust substantive-methodological synergy (Marsh & Hau, 2007) of potential theoretical and practical importance in mathematics-related affect. The focus is on models that include latent variables, multiple indicators, measurement errors, and complex structural relationships such as reciprocal causation. The initial stage involves confirmatory factor analysis conducted to define the validity and reliability of the constructs under study through the measurement of their observed indicators. The specific methods include factor analysis, both exploratory (EFA) and confirmatory (CFA), multi-group analysis, CFA with correlated-uniqueness (CFA-CTCU), as well as non-recursive modeling and the reliability estimates (composite reliability (ω) and Cronbach's alpha (α)).

Both EFA and CFA were used in Studies I and II, together with reliability estimates to validate the VOM constructs in the Ghanaian data set. Confirmatory factor analysis and reliability estimates were used to confirm that the original factor structure of the VOM conformed to the Ghanaian data. There was no support for the confirmatory factor analysis or for the reliability estimates. The subsequent exploratory factor analysis identified a different factor structure in the Ghanaian data set. CFA was then used to confirm the consistency and the structural validity of the new factor structure (Studies I and II). Multi-group confirmatory factor analysis served to establish measurement invariance across student genders (Studies I, II and IV) and countries (Studies III and IV). Moreover, CFA models (e.g., multi-group analysis) incorporating multiple indicators, multiple causes (MIMIC) were used to determine the mean difference in the TIMSS2011 achievement and motivational constructs among the participating African countries. Study IV was based on CFA and non-recursive structural equation modeling, also with TIMSS 2011. Multi-trait-multi-method (MTMM) CFA models (e.g., method effects, see Chapter 3.3.2) were incorporated into the models in Studies III and IV. A more detailed breakdown is given in Tables 2.

3.3.1 Model evaluation and estimation criteria

Several model criteria are combined in this dissertation. First, the overall model fit should be satisfactory. Although the chi-square statistics were reported in all the studies, it was for information purposes only because the measure is very sensitive to sample size. Other statistical indices considered were the Root Mean Square Error of Approximation (RMSEA), the comparative fit index (CFI), and the change in fit between nested measurement invariance models (e.g., ΔCFI) (F. F. Chen, 2007). The CFI varies between zero and one. The fit is considered adequate if the CFI values are > 0.90 (best ≥ 0.95), and the RMSEA ≤ 0.060 for models with a better fit: a RMSEA ≤ 0.080 indicates an adequate fit (Bentler & Bonett, 1980; Hu & Bentler, 1999; West, Taylor, & Wu, 2012). Second, the models were carefully examined from a practical/approximate perspective to assess the change in CFI or RMSEA in the different consecutive models, on the assumption that differences in fit between the less and the more constrained models in the series had to be negligible. According to Chen (2007), if the decrease in model fit in the more parsimonious model is less than or equal to .010 for CFI (Cheung & Rensvold, 2002), or a change in RMSEA of less than .015, then there is reasonable support for the more parsimonious model. In other cases the a *priori* prediction, common sense, a comparison of viable alternative models and a detailed evaluation of the parameter estimates are taken into account.

All the analyses in the present thesis were based on the Mplus robust maximum likelihood estimator (MLR). The Mplus MLR procedure allows tests of fit that are robust in cases of non-normality in the observations, and control for their non-independence. Given the complex design of the TIMSS survey, the model parameters in Studies III and IV were estimated using MLR in combination with the “complex” option in Mplus to account for the clustered design and to adjust for standard errors. The student’s class was used as the clustering variable, and the sampling weights were taken into account (*weighting variable supplied with the data*).

The Mplus feature of full information maximum likelihood (FIML) took care of the missing data in Studies I, II and IV, and in study III missing data was treated using the Mplus multiple imputation procedure (Asparouhov & Muthén, 2010; Graham, 2009; Schafer & Graham, 2002), allowing for the use of the full sample size. Five imputation datasets were used to remove any uncertainty in the imputation process.

3.3.2 Method effects: Negative items

Method effects, or method variance associated with negatively worded items seem to be a common phenomenon in personality assessment and psychological measurement (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Quilty, Oakman, & Risko, 2009; Tomǭs, Oliver, Galiana, Sancho, & Lila, 2013). For instance, self-belief and self-esteem constructs, in particular, are susceptible to response bias, or the method effect. Method variance, in other words—variance that is attributable to the measurement method rather than what the constructs were intended to measure, is a known problem in behavioral research (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003, p. 879). Marsh and colleagues define method effects as “nontrait effects associated with idiosyncratic aspects of particular items or methods of data collection” (Marsh et al., 2013, p. 112). In other words, individuals’ tendencies to respond to questionnaires on a basis other than what the items were designed to measure and resulting in systematic irrelevant error scores in the study (American Education Research Association, 1999; Lindwall et al., 2012; Quilty et al., 2009). One common source of the method effect is the “balance scale”, meaning the use of negatively and positively worded items (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Vautier & Pohl, 2009).

It is argued in the literature that negatively worded or reverse-coded items are included in research instruments to reduce the effect of response-pattern bias (e.g., acquiescence). According to Podsakoff and colleagues, “reverse-coded items are like cognitive “speed bumps” that require respondents to engage in more controlled, as opposed to automatic, cognitive processing” (Podsakoff et al., 2003). Some research studies posit that including both negatively and positively worded items in a construct is a necessity (Quilty et al., 2009), on the assumption that both measure the same construct (e.g., Marsh, 1996). In the light of such observations researchers began to examine whether the method effect reflected systematic measurement errors or a response style with substantive meaning, or was just an irrelevant measurement artifact (Distefano & Motl, 2006; DiStefano & Motl, 2009; Lindwall et al., 2012; Quilty et al., 2009; Urbán, Szigeti, Kökönyei, & Demetrovics, 2014; Ye & Wallace, 2013). For instance, responses to negatively worded items have been found to yield systematic variance that is irrelevant to the content under study, irrespective of the age group (Distefano & Motl, 2006; DiStefano & Motl, 2009; Hooper, Arora, Martin, & Mullis, 2013; Lindwall et al., 2012; Marsh, 1986; Urbán et al., 2014). Cote & Buckley (1987), in turn, found in their review that at least 29.8 percent of the variance in attitude measures across the 70 studies was attributable to systematic sources of measurement error such as common method bias.

On the other hand, Schmitt and Stults (1985) found that if fewer than 10 percent of the respondents in a study respond to negatively worded items differently, or fail to recognize that some items are negatively worded, it can be a possible source of the method effect (Woods, 2006). Indeed, even two negatively worded items within a construct can lower the reliability and validity of the scale (Marsh et al., 2014; Roszkowski & Soven, 2010; Urbán et al., 2014; Woods, 2006). However, method effects can lead to faulty inference in suppressing or inflating the relationships between variables, contributing to Type-I or Type-II errors if not incorporated into the measurement model (Bagozzi, 1993; Marsh, Scalas, & Nagengast, 2010).

It has been shown that when method effects are not accounted for they affect the goodness-of-fit statistics and lead to biased parameter estimates (M.-S. Chiu, 2008, 2012a; DiStefano & Motl, 2009; Marsh, 1994; Marsh et al., 2013, 2014; Marsh, Nagengast, & Morin, 2012; Pohl & Steyer, 2010). Many studies on construct validation (e.g., even for older students) also indicate that positively and negatively worded items in a survey tend to influence the uni-dimensionality of the measurement scales, resulting in two distinct factor structures in the factor analysis (Idaszak & Drasgow, 1987; Marsh, 1986, 1996; Roszkowski & Soven, 2010). A large proportion of the common variance in cases supporting the uni-dimensionality of the constructs is attributable to the method effect due to the negatively worded items (Urbán et al., 2014), which may have a significant impact on the reliability of the construct. However, the negatively worded factors disappear when the reverse-coded items are rewritten positively (Idaszak & Drasgow, 1987; Roszkowski & Soven, 2010).

The strategy commonly used to accommodate the method effect falls within the multi-trait–multi-method (MTMM) framework in the confirmatory factor analysis (CFA) model (Bagozzi, 1993). The most common approaches are correlated traits, correlated uniqueness (CTCU) or simply correlated uniqueness (CU), and the latent-method factor (LMF) or the correlated trait-correlated method (CTCM) framework (Bagozzi, 1993; Marsh & Grayson, 1995) The CTCU framework treats the effect of the negatively and/or positively worded items as a methodological artifact rather than a distinct factor in allowing uniqueness related to the negatively worded and/or the positively worded items to be correlated (Bachman & O'Malley, 1986; Marsh & Grayson, 1995; Ye & Wallace, 2013). The LMF, in contrast, allows the wording effect to be a distinct latent factor (a distinct factor for positively and negatively worded items) along with a substantive latent factor (Lindwall et al., 2012). Thus, the CTCU framework allows the method-wording effect to be removed from the construct in not allowing the wording effect to be examined as a unique factor (Distefano & Motl, 2009; Lindwall et al., 2012; Marsh & Grayson, 1995; Marsh, 1996; Quilty et al., 2009). This helps to eliminate any irrelevant empirical relationship with the constructs or variables used in the study. It has been found in numerous studies adopting this strategy that models incorporating the method effect generally fit the data set better than those that do not (Bagozzi, 1993; DiStefano & Motl, 2009; Horan, DiStefano, & Motl, 2003; Lindwall et al., 2012; Marsh, 1996; Marsh et al., 2013; Quilty et al., 2009; Tomas & Oliver, 1999; Tomǭs et al., 2013). This phenomenon has been identified in large-scale assessment studies such as TIMSS 2003 (M.-S. Chiu, 2008, 2012a, 2012b), TIMSS 2007 (Marsh et al., 2013, 2014), and the National Educational Longitudinal Survey (NELS: 88) (Marsh, 1994) for the self-concept variate.

The construct-validity process used in Studies I and II indicated a method effect associated with negatively worded items; in study III, four of the nine items on the mathematics-self-confidence construct, and two of the five on the like-mathematics scale were negatively worded, as were three of the five items on the mathematics-self-concept scale in study IV. We took a CFA-CTCU approach as advocated in Bachman & O'Malley, (1986); Marsh & Grayson, (1995); Marsh, (1994) to accommodate the method effects associated with negatively worded items and to obtain acceptable goodness of fit, unbiased standard errors and parameter estimates.

3.3.3 Reliability

Reliability is an important concept for scholars in the social and behavioral sciences, helping researchers to gain real insights into relationships among observed phenomena and the individual differences in a specific latent construct (Raykov, 2012). Although it is subject to much misuse, misunderstanding and confusion (Sijtsma, 2009), Cronbach's alpha (α) has long been the most widely used quality indicator in test statistics (e.g., Novick and Lewis 1967). It is acknowledged to be an inconsistent estimator of reliability; it either underestimates or overestimates (Bentler, 2009; Geldhof, Preacher, & Zyphur, 2014; Green & Yang, 2009; Novick & Lewis, 1967; Peterson & Kim, 2013; Raykov, 1998, 2001; Zimmerman, Zumbo, & Lalonde, 1993).

Some of the controversy over the use of Cronbach's alpha stems from the different descriptions of the alpha coefficient in the literature (Cortina, 1993). For instance, it has been described as the mean of all split-half reliabilities (Cronbach, 1951) and as the lower bound of true reliability (Novick & Lewis, 1967; Peterson & Kim, 2013; Raykov, 1998, 2001; Ten Berge & Zegers, 1978), and is equal to reliability in conditions of essential tau-equivalence—linearly related and differing only by a constant (Cortina, 1993; Novick & Lewis, 1967; Ten Berge & Zegers, 1978). The Cronbach's alpha reliability coefficient is also highly dependent on the number of items on a construct (Cortina, 1993), and assumes uncorrelated error scores, which is contentious among researchers (Green & Hershberger, 2000; Green & Yang, 2009; Novick & Lewis, 1967; Raykov, 1998; Yanyun Yang & Green, 2011; Zimmerman et al., 1993). Zimmerman and colleagues, for instance, examined the effect of correlated errors on the alpha coefficient and concluded that they produced a large bias in the reliability of the estimates, the coefficient being markedly overestimated even for relatively small violation of the assumption of uncorrelated error scores.

It is suggested in the literature that if the assumption of uncorrelated error scores is violated methods that incorporate structural equation modeling should be used in estimating the reliability of test scores (Green & Yang, 2009; Yanyun Yang & Green, 2011). However, when the set of items on a construct is unidimensional (e.g., congeneric), with uncorrected errors and has uniformly high loadings (e.g., above .6 on a 0-1 metric), then α is not underestimating by much composite reliability (Raykov, 2012, p. 479). In the present dissertation, the method effect associated with negatively worded items (see CTCU above) was incorporated into all the models that necessitated the use of composite reliability (ω).

The composite reliability measure (ω), (Geldhof et al., 2014; Raykov, 2012), which is usually associated with SEM, was estimated to complement the α estimates. Composite reliability (ω) incorporates the computed factor loadings and the method effect (if any, e.g., correlated uniqueness associated with negatively worded items), and produces more precise estimates of reliability than α does (Geldhof et al., 2014; Raykov, 2012). Composite reliability estimates are known to overcome some of the limiting assumptions of the alpha coefficient and estimates of true reliability than alpha under identical research conditions (Peterson & Kim, 2013; Raykov, 2001). Composite and Cronbach's alpha reliability are interpreted in the same way, with values of .600 to .700 being acceptable in exploratory research (Hair et al., 2010).

Studies have shown that, regardless of their high reliabilities in the original settings, imported constructs tend to show lower reliabilities in different cultural settings (e.g., Tuohilampi et al., 2014). Moreover, given that the TIMSS 2011 constructs come largely from Western research, and that empirical evidence based on previous TIMSS research (Metsämuuronen, 2012a, 2012b; Rutkowski & Rutkowski, 2010) indicates lower reliabilities among participants from developing countries, the reliability estimates for some of the constructs may well be lower in our sub-population. The lower reliabilities may attenuate the validity of the interpretations of our findings based on manifest scale scores, and weaken the statistical power as well as the effect sizes (Raykov, 2012; Schmidt & Hunter, 1996; N. Schmitt, 1996), thus necessitating the use of latent-variable models that account for unreliability, bias, and measurement errors (Cole & Preacher, 2014).

Cronbach's alpha was first estimated to make meaningful comparisons with the original construct in Studies I and II, and both composite and Cronbach's alpha reliabilities were subsequently computed and compared. Both reliability estimates were computed for each country in Studies III and IV. The composite reliabilities were necessary given the need to incorporate the correlated uniqueness (the method effect) associated with the negatively worded items in the computation.

3.3.4 Measurement invariance

The most effective and efficient tool for collecting cross-cultural data thus far is the survey instrument. Researchers commonly compare scores across groups in mathematics related affect studies with little concern about whether the assessment instrument (e.g., anxiety, self-efficacy, self-concept, self-confidence) is operating in the same way or the underlying construct has the same theoretical structure across the groups.

The findings from these studies are presented under the vague assumption of measurement invariance without any empirical tests. For example, there is strong evidence that mathematics-related affect is positively/negatively related to student achievement (M.-S. Chiu, 2009; Grootenboer & Hemmings, 2007; Marsh et al., 2005; Whitebread & Chiu, 2004). However, researchers tend to investigate the differences in these mathematics-related affect constructs across groups, say, by cultures/nations/educational system, gender, ethnicity, and other grouping variables without drawing any inferences that they may be invariable across the groups. If measurement invariance does not hold, any inference drawn about group differences may be flawed, meaningless and ambiguous (Brunner et al., 2013; Cheung & Rensvold, 2002; Cheung & Rensvold, 1999; Lubke, Dolan, Kelderman, & Mellenbergh, 2003; Sass, 2011). In fact, most research on cross-cultural comparison advocates the use of measurement invariance to ease cross-cultural generalizability of the measured models (Little, 1997; Steenkamp & Baumgartner, 1998; Vandenberg & Lance, 2000)

There are several levels of measurement invariance. Three levels are of significance in cross-cultural comparison or the comparison of factor means across groups or countries: configural, metric (also referred to as weak factorial invariance), and scalar (or strong factorial invariance) invariance. This approach involves a series of hierarchical stepwise operations based on multiple-group factor-analytical models beginning with a well-fitting baseline model and further logically defined and increasingly restrictive models (Byrne & van de Vijver, 2010; Little, 1997; Lubke et al., 2003; Steenkamp & Baumgartner, 1998; Vandenberg & Lance, 2000).

The process usually begins with a baseline model, often called the configural model, in which all the parameters are freely estimated across the groups, in other words, the same number of factors holds for each group and the same variables define each factor across groups. A good fit indicates that the same number of factors best represent the data: the next model to test is one that requires the factor loadings to be equal across the groups. The term metric or weak measurement invariance is normally used when the factor loadings are equally held. If metric invariance holds, one could conclude that the constructs are manifested in the same way in each of the groups, and any differences in the covariance between the measured variables are attributable to the common factors (Cheung & Rensvold, 2002; Cheung & Rensvold, 1999; Meredith, 1993; Millsap & Olivera-Aguilar, 2012). Metric invariance is a necessary condition for all other forms of measurement invariance, and a prerequisite for any meaningful group or cross-cultural comparisons (Cheung & Rensvold, 2002; Cheung & Rensvold, 1999). The next level is scalar invariance, meaning that invariance constraints are placed on the measurement intercepts and the factor loadings (metric invariance). Scalar invariance is significant because it implies that differences in the means of the observed variables must be attributable to the influence of the common factors or the underlying construct(s) (Cheung & Rensvold, 2002; Millsap & Olivera-Aguilar, 2012). Finally, strict measurement invariance requires invariance of item uniqueness across groups in addition to metric invariance. Support for strict invariance implies that the scale items measure the latent constructs with the same degree of measurement error, and is an indication of comparable reliability across groups (Cheung & Rensvold, 2002; Cheung & Rensvold, 1999). It is not a requirement for testing differences in factor structure or latent means. With support for configural, metric and scalar invariance, one can confidently assume that the scores are not biased and that the means can be compared across groups.

Achieving scalar invariance has been reported as problematic in comparisons comprising large-scale cross-cultural studies or when many groups are involved. In the case of non-scalar invariance, which is an indication that factor means cannot be compared across groups, two approaches discussed in the literature allow the comparison of factor means: partial measurement invariance (Byrne, Shavelson, & Muthén, 1989) and alignment optimization (implemented only in the Mplus SEM program) (Asparouhov & Muthén, 2014). In the case of partial scalar invariance, some measurement intercepts are invariant and others are not, meaning that the intercepts of the non-invariant items are allowed to vary in the group comparisons. Modification indices (MIs) and expected parameter changes (EPCs) provided by the SEM program(s) are particularly useful in this context. Partial scalar invariance thus allows for cross-cultural comparisons that would otherwise not be practical due to the lack of full scalar invariance. Although this technique works well with small groups it becomes more complex when the groups are larger and more diverse, and many large modification indices are reported.

Other levels of invariance such as variance and covariance structures can be tested for other purposes. This is often called structural invariance, and is of importance in studies involving the validity of multidimensional constructs in which constructs including relations with other constructs are considered necessary for the generalization of the underlying constructs. Invariance of factor variances and covariances is used as an indicator that correlations between latent constructs are invariable across groups.

Multi-group CFA (configural, metric, and structural invariance) and cross-validation were used to test the validity and multidimensionality of the VOM construct in Studies I and II. The sample was divided into halves for the cross-validation purpose. Multi-group CFA measurement invariance was used to cross-validate the VOM construct with the validation sample, and gender invariance was assessed on the final construct. Given our interest in Study III in comparing countries' latent means on the TIMSS2011 motivational-belief constructs, we first conducted a multiple-group factor analysis to see whether the standard model and the nested-factor model showed the appropriate level of measurement invariance across the five countries. Similarly, we used multiple-group analysis in Study IV to examine the measurement-invariance (e.g., mathematics self-concept) construct before introducing a non-recursive CFA model across the five countries.

Table 2. Overview of the Aims, Participants, Measures and Analyses in the Respective Studies.

Study	Participants	Main aims	Measures and reliabilities	Analytical method
Study I	2034 twelfth-graders (58.2% girls).	To explore the methodological implications and challenges related to the ways in which instrument adaptation is undertaken in mathematics-related research on affect; to give a detailed account of the processes involved in applying SEM to validate the instrument and arrive at an alternative factorial structure for the Ghanaian sample.	mathematics self-concept mathematics self-confidence family encouragement teacher quality	EFA CFA Multi-group CFA: Measurement invariance Reliability
Study II	2034 twelfth-graders (58.2% girls).	To give a detailed cross-cultural view of the instruments and possible causes of different factor structures: to discuss the issue of mathematics-related affect in relation to the validity of	mathematics self-concept mathematics self-confidence family encouragement teacher quality	EFA CFA Multi-group CFA Reliability

		cross-cultural construct adaptation, and to further discuss the theoretical implications of the Ghanaian factor structure.		
Study III	38,806 students from five African Countries: 21,437 eighth-graders, 17,369 ninth-graders (see Table 1).	First, to investigate the psychometric properties of the TIMSS 2011 motivational-belief constructs in the five educational systems/nations/cultures, and specifically, to ascertain whether the psychometric properties cut across these countries. Second, to investigate any systematic country/cultural differences in the latent factor means for these constructs and mathematics achievements. Finally, to investigate the relationship between these constructs and other background variables such as parental education, students' educational aspirations, and mathematics achievement	mathematics achievement confidence in mathematics liking mathematics valuing of mathematics teacher responsiveness parental involvement parental education long-term educational aspirations gender	Reliability CFA Multi-group CFA Correlational analysis CFA- multi-trait-multi-method (MTMM) – Method effect Multiple indicators, multiple causes (MIMIC)
Study IV	38,806 students from five African Countries: 21,437 eighth-graders, 17,369 ninth-graders (see Table 1).	To investigate the bidirectional cause–effect between affect and achievement using non-recursive structural equation models.	mathematics achievement mathematics self-concept socioeconomic status long-term educational aspirations gender	Reliability CFA Multi-group CFA Non-recursive modeling
Note: CFA = confirmatory factor analysis, EFA = Exploratory factor Analysis.				

4 Overviews of the original studies

4.1 Studies I and II

Participants, Measures and Methodology
The sample comprised 2,034 12th-grade Ghanaian students (mean age = 18.49 years, standard deviation = 1.25 years; 58.2% girls: see Chapter 3.1 for details about the participants and the sample). A survey questionnaire was used to elicit responses from students concerning their views on mathematics (VOM: see Chapter 3.2.1 for details about the VOM measures). The robust maximum likelihood estimator (MLR) was used to control for non-normality (skewedness or kurtoses) in the data set. The data were analyzed using the reliability, EFA, and multi-group CFA functions in Mplus 7.11 to assess the fit of the VOM to the non-Western context. For cross-validation purpose, the data set was split into two halves.

4.2 Study I

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

There is a strong emphasis on using standardized and validated research instruments in mathematics-related research. This article illustrates the problem of importing survey instruments developed in one cultural setting mostly in the western hemisphere into a new cultural setting regardless of their acceptable psychometric properties in the original settings. In other words, it describes in detail the process of construct validation involving reliability and SEM (exploratory, confirmatory and multi-group analysis).

Specifically, the study assesses the fit of the seven-factor conceptualization of the VOM scale (Roesken et al., 2011) to a non-Western context, and proposes and tests an alternative structure in view of the misfit between the original theoretical seven-factor model and the Ghanaian data set. The article also explores the methodological implications and challenges related to the process of instrument adaptation in the research on mathematics-related affect.

The reliability and the CFA did not fit the Ghanaian sample. EFAs were therefore used to establish a new factorial structure of the VOM constructs for the Ghanaian data set. Following a process of sequential EFAs, the four-factor structure was deemed the best: *mathematics self-concept, mathematics self-confidence, teacher quality* and *family encouragement*. The proposed VOM constructs were based on 37 items with seven factors (ability, effort, teacher quality, family encouragement, enjoyment, difficulty and success), and the acceptable four-factor structure was based on 25 items. The findings indicate that the structure of the VOM scale is empirically and theoretically different from the original structure in the Ghanaian context. The failure of the original scale highlights the importance of adapting mathematics-related affect instruments to different cultural settings. Using SEM facilitated the detection of measurement error and bias. For instance, it was necessary to accommodate the method effect (the correlated uniqueness of two negatively phrased items on the mathematics self-concept scale were allowed to covary) associated with the use of negatively and positively worded items in a survey instrument for the CFA to fit the data set.

After establishing a baseline factor structure, support was found for the configural, metric, and structural invariance across the two samples (validated sample). Gender invariance was also confirmed, supporting the validating of the construct. The findings indicate that the original instruments were not directly applicable in the new settings. More importantly, multi-group analysis confirmed that the four-factor structure was best for the Ghanaian data set. This study explicitly focuses on the methodology associated with construct validation with regard to mathematics-related affect. In light of the very robust approach used, the four-factor structure with its very strong psychometric properties theoretically supported the dimensions of mathematics related affect. The outcome is a robust and reliable approach to construct validation in the research on mathematics-related affect. Cross-cultural disparities were also detected on the VOM structure. However, solid theoretical and substantive subject-area knowledge should guide this approach.

4.3 Study II

Bofah, E. A., & Hannula, M. S. (2015). Studying the factorial structure of Ghanaian twelfth-grade students' views on mathematics. In B. Pepin & B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education: Exploring a mosaic of relationships and interactions* (pp. 355–381). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-06808-4_18

The aims of Studies I and II were similar. Both articles discuss the issues associated with the importation of a mathematics-affect instrument from one culture to another and the associated methodological challenges. Whereas the first study gives a detailed account of the processes involved in applying SEM to validate the instrument and arrive at an alternative factorial structure for the Ghanaian sample, the second study presents a cross-cultural view of the instruments and possible reasons for the different factor structures, focusing on the issue of mathematics-related affect in relation to cross-cultural construct adaptation, construct validity, and the interpretation of findings. It further discusses the extent to which the new factor structures fit the theoretical discussions in the literature.

The design and methodology were the same in both studies. A further look at the theoretical factor structure based on the Finnish sample and the Ghanaian factor structure reveals an interesting pattern. With respect to the new structure, two scales from the original theoretical model were partially confirmed and one was fully confirmed. The reliabilities (both composite (ω) and Cronbach's alpha (α)) were acceptable, although α underestimated *teacher quality, family encouragement* and *mathematics self-confidence* and overestimated the *mathematics-self-concept* construct. The *ability*, *enjoyment* and *effort* scales were discrete concepts in the Finnish sample, which supported the separation of emotions and motivation from cognitive belief, whereas in Ghana, the *ability*, *success*, and *enjoyment* factors were empirically single construct. The *effort*, *enjoyment* and *difficulty* constructs were not confirmed in the Ghanaian data set. The *effort* scale was a unique characteristic feature in the Finnish data set in that other studies validating the VOM could not confirm it.

The differences between the Ghanaian and the Finnish constructs reflect the cultural dimension of students' views on mathematics in the two countries. Given the dramatic cultural differences, it was not surprising that the mathematics self-concept had an underlying structure in the Finnish sample whereas it was a single entity in the Ghanaian sample. However, the construct difference could also be partially attributed to the different methodology used in determining the relevant categories of the belief systems. As such, the conclusions in this article are more reliable. In addition, the use of multi-group CFAs validated the four-factor structure in an independent sample, and there was gender invariance across student groups. This implies strong empirical and theoretical support for the new four-factor structure: the same variables define each factor across all subsamples, the latent constructs have the same relationship within the sample, and any differences in the covariance between the measured variables are due to the common factors. The support of measurement invariance indicates that any further mean comparison within the belief structure could be interpreted as representing the underlying mean differences in the Ghanaian data set.

Moreover, the findings emphasize the importance of ensuring that concepts within an adapted instrument are the same for the original and the target languages and contexts. Furthermore, the cultural and linguistic features of the new target population should be thoroughly taken into account. This would apply to the sample in this study: the official language of instruction in schools and of the survey instrument is not the mother tongue of the learner. Hence, researchers should be conscious of the problems of construct importation and adaptation: aspects including the wording and order of survey questions, negatively worded items, translation-content overlap, similarly worded items, and the reference period of the questions may have a strong impact on issues such as measurement error. One implication of the study is that researchers focusing on cross-cultural mathematics-related affect should acquire both the theoretical and the practical ability to address these issues using appropriate tools such as structural equation modeling. This study has laid a solid foundation on which to base future research in this field, particularly in Ghana: it offers a readily available approach as well as a valid, reliable and applicable instrument.

4.4 Studies III and IV

Participants and methodology

The sample included all the five African countries that participated in TIMSS 2011. Details of the participating countries and their respective sample characteristics are given in Table 1 and described in detail in Chapter 3.1. Mplus 7.3 was used for all the analyses in the study. The robust maximum likelihood estimator (MLR) with standard errors and a chi-square test statistic that is robust to non-normality and non-independence of observations (to account for the dependency issue of students nested within classrooms) were used to control for non-normality (skewedness or kurtoses) in the data set (Muthén & Muthén, 1998-2012: see Chapter 3.3.1 for more details on model-evaluation strategies).

4.5 Study III

Bofah, E. A., & Hannula, M. S. (2015). TIMSS Data in an African Comparative Perspective: Investigating the Factors Influencing Achievement in Mathematics and their Psychometric. *Large-scale Assessments in Education, 3*(1), 4 doi: 10.1186/s40536-015-0014-y.

The article reports on the analysis of a cross-sectional data set of the five African countries that participated in TIMSS 2011. The general purpose of the study was to investigate the psychometric properties (factor structure, reliabilities, method effect, and measurement invariance—country and gender) of the TIMSS 2011 motivational-beliefs constructs across the five educational systems/nations/cultures, and specifically to ascertain the extent to which these properties cut across the countries. Further aims were to find out if there were systematic country or cultural differences in the latent factor means for these constructs, to compare achievements in mathematics, and to explore the relationship between these motivational constructs and other background variables such as parental education, students' long-term educational aspirations, and mathematics achievement. The variables included the Students' Confidence in Mathematics scale (SCM), the Students Value Mathematics scale (SVM), the Students Like Learning Mathematics scale (SLM), the Parental involvement/positive parenting (PIV) scale, and the Teacher Responsiveness (TRES) scale. Three demographic variables were used as indicators of the students' background variables: gender, long-term educational aspirations, and parental education.

Confirmatory factor analysis of the motivational constructs revealed a very low degree of model fit to the combined data set of responses from Ghana, Botswana, South Africa, Morocco, and Tunisia. However, the fit improved substantially after controlling for the method effect (non-trait effects associated with particular items or methods of data collection) associated with the negatively worded items. Following the establishment of a baseline factor structure, a multi-group CFA test of measurement equivalence indicated support for configural, metric and scalar invariance (partial invariance: intercept of non-invariant items are allowed to vary).The relative percentage of item-intercept invariance across the groups was too small to influence any latent mean comparison across the five educational systems. In addition, there was evidence of gender invariance, supporting the validity and reliability of the construct.

Table 3. The Reliabilities of the TIMSS Motivational Constructs.

Country	**PIV**		**SCM**			**SLM**			**SVM**		**TRES**	
	α	ω	α	ω	ωp	α	ω	ωp	α	ω	α	ω
Botswana	.77	.77	.77	.67	.77	.79	.75	.78	.70	.70	.63	.65
Ghana	.79	.73	.73	.51	.70	.60	.48	.56	.70	.71	.56	.57
Morocco	.82	.75	.75	.59	.73	.73	.68	.73	.71	.71	.56	.56
South Africa	.77	.75	.75	.56	.73	.74	.67	.73	.74	.74	.68	.69
Tunisia	.76	.79	.79	.67	.79	.84	.82	.85	.74	.75	.64	.64
Sample	.79	.76	.76	.60	.74	.74	.68	.71	.72	.73	.62	.63

Note: Parental involvement scale (PIV), The Students Value Mathematics scale (SVM), The Students Like Learning Mathematics scale (SLM), The Students Confidence in Mathematics (SCM), The Teacher Responsiveness scale (TRES), Cronbach's Alpha (α), Composite reliability (ω); ωp = composite reliability without correlated uniqueness values.

Estimates of both Cronbach's alpha (α) and the composite (ω) reliabilities indicated that the constructs were poorly specified for SCM and SLM when the method effect (correlated uniqueness associated with negatively worded items) was not incorporated into the reliability estimates, i.e. composite reliability. The reliability seemed to be high initially, in terms of Cronbach's alpha, but once the model was properly specified, the poor reliability became evident (the composite reliability estimates: Table 3). Overall, the sample reliabilities were generally acceptable for the Cronbach's alpha estimates but some of the constructs were far below the acceptable limit for the composite reliability estimates. It was consistently found during the process of reliability estimation that α was not a dependable estimate of scale reliability.

This study produced evidence to suggest that including negatively phrased items in survey instruments can attenuate the reliability and validity of a measure. In other words, the method effect associated with negatively phrased items could obscure the underlying structure of the scale and possibly bias the outcome. The findings also imply a strong relationship of fit between the country-level construct measurements attributable to the method effect associated with the negatively worded items and country-level achievement. In other words, countries in which the method effects were the strongest tended to have lower achievement levels. Given the problems associated with α in the measurement of scale reliability, the use of ω in estimating construct reliability in cross-cultural research is recommended because it incorporates scale factor loading, item variance, and correlated uniqueness, if any (in this case the method effect associated with the negatively phrased items).

The mean comparison was based on the partial scalar measurement invariance model. The Ghanaian sample constituted the reference group in the analysis of the mean comparison of the motivational construct; as such, the latent means for the measurement constructs were fixed at zero in the Ghanaian sample so that the size and direction of the differences in all the remaining four countries could be evaluated in relation to the Ghanaian sample. Consequently, students in all the other countries reported statistically significantly lower values on the mathematics-confidence, liking-mathematics, valuing-mathematics, and teacher-responsiveness constructs. Thus, whereas students in Ghana are more confident in their mathematics ability, students in the other countries place less value on mathematics and like it less than the Ghanaian students do, and think their teachers are less responsive. With regard to parental involvement, students in Botswana, Morocco, and Tunisia reported statistically significantly lower values than their Ghanaian counterparts, whereas students in South Africa reported statistically significantly higher values. The implication is that parents are more involved in their children's education in South Africa than in Ghana, whereas Ghanaian parents are more involved in their children's education than parents in Botswana, Morocco and Tunisia.

The findings of this study seem paradoxical and perplexing in the sense that students from Ghana perform abysmally in mathematics but their motivational beliefs were the highest of all the countries in the sample. There were significant country differences in the relationship between motivational beliefs and mathematics performance. The relationship between achievement and value was the strongest in three of the five countries and strongest for mathematics self-confidence in the other two. The relationship between parental involvement and mathematics was culturally non-invariant: it was positive in Ghana and Morocco, for instance, negative in Botswana and South Africa, and non-existent in Tunisia. A similar outcome was found for the relationship between teacher responsiveness and achievement: it was positive in Ghana, Botswana, and Morocco, but there was no relationship in South Africa and Tunisia.

Achievement scores were standardized, a positive score indicating better-than-average mathematics achievement across the five countries whereas negative scores reflect lower-than-average achievement. The scores of the Tunisian and Botswana students on mathematics achievement were statistically significantly above the mean, whereas the Ghanaian and South African students scored statistically significantly below the mean. The scores of the Moroccan students did not differ statistically significantly from the five-country mean, although it was slightly below it.

Students' long-term educational aspirations (LEA) were statistically significantly lower in Botswana and South Africa and statistically significantly higher in the two Arab countries (Morocco and Tunisia) and Ghana: the relationship with student achievement was the strongest among all the variables in the study. The relationship between LEA and how students value mathematics was higher than all the other motivational measures in all the countries. There was a statistically significant relationship between students' LEA and all the measures of motivational belief in Ghana, Morocco and Tunisia, whereas it was not statistically related to parental involvement and student confidence in mathematics in South Africa. A positive relationship between students' LEA and all the motivational measures except parental involvement was reported in Botswana.

Parental education was statistically significantly related to students' mathematics achievement in all the countries. Whereas the relationship between parental education and achievement was cross-culturally universal, that between parental education and students' motivational beliefs was culture-specific, being negative or positive in some countries and unrelated in others. The relationship between parental education and teacher responsiveness was not statistically significant in any of the countries.

Gender differences favored boys on the motivational-belief constructs and girls in the background variables. Among the motivational constructs, there were seven statistically significant differences favoring girls and 10 favoring boys, the largest difference being in the mathematics-confidence scale. In terms of mathematics achievement, boys outperformed girls in Ghana and Tunisia, whereas girls outperformed boys in Botswana. There was no difference in achievement between boys and girls in Morocco and South Africa. All in all, teachers were more responsive to girls than to boys. Student long-term educational aspirations were higher among girls in all the countries except Ghana.

In sum, the findings indicate that there were systematic psychometric problems associated with the negatively worded items in some of the TIMSS 2011 motivational constructs in the African context that need to be controlled for any meaningful cross-cultural analysis to be made. This affected the reliability and validity of the construct. There was also evidence of a country-level construct-measurement fit due to the method effect associated with negatively worded items. With regard to the relationship linking student motivation with achievement, parental education, educational aspirations and gender, the findings indicate cross-cultural variations as well as cross-cultural universalism. For instance, the associations of higher parental education with higher achievement, long-term educational aspirations and parental involvement were culturally universal, whereas that between affect and achievement was culture-specific. Figure 1 shows the structure of the cross-cultural variations.

4.6 Study IV

Bofah, E. A. (2015). Reciprocal determinism between students' mathematics self-concept and achievement in an African context. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education (CERME 9)* (pp. 1688-1694). Prague, Czech Republic. https://hal.archives-ouvertes.fr/hal-01287995

The aim of the study was to test three alternative propositions concerning the relationship between mathematics self-concept (MSC) and achievement that have been discussed in the literature. The propositions investigated are as follows: (a) there is a unidirectional influence of affect on achievement; (b) there is a unidirectional influence of achievement on affect; and (c) there is a reciprocal relationship between affect and achievement.

The design involves non-recursive structural equation modeling based on the TIMSS 2011 cross-sectional data set. Three variables were used as instrumental indicators: gender, long-term educational aspirations, and socio-economic background. Each model was analyzed separately within each country with a composite set of the plausible values. The result of a confirmatory factor analysis (CFA) verified a clear factor structure after incorporating the method effect associated with combining both negative and positive items in a survey. Moreover, multi-group CFA supported the measurement invariance (e.g., metric invariance) of the MSC construct across the five educational/cultural groups. In spite of the low reliabilities, the CFA multi-group analysis indicated that MSC has the potential to measure the same traits in the same way across diverse groups if the method effect associated with the use of negatively and positively worded items in a survey instrument are incorporated into the models. The outcome of the study supported the factorial structure and subsequent measurement invariance across the different educational/cultural groups of the MSC construct.

On the relationship between affect and achievement, the analysis indicated that reciprocal determinism between mathematics self-concept and achievement is dependent on the national context. The findings from four of the five countries did not support the feedback relationship: mathematics self-concept significantly predicted achievement in Ghana and Botswana, achievement significantly predicted mathematics self-concept in Morocco, and neither achievement nor self-concept predicted each other in South Africa.

Males reported a significantly high mathematics self-concept across all the educational systems, and students' long-term educational aspirations significantly predicted achievement. The relationship between student socioeconomic background and achievement was less consistent, but was evident in more than half the countries: socioeconomic background was a significant predictor of student performance in South Africa, Morocco and Tunisia, for example, and a significant predictor of mathematics self-concept in Ghana, Botswana and Morocco.

The outcome of the study highlights the fact that reciprocal determinism is culture-specific depending on the national context. The effects of mathematics self-concept on achievement and of achievement on self-concept varied between the five educational systems, and in four of the five, the estimates obtained did not support reciprocal determinism. The findings show strong evidence of a contribution factor linking self-belief to achievement, and vice versa.

Gender significantly predicts students' mathematics self-concept, with males reporting higher MSC levels in all five countries. Students' long-term educational aspirations also significantly predict mathematics achievement, higher long-term aspirations predicting higher achievement in all countries. These results give clear evidence of cross-cultural generalization in the relationship between gender and MSC, as well as between long-term educational aspirations and performance. Indeed, the gender effects found in the study are consistent with gender stereotypes (a higher mathematics self-concept in males) identified in cross-sectional and longitudinal studies.

The effects of SES on students' mathematics achievement and MSC were less consistent, but were evident in more than half of the countries. Socioeconomic status predicted achievement in South Africa, Morocco and Tunisia, but was not a significant predictor of mathematics achievement in Ghana and Botswana. Finally, SES was a significant predictor of students' MSC in Ghana, Botswana and Morocco.

5 Main Findings and Discussion

5.1 Brief: Summary and outcome

The foci of the present dissertation are pragmatic, emphasizing substantive-methodological synergy in order to address substantively important issues in relation to measures of mathematics affect and the possible relationship with mathematics achievement. The study addresses the problems involved in the importation of mathematics-related survey instruments developed in one culture for use in another. Methodologically it is based on a robust approach—structural equation modeling—that accounts for measurement error, and controls for bias (e.g., responses to negatively worded items). The methodological focus is to address a major limitation associated with mathematics-related affect research whereby the validity and reliability of the constructs tend to be problematic. On the practical level, the investigation was tailored to a less commonly studied population using a very strong and large data set, and applied evolving models in a methodologically robust manner.

More specifically, the first major aim of the present dissertation was to determine the psychometric properties of mathematics-related affective measures such as the theoretical factor structure, reliabilities, method effect, and measurement invariance, to see how these properties transcend cultures, and the methodological challenges associated with the process. The second major goal was to investigate the relationship between these mathematics-related affective constructs on a cross-cultural level, and the third aim was to explore the relationship between mathematics-related affect, students' background variables, and mathematics achievement in an African context.

Factor analysis was used to model the dimensionality of the constructs. Confirmatory factor analysis (CFA) was applied to model the cross-cultural validity and reliability of the mathematics-related affective constructs. Other CFA models such as multiple indicators, multiple causes (MIMIC) or CFA with covariates, multiple-group CFA, and the correlated methods model were utilized to examine cross-cultural variations in mathematics-related affect and student achievement. Other social factors such as students' long-term educational aspirations, parental education, and socioeconomic background were included as covariates in the MIMIC models.

According to the results of these studies, the psychometric properties of measures of mathematics-related affect developed in the Western context are problematic irrespective of the fact that they are acceptable in the original settings. There were substantial reliability differences in the mathematics-related construct especially when they included negative items. A series of confirmatory factor analyses revealed the need to control for the method effect associated with negatively worded items, which according to the findings could obscure the underlying structure of the scale by influencing its validity and reliability. For instance, without accounting for the correlated errors in the computation of the reliabilities for some of the constructs, the common factor "pick up the slack" and the reliability was overestimated. It was also evident that responses to negatively worded items are systematically culture-specific, and the method effect associated with them appears to diminish with increasing mathematics ability. The outcome of the method effect indicates that differences in responses to negatively worded items are worth studying in their own right, and are not merely substantively irrelevant "noise" that is theoretically uninteresting and needs to be eliminated.

A further goal of the present studies was to determine the methodological feasibility of importing measures of mathematics-related affect developed in one cultural setting (Finland) into a new one (Ghana). This required a series of factor analyses and reliability estimates to confirm the theoretical factor structure of students' views on mathematics and statistically test an alternative structure in the event of a model misfit. The findings from the first set of studies (I & II) provide more details of the cross-cultural dimensions in mathematics-related affect, and reveal disparities in the seven-factor conceptualization of VOM in a non-Western context. Both articles explore the methodological implications and challenges related to the ways in which instrument adaptation is undertaken in the research on mathematics-related affect. According to the findings of Studies I and II, with the right methodology it is easier to control for measurement error and bias, and to enhance understanding of disparities in the constructs. Factor analysis (combined EFAs and CFAs) revealed the best and correct factor structure for the Ghanaian sample. The cross-validation and multi-group invariance over gender add reliability to the outcome of the Ghanaian structure compared with previous studies involving VOM. Studies I and II provide researchers with a solid foundation on which to construct valid, reliable, and applicable questionnaires, and offer a robust step-by-step approach to cross-validating survey instruments from one cultural context to another. Study II further revealed that the most difficult challenge in cross-cultural research on mathematics-related affect relates to validity and comparability. At least three important questions arose. Is there a need to adopt survey instruments? Do the constructs measure what they are intended to measure? How should one interpret the outcomes in new cultural settings? The conclusion in each case is that cultural and linguistic differences need to be factored in to arrive at any meaningful interpretation.

A number of culture-specific and universal conclusions regarding the relationships between affective dispositions and achievement among students could also be drawn from the second set of studies (III and IV) presented in this dissertation. Study III revealed for the first time the psychometric properties of the TIMSS affective measure of motivational beliefs, which was thoroughly investigated in the African context.

It could be concluded from Studies III and IV that there is no simple, universal relationship between affect and student achievement. The relationship was similar in many education systems, but in some the pattern was different. The findings reported in Study III indicate that student task-value (extrinsic motivation), self-confidence and long-term educational aspiration are among the most reliable predictors of performance, as opposed to intrinsic value as reported in studies conducted in the West. This pattern was universal across different educational systems. Country-specific characteristics were relevant and active. It was found in Study IV that the causal relationship between affect and achievement varied across education systems, with varying outcomes across all five African countries in TIMSS 2011.

Study IV was the first to confirm the cross-cultural dimension of the reciprocal determinism between affect and achievement involving international large-scale cross-sectional data in the African context, thereby providing significant new evidence of the generalizability of the uni- and bidirectional relationship between mathematics self-concept and mathematics achievement among students. The findings revealed the cross-cultural specificity of this causal relationship. Students' gender, long-term educational aspirations, and socioeconomic background were used as instrumental variables, which turned out to have a strong impact on achievement and self-concept. It was found in Studies III and IV that the relationships between gender and both mathematics self-concept and long-term educational aspirations were universal and cross-cultural, girls scoring lower on mathematics self-concept, whereas higher long-term educational aspirations were associated with higher achievement. There were varying culture-specific outcomes concerning the relationship between gender and achievement. The varying gender differences in the present cross-country investigation may also reflect differences in school structure, confounded with cultural differences in gender norms.

It was also found in Study III that the relationship between the affective construct and its magnitude are cross-culturally specific. For instance, parental involvement was associated with positive outcomes in terms of the student's motivational belief and performance, but the magnitude and direction of the relationship were culturally non-invariant. All in all, there are plausible theoretical and practical implications regarding the relationships between the measures of motivational belief and performance. By way of synthesizing the studies comprising this dissertation, the following sections discuss the theoretical and practical implications. Finally, the limitations and challenges for future studies are assessed.

5.2 Theoretical implications

The findings reported in this dissertation support as well as refute some of the theoretical perspectives adopted. In terms of construct validation, there are implications for current theories concerning the development and structural composition of mathematics-related affect. The implications also extend to some of the methodological challenges in comparative studies of mathematics-related affect. A number of studies have documented the effect of unreliable psychometric properties in survey instruments and of the relationship between the measures (Bagozzi, 1993; Marsh et al., 2010). However, until now all these studies have been done in the West and in some Asian countries. There has been no research linking the psychometric properties of these measures to students' performance in the African context, and minimal attention has been paid to the effect of negatively worded items on construct validity and reliability.

The results discussed in this dissertation may have a broad influence on researchers using both negatively and positively worded items in surveys in general, and specifically in those focusing on mathematics-related affect. The findings confirm the effect of negatively worded items on the validity and reliability of the measurement instrument. Furthermore, most psychometric properties of mathematics-related constructs have thus far been dependent on Cronbach's alpha as the best measure of construct reliability. The studies comprising this dissertation support the theoretical analogy that, under certain conditions such as when there are correlated errors among the items in a scale, the values for the alpha coefficient exceed the values for composite reliability (Bentler, 2009; Green & Yang, 2009; Peterson & Kim, 2013; Raykov, 2001). The results also show a clear relationship between negatively worded items and student achievement, and indicate that negatively worded items are a "cognitive nuisance" for low achieving students cross-culturally, and as such are worth studying in their own right.

The empirical part of this work confirms that there are substantial cross-cultural differences in the structure of affect in the African context, and in its relationship with achievement. The findings also imply that it would be easy to develop a more reliable measure of student performance than the measure of mathematics-related affect used in cross-cultural research. However, the strong factorial invariance (Cheung & Rensvold, 2002; Cheung & Rensvold, 1999; Meredith, 1993; Millsap & Olivera-Aguilar, 2012) indicates that the constructs are fundamentally similar in all cultural settings (i.e., comparable), hence the hypotheses posited in this dissertation can be meaningfully tested.

Hannula (2001, 2015a, 2015b) succeeded in describing the structure of mathematics-related affect in detail in his comprehensive literature review, and produced empirical evidence supporting an affect cluster (Roesken et al., 2011). The present study was successful in further supporting this approach empirically. However, there was no clear factor structure separating motivation, emotions, and cognitive affect in the Ghanaian data set, as in the original theoretical model. It could therefore be argued that it is also necessary to consider cultural variation in investigations of children's views on mathematics, as well as the methodological approaches involved in the theoretical settings. As Clarke (2013, p. 1863) put it: "the interpretation and application of international comparative research is critically contingent on researchers' capacity to address those "dilemmas" pertinent to their particular design".

These differences have been found in multiple cultural contexts, across a wide age range and among various researchers (Kaldo & Hannula, 2012; Tuohilampi et al., 2014), implying that mathematics-related measures are culture-specific. Given the cultural variations in the VOM construct, it would be important to determine which features are most salient to children and how they relate to mathematics performance. This could help to shed new light on how mathematics-related affect influences individual differences on the cultural level. Studies III and IV addressed that question to some extent.

The findings also support as well as refute some of the theoretical perspectives on motivational belief and achievement adopted here. They give new insights into the relationship between, as well as the predictive nature of mathematics-related affect, achievement, and students' background variables. Overall, the results challenge as well as uphold the implicit assumption that students' affective disposition impacts performance, and vice versa. The implications extend to current theories on the relationship between affect and achievement, including the causal relationship. For instance, in demonstrating the presence of relationships linking motivational beliefs, the long-term educational aspirations and socioeconomic background of students, teacher and parental involvement, and achievement, the results give strong support for modern expectancy value theory and extend the substantial evidence that attests to its effect on students' achievement-related behaviors.

The results also contribute theoretically to the literature on mathematics self-concept regarding the reciprocal relationship between affect and achievement in indicating that these predictions are culture-specific. Another finding reported here is that, whereas most research in Western cultures indicates that intrinsic motivation is more strongly related to performance than extrinsic motivation, the opposite was the case: extrinsic motivation related more strongly to performance. The implication is thus that the association of intrinsic and extrinsic motivation with performance is culture-specific. In terms of gender, the results are consistent with gender stereotyping (e.g., higher mathematics self-concepts and values among males) as described in cross-sectional and longitudinal studies (e.g., Jacobs et al., 2002; Marsh et al., 2006; Nagy et al., 2010; Watt, 2004). With regard to gender differences in mathematics achievement, it was found that females sometimes outperformed males, males sometimes outperformed females, or there were no differences, depending on the culture or country.

Although the finding of a relationship between affect and achievement is certainly not unique in the literature, it is the first of its kind involving structural equation modeling accounting for measurement errors and bias in large-scale assessment in an African context. The evidence of a reciprocal relationship between affect and achievement in mathematics is a good indication of the "struggle" for supremacy between affect and achievement, as indicated in the various educational systems. The TIMSS data set (e.g., mathematics-related motivational belief constructs) was used in these studies as a traditional testing ground for new and evolving theoretical models in the research on mathematics-related affect, as well as a major focus of critical debate on the relationship among these constructs.

All in all, the study makes a significant contribution to indigenous research that integrates cross-cultural perspectives and is crucial to the establishment of more useful and universal theories.

5.3 A proposed model

The model shown in Figure 1 summarizes the ways in which the variables considered in this dissertation appear to interact. The dotted lines indicate relationships that were culturally dependent, and dotted lines with double-headed arrows indicate a possible reciprocal relationship. The relationships linking parental involvement, teacher responsiveness, mathematics-related affect (e.g., self-concept), and achievement were culture-specific, although a reciprocal relationship was evident in some cultures. For instance, the magnitude of the relationship between achievement and "students like mathematics" or "students value mathematics" was different in the different nations: in some, it was strong with regard to liking mathematics, whereas in others it was stronger with regard to value. Gender had a significant but varying effect on some of the factors that influence mathematics achievement: there were gender differences favoring boys for mathematics confidence and liking mathematics, and favoring girls for teacher responsiveness, and cultural differences in the relationship with valuing mathematics.

The dashed and dotted lines indicate inconclusive findings based on this dissertation. All other interconnections indicate universal patterns. The shaded shapes indicate constructs that were culture-specific. Given the cultural dimension in the diagram (Figure 1), further research is needed to explore the complex interplay between these variables. More extensive studies including more countries with different cultural dimensions and involving international large-scale research could shed more light on the findings reported here.

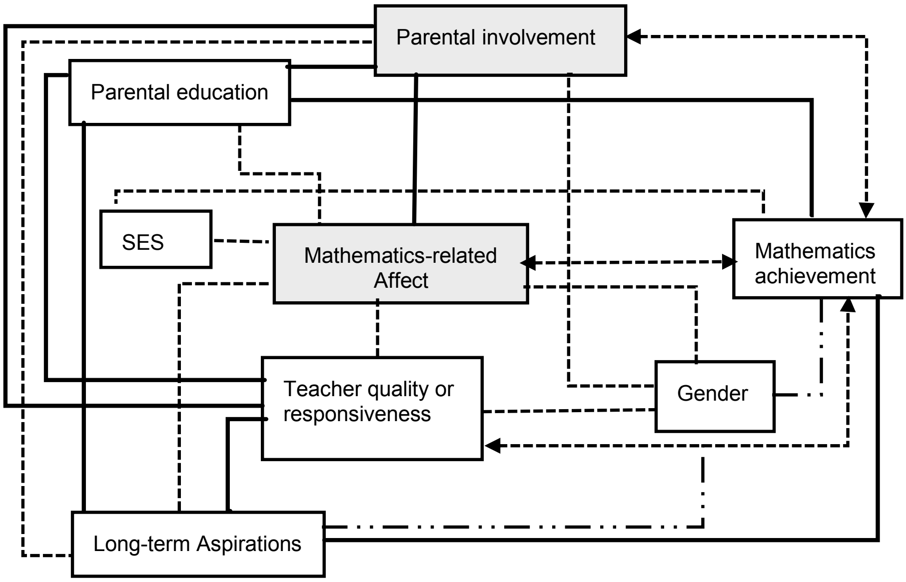

Figure 1. The relationships among the variables considered in this dissertation. The dotted lines indicate relationships that were culturally dependent, whereas the dotted lines with double-headed arrows indicate a possible reciprocal relationship. The dashed and dotted lines indicate inconclusive findings based on this dissertation. The dark lines indicate universal outcomes. The shaded shapes indicate constructs that are cultural dependent.

5.4 Practical implications

This dissertation contributes to the literature on mathematics-related affect in several ways. The implications extend to test development and validation, mathematics instruction and policy-making. More specifically, the findings enhance understanding of the conceptual foundations and statistical procedures underlying the validation of mathematics-related-affect instruments and their development. The robust approach to construct validation and revision, and to the underlying latent structure of multi-component instruments, adds new insights. The development of an effective method for revising a scale in order to achieve higher psychometric standards and better ways of dealing with scale-validation issues is a further contribution.

The failure of the original VOM scale highlights the importance of adapting instruments to different cultural settings. However, the validity of the four-factor VOM model supports the multifaceted nature of mathematics-related affect. This is a clear indication that researchers and policy makers should be careful when making direct comparisons between nations, cultures or educational systems based on manifest variables. I am not suggesting that this conceptualization of VOM represents a comprehensive overview of all the indicators because of the sampling and the structural dimension identified via the new measures, however.

Moreover, because the observed variables are prone to measurement error, modeling the means of latent variables may be more meaningful than using the observed or manifest variables in any comparison of means because manifest variables are known to reflect both the common and specific sources of variance: as such, any culture-related effects may influence not only the common construct-related variance of a set of variables but also the specific variance of one or more of them (Little, 1997). Although the studies discussed here indicate that the strength of the method effect varies across different educational systems and contexts, a careful examination of the literature suggests that common method variance is often a problem and researchers should do whatever they can to control for it (Podsakoff et al., 2003). As I have emphasized throughout this dissertation, this requires implementing the right procedure and statistical methods to control for the method effect. Although this work has not resolved all these issues, I hope it provides some useful suggestions and, perhaps more importantly, a framework that researchers can use when validating and evaluating the potential biasing effects of method variance associated with negatively phrased items. I suggest that researchers should be sufficiently aware of the potential bias associated with survey measures involving negatively worded items to take the necessary steps to deal with it appropriately. It could be concluded from the studies comprising this dissertation that TIMSS should place more emphasis on positively worded items, as other researchers have proposed (Corwyn, 2000; Marsh, 1996). However, this will entail a lack of control for the effect of response-pattern bias (e.g., acquiescence).

With regard to measurement invariance, the studies considered here indicate that the cultural background of the respondents can strongly affect the specific components of an indicator (e.g., when a facet of an item is perceived differently across cultures, or the item is inappropriate or poorly translated), resulting in non-invariance. Although measurement non-invariance may be a meaningful analytical outcome in cross-cultural research if it is sufficiently large (e.g., scaler invariance), it disallows any meaningful quantitative construct comparisons. An important aspect of measurement invariance (e.g., metric invariance) is that the constructs are defined in precisely the same way in each group; hence, any meaningful group comparison can be done with quantitative precision.

The reliabilities of the constructs were problematic although acceptable in all four studies discussed here. Greater emphasis should be given to improving the reliabilities of the motivational-belief constructs in low-achievement countries such as all those participating in this study, otherwise policy makers and researchers cannot posit any meaningful links between student motivation and achievement.

Given the findings associated with the motivational constructs, achievement and the background variables, researchers, educators and policy makers should be more cautious in generalizing the relationship between affect and achievement on a cross-cultural level. In the context of educational assessment, the current work attests the importance of students' motivational beliefs as a component of their achievement in mathematics. Findings such as this have implications for teachers, policy makers and test organizations in terms of taking into account students' motivational beliefs in assessing their academic performance. As a result, they may make better-informed decisions about educational policies and curriculum changes normally implemented based on students' performance in large-scale assessments.

With respect to instructional practices, the relationship between mathematics-related affect, especially the reciprocal effect, indicates that intervention directed towards strengthening the relationship between students' self-belief and achievement should be bi-directional to have any positive impact on both measures: self-enhancement and skills-development models should be integrated into any intervention programs. However, these interventions should be culture-specific based on self-enhancement or skills development. For instance, efforts aimed at improving students' achievement (such as school reforms or curriculum reviews), which are likely to improve self-belief, should be implemented if the aim is to improve skills development. Interventions that promote positive self-beliefs with respect to mathematics learning and achievement should also be encouraged. However, as indicated above, one should be sensitive to the possibilities and limitations associated with educational development in mathematics in African countries to avoid disappointments in the impact of educational reform.

Teacher responsiveness had a stronger association with students' intrinsic and extrinsic motivation and self-confidence than parental involvement. Teachers and parents should thus be encouraged to create or promote measures of motivational belief such as initiating or supporting school or classroom interventions whereby students are free to interact with teachers and ask questions (Guo et al., 2015; Midgley et al., 1989; O'Dwyer, Wang, & Shields, 2015; M.-T. Wang & Eccles, 2012; M.-T. Wang, 2012). Mathematics teachers could also devise ways of arousing students' interest in mathematics by implicitly or explicitly showing its relevance and giving feedback in order to raise their expectations of success. Given the results of this dissertation, mathematics teachers are recommended to devise appropriate interventions in their instruction to foster student's motivational beliefs, which in turn will enhance their performance. Finally, teachers and parents should aim at instilling positive attitudes toward mathematics at school and at home (Areepattamannil & Freeman, 2008; Burden, 1995).

In sum, the results discussed here indicate that students who are motivated to achieve and who have high motivational beliefs are more likely to achieve better scores in mathematics, and vice versa (in respect of mathematics self-concept). Therefore, appropriate programs should be tailored to positively influence these psychological indicators. Encouraging teachers to use instructional strategies such as subject-domain-specific learning activities and teaching components that reflect executive learning processes (e.g., mathematics problem solving), social-context experiences (e.g., discussion with teachers and parents), and allowing time for learning, regulation and monitoring (e.g., teaching students self-regulation and monitoring strategies) could help to improve students' motivation and confidence (Organisation for Economic Co-operation and Development [OCED], 2011; Seidel & Shavelson, 2007). Collectively, the studies discussed here indicate that students are sensitive to and aware of the cultural environment around them, and adjust their motivational beliefs and themselves as learners of mathematics to reflect prevailing societal norms. The findings reveal the apparent impact of the socio-cultural climate on students' motivational-belief structures.

This dissertation should be useful to researchers wishing to explore the relationship between mathematics-related affect and achievement in an African context. Although there have been several studies examining cross-national and cross-cultural differences in mathematics achievement, mathematics-related affect, and parental and teacher involvement (e.g., Else-Quest et al., 2010), the involvement of African countries has been limited, and in some cases non-existent.

5.5 Strengths, limitations, and future directions

Some strengths and limitations of the studies have to be considered in interpreting the findings. Despite the potential limitations, the TIMSS sampling design provided a useful yardstick for the present motivational study in terms of generalization to specific educational systems. The TIMSS survey is one of the few cross-sectional large-scale-assessment data sets to provide diverse motivational constructs based on multiple items as well as both cognitive and non-cognitive measures, although the psychometric properties were weak in our sample. The main characteristics associated with the TIMSS data set coupled with using the right methodology, as discussed in this dissertation, enabled us to investigate the relationship between dimensions of mathematics-related affect among students, corrected for measurement error, and their mathematics achievement across the different educational systems. The available data also enabled us to develop a better theoretical understanding of the roles that mathematics-related motivational beliefs and students' background variables play in shaping their mathematics achievement, and vice versa. Indeed, the TIMSS data set has been used as a traditional testing ground for new and evolving theoretical models of mathematics-related affect, and in the debate on the relationship between these constructs and student achievement. Given the regular cycle of studies (TIMSS 1995, TIMSS 1999, TIMSS 2003, TIMSS 2007, TIMSS 2011, TIMSS 2015), the data is well suited to testing the long-term implications of measures of students' motivational belief, which are among the key aspects of modern expectancy-value theory.

However, the sampling procedure does not justify the confident generalization of the findings of Studies I and II to other educational contexts. Specifically, the nine schools participating in the study in no way represent the full range of secondary schools in Ghana, but were nevertheless chosen to represent typical types of high school in terms of social intake, disciplines offered, and rates of academic success and failure. Therefore, the results cannot be generalized to students in all schools, although they are representative of students in a range that is "typical" of the secondary-school system in Ghana.

There are two very common problems associated with surveys such as the TIMSS standardized assessment. The first concerns the school subject (e.g., mathematics or science) and the second is the period of assessment. In that respect the assessment is both a snapshot of what is really learned in school as well as an ongoing dynamic process. In the end, it is skewed towards a few school subjects without cognizance of the fact that there are other major subject areas in the school curriculum as well.

A further problem with large-scale assessments such as TIMSS is that the participating countries may have curricula that are misaligned with the contents of the study (Howie & Hughes, 2000; Ndlovu & Mji, 2012). There is evidence that the greater the misalignment of items in the survey, the poorer is the students' performance in those countries (Altinok, 2008).

Another limitation is that the point of large-scale international surveys is to generalize and/or transpose any relationships within a particular country to another. For instance, as demonstrated in this and other studies (e.g., Wilkins et al., 2002), parental involvement influences pupil performance, but the relationship is culture-specific. These problems are even more serious in cross-cultural comparisons because the issue of parental involvement in children's education is a complex phenomenon that often transcends the geographical boundaries of home and school. Another cross-cultural dimension concerned the relationship between the value of mathematics and achievement, which was culture-specific although males in general reported higher values than females.

The low reliabilities associated with some of the measures of mathematics-related affect in this dissertation imply substantial measurement error and/or limited true individual differences, and this may attenuate the validity of interpretations of the findings based on manifest scale scores, and weaken the statistical power. The use of latent-variable models that account for unreliability, bias, and measurement error, such as those used in all the studies comprising this dissertation, is therefore recommended. However, the differences between the latent means of the motivational constructs and achievement should be interpreted with caution, given the low reliabilities of some of the constructs, the method effect associated with the negatively phrased items, and the low performance in mathematics achievement among students in TIMSS 2011 in the five participating countries. The eighth graders' achievement in Morocco and Ghana was significantly low (percentage of students with achievement too low for estimation is over 25 percent), as was that of the ninth-graders in South Africa and Botswana (percentage of students with achievement too low for estimation is over 15 but not exceeding 25 percent) (Mullis et al., 2012).

Modeling reciprocal determinism on cross-sectional data is a very challenging exercise, which requires theoretical–substantive assumptions not subject to any test that the system has reached equilibrium, such as about fixed zero effects. Given the challenges with non-recursive models in causality analysis, replicating this study with a longitudinal data set should bolster the outcome. However, international large-scale longitudinal studies, if not impossible will be administratively difficult and expensive, hence cross-sectional data is the only option. The outcome of this study implies that the feedback loop, or the relationship between affect and achievement, could extend far beyond that in previous studies limited to the strength of the relationships between these measures. However, it should be noted that the outcome is a *snapshot* of an ongoing dynamic system and may vary depending on when it was observed.

Furthermore, scalar measurement invariance is frequently not satisfied in practice in cross-cultural research, and partial measurement invariance is often used as a “compromise” between full and a complete lack of invariance (Byrne et al., 1989). The number of invariance intercepts in this study was the minimal necessary for comparisons of cross-national differences in factor means to be meaningful (Steenkamp & Baumgartner, 1998).

Finally, we examined the less commonly studied population using a nationally representative sample that allowed us to generalize the results to a larger population. However, to gain a better sense of the influence of students’ mathematics-related affect as measured in this study; future research should incorporate the effects of the imposed school structure and cultural norms. Models presented in this dissertation are not uniformly “best”, but for any specific form of CFA model that is influenced by the method effect; there is always a preferred option, which was considered.

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*(3), 219–234.

Abu-Hilal, M. M. (2001). Correlates of achievement in the United Arab Emirates: A sociocultural study. In D. M. McInerney & V. S. Etten (Eds.), *Research on Sociocultural Influences on Motivation and Learning* (Vol. 1, pp. 205–230). Greenwich, CT: Information Age.

Aiken, L. R. (1971). Verbal factors and mathematics learning: A review of research. *Journal for Research in Mathematics Education*, *2*(4), 304–313.

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 Countries. *Educational Researcher*, *36*(7), 369–387. doi:10.3102/0013189X07308739

Altinok, N. (2008). An international perspective on trends in the quality of learning achievement (1965-2007). Education for all Gobal Monitoring Report 2009, Overcoming inequality:Why governance matters. UNESCO. Retrieved from unesco.org

Ambady, N., Paik, S. K., Steele, J., Owen-Smith, A., & Mitchell, J. P. (2004). Deflecting negative self-relevant stereotype activation: The effects of individuation. *Journal of Experimental Social Psychology*, *40*(3), 401–408. doi:10.1016/j.jesp.2003.08.003

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

Andrews, P., Diego-Mantecón, J., Eynde Op't, P., & Sayers, J. (2007). Evaluating the sensitivity of the refined mathematics-related beliefs questionnaire to nationality, gender and age. In D. Pitta-Pantazi & G. Philippou (Eds.), *Proceedings of the fifth congress of the European Society for Research in Mathematics Education* (pp. 209–218). larnaca, Cyprus: Department of Education, University of Cyprus.

Areepattamannil, S., & Freeman, J. G. (2008). Academic achievement, academic self-concept, and academic motivation of immigrant adolescents in the greater Toronto area secondary schools. *Journal of Advanced Academics*, *19*(4), 700–743.

Aronson, J., & Good, C. (2003). The development and consequences of stereotype vulnerability in adolescents. In F. Pajares & T. Urdan (Eds.), *Adolescence and education. Academic motivation of adolescents* (pp. 299–330). Greenwich, CT: Information Age Publishing.

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science*, *11*(5), 181–185.

Asparouhov, T., & Muthén, B. O. (2010). Multiple imputation with Mplus. *Mplus Technical appendices.Los Angeles, CA: Muthen & Muthen*. Retrieved from http://statmodel2.com/download/Imputations7.pdf

Asparouhov, T., & Muthén, B. O. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 495–508. doi:10.1080/10705511.2014.919210

American Education Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

Astone, N. M., & McLanahan, S. S. (1991). Family Structure, Parental Practices and High School Completion. *American Sociological Review*, *56*(3), 309-320. doi:10.2307/2096106

Bachman, J. G., & O'Malley, P. M. (1986). Self-concepts, self-esteem, and educational experiences: The frog pond revisited (again). *Journal of Personality and Social Psychology*. doi:10.1037/0022-3514.50.1.35

Bagozzi, R. P. (1993). Assessing construct validity in personality research: Applications to measures of self-esteem. *Journal of Research in Personality*. doi:10.1006/jrpe.1993.1005

Barkatsas, A. N., Forgasz, H. J., & Leder, G. C. (2001). The gender stereotyping of mathematics: Cultural dimensions. In *24th Annual MERGA Conference, Sydney* (Vol. 3, pp. 79–86). Sydney.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(5), 1860–1863. doi:10.1073/pnas.0910967107

Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, *74*(1), 137–143. doi:10.1007/s11336-008-9100-1

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. doi:10.1037/0033-2909.88.3.588

Berry, J. W. (1985). Learning mathematics in a second language: Some cross-cultural issues. *For the Learning of Mathematics* *5*(2), 18-23.

Bhanot, R. T., & Jovanovic, J. (2009). The links between parent behaviors and boys' and girls' science achievement beliefs. *Applied Developmental Science*, *13*(1), 42–59. doi:10.1080/10888690802606784

Bharadwaj, P., De Giorgi, G., Hansen, D., & Neilson, C. (2012). *The Gender Gap in Mathematics: Evidence from Low- and Middle-Income Countries* (No. 18464). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w18464

Biggs, J. B. (1996). Western misconceptions of the Confucian-Heritage learning culture. In A. Watkins & J. B. Biggs (Eds.), *The Chinese learner:Cultural, pychological and contextual influrence* (pp. 45–67). Hong Kong: Comparative Education Research Centre.

Boaler, J., & Sengupta-Irving, T. (2006). Nature, neglect and nuance: Changing accounts of sex, gender and mathematics. In C. Skelton, B. Francis, & L. Smulyan (Eds.), *The SAGE handbook of gender and education* (pp. 207–220). Thousand Oaks, CA: SAGE.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40.

Bos, K., & Kuiper, W. (1999). Modelling TIMSS Data in a European Comparative Perspective : Exploring Influencing Factors on Achievement in Mathematics in Grade 8. *Educational Research and Evaluation: An International Journal on Theory and Practice*, *5*(2), 157–179.

Brown, M., Brown, P., & Bibby, T. (2008). “I would rather die”: reasons given by 16-year-olds for not continuing their study of mathematics. *Research in Mathematics Education*, *10*(1), 3–18. doi:10.1080/14794800801915814

Brunner, M., Gogol, K. M., Sonnleitner, P., Keller, U., Krauss, S., & Preckel, F. (2013). Gender differences in the mean level, variability, and profile shape of student achievement: Results from 41 countries. *Intelligence*, *41*(5), 378–395. doi:10.1016/j.intell.2013.05.009

Burden, P. R. (1995). Classroom management and discipline: Methods to facilitate cooperation and instruction. White Plains, NY: Longman.

Byrne, B. M. (2000). Measuring self-concept across culture: Issues, caveats, and practice. In H. W. Marsh & R. G. Craven (Eds.), *Self-concept theory, research and pactice: Advances for the New Millennium* (pp. 30–41). SELF Research Centre, University of Western Sydney.

Byrne, B. M., Shavelson, R. J., & Muthén, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–466.

Byrne, B. M., & van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*. doi:10.1080/15305051003637306

Calsyn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: cause or effect of academic achievement? *Journal of Educational Psychology*, *69*(2), 136–45. doi:10.1037/0022-0663.69.2.136

Campbell, J. R., & Mandel, F. (1990). Connecting math achievement to parental influences. *Contemporary Educational Psychology*. doi:10.1016/0361-476X(90)90006-M

Cates, G. L., & Rhymer, K. N. (2003). Examining the relationship between mathematics anxiety and mathematics performance: An instructional hierarchy perspective. *Journal of Behavioral Education*, *12*(1), 23–34. doi:10.1023/A:1022318321416

Chaman, M. J., Beswick, K., & Callingham, R. (2014). Factors influencing mathematics achievement among secondary school students: A review. In N. Fitzallen, R. Reaburn, & F. Fan (Eds.), *The future of educational research: Perspectives of beginning researchers* (pp. 227–238). Rotterdam, The Netherlands: Sense Publishers.

Chen, C., & Stevenson, H. W. (1995). Motivation and Mathematics Achievement: A Comparative Study of Asian-American, Caucasian-American, and East Asian High School Students. *Child Development*, *66*(4), 1215–1234.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. doi:10.1080/10705510701301834

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology*, *95*(5), 1005–18. doi:10.1037/a0013193

Chen, P. P. (2002). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences*, *14*(1), 79–92. doi:10.1016/j.lindif.2003.08.003

Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., & Lin, S. S. J. (2013). The relationship between academic self-concept and achievement: A multicohort–multioccasion study. *Learning and Individual Differences*, *23*, 172–178.

Cheung, G. W., & Rensvold, R. B. R. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255. doi:10.1207/S15328007SEM0902_5

Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups : A reconceptualization and proposed new method. *Journal of Management*, *25*(1), 1–27.

Chiu, M. M., & Klassen, R. M. (2010). Relations of mathematics self-concept and its calibration with mathematics achievement: Cultural differences among fifteen-year-olds in 34 countries. *Learning and Instruction*, *20*(1), 2–17.

Chiu, M. M., & Xihua, Z. (2008). Family and motivation effects on mathematics achievement: Analyses of students in 41 countries. *Learning and Instruction*, *18*(4), 321–336. doi:10.1016/j.learninstruc.2007.06.003

Chiu, M.-S. (2008). Achievements and self-concepts in a comparison of math and science: exploring the internal/external frame of reference model across 28 countries. *Educational Research and Evaluation*, *14*(3), 235–254. doi:10.1080/13803610802048858

Chiu, M.-S. (2009). Affective, cognitive, and social factors in reducing gender differences in measurement and algebra achievements. In M. Tzekaki, M. Kaldrimidou, & C. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education* (pp. 321–328). Thessaloniki, Greece: PME.

Chiu, M.-S. (2012a). Differential psychological processes underlying the skill-development model and self-enhancement model across mathematics and science in 28 Countries. *International Journal of Science and Mathematics Education*, *10*(3), 611–642.

Chiu, M.-S. (2012b). The internal/external frame of reference model, big-fish-little-pond effect, and combined model for mathematics and science. *Journal of Educational Psychology*, *104*(1), 87–107. oi:10.1037/a0025734

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 eigth Congress of the European Society for Research in Mathematics Education* (pp. 1855–1864). Antalya, Turkey: ERME.

Cogan, L. S., & Schmidt, W. H. (2002). " Culture Shock " – eighth- grade mathematics from an international perspective. *Educational Research and Evaluation: An International Journal on Theory and Practice*, *8*(1), 13–39.

Cole, D. a, & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*(2), 300–15. doi:10.1037/a0033805

Coley, R. J. (2002). *An uneven start: Indicators of inequality in school readiness*. Princeton, NJ: Educational Testing Service.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104.

Corwyn, R. F. (2000). The factor structure of global self-esteem among adolescents and adults. *Journal of Research in Personality*, *34*, 357–379.

Cote, J. A., & Buckley, R. M. (1987). Estimating trait, method, and error Variance: Generalizing Across 70 construct validation studies. *Journal of Marketing Research*, *24*(3), 315–318. doi:Doi 10.2307/3151642

Cretchley, P. C. (2008). Advancing research into affective factors in mathematics learning: Clarifying key factors, terminology and measurement. In M. Goos, R. Brown, & K. Makar (Eds.), *Proceedings of the 31st Annual*

Conference of the Mathematics Education Research Group of Australasia (pp. 147–153). Brisbane, QLD: MERGA.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334. doi:10.1007/BF02310555

Davis, H. A. (2003). Conceptualizing the Role and Influence of Student-Teacher Relationships on Children's Social and Cognitive Development. *Educational Psychologist*, *38*(4), 207–234.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology*, *19*(2), 294.

Davis-Kean, P., Jacobs, J., Bleeker, M., Eccles, J. S., & Melanchuk, O. (2007). How Dads Influence Their Daughters' Interest In Math. *ScienceDaily, University Of Michigan*. Retrieved April 10, 2015, from http://www.sciencedaily.xn--com-4n0a/releases/2007/06/070624143002.htm

Deaux, K., & Major, B. (1987). Putting gender into context: An interactive model of gender-related behavior. *Psychological Review*, *94*(3), 369–389.

DeBellis, V. A., & Goldin, G. A. (1997). The affective domain in mathematical problem-solving. In E. Pehkonen (Ed.), *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 209–16). Helsinki: University of Helsinki.

DeBellis, V. A., & Goldin, G. A. (2006). Affect and meta-affect in mathematical problem solving: A representational perspective. *Educational Studies in Mathematics*, *63*(2), 131–147. doi:10.1007/s10649-006-9026-4

Desimone, L. (1999). Linking parent involvement with student achievement: Do race and income matter? *The Journal of Educational Research*, *93*(1), 11.

Distefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*, *13*(3), 440–464. doi:10.1207/s15328007sem1303_6

Distefano, C., & Motl, R. W. (2009). Methodological artifact or substance? Examinations of wording effects associated with negatively worded items. In T. Teo & K. M. Swe (Eds.), *Structural Equation Modeling in Educational Research: Concepts and Applications* (pp. 59–77). Boston: Sense Publishers.

DiStefano, C., & Motl, R. W. (2009). Personality correlates of method effects due to negatively worded items on the Rosenberg Self-Esteem scale. *Structural Equation Modeling in Educational Research: Concepts and Applications*, *46*(3), 309–313. doi:10.1016/j.paid.2008.10.020

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly*, *18*, 585–609.

Eccles, J. S. (2009). Who Am I and What Am I Going to Do With My Life? Personal and Collective Identities as Motivators of Action. *Educational Psychologist*, *44*(2), 78–89. doi:10.1080/00461520902832368

Eccles, J. S. (2011a). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development*. *35*(3), 195-201. doi:10.1177/0165025411398185

Eccles, J. S. (2011b). Understanding educational and occupational choices. *Journal of Social Issues*. *67*(3), 644-648. doi:10.1111/j.1540-4560.2011.01718.x

Eccles, J. S., & Jacobs, J. E. (1986). Social forces shape math attitudes and performance. *Signs: Journal of Women in Culture and Society*, *11*(2), 367-380. doi:10.1086/494229

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues*, *46*(2), 183–201. doi:10.1111/j.1540-4560.1990.tb01929.x

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132.

Eccles[Parsons], J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectations, values, and academic behavious. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 76–146). San Francisco, CA: W. H. Freman and Company.

Eccles[Parsons], J. S., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development*, *53*(2), 310–321.

Eccles[Parsons], J. S., Adler, T., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, *46*(26--43), 26–43. doi:10.1037/0022-3514.46.1.26

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103. doi:10.1037/a0018053

Else-Quest, N. M., Mineo, C. C., & Higgins, A. (2013). Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychology of Women Quarterly*, *37*, 293–309. doi:10.1177/0361684313480694

Epstein, J. L. (2010). School/Family/Community partnerships: Caring for the children we share. *Phi Delta Kappan*, *92*(3), 81–96.

Erberber, E., Stephens, M., Mamedova, S., Ferguson, S., & Kroeger, T. (2015). *Socioeconomically disadvantaged students who are academically successful: Examining academic resilience cross-nationally* (No. 5). *IEA'S Policy Brief Series*. Amsterdam, The Netherlands: IEA. Retrieved from http://www.iea.nl/policy_briefs.html

Fan, W., & Williams, C. M. (2010). The effects of parental involvement on students' academic self-efficacy, engagement and intrinsic motivation. *Educational Psychology*, *30*(1), 53–74. doi:10.1080/01443410903353302

Fan, X. (2001). Parental involvement and students' academic achievement: A growth modeling analysis. *The Journal of Experimental Education*, *70*(1), 27–61. doi:10.1080/00220970109599497

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review*, *13*(1), 1–22.

Fennema, E., & Sherman, J. A. (1976). Fennema-Sherman mathematics attitudes scales. *JSAS Catalogue of Selected Documents in Psychology*, *6*, 31.

Filmer, D., & Pritchett, L. (1999). The effect of household wealth on educational attainment: evidence from 35 countries. *Population and development1*, *25*(1), 85–420. doi:10.1111/j.1728-4457.1999.00085.x

Flore, P. C., & Wicherts, J. M. (2014). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology*. doi:10.1016/j.jsp.2014.10.002

Forgasz, H. J., & Mittelberg, D. (2008). Israeli Jewish and Arab students' gendering of mathematics. *ZDM*, *40*(4), 545–558. doi:10.1007/s11858-008-0139-3

Forgasz, H., Leder, G., Mittelberg, D., Tan, H., & Murimo, A. (2015). Affect and Gender. In B. Pepin & B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education* (pp. 245–268). Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-06808-4_12

Foy, P., Brossman, B., & Galia, J. (2013). *Scaling the TIMSS and PIRLS 2011 Achievement Data*. Chestnut Hill, MA. Retrieved from http://timssandpirls.bc.edu/methods/index.html

Frade, C., Roesken, B., & Hannula, M. S. (2010). Identity and affect in the context of teachers' professional development. In M. M. F. Kawasaki & T. F. Pinto (Eds.), *Proceedings of the 34th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 247–9). Belo Horizonte, Brazil: PME.

Fredricks, J. a, & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: growth trajectories in two male-sex-typed domains. *Developmental Psychology*, *38*(4), 519–533.

Frempong, G. (2010). Equity and quality mathematics education within schools : findings from TIMSS data for Ghana. *African Journal of Research in Mathematics*, *14*(3), 50–62. doi:10.1080/10288457.2010.10740691

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics —A "hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*. *22*(4), 497-514. doi:10.1007/BF03173468

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, *2*(2), 210–240.

Furinghetti, F., & Pehkonen, E. (2002). Rethinking characterizations of beliefs. In G. C. Leder, E. Pehkonen, & G. Toörner (Eds.), *In Beliefs:A hidden variable in mathematics education?* (pp. 39–58). Dordrecht, The Netherlands: Kluwer.

Ganley, C. M., & Vasilyeva, M. (2011). Sex differences in the relation between math performance, spatial skills, and attitudes. *Journal of Applied Developmental Psychology*, *32*(4), 235–242.

Gates, P., & Vistro-Yu, C. P. (2003). Is mathematics for all? *Second International Handbook of Mathematics Education*, *10*, 31–73.

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. doi:10.1037/a0032138

Gil-Flores, J., Padilla-Carmona, T. M., & Suárez-Ortega, M. (2011). Influence of gender, educational attainment and family environment on the educational aspirations of secondary school students. *Educational Review*, *63*(3), 345–363. doi:10.1080/00131911.2011.571763

Gonzalez-DeHass, A. R., Willems, P. P., & Holbein, M. F. D. (2005). Examining the relationship between parental involvement and student motivation. *Educational Psychology Review*, *17*(2), 99–123. doi:10.1007/s10648-005-3949-7

Graham, J. W. (2009). Missing data analysis: making it work in the real world. *Annual Review of Psychology*, *60*, 549–576.

Green, S. B., & Hershberger, S. L. (2000). Correlated errors in true score models and their effect on Coefficient Alpha. *Structural Equation Modeling: A Multidisciplinary Journal*, *7*(2), 251–270.

Green, S. B., & Yang, Y. (2009). Commentary on coefficient alpha: A cautionary tale. *Psychometrika*, *74*, 121–135. doi:10.1007/s11336-008-9098-4

Grootenboer, P., & Hemmings, B. (2007). Mathematics Performance and the Role Played by Affective and Background Factors. *Mathematics Education Research Journal*, *19*(3), 3–20.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*(1), 124–136.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, Gender, and Math. *Science*, *320*(5880), 1164–5. doi:10.1126/science.1154094

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The Role of Parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*(3-4), 153–166. doi:10.1007/s11199-011-9996-2

Guo, J., Marsh, H. W., Morin, A. J. S., Parker, P., & Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations and attainment over eight years: A multiwave, longitudinal study. *American Educational Research Journal*, 0002831214565786. doi:10.3102/0002831214565786

Haag, N., Heppt, B., Stanat, P., Kuhl, P., & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction*, *28*, 24–34.

Hannula, S. M. (2002). Attitude towards mathematics: Emotions expectations and values. *Educational Studies in Mathematics*, *49*(1), 25-46

Hannula, S. M. (2006). Motivation in mathematics: Goals reflected in emotions. *Educational Studies in Mathematics*, *63*(2), 165–178. doi:10.1007/s10649-005-9019-8

Hannula, S. M. (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: University of Rzeszów, Poland.

Hannula, S. M. (2012). Exploring new dimensions of mathematics-related affect: embodied and social theories. *Research in Mathematics Education*, *14*(2), 137–161. doi:10.1080/14794802.2012.694281

Hannula, M. S. (2015a). Emotions in problem solving. In J. S. Cho (Ed.), *Selected Regular Lectures from the 12th International Congress on Mathematical Education* (pp. 269–288). Cham: Springer International Publishing. doi:10.1007/978-3-319-17187-6_16

Hannula, S. M. (2015b). Reaction to section 2: 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 SE - 13* (pp. 269–277). Cham, Switzerland: Springer International Publishing.

Hannula, S. M., Bofah, E. A., Tuohilampi, L., & Metsämuuronen, J. (2014). A longitudinal analysis of the relationship between mathematics-related affect and achievement in Finland. In S. Oesterle, P. Liljedahl, C. Nicol,

& D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (pp. 249–256). Vancouver, Canada, Canada: PME.

Hannula, S. M., Maijala, H., Pehkonen, E., & Nurmi, A. (2005). Gender comparisons of pupils ' self-confidence in mathematics learning. *Nordic Studies in Mathematics*, *10*(3), 29–42.

Hansford, B. C., & Hattie, J. A. (1982). The Relationship Between Self and Achievement/Performance Measures. *Review of Educational Research*, *52*(1), 123–142. doi:10.3102/00346543052001123

Hassan, M. M., & Khalifa, A. (1999). Sex differences in science achievement across ten academic years among high school students in the United Arab Emirates. *Psychological Reports*, *84*, 747–757.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses related to achievement*. London: Routledge Taylor and Francis Group.

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, *33*(4), 1829–1878.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, *21*(1), 33–46. doi:10.2307/749455

Herbert, J., & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of Applied Developmental Psychology*, *26*(3), 276–295.

Herring, M., & Wahler, R. G. (2003). Children's cooperation at school: The comparative influences of teacher responsiveness and the children's home-based behavior. *Journal of Behavioral Education*, *12*(2), 119–130.

Hooper, M., Arora, A., Martin, M. O., & Mullis, I. V. S. (2013). Examining the behavior of "reverse directional" items in the TIMSS 2011 context questionnaire scales. In *IEA International Research Conference*. Nanyang, Singapore.

Horan, P. M., DiStefano, C., & Motl, R. W. (2003). Wording effects in self-esteem scales: methodological artifact or response style ? *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(3), 435–455.

House, J. D. (2006). Mathematics beliefs and achievement of elementary school students in Japan and the United States: results from the Third International Mathematics and Science Study. *The Journal of Genetic Psychology*, *167*(1), 31–45. doi:10.3200/GNTP.167.1.31-45

Howie, S. J. (2013). Language and other background factors affecting secondary pupils ' performance in mathematics in South Africa. *African Journal of Research in Mathematics, Science and Technology Education*, *7*(1), 1–20.

Howie, S. J., & Hughes, C. (2000). South Africa. In D. Robitaille, A. Beaton, & T. Plomb (Eds.), *The Impact of TIMSS on the Teaching and Learning of*

Mathematics and Science (pp. 139–145). Vancouver, BC: Pacific Educational Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*. *6*(1), 1-55. doi:10.1080/10705519909540118

Hughes, J., & Kwok, O.-M. (2007). Influence of Student-Teacher and Parent-Teacher Relationships on Lower Achieving Readers' Engagement and Achievement in the Primary Grades. *Journal of Educational Psychology*, *99*(1), 39–51. doi:10.1037/0022-0663.99.1.39

Hughes, J. N., Gleason, K. a, & Zhang, D. (2005). Relationship influences on teachers' perceptions of academic competence in academically at-risk minority and majority first grade students. *Journal of School Psychology*, *43*(4), 303–320. doi:10.1016/j.jsp.2005.07.001

Huntsinger, C. S., & Jose, P. E. (2009). Parental involvement in children's schooling: Different meanings in different cultures. *Early Childhood Research Quarterly*, *24*(4), 398–410. doi:10.1016/j.ecresq.2009.07.006

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities charactize math performance. *Science*, *321*(July), 494–495. doi:10.1126/science.1160364

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(22), 8801–8807.

Ice, C. L., & Hoover-Dempsey, K. V. (2011). Linking parental motivations for involvement and student proximal achievement outcomes in homeschooling and public schooling settings. *Education and Urban Society*, *43*(3), 339–369. doi:10.1177/0013124510380418

Idaszak, J. R., & Drasgow, F. (1987). A revision of the Job Diagnostic Survey: Elimination of a measurement artifact. *Journal of Applied Psychology*, *72*(1), 69–74. doi:10.1037/0021-9010.72.1.69

International Association for the Evaluation of Educational Achievement. (2014). About TIMSS and PIRLS. Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College, and International Association for the Evaluation of Educational Achievement. Retrieved from http://timssandpirls.bc.edu/

International Mathematical Union. (2009). *Mathematics in Africa: Challenges and opportnities. A report to the John Templeton Foundation from the developing countries strategies group International mathematical union*. Retrieved from www.mathunion.org

Jacobs, J. E., & Bleeker, M. M. (2004). Girls' and boys' developing interests in math and science: do parents matter? *New Directions for Child and Adolescent Development*, (106), 5–21. doi:10.1002/cd.113

Jacobs, J. E., & Eccles, J. S. (2000). Parents task values and real-life achievement-related choices. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (14th ed., pp. 405–439). San Diego, CA: Academic Press.

Jacobs, J. E., Lanza, S., Osgood, W. D., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*(2), 509–527.

Jansen, M., Schroeders, U., & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences*, *30*, 11–21.

Jeynes, W. H. (2003). A Meta-Analysis: The effects of parental involvement on minority children's academic achievement. *Education and Urban Society*, *35*(2), 202–218. doi:10.1177/0013124502239392

Jeynes, W. H. (2007). The relationship between parental involvement and urban secondary school student academic achievement: A Meta-Analysis. *Urban Education*, *42*(1), 82–110. doi:10.1177/0042085906293818

Jones, P. L. (1982). Learning mathematics in a second language: A problem with more and less. *Educational Studies in Mathematics*, *13*(3), 269–287. doi:10.1007/BF00311245

Jurdak, M. (2014). Socio-economic and cultural mediators of mathematics achievement and between-school equity in mathematics education at the global level. *ZDM*, *46*, 1025–1037. doi:10.1007/s11858-014-0593-z

Jurdak, M. E. (1977). Structural and linguistic variables in selected inference patterns for bilinguals in grades six to ten. *Educational Studies in Mathematics*, *8*(2), 225–238. doi:10.1007/BF00241027

Jurdak, M. E. (2014). Religion and Language as Cultural Carriers and Barriers in Mathematics Education-Revisited. *Journal of Humanistic Mathematics*, *2*(4), 47–57. doi:10.5642/jhummath.201402.05

Keith, T. Z. T., Keith, P. B. P., Troutman, G. C., Bickley, P. G., Trivette, P. S., & Singh, K. (1993). Does parental involvement affect eighth-grade student achievement? Structural analysis of national data. *School Psychology Review*, *22*(3), 474–496.

Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A. M., & Patrick, H. (2006). Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology*, *42*(1), 11–26. doi:10.1037/0012-1649.42.1.11

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes and women's math performance: How implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *Journal of Experimental Social Psychology*, *43*, 825–832. doi:10.1016/j.jesp.2006.08.004

Klebanov, P. K., Brooks-Gunn, J., & Duncan, G. J. (1994). Does neighborhood and family poverty affect mothers' parenting, mental health, and social support? *Journal of Marriage and the Family*, *56*, 441–455.

Kloosterman, P., & Stage, F. F. K. (1992). Measuring beliefs about mathematical problem solving. *School Science and Mathematics*, *92*(3), 109–115.

Kupari, P., & Nissinen, K. (2013). Background factors behind mathematics achievement in Finnish education context: Explanatory models based on TIMSS 1999 and TIMSS 2011 data. In *the 5th IEA International Research Conference (IRC-2013)*. National Institute of Education, Nanyang Technological University, Singapore.

Langhout, R. D., Drake, P., & Rosselli, F. (2009). Classism in the university setting: Examining student antecedents and outcomes. *Journal of Diversity in Higher Education*. doi:10.1037/a0016209

Lareau, A. (1987). Social class differences in family-school relationships: The importance of cultural capital. *Sociology of Education*, *60*(2), 73–85.

Leaper, C., Farkas, T., & Brown, C. S. (2012). Adolescent girls' experiences and gender-related beliefs in relation to their motivation in math/science and English. *Journal of Youth and Adolescence*, *41*(3), 268–282.

Leder, G., & Forgasz, H. (2008). Mathematics education: new perspectives on gender. *ZDM*, *40*(4), 513–518. doi:10.1007/s11858-008-0137-5

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*(3), 355–365.

Lee, J., & Stankov, L. (2013). Higher-order structure of noncognitive constructs and prediction of PISA 2003 mathematics achievement. *Learning and Individual Differences*, *26*, 119–130. doi:10.1016/j.lindif.2013.05.004

Leung, F. K. S. (2002). Behind the high achievement of East Asian students. *Educational Research and Evaluation*, *8*(March 2015), 87–108.

Levpuscek, M. P., & Zupancic, M. (2008). Math achievement in early adolescence: The role of parental involvement, teachers' behavior, and students' motivational beliefs about math. *The Journal of Early Adolescence*, *29*(4), 541–570. doi:10.1177/0272431608324189

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: a review. *Educational Research*. doi:10.1080/0013188990410106

Lim, S. Y., & Chapman, E. (2012). An Investigation of the Fennema-Sherman Mathematics Anxiety Subscale. *Measurement and Evaluation in Counseling and Development*, *46*(1), 26–37.

Lindwall, M., Barkoukis, V., Grano, C., Lucidi, F., Raudsepp, L., Liukkonen, J., & Thøgersen-Ntoumani, C. (2012). Method effects: The problem with negatively versus positively keyed items. *Journal of Personality Assessment*, *94*(2), 196–204. doi:10.1080/00223891.2011.645936

Little, T. D. (1997). Mean and Covariance Structures (MACS) Analyses of Cross-Cultural Data: Practical and Theoretical Issues. *Multivariate Behavioral Research*, *32*(1), 53-76. doi:10.1207/s15327906mbr3201_3

Liu, S., & Meng, L. (2010). Re-examining factor structure of the attitudinal items from TIMSS 2003 in cross - cultural study of mathematics self-concept. *Educational Psychology: An International Journal of Experimental Educational Psychology*, *30*(6), 699–712.

Lubke, G. H., Dolan, C. V., Kelderman, H., & Mellenbergh, G. J. (2003). On the relationship between sources of within- and between-group differences and measurement invariance in the common factor model. *Intelligence*, *31*(6), 543–566. doi:10.1016/S0160-2896(03)00051-5

Ma, X. (2011). A Meta-Analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, *30*(5), 520– 540.

Ma, X., & Johnson, W. (2008). Mathematics as the critical filter: Curricular effects on gendered career choices. In H. M. G. Watt & J. S. Eccles (Eds.), *Gender and occupational outcomes:Longitudinal assessment of individual, social, and cultural influences* (pp. 55–83). Washington, DC: American Psychological Association.

Ma, X., & Kishor, N. (1997). Attitude towards self, social factors, and achievement in mathematics: A meta-analytic Review. *Educational Psychology Review*, *9*(2), 89–120.

Ma, X., Kishor, N., & Kisor, N. (2014). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for Research in Mathematics Education*, *28*(1), 26–47.

Ma, X., & Xu, J. (2004a). Determining the causal ordering between attitude toward mathematics and achievement in mathematics. *American Journal of Education*. doi:10.1086/383074

Ma, X., & Xu, J. (2004b). The causal ordering of mathematics anxiety and mathematics achievement: a longitudinal panel analysis. *Journal of Adolescence*, *27*(2), 165–79. doi:10.1016/j.adolescence.2003.11.003

Marchant, G. J., Paulson, S. E., & Rothlisberg, B. A. (2001). Relations of middle school students' perceptions of family and school contexts with academic achievement. *Psychology in the Schools*, *38*(6), 505–519.

Marjoribanks, K. (2003). Learning environments, family contexts, educational aspirations and attainment: A moderation-mediation model extended. *Learning Environments Research*, *6*(3), 247–265.

Marks, G., McMillan, J., Hillman, K., & Australian Council for Educational Research. (2001). *Tertiary entrance performance: The role of student background and school factors* (No. 22). *The Longitudinal Surveys of Australian Youth (LSAY) Research Reports*. Melbourne.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224–253. doi:10.1037/0033-295X.98.2.224

Marsh, H. W. (1986). Negative item bias in ratings scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology*, *22*(1), 37–49. doi:10.1037/0012-1649.22.1.37

Marsh, H. W. (1993). Academic self-concept: Theory, research, and measurement. In J. Suls (Ed.), *The self in social perspective. Psychological perspectives on the self* (Vol 4., pp. 59–98). Hillsdale, NJ: :Lawrence Erlbaum Associates, Inc.

Marsh, H. W. (1994). Using the national longitudinal study of 1988 to evaluate theoretical models of self-concept: The self-description questionnaire. *Journal of Educational Psychology*, *86*(3), 439–456.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology*, *70*(4), 810–819. doi:10.1037/0022-3514.70.4.810

Marsh, H. W. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, UK: British Psychological Society.

Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J. S., Abdelfattah, F., Leung, K. C., ... Parker, P. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology*, *105*(1), 108–128. doi:10.1037/a0029907

Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., & Nagengast, B. (2014). The Big-Fish-Little-Pond Effect in Mathematics: A Cross-Cultural Comparison of U.S. and Saudi Arabian TIMSS Responses. *Journal of Cross-Cultural Psychology*, *45*(5), 777–804.

Marsh, H. W., & Grayson, D. (1995). Latent variable models of multitrait–multimethod data. In R. H. Hoyle (Ed.), *Structural equation modeling:Concepts, issues and application* (pp. 177–198). Thousand Oaks, CA: Sage.

Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology*, *96*(1), 56–67. doi:10.1037/0022-0663.96.1.56

Marsh, H. W., & Hau, K.-T. (2007). Applications of latent-variable models in educational psychology: The need for methodological-substantive synergies. *Contemporary Educational Psychology*, *32*, 151–170.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis- testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341.

Marsh, H. W., & Köller, O. (2004). Unification of theoretical models of academic self-concept/achievement relations: Reunification of east and west German school systems after the fall of the Berlin Wall. *Contemporary Educational Psychology*, *29*(3), 264–282. doi:10.1016/S0361-476X(03)00034-1

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2012). Measurement invariance of Big-Five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and La Dolce Vita effects. *Developmental Psychology*. doi:10.1037/a0026913

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: unidimensional and multidimensional perspectives of self-concept. *Personality & Social Psychology Bulletin*, *34*(4), 542–52.

Marsh, H. W., Scalas, L. F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem Scale: traits, ephemeral artifacts, and stable response styles. *Psychological Assessment*, *22*(2), 366–381. doi:10.1037/a0019225

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. doi:10.1111/j.1467-8624.2005.00853.x

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of Personality*. doi:10.1111/j.1467-6494.2005.00380.x

Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology*, *89*(1), 41–54.

Martin, M. O., & Mullis, I. V. S. (Eds.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://pirls.bc.edu/methods/index.html

Martiniello, M. (2008). Math Word Problems. *Harvard Educational Review*, *78*(2), 333–369.

Martiniello, M. (2009). Linguistic complexity, schematic representations, and differential item functioning for English language learners in math tests. *Educational Assessment*, *14*(3-4), 160–179.

May, S. (2008). Introduction to volume I: Language policy and political issues in education. In S. May & N. Hornberger (Eds.), *Encyclopedia of Language and Education* (2nd ed., Vol. 1, pp. xiii–xviii). New York: Springer.

McGraw, R., Lubienski, S. T., & Strutchens, M. E. (2006). A closer look at gender in NAEP mathematics achievement and affect data: Intersections with achievement, race/ethnicity, and socioeconomic status. *Journal for Research in Mathematics Education*, *37*(2), 129–150.

McKenzie, D. J. (2005). Measuring inequality with asset indicators. *Journal of Population Economics*, *18*(2), 229–260. doi:10.1007/s00148-005-0224-7

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

McLeod, D. B. (1994). Research on affect and mathematics learning in the JRME: 1970 to the present. *Journal for Research in Mathematics Education*, *25*(6), 637–647.

McNeal Jr., R. B. (1999). Parental involvement as social capital: differential effectiveness on science achievement, truancy, and dropping out. *Oxford University Press*, *78*(1), 117–144. doi:10.2307/3005792

Meece, J. L., Glienke, B. B., & Burg, S. (2006). Gender and motivation. *Journal of School Psychology*, *44*(5), 351–373. doi:10.1016/j.jsp.2006.04.004

Meece, J. L., Parsons, J. E., Kaczala, C. M., & Goff, S. B. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin, 91*(2), 324–348. http://doi.org/10.1037/0033-2909.91.2.324

Meelissen, M., & Luyten, H. (2008). The Dutch gender gap in mathematics: Small for achievement, substantial for beliefs and attitudes. *Studies in Educational Evaluation*, *34*(2), 82–93.doi:10.1016/j.stueduc.2008.04.004

Mendick, H. (2006). *Masculinities in mathematics.* New York, NY: Open University Press.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543. doi:10.1007/BF02294825

Metsämuuronen, J. (2012a). Challenges of the Fennema-Sherman Test in the International Comparisons. *International Journal of Psychological Studies*, *4*(3). doi:10.5539/ijps.v4n3p1

Metsämuuronen, J. (2012b). Comparison of mental structures of eighth-graders in different countries on the basis of Fennema-Sherman Test. *International Journal of Psychological Studies*, *4*(4). doi:10.5539/ijps.v4n4p1

Michaelowa, K., Filmer, D., & Pritchett, L. (2001). Primary education quality in Francophone Sub-Saharan Africa : Determinants of learning achievement and efficiency considerations. *World Development*, *29*(10), 1699–1716.

Middleton, J. A., & Spanias, P. (1999). Motivation for achievement in Mathematics : Findings, generalizations, and criticisms of the research. *Journal for Research in Mathematics Education*, *30*(1), 65–88.

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Student/Teacher Relations and Attitudes towards Mathematics Before and After the Transition to Junior High School. *Child Development*.

Miller, H., & Bichsel, J. (2004). Anxiety, working memory, gender, and math performance. *Personality and Individual Differences*, *37*(3), 591–606. doi:10.1016/j.paid.2003.09.029

Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle (Ed.), *Hanbook of structural equation modeling* (pp. 380–392). New York, NY: The Guilford Press.

Moller, J., Streblow, L., Pohlmann, B., & Köller, O. (2006). An extension to the internal/external frame of reference model to two verbal and numerical domains. *European Journal of Pyschology of Education*, *XXI*(4), 467–487.

Morony, S., Kleitman, S., Lee, Y. P., & Stankov, L. (2013). Predicting achievement: Confidence vs self-efficacy, anxiety, and self-concept in Confucian and European countries. *International Journal of Educational Research*, *58*, 79–96. doi:10.1016/j.ijer.2012.11.002

Mullis, I.V.S., Martin, M.O., & Foy, P. (with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J. . (2008). *TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

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.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 Assessment Frameworks*. Amsterdam, the Netherlands: TIMSS & PIRLS International Study Center Lynch School of Education, Boston College.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide: Seventh edition* (7th ed.). Los Angeles, CA: Muthén and Muthén.

Nagengast, B., & Marsh, H. W. (2012). Big Fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology*. doi:10.1037/a0027697

Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, *12*(4), 323–345. doi:10.1080/13803610600765687

Nagy, G., Watt, H. M. G., Eccles, J. S., Trautwein, U., Lüdtke, O., & Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: Different countries, different trajectories? *Journal of Research on Adolescence*, *20*(2), 482–506. doi:10.1111/j.1532-7795.2010.00644.x

National Center for Education Statistics. (2008). Percentage of high school dropouts among persons 16 through 24 years old (status dropout rate), by income level, and percentage distribution of status dropouts, by labor force status and educational attainment: 1970 through 2007. Retrieved from http://nces.ed.gov/programs/digest/d08/tables/dt08_110.asp

Ndlovu, M., & Mji, A. (2012). Alignment between South African mathematics assessment standards and the TIMSS assessment frameworks. *Pythagoras*, *33*(3), 9 pages. doi:10.4102/pythagoras.v33i3.182

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math not = me. *Journal of Personality and Social Psychology*, *83*, 44–59. doi:10.1037/0022-3514.83.1.44

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., ... Greenwald, A. G. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. In *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 106, pp. 10593–10597).

Novick, M. R., & Lewis, C. (1967). Coefficient alpha and the reliability of composite measurements. *Psychometrika*, *32*, 1–13.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How Large Are Teacher Effects? *Educational Evaluation and Policy Analysis*. doi:10.3102/01623737026003237

O'Dwyer, L. M., Wang, Y., & Shields, K. A. (2015). Teaching for conceptual understanding: A cross-national comparison of the relationship between teachers' instructional practices and student achievement in mathematics. *Large-Scale Assessments in Education*, *3*(1). doi:10.1186/s40536-014-0011-6

Organisation for Economic Co-operation and Development . (2011). Against the odds: Disadvantaged students who succeed in school. *OECD Publishing*. Retrieved from http://dx.doi.org/10.1787/9789264090873-en

Organisation for Economic Co-operation and Development . (2015). PISA 2012 results: The ABC of gender equality in education: aptitude, behaviour, confidence. *PISA, OECD Publishing*. Retrieved March 13, 2015, from http://www.oecd.org/pisa/keyfindings/pisa-2012-results-gender.htm

Op 't Eynde, P., Corte, E. De, & Verschaffel, L. (2002). Framing Students' Mathematics- Related Beliefs: A Quest for Conceptual Clarity and a Comprehensive Categorization. In G. Leder, E. Pehkonen, & G. Törner (Eds.), *In Beliefs: A Hidden Variable in Mathematics Education?* (pp. 13–38). Dordrecht: Kluwer.

Op 't Eynde, P., De Corte, E., & Verschaffel, L. (2006). Epistemic dimensions of students' mathematics-related belief systems. *International Journal of Educational Research*, *45*(1-2), 57–70. doi:10.1016/j.ijer.2006.08.004

Orr, A. (2003). Black-white differences in achievement: The importance of wealth. *Sociology of Education*, *76*, 281–304.

Osborn, M. (2004). New methodologies for comparative research? Establishing "constants" and "contexts" in educational experience. *Oxford Review of Education*, *30*(2), 265–285. doi:10.1080/0305498042000215566

Ottevanger, W., van den Akker, J., & de Feiter, L. (2007). *Developing science, mathematics, and ICT education in Sub-Saharan Africa: Patterns and promising practices* (No. 101). *World Bank working paper:Africa human development series*. Washington, D.C. doi:10.1596/978-0-8213-7070-4

Pajares, F., Britner, S. L. S., & Valiante, G. (2000). Relation between Achievement Goals and Self-Beliefs of Middle School Students in Writing and Science. *Contemporary Educational Psychology*, *25*(4), 406–422.

Park, H. (2008). The varied educational effects of parent-child communication: A comparative study of fourteen Countries. *Comparative Education Review*, *52*(2), 219–243. doi:10.1086/528763

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, *34*(1), 29–48.

Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multicontext study. *Developmental Psychology*, *48*(6), 1629–42. doi:10.1037/a0029167

Paxson, C., & Schady, N. (2007). Cognitive development among young children in Ecuador the roles of wealth, health, and parenting. *The Journal of Human Resources*, *42*(1), 49–84. doi:10.3368/jhr.XLII.1.49

Peterson, R. A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. *The Journal of Applied Psychology*, *98*(1), 194–8. doi:10.1037/a0030767

Pfautz, H. W., Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., … York, R. L. (1967). Equality of Educational Opportunity. *American Sociological Review*. Washington, DC.: U.S. Government Printing Office. doi:10.2307/2091096

Pietsch, J., Walker, R., & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology*, *95*(3), 589–603.

Plante, I., de la Sablonnière, R., Aronson, J. M., & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology*, *38*(3), 225–235. doi:10.1016/j.cedpsych.2013.03.004

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, *88*(5), 879–903. doi:10.1037/0021-9010.88.5.879

Pohl, S., & Steyer, R. (2010). Modeling common traits and method effects in multitrait-multimethod analysis. *Multivariate Behavioral Research*. doi:10.1080/00273170903504729

Quilty, L. C., Oakman, J. M., & Risko, E. (2009). Correlates of the Rosenberg Self-Esteem Scale Method Effects. *Structural Equation Modeling*, *13*(1), 99–117. doi:10.1207/s15328007sem1301_5

Ramsey, L. R., & Sekaquaptewa, D. (2011). Changing stereotypes, changing grades: A longitudinal study of stereotyping during a college math course. *Social Psychology of Education*, *14*, 377–387. doi:10.1007/s11218-010-9150-y

Raykov, T. (1998). Coefficient Alpha and Composite Reliability With Interrelated Nonhomogeneous Items. *Applied Psychological Measurement*. doi:10.1177/014662169802200407

Raykov, T. (2001). Bias of Coefficient α for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, *25*(1), 69–76. doi:10.1177/01466216010251005

Raykov, T. (2012). Scale construction and development using structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 472–492). New York, NY: The Guilford.

Reyes, L. (1984). Affective variables and mathematics education. *The Elementary School Journal*, *84*(5), 558–581.

Richardson, F., & Suinn, R. (1972). The Mathematics Anxiety Rating Scale: Psychometric data. *Journal of Counseling Psychology*, *19*(6), 551–554.

Robinson, M. (2003). Student Enrollment in High School AP Sciences and Calculus: How does it Correlate with STEM Careers? *Bulletin of Science, Technology and Society*, *23*(4), 265–273. doi:10.1177/0270467603256090

Roesken, B., Hannula, M. S., & Pehkonen, E. (2011). Dimensions of students' views of themselves as learners of mathematics. *ZDM-The International Journal on Mathematics Education*, *43*(4), 497–506. doi:10.1007/s11858-011-0315-8

Rosenberg, M., Schoenbach, C., Carmi, S., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem : Different concepts, different outcomes. *American Socioogical Review*, *60*(1), 141–156.

Roszkowski, M. J., & Soven, M. (2010). Shifting gears: consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education*, *35*(1), 113–130. doi:10.1080/02602930802618344

Rutkowski, L., & Rutkowski, D. (2010). Getting it "better": the importance of improving background questionnaires in international large-scale assessment. *Journal of Curriculum Studies*, *42*(3), 411–430.

Ryan, R. M., Stiller, J. D., & Lynch, J. H. (1994). Representations of relationships to teachers, parents, and friends as predictors of academic motivation and self-esteem. *The Journal of Early Adolescence*, *14*(2), 226–249. doi:10.1177/027243169401400207

Saha, L. J. (1994). Aspirations and expectations of students. In T. Husen & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (pp. 354–358). Oxford, UK: Pergamon.

Sahin, A., Morgan, J. R., & Erdogan, N. (2012). Do high school computer and AP courses, and SAT test scores help students choose STEM majors in college? In *American Society for Engineering Education*. San Antonio, United States: ASEE. Retrieved from http://www.asee.org

Sanders, C. E., Field, T. M., & Diego, M. A. (2001). Adolescents' academic expectations and achievement. *Adolescence*, *36*(144), 795–802.

Sandman, R. S. (1980). The Mathematics Attitude Inventory: Instrument and User's Manual. *Journal for Research in Mathematics Education*, *11*(2), 148–49. Retrieved from http://eric.ed.gov/?id=EJ218497

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment*, *29*(4), 347–363.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, *7*(2), 147.

Schmader, T., Johns, M., & Barquissau, M. (2004). The Costs of Accepting Gender Differences : The Role of Stereotype Endorsement in Women ' s Experience in the Math Domain. *Sex Roles*, *50*, 835–850.

Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, *1*(2), 199–223. doi:10.1037/1082-989X.1.2.199

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, *89*, 623–642. doi:10.1037/0022-3514.89.4.623

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*(4), 350–353.

Schmitt, N., & Stults, D. M. (1985). Factors defined by negatively keyed items: the result of careless respondents? *Applied Psychological Measurement*, *9*, 367–373. doi:10.1177/014662168500900405

Schoenfeld, A. H. (1983). Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance. *Cognitive Science*, *7*, 329–363.

Schoenfeld, A. H. (1985). *Mathematical problem-solving*. New York, NY: Academic Press.

Schulz, W. (2005). Measuring the socio-economic background of students and its effect on achievement in PISA 2000 and PISA 2003. In *Annual Meetings of the American Educational Research Association (AERA)*. San Francisco: AERA.

Seaton, M., Parker, P., Marsh, H. W., Craven, R. G., & Yeung, A. S. (2014). The reciprocal relations between self-concept, motivation and achievement: juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology*, *34*(1), 49–72. doi:10.1080/01443410.2013.825232

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*(4), 454–499.

Sells, L. (1973). Developing opportunities for minorities in graduate education. In R. T. Thomas (Ed.), *High school mathematics as the critical filter in the job market* (pp. 37–39). Berkeley: University of California Press.

Semela, T. (2010). Who is joining physics and why? Factors influencing the choice of physics among Ethiopian university students. *International Journal of Environmental & Science Education*, *5*(3), 319–340.

Sewell, W. H., Hauser, R. M., & Wolf, W. C. (2011). Status and occupational sex , schooling ,. *American Journal of Sociology*, *86*(3), 551–583.

Sewell, W. H., & Shah, V. P. (1967). Socioeconomic Status, Intelligence, and Attainment of Higher Education. *Sociology of Education*, *40*(1), 1–23.

Sharp, R. M. (1995). *Scribble Scrabble: Ready-in-a-minute math games*. Blue Ridge Summit, PA: TAB Books.

Sheldon, S. B. (2002). Parents' social network and beliefs as predictors of parent involvement. *The Elementary School Journal*, *102*(4), 301–316.

Sheldon, S. B., & Epstein, J. L. (2005). Involvement counts: Family and community partnerships and mathematics achievement. *The Journal of Educational Research, 98*(4), 196–206.

Shen, C., & Pedulla, J. J. (2000). The Relationship between students' achievement and their self-perception of competence and rigour of mathematics and science: A cross-national analysis. *Assessment in Education: Principles, Policy & Practice*, *7*(2), 237–253.

Shen, C., & Tam, H. P. (2008). The paradoxical relationship between student achievement and self-perception: a cross-national analysis based on three waves of TIMSS data. *Educational Research and Evaluation: An International Journal on Theory and Practice*, *14*(1), 87–100.

Shoemaker, D. M. (1971). *Principles and Procedures of Multiple Matrix Sampling*. (No. ED057100). *ERIC*. Southwest Regional Educational Lab., Inglewood, CA. Retrieved from http://eric.ed.gov/?id=ED057100

Siegler, R. S. (2009). Improving the numerical understanding of children from low-income families. *Child Development Perspective*, *3*(2), 118–124.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika*, *74*(1), 107–120.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and Science Achievement: Effects of Motivation, Interest, and Academic Engagement. *The Journal of Educational Research*, *95*(6), 323–332. doi:10.1080/00220670209596607

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, *75*(3), 417–453. doi:10.3102/00346543075003417

Skaalvik, E. M., & Skaalvik, S. (2008). Self-concept and self-efficacy in mathematics: Relation with mathematics motivation and achievement. In *New developments in the psychology of motivation.* (pp. 105–128).

Skaalvik, E. M., & Valås, H. (1999). Relations among achievement, self-concept, and motivation in mathematics and language Arts: A longitudinal study. *The Journal of Experimental Education*, *67*(2), 135–149. doi:10.1080/00220979909598349

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology*, *35*, 4–28. doi:10.1006/jesp.1998.1373

Stankov, L., & Lee, J. (2014). Quest for the best non-cognitive predictor of academic achievement. *Educational Psychology*, *34*(1), 1–8.

Stankov, L., Lee, J., Luo, W., & Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences*, *22*(6), 747–758.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613–629.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811. doi:10.1037/0022-3514.69.5.797

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*(1), 78–90.

Steinberg, J. R., Okun, M. A., & Aiken, L. S. (2012). Calculus GPA and math identification as moderators of stereotype threat in highly persistent women. *Basic and Applied Social Psychology*, *34*, 534–543.

Stiller, J. D., & Ryan, R. M. (1992). Teachers, parents, and student motivation: The effects of involvement and autonomy support. In *the Annual Meeting of the American Educational Research Association*. San Francisco, CA. Retrieved from http://eric.ed.gov/?id=ED348759

Stipek, D. J., & Gralinski, J. H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology*, *83*(3), 361–371.

Stromquist, N. P. (2007). Gender equity education globally. In *Handbook for achieving gender equity through education* (2nd ed., pp. 33–42). Mahwah, NJ: Erlbaum.

Sui-Chu, E. H., & Willms, D. J. (1996a). Effects of Parental Involvement on Eighth-Grade Achievement. *Sociology of Education*, *69*(2), 126–141.

Sui-Chu, E. H., & Willms, J. D. (1996b). Effects of Parental Involvement on Eighth-Grade Achievement. *Sociology of Education*, *69*(2), 126–141.

Swann Jr., W. B. (1997). The trouble with change: Self-verification and allegiance to the self. *Psychological Science*, *8*(3), 177–180.

Tapia, M., & Marsh, G. E. (2004). An Instrument to Measure Mathematics Attitudes. *Academic Exchange Quarterly*, *8*(2), 16.

Teachman, J. D. (1987). Family background, educational resources, and educational attainment. *American Socioogical Review*, *52*, 548–557.

Teachman, J. D., & Paasch, K. (1998). The family and educational aspiration. *Journal of Marriage and the Family*, *60*(3), 704–714.

Tella, A. (2007). The Impact of motivation on student's academic achievement and learning outcomes in mathematics among secondary school students in Nigeria. *Eurasia Journal of Mathematics, Science & Technology Education*, *3*(2), 149–156.

Ten Berge, J. M., & Zegers, F. E. (1978). A series of lower bounds to the reliability of a test. *Psychometrika*, *43*(4), 575–579.

Tighezza, M., Tighezza, & M'hamed. (2014). Modeling relationships among learning, attitude, self-perception, and science achievement for grade 8 saudi students. *International Journal of Science and Mathematics Education*, *12*(4), 721–740. doi:10.1007/s10763-013-9426-8

Tocci, C. M., & Engelhard Jr, G. (1991). Achievement, parental support and gender differences in attitudes towards mathematics. *The Journal of Educational Research*, *84*(5), 280–286.

Tomas, J. M., & Oliver, A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 84–98. doi:10.1080/10705519909540120

Tomǭs, J. M., Oliver, A., Galiana, L., Sancho, P., & Lila, M. (2013). Explaining method effects associated with negatively worded items in trait and state global and domain-specific self-esteem scales. *Structural Equation Modeling*, *20*(2), 299–313. doi:10.1080/10705511.2013.769394

Topor, D. R., Keane, S. P., Shelton, T. L., & Calkins, S. D. (2010). Parent involvement and student academic performance: A multiple mediational analysis. *Journal of Prevention and Intervention in the Community*, *38*(3), 183–197. doi:10.1080/10852352.2010.486297

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763–777. doi:10.1037/a0027470

Trudell, B. (2011). The making of a killer (language): Language contact and language dominance in sub-Saharan Africa. In *European Conference on African Studies*. Uppsala, Sweden. Retrieved from http://www.aegis-eu.org/archive/ecas4/ecas-4/panels/141-156/panel-149/Barbara-Trudell-full-paper.pdf

Tuohilampi, L., Hannula, M. S., Varas, L., Giaconi, V., Laine, A., Näveri, L., & i Nevado, L. S. (2014). Challenging the 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

United Nations Educational, Scientific and Cultural Organization. (2013). *The Education for All Global Monitoring Report-Fact Sheet. EFA Global Monitoring Report*. UNESCO.

UNESCO International Bureau of Education. (2011). World Data on Education: Seventh edition. Retrieved July 28, 2015, from http://www.ibe.unesco.org/en/services/online-materials/world-data-on-education/seventh-edition-2010-11.html

United Nations Educational, Scientific and Cultural Organization. (2010). *Global monitoring report 2010:Reaching the marginalized. Education for All*. Oxford, U.K: Oxford University Press and UNESCO.

Urbán, R., Szigeti, R., Kökönyei, G., & Demetrovics, Z. (2014). Global self-esteem and method effects : Competing factor structures, longitudinal invariance, and response styles in adolescents. *Behavior Research Methods*, *46*(2), 488–498. doi:10.3758/s13428-013-0391-5

Ursini, S., & Sánchez, G. (2008). Gender, technology and attitude towards mathematics: a comparative longitudinal study with Mexican students. *ZDM*, *40*(4), 559–577. doi:10.1007/s11858-008-0120-1

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement : A meta-analytic review . *Educational Psychologist*, *39*(2), 111–133. doi:10.1207/s15326985ep3902

van de Vijver, F. J. R., & Leung, K. L. (2000). Methodological issues in psychological research on culture. *Journal of Cross-Cultural Psychology*, *31*(1), 33–51. doi:10.1177/0022022100031001004

Van Fleet, J. W. (2012). Africa's education crisis: In school but not learning. Center for Universal Education at Brookings Africa Learning Barometer. Retrieved December 29, 2015, from http://www.brookings.edu

Van Fleet, J. W., & Watkins, K. (2012). Africa Learning Barometer. Center for Universal Education at Brookings Africa Learning Barometer. Retrieved November 3, 2015, from http://www.brookings.edu/research/interactives/africa-learning-barometer

Vandenberg, R. J., & Lance, C. E. (2000). A Review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4–70. doi:10.1177/109442810031002

Vautier, S., & Pohl, S. (2009). Do balanced scales assess bipolar constructs? The case of the STAI scales. *Psychological Assessment*, *21*(2), 187–193.

von Davier, M., Gonzalez, E., & Mislevy, R. J. (2009). Plausible values: what are they and why do we need them?. *IERI Monograph Series*, *2*, 9–36.

Vukovic, R. K., Roberts, S. O., & Green Wright, L. (2013). From parental involvement to children's mathematical performance: The role of mathematics anxiety. *Early Education & Development*, *24*(4), 446–467.

Wang, J. (2007). A trend study of self-concept and mathematics achievement in a cross-cultural context. *Mathematics Education Research Journal*, *19*(3), 33–47. doi:10.1007/BF03217461

Wang, J., & Lin, E. (2008). An Alternative Interpretation of the Relationship between Self-Concept and Mathematics Achievement: Comparison of Chinese and US Students as a Context. *Evaluation & Research in Education*, *21*(3), 154–174. doi:10.1080/09500790802485203

Wang, L., Li, X., & Li, N. (2014). Socio-economic status and mathematics achievement in China: A review. *ZDM*, *46*, 1051–1061. doi:10.1007/s11858-014-0617-8

Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology*. doi:10.1037/a0027247

Wang, M.-T., & Eccles, J. S. (2012). Adolescent behavioral, emotional, and cognitive engagement trajectories in school and their differential relations to educational success. *Journal of Research on Adolescence*, *22*(1), 31–39.

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th- through 11th-grade Australian students. *Child Development*, *75*(5), 1556–1574.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York, NY: The Guilford Press.

Whitebread, D., & Chiu, M.-S. (2004). Patterns of Children'S Emotional Responses To Mathematical Problem-Solving. *Research in Mathematics Education*, *6*(1), 129–153. doi:10.1080/14794800008520134

Wiest, L. R. (2010). Out-of-school-time (OST) programs as mathematics support for females. In H. J. Forgasz, J. R. Becker, K.-H. Lee, & O. B. Steinthorsdottir (Eds.), *International perspectives on gender and mathematics education* (pp. 55–86). Charlotte, NC: Information Age Publishing.

Wigfield, A. (1994). Expectancy-Value Theory of Achievement Motivation: A Develpmental Perspective. *Economics of Education Reviewducational Psychology*, *6*(1), 49–78. doi:10.1006/ceps.1999.1015

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81.

Wilkins, J. L. M. (2004). Mathematics and Science Self-Concept: An International Investigation. *The Journal of Experimental Education*, *72*(4), 331–346. doi:10.3200/JEXE.72.4.331-346

Wilkins, J. L. M., Zembylas, M., & Travers, K. J. (2002). Investigating correlates of mathematics and science literacy in the final year of secondary school. In D. F. Robitaille & A. E. Beaton (Eds.), *Secondary Analysis of the TIMSS Data* (pp. 291–316). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Williams, T., & Williams, K. (2010). Self-efficacy and performance in mathematics: Reciprocal determinism in 33 nations. *Journal of Educational Psychology*, *102*(2), 453–466. doi:10.1037/a0017271

Wolf, M. K., Chang, S. M., Jung, H., Farnsworth, T., Bachman, P. L., Nollner, J., & Shin, W. H. (2008). *An investigation of the language demands in states' content-area and English language proficiency tests* (CRESST Report No. 738). (M. K. Wolf, J. L. Herman, J. Kim, J. Abedi, S. Leon, N. Griffin, ... CRESST/ University of California, Eds.) *Providing validity evidence to improve the assessment of English language learners*(pp. 9-

53). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, *28*(3), 189–194. doi:10.1007/s10862-005-9004-7

Wright, S. (2004). *Language Policy and Language Planning: From Nationalism to Globalisation*. New York: Palgrave.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*, 57–67.

Yanyun Yang, & Green, S. B. (2011). Coefficient Alpha: A reliability coefficient for the 21st Century? *Journal of Psychoeducational Assessment*, *29*, 377–392.

Ye, F., & Wallace, T. L. (2013). Psychological sense of school membership scale: Method effects associated with negatively worded items. *Journal of Psychoeducational Assessment*, *32*(3), 202–215.

Yeung, W. J., & Conley, D. (2008). Black-white achievement gap and family wealth. *Child Development*, *79*(2), 303–324. doi:10.1111/j.1467-8624.2007.01127.x

Yeung, W. J., Linver, M. R., & Brooks-Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child Development*, *73*(6), 1861–1879. doi:10.2307/3696422

Zhang, M. (2003). Links between school absenteeism and child poverty. *Pastoral Care in Education*, *21*(1), 10–17.

Zhu, Y., & Leung, F. K. S. (2011). Motivation and achievement: Is there an East Asian model? *International Journal of Science and Mathematics Education*, *9*(5), 1189–1212.

Zimmerman, D. W., Zumbo, B. D., & Lalonde, C. (1993). Coefficient alpha as an estimate of test reliability under violation of two assumptions. *Educational and Psychological Measurement*, *53*, 33–49.