UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Examining the Effects of Non-Cognitive Factors on Mathematics Achievement Across National Groups: USA, Germany, Japan, and Korea

Permalink

https://escholarship.org/uc/item/8rf4j6rf

Author

Walker, Soung Hwa

Publication Date

2017

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Examining the Effects of Non-Cognitive Factors on Mathematics Achievement Across National Groups: USA, Germany, Japan, and Korea

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Education

by

Soung Hwa Walker

June 2017

Dissertation Committee:
Dr. Keith F. Widaman, Chairperson
Dr. Thomas M. Smith
Dr. Robert Rosenthal
Dr. Soojin Park

Copyright by
Soung Hwa Walker
2017

The Dissertation of Soung Hwa Walker is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

My deepest gratitude goes to my advisor and mentor, Dr. Keith F. Widaman. Without his keen wisdom and kind guidance in each step, my doctoral dissertation would not have been possible. Dr. Widaman's unwavering support has been my core motivation to stay strong in order to finish my life-long greatest academic challenge – completing dissertation. He is a rare scholar who values students' own research interests, and his doors are always open to all, even to non-education-major students. He listens to students' struggles patiently, explains data related issues very thoroughly, and guides us and pushes us to the next level; the best advisor that any grad student can hope for.

I can honestly say that it has been a tremendous honor to learn Structural Equation Modeling (SEM) from Dr. Widaman because of his enviable ability to analyze data and explain SEM issues in a simple manner. I thank you so much, Dr. Widaman, for allowing me to use your own-rich-library of books – your generosity can never be expressed enough in words. I have tried to pay your generosity forward to my students in my own way…

My other special thanks go to many other remarkable professors – (1) Dr. Thomas Smith (helped me and believed in my study, and sharped my thinking to grow academically), (2) Dr. Robert Rosenthal, and (3) Dr. Soojin Park who made invaluable contributions to improve the quality of this project. Dr. Rosenthal – I admire your brilliant mind. You have believed in me and expected the best out of me and always treated me like a young scholar; so, I have been thinking like one, acting like one, and perhaps, will become one. You are the living evidence of what positive expectation is all about – so, thank you…

My utmost honor, respect, and thanks goes to my parents who sacrificed themselves so much for my brother, my sisters, and me; their unconditional love became a true inspiration in my life, and will never be forgotten. Their deep-rooted belief/trust in me has always been my true inner-fuel to do well beyond the academic realm. Yes, father, I became a teacher, too – I am sure you are proud of what I have become. And yes, mother, anything's possible - not perfect, but you read now! I thank you so much that you taught me the joy of giving...

Furthermore, I thank the Lord, Jesus Christ, for my precious family (Deron, John, & Christy). Their endless love, support, and, of course, lots of prayers have helped me to endure and cherish many ups and downs during my graduate school years. Thank you for making me laugh when I needed to the most. I know I am far from perfect – thank you for loving me no matter what. For the past 5 years, Deron, a strong believer in me as doctoral material - you have been a dad and a mom for our kids; I thank God for our 20+ years together! Hey Christy, my precious girl, you are beautiful inside and out, and smart, I thank God for you every day; I appreciate your maturity and understanding. Hey John, my young man who is kind, intelligent, God-fearing, and sweet-hearted, I thank God for you daily; I can never tell you enough how much you both mean to me! My own blood and flesh, you both are truly miracles from GOD – incredibly awesome; and I know for sure you will surpass me and I will be delighted!

And I also thank God for my many friends from schools, bowling leagues, and churches (too many to name them all) for their countless prayers and, definitely, many laughs, cries, and talks, which include my frequent monologues about stats. I am super

grateful for Yeonhee and Hyunju -- our priceless friendships (countless international calls & text messages in the middle of the many-nights which brought me tears and sometimes made me laugh insanely), and for our cell-group members, Korean Riverside Baptist Church – lots of prayers, loves, and delicious foods. Thank you for praying for me to be strong, so that I could focus on completing this enormous-worthy project; thank you for listening to me, being truly happy for my academic success, and empathizing with my struggles without being judgmental so that I could grow in Christ.

Thank you, Jesus, for saving my soul, forgiving my sins, and for bringing me up this far – YOUR NAME IS GLORIFIED, PRAISED, AND WORSHIPED FOREVER!

DEDICATIONS

FOR MY LORD, JESUS CHRIST

MY MASTER, MY STRENGTH

MY PARENTS, MY FAMILY, and MR. FRANKLIN & MRS. JANIE HARKINS

Light of the World
You stepped down into darkness
Open my eyes
Let me see
…
Here I am to worship
Here I am to bow down
Here I am to say that You're my God
…
By Tim Hughes

****

I have a home
Eternal home
But for now I walk this broken world
You walked it first
You know *my* pain
But You show hope can rise again up from the grave
Abide with me
Abide with me
Don't let me fall
Don't let go
…
By Matt Redman

Have I not commanded you? Be strong and courageous. Do not be afraid; do not be discouraged, for the Lord your God will be with you wherever you go (Joshua 1:9)

ABSTRACT OF THE DISSERTATION

Examining the Effects of Non-Cognitive Factors on Mathematics Achievement Across National Groups: USA, Germany, Japan, and Korea

by

Soung Hwa Walker

Doctor of Philosophy, Graduate Program in Education
University of California, Riverside, June 2017
Dr. Keith F. Widaman, Chairperson

Since there is limited research on the applicability of the Theory of Planned Behavior (TPB) model in educational contexts, the current cross-national comparison study aimed to investigate how non-cognitive factors, specifically, the TPB model components, affect students' math outcomes in USA and their peers in three other countries, Germany, Japan, and Korea. Using internationally representative samples from the 2012 Programme for International Student Assessment (PISA) data, various structural equation models (SEM) were employed to examine the TPB applicability in the academic domain across four national groups.

One key strength of the current investigation was the use of two random samples from each country (the derivation and cross-validation samples). The results of the factor structural analyses from both the derivation (N = 10681) and cross-validation (N =

10682) samples indicated that the measurement of the TPB model components were invariant across four countries. From both 1st/2nd samples, two findings were consistent with "West vs. Asia" contrast: (a) the influence of Attitude, among the TPB model components, on Math Achievement was much stronger in Japan and Korea than USA and Germany; (b) the estimated effect size of Math Achievement was much greater in Japan and Korea than USA and Germany.

However, most other findings were contrasts showing that one Western country or one Asian country differed from the other three countries, inconsistent with a generic "West vs. Asia" cultural comparison model: (a) the influence of Control on Math Intention in USA was distinctively greater than for the other three countries; (b) the direct effect of Control on Math Achievement was strong in USA, Germany, and Korea, but weak in Japan; (c) lastly, the influence of Attitude on Math Intention and Math Intention on Math Achievement in Korea were much stronger than in the other three countries. In addition, among three Subjective Norm factors (Friend, Parent, and Teacher), the effect of SN-Teacher on Math Intention was much stronger than the other two SN factors in all four countries. The present cross-national comparison study concluded with implications for math education and future research directions.

TABLE OF CONTENTS

ABSTRACT ……..…… viii
LIST OF TABLES …….… xii
LIST OF FIGURES xiv

CHAPTER ONE: **INTRODUCTION**

Why Mathematics? …….……..... 1
Research on Mathematics education …….. 2
Statement of the Problem …… 4
U.S. students' math performance in global perspective ……..... 4
Asian students' math performance in global perspective ……..... 6
What's new? ……... 7
Non-cognitive attributes and math achievement ……..... 7
Significance of the study ...……. 7

CHAPTER TWO: **REVIEW OF LITERATURE**

Overview …… 9
Attitude-behavior relations ....……. 10
Theory of Reasoned Action (TRA) ……..…….. 10
Theory of Planned Behavior (TPB) ……...……. 14
TPB model components in educational contexts …….…….. 17
Attitude factor …….…….……. 17
Attitude factor: Global perspectives …….. 20
Subjective norm factor …….…….. 21
Subjective norm factor: Global perspectives ……. 24
Perceived Behavioral Control (PBC) factor ……... 25
PBC factor: Global perspectives …… 27
Research Questions ……. 29

CHAPTER THREE: **METHOD**

Data Source: 2012 PISA …… 31
Characteristics of PISA ..…… 31
Sampling procedures of PISA …… 32
Plausible values in PISA …… 32
Subgroup dataset for the present research ……. 34
Measures …… 34
Data Analyses …… 37
Why Structural Equation Modeling (SEM)? ……. 37
Measurement and structural model …… 39
Multiple-Group analyses ……... 41
Measurement invariance testing …… 42

CHAPTER FOUR: **RESULTS**

Preliminary analyses …….. 46
TPB measurement models ……..…… 53
Confirmatory factor analyses ……...……. 53
Testing measurement invariance of the TPB model …….……. 57
TPB Structural models …… ……...…… 66
Structural relations of the TPB model comparisons ……..…… 66
Multiple-group analyses of the TPB model components ……... 75
Research Questions …… 81

CHAPTER FIVE: **DISCUSSIONS**

Overview …… 84
Cross-national comparisons of non-cognitive factors ……... 85
Limitations ……. 89
What's next? 90

REFERENCES ……. 92

LIST OF TABLES

Tables		Page
1.	Descriptions of Constructs of the Proposed Conceptual Model ……	36
2.	Types of Factorial Measurement Invariance Testing ….……....	44
3.	Means and Standard Deviations for Scale Scores of All Countries …...……	46
4.	Correlations among Scale Scores for Latent Factors of All Countries ……..	47
5.	Means and Standard Deviations of Scale Scores by Each Country ……...	48
6.	Descriptive Statistics of 23 Observed Variables for the Entire Sample …….	49
7.	Means and Standard Deviations for 23 Observed Variables by Country ……...	50
8.	Standardized Factor Ladings of 23 Indicators for the Entire Sample ……	53
9.	Standardized Factor Loadings for 14 Indicators of the Entire Sample ……..	55
10.	Estimated Correlations among 14 Indicators for the Entire Sample ……..	56
11.	Summary of Measurement Invariance Testing for the Derivation Sample ……	58
12.	Summary of an Overall Sequential TPB Measurement Model or both the Derivation and Cross-Validation Samples …….	60
13.	Estimated Intercepts and Unstandardized Factor Loadings for the Derivation Sample ……	62
14.	Means and Variances on Latent Variables for the Most Optimal Model from the Partial Strong Invariance (Model 4) using the first half sample by Each Country …	64
15.	Correlations among Latent Factors from Final Partial Strong Invariance Model for the First Half (Derivation) Sample by Each Country ……	65
16.	Structural Model Comparisons from the First Half Sample (Derivation) …….	67
17.	Sequential Analyses of Each Regression Path of the TPB Structural Models …..…	72
18.	Overall Sequential Structural Model Comparisons for both the Derivation and Cross-Validation Samples …….	73

19. Regression Analysis for the TPB Structural Relations and the Estimated Effect Size for Both the Derivation and the Cross-Validation Samples ……... 75

20. Correlations among Predictor Variables from both the First Half (derivation) and the Second Half (Cross-validation) Samples ……. 79

LIST OF FIGURES

Figure		Page
1	2012 PISA math scores for selective countries …….	4
2	The 5th TIMSS math scores of 4th and 8th grade for selective countries …….	5
3	TRA model by Fishbein and Ajzen (1975) ……	10
4	Alternative model 1 of TRA (Bentler & Spekart, 1979) ……	12
5	Alternative model 2 of TRA (Bentler & Spekart, 1979) ……	12
6	Theory of Planned Behavior (Ajzen, 1991) ……...	14
7	Students' self-reports on math ability – 1st example: Western vs. Asian countries ……...	27
8	Students' self-reports on math ability – 2nd example: Western vs. Asian countries ……...	28
9	Conceptual model of the present research – an extended TPB …….….	29
10	Structural model comparisons from all free (Model A) to all constrained (Model E)	67
11	Estimated regression coefficients - USA, Germany, Japan, and Korea …….	78

CHAPTER ONE

INTRODUCTION

Why Mathematics?

A strong foundation in mathematics is an essential aspect of students' cognitive functioning and critical thinking skills, and proficiency in mathematics can have broad effects on overall school success by students. Students' early math proficiency frequently influences their later choices to pursue education in fields of science, technology, engineering, and mathematics (STEM). STEM-based education has become a vital tool for 21st century societies because mathematics represents a core path for research in science, technology, and engineering, as well as medical industries.

Furthermore, students' strong math skills often play critical roles in advancing their future employment opportunities, especially by providing more career options, promotions, and wage increases, which also impact a nation's economy (Geary, 1996; Organization for Economic Cooperation and Development [OECD], 2014a). More specifically, "across OECD countries, more than three out of four students agree or strongly agree that learning mathematics will improve their career prospects" (OECD, 2013, p. 56). Having recognized the increasingly important role of math on both individual and national levels, educational efficiency in mathematics instruction has recently drawn more attention from policy makers, researchers, and practitioners globally.

Research on Mathematics Education

U.S. students' unsatisfactory achievement outcomes in mathematics (and science) is not merely a recent concern. In 1983, A Nation at Risk was published by the National Commission on Excellence in Education (NCEE) under President Reagan's administration, and this report detailed the lack of rigorous curricula in mathematics (and science) among other issues (for more details, see Smith, 2004). Since then, the number of mathematics (and science) course requirements for high school graduation has increased in many states as one part of educational reform efforts (NCEE, 1983; Smith, 2004).

In addition to policy and curriculum efforts, countless studies have sought to identify factors that are important components if students are to enhance their mathematics competence. One such factor that has been identified is working memory (WM), which refers to a cognitive system that holds a limited amount of information for a short period of time (Alloway, 2011). Baddeley (1992) defines working memory as, "…the brain system that provides temporary storage and manipulation of the information necessary for such complex cognitive tasks as language, comprehension, learning, and reasoning" (p. 556). Studies have found that students with strong WM capacity tend to show greater academic achievement. Thus, WM research has emphasized strategies to manipulate or monitor information processing in learners' brains (i.e., executive function skills) or techniques to help boost memory recall or retention (e.g., Cragg & Gilmore, 2013).

Nevertheless, techniques to expand WM alone may not hold long-term benefits for competence in mathematics because temporarily accumulated knowledge in WM does not last very long, especially when students do not internalize the value of learning math. Stated differently, learning mathematics itself is a cognitive process; thus, cognitive intervention programs and brain-based strategies are useful to help improve math achievement, in which explicit instructional methods are quite valuable. At the same time, it is important to note that math outcomes might be limited if students strongly dislike anything to do with math and do not believe that they are actually capable of learning math (Ma & Kishor, 1997; McLeod, 1992; Neale, 1969; Zan & Di Martino, 2007).

Moreover, research findings suggest that affective factors such as attitude (e.g., like/dislike, emotion) and perceived belief regarding cognitive ability play critical roles in students' intellectual engagement and cognitive performance (Ma & Kishor, 1997; Neale, 1969; Zan & Di Martino, 2007). Students' social norm influences also impact how much time and effort they are willing to invest in what they are supposed to be studying. Considering social and psychological influences on math learning outcomes, research on the relations between attitude toward math and achievement in math has been increasingly emphasized (e.g., Chen & Stevensen, 1995; Lipnevich, MacCann, Krumm, Burrus, & Roberts, 2011; Ma & Kishor, 1997; Minato & Yanase, 1984).

In short, no single-particular-factor holds the ultimate answer for improving math learning outcomes exclusively or could explain the complexity of academic success or failure in mathematics for all students. Instead, many factors may have a role in

promoting interest and skill in mathematics. That is, learning math, as with other academic subjects, is a complicated cognitive process involving psychological, cultural, environmental, and social influences (Hallinan, 2008). Only if we understand the role that each of these domains play in affecting student interest in mathematics and dedication to learning mathematics will we have a more adequate conception of how we can intervene to improve mathematics proficiency in students.

Statement of the Problem

U.S. students' math performance in global perspective

A significant amount of research has been conducted to identify factors that could help to improve students' math skills, and numerous evidence-based intervention programs have been developed and implemented in schools to enhance math education. Despite promising research on math education, U.S. students' math performance has been falling behind other countries, especially Asian countries (OECD, 2014a). A recent report has indicated that American youths' math performance is below the OECD average, which places the United States 27th among the 34 OECD countries that participated in the 2012 Programme for International Student Assessment (PISA). According to the 2012 PISA data, the U.S. average math score is 481 compared to the OECD average of 494. The highest performing Asian countries' scores are 613 for China, 573 for Singapore, 560 for Taiwan, 554 for Korea, and 536 for Japan, respectively (PISA; OECD, 2014b).

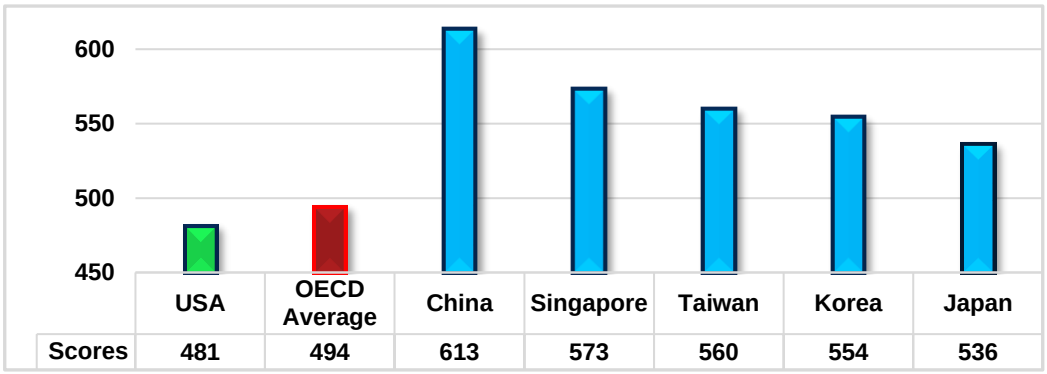
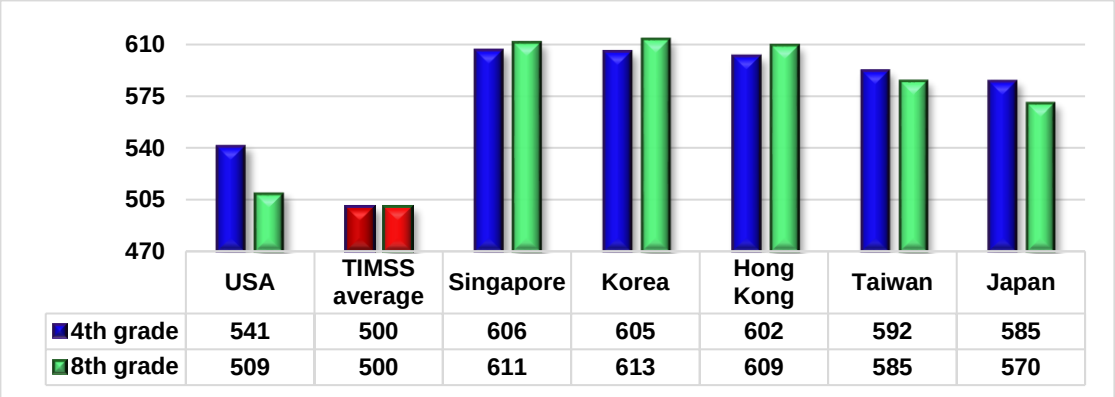


Figure 1. *2012 PISA math scores for selective countries*

In addition, the Trends in International Mathematics and Science Study (TIMSS) also indicates that students from many nations and economies around the world perform better than U.S. students in math and science. TIMSS represents another international student assessment that has been used for global comparisons of U.S. students' math and science performance and progress relative to students in other countries. TIMSS began in 1995, and the TIMSS assessments are repeated every 4 years (National Center for Education Statistics, 2011).

In a recent evaluation, the 5th TIMSS data reports show that U.S. 4th and 8th grade average mathematics scores are 541 and 509, respectively, which are higher than the TIMSS scaled average scores of 500 for both 4th grade (among 52 participating countries) and 8th grade (among 45 participating countries). Thus, U.S. students' average math performances are higher than the TIMSS average in mathematics, but still fall behind performance by peers from many Asian countries. According to the TIMSS report, the top five highest-performing countries for 4th grades are Singapore, South Korea, Hong Kong, Taiwan, and Japan, and for 8th grades are South Korea, Singapore, Taiwan, Hong

Kong, and Japan, respectively. Detailed scores are listed below for convenience (TIMSS; National Center for Education Statistics, 2011).



	USA	TIMSS average	Singapore	Korea	Hong Kong	Taiwan	Japan
4th grade	541	500	606	605	602	592	585
8th grade	509	500	611	613	609	585	570

Figure 2. *The 5th TIMSS math scores of 4th and 8th grade for selective countries*

Asian students' math performance in global perspective

International education assessments from PISA and TIMSS have reported that Asian students consistently outperform American students and students from other countries around the world. Studies also show that Asian students frequently do better in American schools when compared to non-Asian students (e.g., Chen & Stevenson, 1995; Peng & Wright, 1994; Sun, 1998). A significant amount of research has attempted to identify contributing factors to explain Asian students' high academic achievement. Among such research investigations, the influences of culture and parental aspirations on education that are deeply rooted from Confucian heritage have been frequently examined.

For example, Mau (1997) reported that parental educational involvement, student's perceived parental expectation, and educational effort are strongly related to Asian students' high academic achievement. In order to better understand how parental-influences contribute to shaping Asian students' study habits and academic goals, Kim,

Schallert, and Kim (2010) conducted a study and found that Korean students' different goal orientations are predicted by their perceptions of parental goals and motivation styles.

What's new?

Non-cognitive attributes and math achievement

The purpose of the present research was to comprehensively examine multiple affective attributes (i.e., non-cognitive factors) that are associated with students' math outcomes in the U.S. and their peers in other countries (Germany, Japan, and Korea) using internationally representative large-scale datasets from the 2012 PISA. Specifically, the present study seeks to understand how math-related intentional behavior is predicted by non-cognitive attributes across different national groups. To accomplish the current study goals, Ajzen's (1991) theory of planned behavior (TPB) was utilized. According to the TPB model, Intention is predicted by Attitude, Subjective Norm, and Perceived Control; Perceived Control and Intention, together, predict behavioral outcomes (Details will be discussed in Chapter 2).

Significance of the Study

Previous studies that used the TPB model to understand people's volitional behaviors have typically treated the Subjective Norm (i.e., Friend and Parent) variable as a single factor. The current investigation, however, aims to examine the Subjective Norm variable on a deeper level by treating 'Friend' and 'Parent' as separate latent factors, which is new to the existing literature in TPB research. Additionally, taking academic contexts into consideration, the teacher variable will also be added to the original TPB

model as a part of the Subjective Norm variables. Examining these Subjective Norm variables separately will allow more precise determination of any cultural differences between Western and Asian countries in terms of identifying the predictability of each Subjective Norm factor (i.e., Friend, Parent, and Teacher) on intentional math behaviors. Furthermore, the present study aims to investigate whether each latent construct of the TPB model is being measured consistently across national groups (i.e., USA, Germany, Japan, and Korea). The results of measurement invariance testing from the present cross-national comparison study may add new insights by broadening the TPB model in terms of theoretical and practical applicability in education on a global scale.

CHAPTER TWO

REVIEW OF LITERATURE

Overview

Cognitive abilities (e.g., IQ, accumulated content knowledge, etc.) constitute a significant but limited picture of academic behaviors and educational outcomes (e.g., mathematics achievement). Having similar prior academic records, some students tend to do well in math while others do not, even if they have the same learning environments where the same cognitive intervention programs are implemented. Differences of this sort in achievement outcomes suggest that non-cognitive characteristics may affect the level of individuals' motivation and/or engagement. Among various affective or non-cognitive attributes/factors that have been investigated in educational contexts, some of the following are most frequently cited; attitude (Lipnevich et al., 2011; Ma & Kishor, 1997; McLeod, 1992; Neale, 1969; Zan & Di Martino, 2007), growth mindset (Dweck, 2002, 2006), grit or perseverance (Duckworth, Peterson, Matthews, & Kelly, 2007), conscientiousness (Noftle & Robins, 2007; Poropat, 2009), self-efficacy (Bandura, 1977, 1986; Pajares, 1996; Schunk, 1985, 2012; Zimmerman, 2000), and social norm influences (e.g., Eccles & Jacobs, 1986).

Utilizing Ajzen's (1991) Theory of Planned Behavior (TPB) model, the current research aims to examine how multiple non-cognitive attributes are related to academic achievement in math. In what follows, the theoretical framework of the present research's

conceptual model (an extended version of TPB model) and its background will be discussed: (1) an overview of the TPB model development and application to various research contexts, and (2) the prior research on each non-cognitive factor of TPB model components (i.e., attitude, social norm influence, perceived belief/control) that are associated with academic outcomes.

Attitude-Behavior Relations

When individuals hold favorable attitudes toward certain tasks/behaviors, they are more likely to engage in those targeted behaviors, and tend to try harder to execute those tasks. In other words, people's intention to participate in the specific behaviors is related to the degree of positive attitude they hold toward the task. In understanding attitude-behavior relations, Fishbein (1967) eloquently pointed out that researchers must measure the construct of attitude that was related to outcome behaviors to accurately examine attitude-behavior relations. For example, measuring attitude toward math and achievement in math would be reasonable to examine whether math attitude is associated with math learning outcomes. Conversely, if one wants to investigate the relationship between attitude toward reading and the level of reading fluency between girls and boys, then measuring more general attitudes toward schooling may not allow one to draw a reasonable conclusion as to whether or not attitude predicts girls' or boys' reading fluency.

Theory of Reasoned Action (TRA)

In an attempt to explain attitude-behavior relations among human actions, Fishbein and Ajzen (1975) proposed a theoretical model called Theory of Reasoned

Action (TRA). According to the TRA model (see *Figure* 3), the immediate antecedent of a behavioral outcome is people's intention to perform the targeted behavior, and intention is predicted by, both, (1) the strength of the attitude that s/he holds toward the behavior and (2) subjective norm connected with the behavior. In this TRA model, no direct associations from attitude to behavior and from subjective norm to behavior are posited. Instead, both attitude and subjective norm factors indirectly predict behavioral outcomes through intention, which serves as a mediator (Fishbein & Ajzen, 1975).

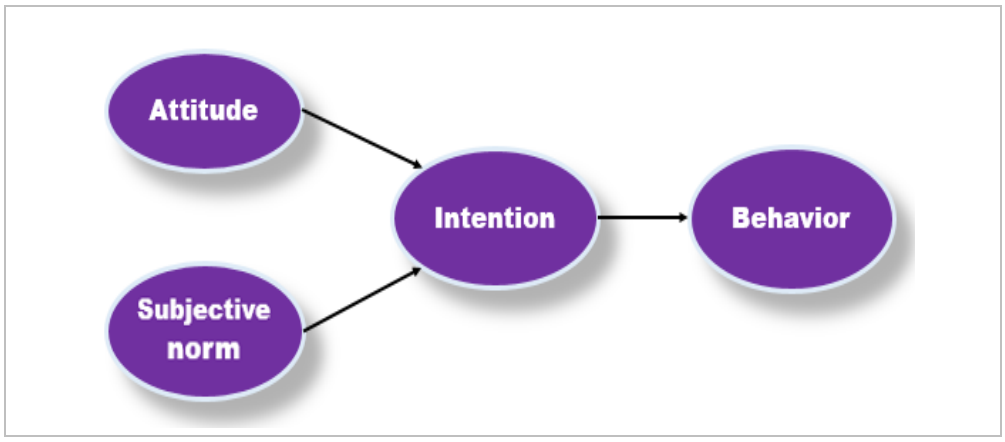


Figure 3. TRA model by Fishbein and Ajzen (1975)

TRA seeks to understand and predict people's voluntary behaviors, assuming that the behaviors being studied are under full volitional control (Madden, Ellen, & Ajzen, 1992). Fishbein and Ajzen (1975) acknowledged that, if people's behavioral outcomes/actions are, at least in part, determined by factors beyond individuals' voluntary control, the TRA model should not be applied. That is, whenever the performance of some action requires knowledge, skills, resources, or others' cooperation, or necessitates overcoming environmental obstacles, the conditions of the model cannot be met (Ajzen, 1991; Fishbein & Ajzen, 1975; Sheppard, Hartwick, &

Warshaw, 1988). Such consumers' actions under the TRA model, which do not require individuals' special skills or knowledge, could include drinking at parties, taking diet pills, using grocery coupons, voting choices, etc.

Strong empirical evidence has accumulated to support the TRA conceptual model in consumers' behavioral intentions and their actions (i.e., actual behavioral outcomes). For example, Sheppard et al. (1988) reviewed 87 independent studies that used Fishbein and Ajzen's model of TRA. Statistically significant frequency-weighted correlations were reported; intention → behavior relation, $r = .53$, $p = .01$, and attitude + subjective norm → intention relation; $r = .66$, $p = .001$ (Sheppard et al., 2008, p. 336). Another interesting finding from their meta-analytical review is that less than 20% of 87 independent studies used the original Fishbein and Ajzen's model; that is, a considerable number of studies measured goals instead of intention, and alternative activity outcomes instead of a single behavioral outcome. The TRA model was initially proposed and developed to investigate a single behavioral outcome that involves explicitly measuring intention related to the targeted behavior (Sheppard et al., 1988), but many modifications to the original TRA model have clearly been made across studies.

Although a significant amount of research shows the applicability of the TRA conceptual model in psychology and social science, indicating that attitude predicts behavior only through intention as a mediator, some research findings have reported otherwise. For example, Bentler and Speckart (1979) proposed two alternative versions of the Fishbein and Ajzen TRA model. The first alternative model (*Figure* 4)

hypothesized that attitude directly predicts both intention and the targeted action (behavioral outcome).

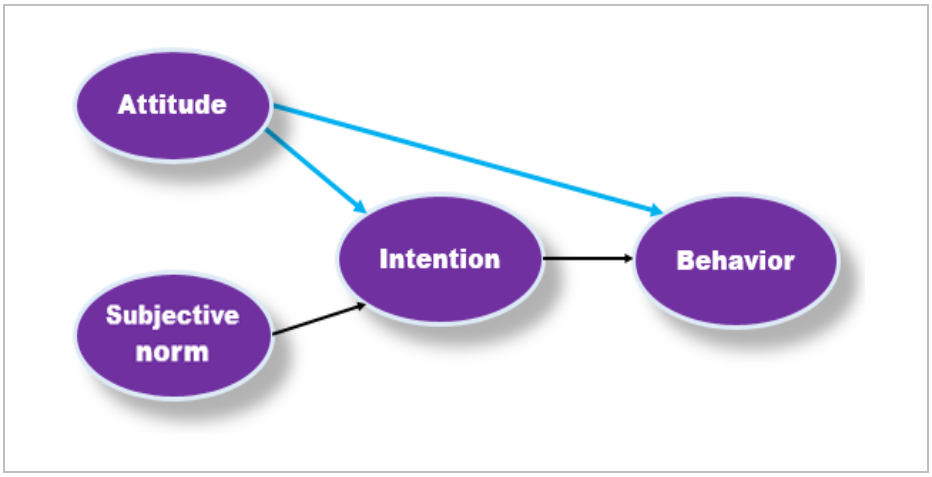


Figure 4. *Alternative model 1 of TRA (Bentler & Spekart, 1979)*

The second hypothesized alternative model of TRA by Bentler and Speckart (1979) posits that prior experienced behavior, if related, may influence that targeted behavioral outcome either directly or indirectly through intention (*Figure* 5).

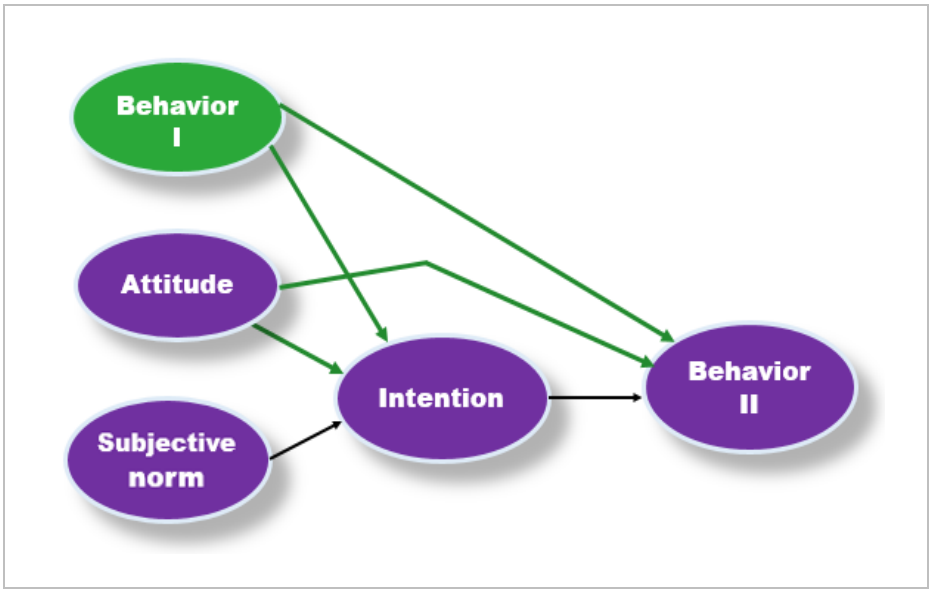


Figure 5. *Alternative model 2 of TRA (Bentler & Spekart, 1979)*

In this model, the targeted behavior (Behavior II) is predicted by multiple factors, i.e., attitude, past behavior (behavior I), and intention. In their 2nd proposed model, Behavior I (along with attitude, subjective norm, and intention) is measured at the beginning of the study – two weeks prior to the Behavior II measurement (Bentler & Speckart, 1979, p. 458). Structural equation modeling is a more optimal way than several multiple regression analyses when evaluating TRA model applicability (Bentler & Speckart, 1979, p. 455).

Using the two different extended models of the original Fishbein and Ajzen's TRA, Bentler and Speckart (1979) found that previous behavioral experiences (Behavior I) and attitude have both direct and indirect influences on the actual action (Behavior II) – the targeted behavioral outcome (Bentler & Speckart, 1979, p. 461). More importantly, prior behavior and attitude account for a significant degree of variability in drug consumption behavior of young adults (i.e., 228 college students) that is not accounted for by the intention factor (p. 458).

Regression weights	Alcohol	Marijuana	Hard drugs
Behavior I → Behavior II	.36	.32	.27
Attitude → Behavior II	.25	.38	.33

Theory of Planned Behavior (TPB)

The Theory of Planned Behavior (TPB) has been "made necessary" to overcome the limitation of the Theory of Reasoned Action (TRA) by adding the construct of perceived behavioral control (Ajzen, 1991, p. 181; Ajzen & Fishbein, 1977). The TRA was designed to predict people's volitional (or voluntary) behaviors, whereas TPB

predicts and explains people's behavioral outcomes that are not under complete volitional control (Ajzen, 1991, p. 181). For both TPB and TRA, the central factor for predicting the targeted behavioral performance is intention. Ajzen (1991) asserted that intention should be an indication of how hard an individual is willing to try and how much of an effort an individual is planning to exert/exercise in order to execute the targeted behavioral outcomes. In the TRA model, two factors (attitude and subjective norm) predict individuals' behavioral intention. But in the TPB, an extension of TRA, three conceptually distinctive determinants -- attitude, subjective norm, and perceived behavioral control – are hypothesized to influence intention (*Figure* 6).

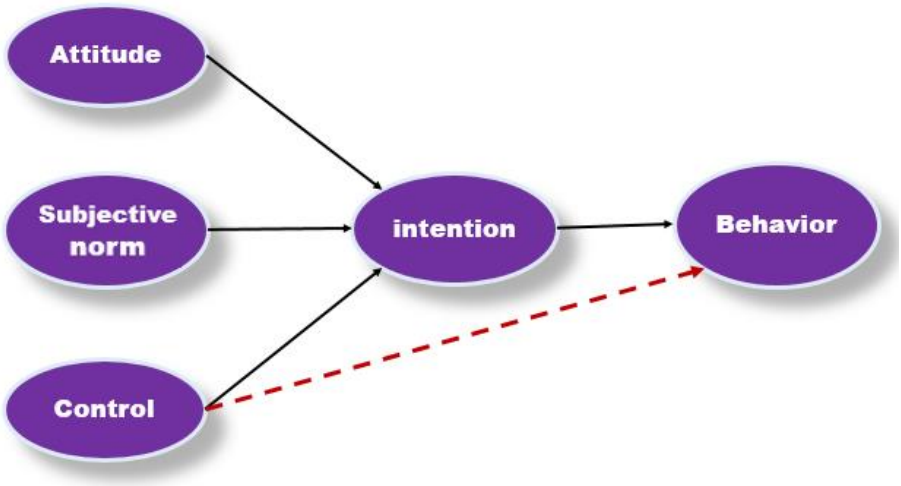


Figure 6. Theory of Planned Behavior (Ajzen, 1991)

Attitude refers to the degree to which an individual maintains an overall un/favorable appraisal toward the given task or behavior. Subjective Norm refers to the perceived social influence, social pressure, or social environment involved in executing the expected behavioral performance (Ajzen, 1991; Bentler & Speckart, 1979). Perceived

Behavioral Control indicates the strength of one's perceived belief about the capability (i.e., easy or difficult) to perform the targeted behavior (Ajzen, 1991), which is a similar construct to self-efficacy belief (Bandura, 1977). Simply put, the TPB model suggests that three conceptually independent antecedents (i.e., Attitude, Subjective Norm, and Perceived Control) predict Intention, and that Intention and Perceived Control, jointly, predict behavioral outcomes.

The Theory of Planned Behavior framework has been widely employed to understand and explain people's intentions to engage and carry out intended target behaviors in various contexts, including teachers' intention for technology usage (Lee, Cerreto, & Lee, 2010; Teo & Tan, 2012), science teachers' intention to incorporate environmental risk education (Zint, 2002), and nutrition-related behavior (e.g., Riebl et al., 2015). For example, Sheeran (2002) conducted a meta-analysis with a total sample size of n = 82,107, based on 422 hypothesis test results, and reported that intention had accounted for 28% of the variance, on average, in intention-behavior studies (p. 3-4). Using 185 independent studies, Armitage and Conner (2001) found that TPB accounted for 27% and 39% of the variances in behavior and intention, respectively. Among three independent determinants of intention in the TPB model, Attitude typically turns out to be the strongest predictor of Intention.

In addition, research regarding the TPB applicability in educational contexts, Lipnevich et al. (2011) investigated the relationship between math attitude and math achievement for middle school students. Participants in their study were 8th grade students from the U.S. (N = 382) and Belarus (N = 339). The authors reported that 62.6%

and 65% of variations in intention are explained by Attitude, Subjective Norm, and Perceived Behavioral Control for U.S. and Belarusian 8th graders, respectively. Noticeable findings, from both countries' samples, revealed that (1) the attitude component of the TPB model turned out to be the strongest predictor of intention, and (2) the structure of the theory of planned behavior model (hypothesizing the significantly positive relationship exists between math attitude and math achievement) was well supported by the data from both U.S. and Belarusian samples (Lipnevich et al., 2011). However, unlike Bentler and Speckart's (1979) study, Lipnevich et al. (2011) did not test whether attitude had a direct effect on the targeted behavioral outcome, i.e., math achievement.

TPB model components in educational contexts

Attitude factor

As stated previously, the attitude variable has been documented as the strongest predictor of behavioral intention among the TPB model components (Armitage & Conner, 2001; Lipnevich et al., 2011). Research on attitude toward mathematics has been focused, primarily, on two distinctive areas; one is concerned with theoretical development of the attitude construct (Hannula, 2002; Neale, 1969; Zan, & Di Martino, 2007) and the other is concerned with development of measuring attitude instruments (Aiken, 1970, 1974; Majeed, Darmawan, & Lynch, 2013; Tapia, 1996; Tapia & Marsh, 2002). Derived from the field of social psychology, the early definition of attitude referred to "a mental and neural state of readiness, organized through experience," that

influences people's action/response to the particular objects or situations (Allport, 1935, p. 810). In terms of math learning, Neale (1969) defined mathematics attitude as "a liking or disliking of mathematics, a tendency to engage in or avoid mathematical activity, a belief that one is good or bad at mathematics, and a belief that mathematics is useful or useless" (p. 632). From this multidimensional definition, one's attitude toward mathematics may be rather complex, involving three components, which are emotional response (dis/like, or enjoy) to math, one's belief about math, and academic behavior related to math (McLeod, 1992; Majeed et al., 2013; Zan, & Di Martino, 2007).

Generally speaking, attitudes are conceptualized as learned predispositions or tendencies to respond to certain objects, situations, or tasks either favorably or unfavorably (Aiken, 1970; Fishbein, 1967). According to this theoretical perspective, an individual's attitude toward math is not something with which s/he is born. For example, students' attitudes toward mathematics tend to be more negative from primary to secondary schooling (e.g., McLeod, 1992). This is because, in part, middle-school and high-school students are expected to learn more advanced mathematics content, which may require more time and effort to master deeper levels in several math sub-domains.

One may ask why a growing body of research examines the role of attitude in relations to math learning. Why not just focus on developing specific cognitive strategies as to how to absorb massive content in the course of enhancing math skills? To answer those questions, it is important to note that any cognitive component has an affective objective, and any affective component has a cognitive objective (Maker, 1982; Reyes, 1984), which implies that cognitive and affective domains are closely intertwined and

that affective factors should not be disregarded entirely. To emphasize these relations – affective and cognitive domains in learning, "the National Council of Teachers of Mathematics (1989) and the National Research Council (1989) have encouraged mathematics educators to incorporate affective factors with cognitive factors in mathematics teaching and learning" (Ma & Kishor, 1997, p. 26). Moreover, in terms of the positive impact of attitude on math learning, Suydam and Weaver (1975) argued that students learn more effectively when they are interested in what they learn, and they achieve better if they like what they learn.

For international context, Lindgren, Silva, Faraco, and Rocha (1964) conducted research on attitude and achievement in math in a sample of Brazilian elementary school students. A significant positive correlation between problem-solving-attitude and the achievement test in arithmetic was found (r = .24) among 4th graders (n = 108) in Brazil. Furthermore, upon surveying 387 high school students from the United Arab Emirates, Khine, Mutawah, and Afari (2015) conceptualized that the three sub-domain attitudinal factors (i.e., liking, value, and confidence) would predict math achievement independently. Khine et al. (2015) found that 36% of the variance in math achievement (R^2 = .36) was explained by the confidence factor (p. 207).

In addition, Ma and Kishor (1997) conducted a meta-analysis to examine the relations between attitude toward math and achievement in math using 107 independent studies from 1966 to 1993. Findings showed an overall weighted mean effect size (r = .12) that was positive but not strong; however, the findings from separate analyses by sex (r = .26 for males and r = .23 for females) are rather substantial. Although the

magnitudes of these effect sizes may be viewed as relatively small in terms of statistical value, the practical implication is still meaningful and important (Cohen,1988; Ma & Kishor, 1997; Rosenthal & Rubin, 1982). That is because the effect size of r = .23, according to the Binomial Effect Size Display (*BESD*), is equivalent to increasing math achievement rate from 38.5% to 61.5% (Rosenthal & Rubin, 1982).

Attitude factor: Global perspectives

A growing body of research has pointed out that Asian students, East-Asian (e.g. Chinese, Japanese and Korean) students in particular, show higher academic performance in math and science than non-Asian students (e.g., Chen & Stevenson, 1995; Schneider & Lee, 1990; Sun, 2002). Having closely examined the highest math performing countries (Singapore, South Korea, Japan, and Hong Kong, respectively) from the 3rd TIMSS survey data, Leung (2002) reported that no particular commonality was found among the top achieving East-Asian countries. He further noted that attitude variables (i.e., importance of learning math, interest in math, liking/disliking math, etc.) turned out to be strongly related to math scores for most countries around the world, not just for the top performing Asian countries.

Although the results of the descriptive statistics analyses from the survey questionnaires do not provide a clear picture of the reasons why East Asian students perform better in math than their peers from the rest of the countries, Leung (2002) suggested that certain cultural attributes might be uniquely associated with math outcomes among East-Asian countries. Some studies have tried to provide possible explanations of Asian students' academic success, but limited research is available on

whether or not attitude directly or indirectly influences Asian students' math achievement when compared to Western countries' students.

Social Norm factor

According to the TPB framework, subjective norm refers to the perceived social influence, social pressure, or social environment to execute the targeted tasks or behavioral performances (Ajzen, 1991; Bentler & Speckart, 1979). Subjective norm is influenced by the normative belief of significant/important others and the motivation to comply with those influential people around them (Ajzen, 1991). In other words, social norm represents the degree to which one's intention to engage (or participate) in certain tasks is impacted by people whom s/he thinks highly of, cares for, or respects the most. Among other factors, research shows that the quality of experiences students has such as parental supports and social interactions with peers and teachers might be critical components that could directly and/or indirectly shape students' approaches to academic engagement and learning outcomes (Phelan, Davidson, & Cao, 1992; Wetzel, 1998).

Developmentally speaking, adolescents tend to pursue more independence socially, psychologically, and emotionally. Thus, healthy social supports from parents, peers, and teachers during these crucial transitions from young adolescent to adulthood might greatly affect students' perceptions of academic motivation and achievement and the long-term view of education. Additionally, experts have argued that learning typically does not occur in complete isolation, and that learning results from an intricate social and interpersonal process (Dewey, 1934; Goodenow, 1992; Vygotsky, 1978). That is, the most efficient learning may take place through social interactions (e.g., child to child,

child to adult, and child to teacher), which stresses the importance of one's surrounding social and cultural environments (Vygotsky, 1978).

Concerning social norm influences on academic motivation, Wentzel (1998) examined relations between perceived social supports and various motivational outcomes with young adolescents (N = 167). The results of her study showed that (1) perceived peer support turned out to be the strongest predictor of prosocial goal pursuit (β = 0.29, p < .001), (2) perceived teacher support was the strongest predictor of school interest (β = 0.33, p < .001), and (3) family cohesion predicted mastery orientation the most (β = 0.23, p < .001) among motivation variables (Wentzel, 1998, p. 206).

In terms of peer influences on educational outcomes, one of the most frequently cited early studies is the Coleman Report (Coleman et al., 1966) – *Equality of Educational Opportunity*. Along with socio-economic status (SES) and home background, the report has indicated that "attributes of other students account for far more variation in the achievement of minority group children than do any attributes of school facilities…" (Coleman et al., 1966, p. 302). For example, suppose a 9th grader enjoys doing math homework and is very interested in continuously taking advanced math courses. However, all his/her close friends express how much they dislike math in their daily conversation, and they constantly ridicule his/her efforts to finish math assignments. Under this hypothetical social circumstance, according to the TPB model, although his/her attitude toward math is positive, this 9th grader may not work as hard as s/he should for the upcoming math test because of the peer pressures, which, in turn, may result in negative math outcomes.

Besides peer influence on education development, research also shows that parental influence has a positive effect on students' academic achievement (Kim et al., 2010; Peng & Wright, 1994). Across 25 empirical studies (cumulative sample size, N = 133,577) from 1987 to 1996, Fan and Chen (2001) conducted a meta-analysis on the relations between parental involvement and students' achievement. The effect size (the overall average correlation coefficient, r = .25) was reported as statistically significant (Fan & Chen, 2001, p. 11).

Regarding effects of parental roles and home environments on achievement, Peng and Wright (1994) found that family income level and parental educational level were positively associated with student achievement across all participating ethnic groups. Peng and Wright (1994) found that Asian parents did not provide any more direct help to their children with homework than other parents did. One significant finding, however, was that Asian parents showed their high educational aspiration explicitly or implicitly to their children more than other parents did. That is, "about 80% of Asian American parents expected their children to have at least a bachelor's degree, compared with 50% of Hispanic, 58% of Black, and 62% of White parents" (Peng & Wright, 1994, p. 350).

Another significant social influence that contributes to students' educational outcome is the teacher. Hallinan's research (2008) has pointed out the crucial role of teachers in student learning, by asserting that "a large part of a student's school day is spent in verbal and nonverbal interactions with teachers" in classrooms (p. 272). Such supportive authority figures as teachers, then, may positively contribute to students' engagement, autonomy, and competence in the learning process, and could help them to

obtain higher grades (Finn & Rock, 1997; Reyes, Brackett, Rivers, White, & Salovey, 2012). Conversely, disengaged and struggling students, who receive no teacher supports, are more likely to fail and drop out of school, especially when they feel alienated or disconnected from their teachers and peers (Finn, 1989; Phelan et al., 1992; Reyes et al., 2012). Becker and Luthar (2002) also found that, when young students develop positive rapport with their teachers, they are more likely to experience enhanced social growth and suffer less from the negative effects of a stressful school environment.

Moreover, research evidence shows that the best learning outcomes typically occur when teachers provide a warm, encouraging, and safe classroom environment (Jones et al., 2014; Reyes, Bailey, & Jacob, 2012). Having established good rapport with their teachers, students tend to be less anxious and seek help when they struggle because they know they are not under attack and feel protected emotionally and intellectually. Furthermore, Hamre and Pianta (2001) defined positive teacher-student relationships as having the presence of closeness, warmth, and positivity. With a longitudinal study of K-8th graders (N = 179), Hamre and Pianta (2001) found that the quality of early teacher-child relationships was associated with a positive trajectory of later schooling.

Subjective Norm factor: Global perspectives

Having examined previous research investigations, the Subjective Norm factor (i.e., Friend, Parent, Teacher) or social environment, among other related attributes, appears to have a moderate impact on student outcomes such as academic achievement, motivation, socio-emotional development (Peng & Wright, 1994; Reyes et al., 2012; Vygotsky, 1978; Wentzel, 1998). What needs to be done, however, is to identify whether different aspects

of subjective norm have differential effects on math achievement for Asian students and their counterparts in Western countries.

Perceived Behavior Control (PBC) factor

According to Ajzen (1991), perceived behavior control (PBC) refers to the magnitude of one's perceived belief about the capability (i.e., easy or difficult) to perform the targeted behavior. The perceived control factor in the TPB model is similar to the construct of self-efficacy belief, which is based on Bandura's (1977) social cognitive theory. Self-efficacy focuses on beliefs or judgments about individuals' performance capabilities for specific tasks in particular situations (Pajares, 1996; Schunk, 1985, 2012; Zimmerman, 2000).

Stated differently, self-efficacy involves people's perception or discernment of their "capabilities to organize and execute courses of action required to attain designated types of performances" (Bandura, 1986, p. 391), which is task-specific activity. Ames (1991) also emphasized the role of self-efficacy in academic contexts by asserting that "self-efficacy is often a critical factor predicting children's task choices, willingness to try and persist on difficult tasks, and even actual performance…" (p. 412). In addition, Linnenbrink and Pintrich (2003) claimed that, among other motivational constructs, "self-efficacy [was]…key to promoting students' engagement and learning" (p. 119) by delineating the role of self-efficacy to facilitate the three specific areas of engagement in the classroom, i.e., students' behavioral, cognitive, and motivational engagement.

In terms of the influence of PBC on behavior intention, Ajzen (1991) stated that the predictability of PBC (i.e., self-efficacy belief) on intention may vary, meaning that it

depends upon the type of behavioral outcomes to be measured and the nature of situations. From reviewing nine independent studies that used the TPB model, the range of multiple correlations (PBC + intention → behavioral outcome) was found from .20 to .78 (average .51). Interestingly, as one might have anticipated, two studies (i.e., losing weight and getting an 'A' in a course) show the lowest predictabilities among other volitional control studies (Ajzen, 1991, p. 187).

Theoretically, the constructs of PBC and self-efficacy beliefs are similar. Thus, in what follows, research on self-efficacy that is associated with academic outcomes will be discussed. For example, Betz and Hackett (1983) investigated how math self-efficacy is related to selection of science-based college majors by females and males. Participants were asked to rate their confidence level in their ability (i.e., self-efficacy) to successfully solve the problem or perform the task. For the college course subscale, participants were asked to rate their confidence level in their ability to complete the course with a grade of "B" or better; a 10-point scale was used (ranging from "0" = no confidence at all to "9" = complete confidence). Undergraduate male students "reported significantly stronger mathematics self-efficacy expectations than did females on all three subscales (i.e., Math Problems, Math Tasks, and College Courses) and on the total scale" (Betz & Hackett, 1983, p. 342).

Multon, Brown, and Lent (1991) conducted a meta-analysis on the relation between self-efficacy and academic performance/persistence, analyzing studies published between 1977 and 1988. Multon et al. (1991) reported that efficacy beliefs are positively associated with academic performance, "r = .38," (Multon et al., 1991, p. 33). More

recently, Motlagh, Amrai, Yazdani, Abderahim, and Souri (2011) examined the relation between self-efficacy and academic achievement. Two-hundred fifty high school students were involved in the study; five sub-categories of the self-efficacy construct (i.e., self-belief, self-regulation, self-evaluation, self-stimulation, and self-monitoring) were measured as predictors of academic achievement. They found that self-evaluation and self-regulation are the strongest factors among five subcategories of the self-efficacy construct that predicted high school students' academic achievement.

PBC factor: Global perspectives

From the PISA 2012 report, across OECD participating countries, most adolescents (i.e., 15-years old) around the world seem to think that PBC is related to math outcomes:

> …92% of students reported that they agree or strongly agree that *they can succeed in mathematics if they put enough effort*, 83% reported that *whether they do well in mathematics or not is completely up to them*, …83% reported that *they agree or strongly agree that they could do well in mathematics if they wanted to*. … (OECD, 2013, p. 62, italicized are mine for emphasis)

From his descriptive analyses using the 3rd TIMSS data, Leung (2002) reported that East Asian students appear lacking in confidence, despite their outstanding achievement. He further noted that the way Asian students respond to others about their own academic capabilities, sometimes, could be seen as having low self-efficacy from non-Asians' point of view, because Asian students are typically taught to be modest and "not to be boastful" (Leung, 2002, p. 106).

The following examples, which are drawn from the PISA 2012 student questionnaire data, illustrate how Asian students respond to the questions about their academic abilities (Y-axis represents the percentile of student responses by countries).

Example.1]	Regarding openness to problem solving, how well does the following statement describes you? **I am quick to understand things**

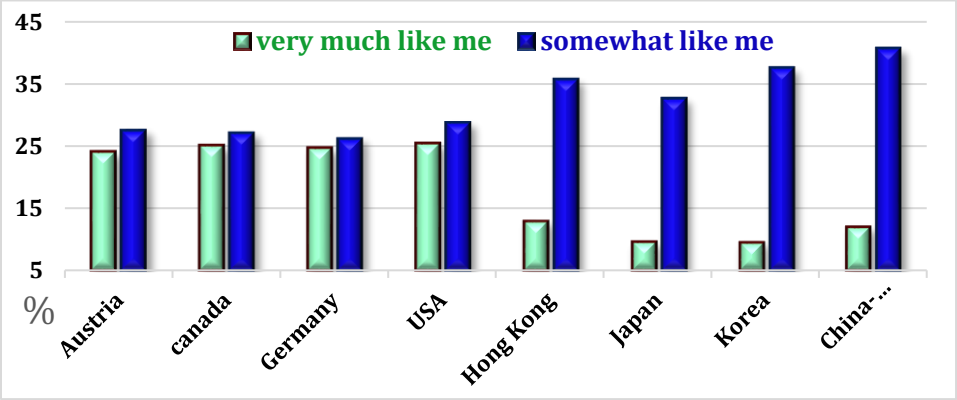


Figure 7. Self-reports on math ability – 1st example: Western vs. Asian countries
Source: ***OECD, PISA 2012 Database***

If one looked only at the green bar graph, one might assume that Asian students in general consider themselves as 'not very smart' or 'slower learners' when compared to their peers from Western countries. However, from the blue graphs from Asian countries (Hong Kong, Japan, Korea, and China-Macao), it suggests that many Asian students do think that they are actually quick learners.

Example.2]	Thinking about studying mathematics: to what extent do you agree with the following statements? **I learn mathematics quickly**

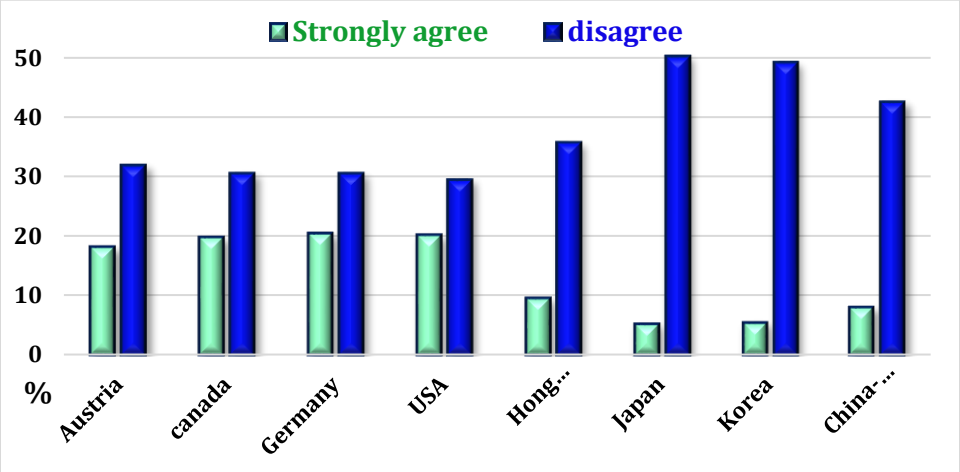


Figure 8. Self-reports on math ability – 2nd example: Western vs. Asian countries
Source: *OECD, PISA 2012 Database*

Again, having examined the 2nd example above, it may not be an accurate interpretation to make a sweeping generalization that Asian students' self-efficacy beliefs are not as strong compared to their peers in Western countries – based on the green bar graph only. Perhaps, responses by Asian students may reflect a cultural issue, being influenced by Confucian heritage. That is, Some Asian students may be taught to be reflective about their own academic abilities by being humble, which might influence them to work harder toward the subject that they are supposed to learn.

Research questions

Although research findings clearly show that the TPB model has been well-established to explain attitude-behavior relations in a number of behavioral domains, there seems to be limited research on the applicability of the TPB model for academic outcomes (except one by Lipnevich et al., 2011). Further, having reviewed the prior

literature on the relations between non-cognitive factors (i.e., attitude, social norm influence, perceived belief, etc.) and academic outcomes, the present study aims to investigate how multiple non-cognitive factors are related to math achievement, by utilizing the TPB model, across national groups – USA, Germany, Japan, and Korea.

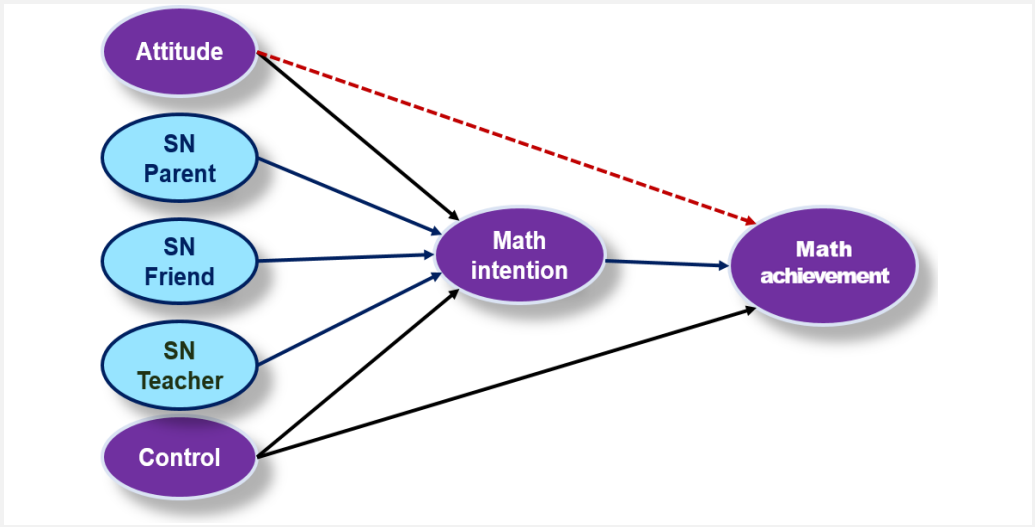


Figure 9. Conceptual model of the present research – an extended version of TPB
Note. Attitude: attitude toward math, SN: Subjective Norm, Control: perceived belief, Math intention: intentional academic behavior related to math (e.g., I study hard for math quizzes)

Research Questions:

1. Do non-cognitive factors (i.e., the TPB model components) convey approximately the same meaning in four different countries – USA, Germany, Japan, and Korea?

2. Would the structure of the TPB model be supported by the 2012 PISA data across the four national groups?

3. Would a substantial amount of the variance in Math Behavior Intention be explained by non-cognitive factors (i.e., TPB model components) across national groups?

4. Are cultural differences evident in the magnitude of effects of each subjective norm factor on intentions to work hard on mathematics?

5. Would a substantial amount of the variance in Math Achievement be explained by Attitude, Intention, and Perceived Control across national groups?

CHAPTER THREE

METHODS

Data Source: 2012 PISA

Characteristics of PISA

The dataset to be used in the current proposed study is drawn from the Programme of International Student Assessment (PISA) of 2012. PISA is a large-scale international educational assessment, developed by Organization for Economic Co-operation and Development (OECD) since 2000, which focuses on reading, mathematics, and science. PISA assessments are repeated every 3 years with an emphasis on specific areas each year – reading in 2000, math in 2003, science in 2006, and so forth (OECD, 2014b). In 2012, math was again the major focus of PISA assessment, which was the reason why the 2012 PISA was used for the present research investigation concerning effects of non-cognitive factors on math achievement. According to OECD, all PISA data are publicly available.

Unlike other large-scale international assessments such as the Trends in International Mathematics and Science Study (TIMSS), PISA seeks to examine how well young adults, toward the end of their compulsory education, can utilize their acquired skills and knowledge that are essential for full participation in modern societies (OECD, 2014b). That is, PISA assessment measures how students can apply skills and knowledge (i.e., how well they can apply what they know in their real-life situations), whereas

TIMSS assessments measure students' curriculum-based knowledge (i.e., how much they know).

Sampling procedures of PISA

Unlike a simple random sampling procedure, PISA assessments employ a two-stage stratified sampling design for practical reasons such as efficiency of measurements and issues with budgeting. The first-stage sampling involves selecting schools where 15-year-olds are attending through systematic sampling procedures (OECD, 2014b). In each country, once eligible schools are identified, the second-stage sampling procedure applies, meaning that 15-year-old students are randomly assigned to take tests from randomly chosen classes within the selected schools. This stratified sampling procedure reduces sampling errors of the population (OECD, 2014b).

In 2012 PISA assessments, about 510,000 students from 65 countries and economies (OECD member countries are 34) participated. The randomly chosen 510,000 students are representative of approximately 28 million 15-year-old students around the world who were enrolled in schools. Those students took cognitive tests (reading, math, and science) and answered survey questionnaires that are related to their personal backgrounds and various educational aspects (OECD, 2014).

Plausible Values in PISA

Large-scale international education assessments such as PISA or TIMSS are typically concerned with population estimates rather than individual estimates. One of the chief purposes of a large-scale study like PISA is to analyze the characteristics of the population (i.e., 15-year-old students among different countries) in terms of progress in

achievement (OECD, 2009; Wu, 2005). That is, interest for large-scale educational assessments like PISA lies in reducing error in making inferences about the target population, not making inferences at the level of individual participants. Thus, student math performance scores in PISA were reported through 5 plausible values (PV) – estimates of math performances, not the raw test scores of math. Based on Rubin's work (1987) on multiple imputations, the plausible value approach was developed by Mislevy and Sheehan (1989) in analyzing large secondary datasets (Monseur & Adams, 2009; Wu, 2005). PVs represent the multiple imputed values for unobservable latent variables (Asparouhov & Muthén, 2010; OECD, 2009; Wu, 2005, p. 114). Students' true cognitive abilities cannot be observed and therefore must be estimated from their answers on tests (Jerrim & Choi, 2014). In other words, students do not complete the entire math tests in PISA; each student answers different sets of math test items.

Simply stated, PVs represent the range of achievement outcomes that individuals might reasonably have, based on their test item responses (Wu, 2005). PVs are random draws or random values from mathematically computing distributions (i.e., probability/posterior distributions) and, thus, should not be identified as individuals' raw scores (OECD, 2009; Wu, 2005). Although 5 plausible values on mathematics performances for each student were reported in 2012 PISA, on average, one plausible value may be sufficient to provide "unbiased population estimates as well as unbiased sampling variances" (OECD, 2009, p. 131).

Subgroup dataset for the present research

The aim for the present cross-national comparison study was to provide a better understanding of how multiple non-cognitive factors were associated with adolescents' math achievement. Among the entire 65 countries that participated in the 2012 PISA assessment, two Western countries (USA: N = 4978 and Germany: N = 5001) and two Asian countries (Japan: N = 6531 and South Korea: N = 5033) were selected to examine whether or not there were cultural differences or national similarities among 15-year-olds' in math achievement. The following items were downloaded from the official OECD website to conduct the present research:

(http://www.oecd.org/pisa/pisaproducts/database-cbapisa2012.htm)

(1) Student Questionnaire File in TXT format (CBA_STU12_MAR31)
(2) Codebook for Student Questionnaire Data file
(3) SASTM Control File – SAS syntax to read Student Questionnaire Data File

From the 2012 PISA database, a subgroup dataset was created, which included four countries (USA, Germany, Japan, and South Korea) along with all variables associated the proposed conceptual model components.

Measures

To evaluate the utility of the TPB model in educational contexts, observed variables (survey questionnaire items) should reflect the corresponding latent factors of the TPB model. In other words, to be able to predict math achievement accurately, an intention factor should measure math-related intentional academic behavior (e.g., *I work hard on my math homework*). That is, based on the TPB framework (Ajzen, 1991), items that were measured regarding a general school-related factor (e.g., *Things are ideal in my*

school) or a general math-related-behavior (e.g., *I play chess*) may not be applicable and should not be used as indicators of the math-intention construct to predict math achievement. As such, based on the TPB theoretical framework (Ajzen, 1991), an attitude factor related to math (e.g., I enjoy reading about math) seems to be more suitable than an attitude factor related to school (e.g., Trying hard at school is important) to examine the applicability of the TPB model on academic achievement in math.

The selected survey items for the present study's conceptual model were extracted from the 2012 PISA Student Questionnaire File, and they were answered on 4-point Likert-type scales (i.e., 1 = strongly agree, 2 = agree, 3 = disagree, and 4 = strongly disagree). However, for the purpose of simplicity and convenience, all measured items were rescaled; 3 as strongly agree to 0 as strongly disagree, which was to show that the higher number would represent the greater numerical value of the students' self-reported survey responses. In what follows, all relevant latent constructs of the TPB model components for the present study are provided in Table 1.

Table 1. *Descriptions of Constructs of the Proposed Conceptual Model*

constructs	**Variable names**	**Variable descriptions**
Math Attitude	29Q01	I enjoy reading about mathematics
	29Q03	I look forward to my mathematics lessons
	29Q04	I do mathematics because I enjoy it
	29Q06	I am interested in the things I learn in mathematics
SN - Friend	35Q01	Most of my friends do well in mathematics
	35Q02	Most of my friends work hard at mathematics
	35Q03	My friends enjoy taking mathematics tests
SN - Parent	35Q04	My parents believe it's important for me to study mathematics
	35Q05	My parents believe that mathematics is important for my career
	35Q06	My parents like mathematics
SN - Teacher	Last math teacher: to what extent do you agree?	
	83Q02	My teacher provides extra help when needed
	83Q03	My teacher helps students with their learning
	How often do these things happen in your mathematics lessons?	
	77Q02	The teacher gives extra help when students need it
	77Q04	The teacher helps students with their learning
Perceived	43Q01	If I put in enough effort I can succeed in mathematics
Control	43Q02	Whether or not I do well in mathematics is completely up to me
In Math	43Q05	If I wanted to, I could do well in mathematics
Intentional	46Q01	I finish my homework in time for mathematics class
math behavior	46Q02	I work hard on my mathematics homework
	46Q03	I am prepared for my mathematics exams
	46Q04	I study hard for mathematics quizzes
Math	PV1math	Plausible value 1 in mathematics
achievement	PV2math	Plausible value 2 in mathematics

Source*: OECD, PISA 2012 Database*

Data Analyses

Why Structural Equation Modeling (SEM)?

The present study aimed to investigate the direct/indirect effects of multiple non-cognitive attributes on math achievement across national groups (i.e., USA, Germany, Japan, and Korea) based on the TPB framework using the large-scale international data from the 2012 PISA. To examine the proposed theoretical model (see *Figure* 9), Structural Equation Modeling (SEM) was used for analyzing the data with Mplus 7 (Muthén & Muthén, 1998-2012). SEM has become a useful method in social and behavioral sciences. SEM allows researchers to be explicit about their theory, and to specify, estimate, and test the hypothesized interrelations among a set of complex latent factors (Crowley & Fan, 1997, p. 510; Kline, 2011). Thus, the goal of SEM techniques for the present research was to match the proposed explicit theory with model (e.g., the TPB model) and the observed data (e.g., four countries of the 2012 PISA data) as closely as possible. The total sample size for the study was 21,363 from four countries (USA: N = 4978, Germany: N = 5001, Japan: N = 6351, and South Korea: N = 5033).

It is helpful to have a big sample size in SEM analyses, and the 2012 PISA dataset is rather large. Because large samples were available for analyses, the sample data from each country were divided into 2 groups. Only the first half of the data from the four countries was used for initial SEM analyses and therefore functioned as derivation samples. The other half sample from each country was used as cross-validation samples to examine if the cross-validation sample would produce similar results to those from the derivation samples. In prior research on cross-national comparative studies, TPB model

studies, or other non-cognitive studies in relation to academic contexts, this replication strategy was rarely (if ever) used in the past.

In what follows, several advantages of using SEM to answer the proposed research questions are discussed, but not limited to:

(1) First, SEM does not designate a specific/single statistical technique. Rather, under one umbrella term, SEM refers to multivariate statistical techniques with combinations of multiple regression analysis and factor analysis, which allows us to examine a set of complex relations simultaneously (Kline, 2011; Nachtigall, Kroehne, Funke, & Steyer, 2003; Ullman, 2006; Widaman & Thompson, 2003).

(2) Second, SEM software programs allow the user to (a) examine those multivariate theories involving latent entities in a single analysis and (b) specify and test "even very complex structural models with relative speed and ease" (Widaman & Thompson, 2003, p. 16). For the present study, SEM techniques were employed to test direct/indirect effects of the multiple non-cognitive attributes on math achievement across four national groups simultaneously in order to compare the cultural differences (Western vs. Asian countries) and/or similarities among four countries.

(3) Third, what distinguishes SEM analysis from other traditional analysis such as ANOVA is that the SEM program not only examines covariance structures but also estimates means of latent variables while "ANOVA is concerned with means of observed variables only" (Kline, 2011, p. 10).

Measurement and Structural Model

SEM involves both measurement model (i.e., relations between observed and latent variables) and structural model (i.e., relations among latent variables themselves). For the measurement model, Exploratory Factor Analysis (EFA) was conducted to identify whether or not each observed variable (i.e., indicators/test items) was significantly loaded on the designated latent constructs (i.e., non-cognitive factors). Observed variables were eliminated if they had loadings close to zero on their intended factor or if they loaded on more than one factor. Once optimal indicators were identified using EFA, then, Confirmatory Factor Analyses (CFA) were conducted to make sure all observed variables (i.e., indicators) had high loadings on each specified latent factor. Table 1 shows the final CFA results, and inspection of Table 1 reveals that all observed variables for each latent construct had high loadings on their respective factors, so further analyses were justified. For structural model, using CFA techniques, multi-group analyses were conducted in order to investigate the complex interrelated non-cognitive factors on academic achievement in math across four national groups. More details regarding multiple-group analyses are discussed later.

In SEM terminology, variables, either latent or observed, that "only exert an 'effect' on other variables are called exogenous variables, while those that receive an effect from any others are called endogenous variables" (Crowley & Fan, 1997, p. 510). Conceptually, the proposed theoretical model (see *Figure* 9) can be understood as follows: Attitude, SN-Friend, SN-Parent, SN-Teacher, and Perceived Control are correlated latent constructs that have several observed variables (i.e., indicators). The

Intention construct is causally affected by all five exogenous latent constructs (Attitude, SN-Friend, SN-Parent, SN-Teacher, and Perceived Control), and Math Achievement construct (i.e., two PVs values were used) is causally affected by Attitude, Intention, and Perceived Control. Meanwhile, it is important to note, regarding the graphic representation of SEM, that "the word cause is meant to provide no philosophical meaning beyond a short-hand designation for a hypothesized unobserved process" (Bentler & Speckart, 1979, p. 456).

In order to account for missing data, full information maximum likelihood estimation (FIML) was used to fit models through Mplus 7 (Muthén & Muthén, 1998-2012). "FIML uses all available data, and only drops cases when information is missing on all data points" (Hernández, Robins, Widaman, & Conger, 2014, p. 8). Using FIML produces less biased and more reliable results compared with listwise or pairwise deletion to deal with missing data (Widaman, 2006). In addition, to assess model fit (i.e., how well the two structures match, meaning that how well the 2012 PISA data and the TPB model would closely align), several goodness of fit statistics were examined, which includes (1) Chi-Square (χ^2) test of significance of model fit and (2) other incremental fit indices.

The likelihood ratio Chi-Square statistic is based on "the discrepancy between the sample moment (or covariance) matrix and the model-implied moment (or covariance) matrix" (Widaman & Thompson, 2003, p. 17). Thus, a non-significant Chi-Square (χ^2) value would imply good fit because it suggests that the sample data – covariances and means – are closely modeled by the structural model, leaving small residual values. Other

incremental fit indices such as the Comparative Fit Index (CFI; Bentler, 1990), the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), and the Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993) with the confidence interval were used to evaluate the best fitting model to data. CFI and TLI values of .95 or greater indicate good fit of the model. On the other hand, an RMSEA value of less than .06 indicates the model has close fit to the data (Hu & Bentler, 1999; Kline, 2011; Raykov & Marcoulides, 2012). Detailed information regarding incremental fit indices in SEM is discussed in Widaman and Thompson (2003).

Multiple-group Analyses

Once factor relations (between observed and latent variables) have been well developed from the entire sample of four countries (i.e., all relevant indicators were identified to each corresponding latent construct), it is imperative to assess whether or not the measurement model for the TPB model components varied across all four countries. Establishing measurement invariance across all groups (USA, Germany, Japan, and Korea) is critical for supporting tests across groups in the differences/similarities in structural relations among latent variables. In other words, if factor scores are not comparable across groups' population, then the differences between groups in mean levels or in the pattern of correlations would be potentially biased and the results may be substantively misleading (Meredith, 1993; Widaman & Reise, 1997; Widaman, Ferrer, & Conger, 2010). Hence, testing measurement invariance across groups (i.e., USA, Germany, Japan, and Korea) is important before evaluating regression path coefficients

among latent factors to examine the cultural differences or similarities among the four nations.

Measurement Invariance Testing

As mentioned earlier, constructs (e.g., cognitive abilities, attitudes, beliefs, etc.) cannot be measured directly, so multiple observable variables (i.e., indicators or survey items that can be numerically scored) are used as indicators of the hypothetical constructs. However, typically, indicators are not perfect due to measurement errors. A core principle of measurement invariance is that latent variables are assumed to be on the same scale if the relations between the indicators and the constructs (i.e., non-cognitive attributes in the present study) are invariant across groups. One way to assess measurement invariance is through factor analytic techniques (Cheung & Rensvold, 2002; Meredith, 1993; Widaman & Riese, 1997; Windle, Iwawaki, & Lerner, 1988). In this factorial approach, the relations between multiple observed variables and multiple common factors in a matrix format can be specified as follows:

$$X = \tau + \Lambda_x \xi + \delta, \quad (1)$$

where X is a vector of observed variables, $\boldsymbol{\tau}$ is a vector of intercepts, Λ_x is matrix of factor loadings, ξ is a vector of common factors, and δ is a vector of unique variables. The principle of multiple-group analysis in CFA as opposed to single-group analysis is to fit factor models in several groups (e.g., all four countries in the current study) simultaneously (Lubke, Dolan, Kelderman, & Mellenbergh, 2003). Measurement invariance is defined as follow:

$$f(Y|\eta, s) = f(Y|\eta) \quad (2)$$

where $\boldsymbol{f}(\cdot)$ is a distribution function, $\boldsymbol{s}$ denotes group membership (e.g., 1 = USA, 2 = German, 3 = Japan, and 4 = Korea) and $\boldsymbol{Y}$ and $\boldsymbol{\eta}$ are observed scores and factor scores, respectively (Lubke et al., 2003, p. 549).

CFA applications for testing measurement invariance were first presented in 1970 by Joreskog (1971). Then, based on work from Thurstone (1947), Horn, McArdle, and Mason (1983), and Meredith (1993), Widaman and Reise (1997) identified four levels of factorial invariance and presented a sequence of nested models; namely, configural, weak, strong, and strict invariance (see also Widaman & Thompson, 2003). Configural invariance serves as a baseline for subsequent factor analytic measurement invariance tests. Within structural modeling, a multiple-group analysis is done by analyzing all four groups simultaneously using a linear CFA model.

In terms of examining the sequential measurement invariance test (from configural to strong factorial invariance) to establish model fit to the data, Chi-Square (χ^2) statistical tests were used to evaluate the overall model fit across groups. However, it is important to note that χ^2 tests are to test the null hypothesis (i.e., no statistically significant differences exist between the observed and the proposed factor structures), which implies non-significant χ^2 is desirable. Also note that Chi-Square is sensitive to sample size. Since the present national comparison study involves a large sample size drawn from the 2012 PISA, additional comparative fit-indices (e.g., RMSEA, TLI, CFI) were used to assess the overall model fit. Based on Widaman and Reise (1997), in the

following, sequential factorial measurement invariance testing, in a simple form, are summarized.

Table 2. *Types of Factorial Measurement Invariance Testing*

Step 1:	Configural	• Same pattern of fixed and free loadings across group • Values of the parameters vary (different in each group) • Serves as a baseline for sequential measurement invariance tests
Step 2:	Weak (Metric)	• Factor loadings ($\mathbf{\Lambda}_G = \Lambda$) are constrained to invariance • Same factor loadings across groups • Evaluates if the construct has the same meaning across groups
Step 3:	Strong (Scalar)	• Intercepts ($\boldsymbol{\tau}_G = \boldsymbol{\tau}$) are constrained to invariance + Step 2 • Same item intercepts across groups • Evaluates latent factor mean differences across group
Step 4:	Strict	• Unique variances ($\boldsymbol{\delta}_G = \boldsymbol{\delta}$) are constrained to invariance + Step 3 • Same unique (error) variance of each indicators across groups • Evaluates how group differences in variances of indicators are attributable only to group differences in variances of latent factors, since error variances are invariant across groups – nevertheless, this occurs hardly ever in empirical research

Although Step 4 (strict invariance model) may be an ideal, it rarely holds in practice. Thus, researchers suggest that the establishment of invariant unique variances over groups is not necessary (Schmitt & Kuljanin, 2008; Widaman & Reise, 1997). From an intensive review on measurement invariance research that was published between 2000 and 2007, Schmitt and Kuljanin (2008) reported that, among 88 studies, "75 conducted empirical analyses of measurement invariance using CFA methods" (p. 212). They also reported that strict invariance was rarely achieved in reality, thus, most studies addressed only configural, weak, and strong invariance.

Having examined thoroughly, Schmitt and Kuljanin (2008) also argued that approximately half of their reviewed studies had conducted partial invariance because full invariance has been difficult to attain in practice. Although experts disagree about whether partial invariance should be allowed, it has been considered reasonable to utilize a partial invariance approach in establishing measurement consistency across groups. As one example in the study of cross-cultural comparisons, Lipnevich et al. (2011) used a partial invariance test for their investigation on the relations between attitude and math achievement among U.S. and Belarus 8th graders. The present research also utilized partial invariance in testing measurement invariance across national groups.

Having identified a sufficiently well-fitting model of measurement invariance across groups using both derivation and cross-validation samples, structural modeling of regression relations among the complex latent factors were conducted for all four nations simultaneously. These final multi-group analyses were crucial for the following reasons. First, it allowed me to examine how each component of the TPB model affected math achievement across groups. Second, it helped me to compare directly any differences among four countries (i.e., USA, Germany, Japan, and South Korea), in terms of effects of non-cognitive factors on math achievement. Lastly, the second half data were used to examine whether results from these cross-validation samples led to similar conclusions as would results from the derivation samples.

CHAPTER FOUR

RESULTS

Preliminary Analyses

From among 65 countries that participated in the 2012 PISA assessments, four countries were selected for the data analyses; USA (N = 4978), Germany (N = 5001), Japan (N = 6351), and Korea (N = 5033). In order to understand better the data from the four countries (N = 21363), several descriptive analyses were conducted. Table 3 shows means and standard deviations of scale scores representing the latent factors in the conceptual model components for the present research (i.e., Theory of Planned Behavior -TPB model). All measured items (observed variables) of latent factors, besides the criterion variable – math achievement, were answered on 4-Likert type scales from 4 (strongly disagree) to 1 (strongly agree), but re-scaled from 0 (strongly disagree) to 3 (strongly agree) for convenience.

Table 3. *Means and Standard Deviations for Scale Scores of All Countries*

Factors	Mean	SD	Min.	Max.
Attitude	4.75	3.17	0	12
SN-Friend	3.96	1.60	0	9
SN-Parent	5.51	1.99	0	9
SN-Teacher	8.22	2.78	0	12
Control	6.69	1.81	0	9
Intention	6.72	2.67	0	12
Math	522.2	96.76	148.55	857.85

Note. SN = subjective Norm; Math = two plausible values (pv1math and pv2math) were used to represent math achievement. Attitude, SN-Teacher, and Intention have 4 indicators, thus, 3 x 4 = 12 (maximum). SN-Friend, SN-Parent, and Control have 3 indicators, thus, 3 x 3 = 9 (maximum).

SN-friend, SN-parent, and Control factors had 3 indicators, and their scale mean scores were 3.96, 5.51, and 6.69, respectively. Attitude, SN-teacher, and Intention factors had 4 indicators, and their scale mean scores were 4.75, 8.22, and 6.72, respectively. To examine how closely latent factors of the TPB model components were associated with one another, correlation analyses were conducted. Pearson correlation coefficients were calculated for the entire sample (N = 21363), which was the total number of participants from the four countries. Table 4 provides the results of the correlations among latent factors.

Table 4. *Correlations among Scale Scores for Latent Factors of All Countries*

	1	2	3	4	5	6	7
1. Attitude	1						
2. SN friend	.28	1					
3. SN parent	.36	.27	1				
4. SN teacher	.28	.25	.17	1			
5. Control	.41	.13	.33	.24	1		
6. Intention	.44	.20	.27	.26	.28	1	
7. Math	.24	**-.01**	.11	.05	.20	.09	1

Note. SN = Subjective Norm. Bolded value indicates statistically non-significant (p = .39)

All scale scores were statistically significantly correlated (p < .001), except one (SN Friend with Math, p = .39). Although the correlations between SN-Teacher with Math (r = .05) and Intention with Math (r = .09) were statistically significant, they were relatively small. Attitude was highly correlated with Intention (r = .44) and Control (r = .41).

Table 5 shows means and standard deviations of scale scores for latent factors by each country. The means of scale scores on the SN Parent factor in Western countries (6.16 in USA and 6.28 in Germany) were higher than in Asian countries (4.61 in Japan and 5.38 in Korea).

Table 5. *Means and Standard Deviations of Scale Scores by Each Country*

	Western Countries			
	USA		**Germany**	
Latent Factors	Mean	SD	Mean	SD
Attitude	5.41	3.15	4.77	3.30
SN-Friend	4.53	1.55	3.89	1.54
SN-Parent	**6.16**	1.72	**6.28**	1.82
SN-Teacher	9.23	2.50	7.15	3.21
Control	6.94	1.68	6.89	1.90
Intention	7.61	2.43	6.90	2.75
Math	481.21	87.6	513.86	95.1
	Asian Countries			
	Japan		**Korea**	
Latent Factors	Mean	SD	Mean	SD
Attitude	4.41	3.13	4.52	3.02
SN-Friend	3.73	1.60	4.10	1.49
SN-Parent	**4.61**	2.02	**5.38**	1.84
SN-Teacher	8.30	2.73	8.03	2.34
Control	6.48	1.83	6.55	1.79
Intention	6.73	2.62	5.68	2.61
Math	535.9	92.39	553.72	97.15

Note. SN = subjective norm, SD = Standard deviation. Bolded values are to emphasize the difference between Western and Asian countries.

Table 6 provides descriptive statistics of all measured items (i.e., 23 observed variables/indicators) for the entire sample of four countries. The skewness and the kurtosis of all 23 observed variables were lower than the absolute value of 3, which indicates the data were fairly normally distributed.

Table 6. *Descriptive Statistics of 23 Observed Variables for the Entire Sample*

Factors	Measured Items	Mean	Std. Dev	Skewness	Kurtosis
Attitude	29q01	**0.96**	0.70	0.55	-0.33
	29q03	1.19	0.82	0.34	-0.67
	29q04	1.19	0.89	0.39	-0.75
	29q06	1.41	0.86	0.08	-0.85
SN Friend	35q01	1.55	0.48	-0.00	-0.23
	35q02	1.61	0.50	-0.14	-0.19
	35q03	**0.80**	0.47	0.65	0.66
SN Parent	35q04	2.15	0.60	-0.71	0.19
	35q05	1.96	0.71	-0.43	-0.47
	35q06	1.41	0.65	0.11	-0.47
SN Teacher	77q02	2.04	0.78	-0.54	-0.58
	77q04	2.07	0.78	-0.63	-0.42
	83q02	2.04	0.66	-0.69	0.16
	83q03	2.07	0.59	-0.69	0.41
Perceived Control	43q01	2.29	0.51	-0.81	0.51
	43q02	2.30	0.53	-0.86	0.52
	43q05	2.10	0.63	-0.61	-0.08
Intention	46q01	1.76	0.78	-0.21	-0.72
	46q02	1.71	0.69	-0.19	-0.52
	46q03	1.77	0.74	-0.25	-0.60
	46q04	1.48	0.71	0.11	-0.58
Math	pv1math	522.34	98.42	-0.06	-0.27
	pv2math	522.05	98.19	-0.06	-0.29

Two mean scores were noticeably low. One item was 29q01 (0.96), which was an indicator of the attitude factor – '*I enjoy reading about mathematics'*; the other item was 35q03 (0.80), which was an indicator of the Subjective Norm Friend factor – '*my friends enjoy taking mathematics tests*.' Table 7 shows results of the separate descriptive statistics by each country. Again, the mean score of item 35q03 (*my friends enjoy taking mathematics tests*) was the low for all countries: USA (0.94), Germany (0.57), Japan (0.91), and Korea (0.72). Interestingly, the means for item 29q01 (*I enjoy reading about mathematics*) were low in Germany (0.79) and Japan (0.81).

Table 7. *Means and Standard Deviations for 23 Observed Variables by Country*

		USA		Germany	
Factors	**Indicators**	**Mean**	**SD**	**Mean**	**SD**
Attitude	29q01	1.19	0.83	**0.79**	0.82
	29q03	1.43	0.89	1.21	0.97
	29q04	1.27	0.93	1.25	1.04
	29q06	1.51	0.89	1.53	0.96
SN Friend	35q01	1.81	0.63	1.55	0.71
	35q02	1.77	0.65	1.37	0.73
	35q03	**0.94**	0.68	**0.57**	0.69
SN Parent	35q04	2.36	0.62	2.38	0.71
	35q05	2.21	0.72	2.22	0.81
	35q06	1.59	0.81	1.67	0.80
SN Teacher	77q02	2.24	0.86	1.98	0.98
	77q04	2.38	0.80	1.64	1.02
	83q02	2.28	0.74	1.80	0.97
	83q03	2.32	0.70	1.73	0.92
Perceived Control	43q01	2.46	0.60	2.42	0.69
	43q02	2.22	0.74	2.26	0.83
	43q05	2.26	0.72	2.20	0.83
Intention	46q01	2.12	0.76	1.88	0.95
	46q02	1.20	0.74	1.59	0.87
	46q03	1.91	0.76	1.92	0.84
	46q04	1.59	0.82	1.52	0.91
Math	PV1	481.03	89.18	513.93	96.74
	PV2	481.39	89.20	513.79	96.26

Table 7. *Continued.*

Factors	Variable	Japan		Korea	
		Mean	**SD**	**Mean**	**SD**
Attitude	29q01	**0.81**	0.81	1.07	0.82
	29q03	1.19	0.90	0.95	0.79
	29q04	1.14	0.92	1.12	0.89
	29q06	1.27	0.93	1.39	0.91
SN	35q01	1.31	0.66	1.59	0.66
Friend	35q02	1.51	0.70	1.80	0.67
	35q03	**0.91**	0.67	**0.72**	0.62
SN	35q04	1.82	0.83	2.16	0.75
Parent	35q05	1.58	0.85	1.95	0.80
	35q06	1.21	0.80	1.27	0.72
SN	77q02	2.03	0.85	1.90	0.82
Teacher	77q04	2.09	0.81	2.10	0.79
	83q02	2.08	0.76	1.97	0.71
	83q03	2.10	0.72	2.06	0.65
Perceived	43q01	2.17	0.78	2.17	0.69
Control	43q02	2.42	0.67	2.26	0.66
	43q05	1.89	0.82	2.12	0.71
Intention	46q01	1.65	0.90	1.46	0.77
	46q02	1.73	0.83	1.52	0.79
	46q03	1.91	0.86	1.34	0.81
	46q04	1.45	0.84	1.37	0.77
Math	PV1	536.06	94.15	554.23	98.77
	PV2	535.73	93.92	553.21	99.04

Note. SN = Subjective Norm, PV = plausible values in math. Bolded values indicate the low mean scores – below "1.00"

TPB Measurement Models

Confirmatory Factor Analyses

Before CFAs, exploratory factor analyses (EFAs) were conducted to detect any unusual factor loading patterns. EFAs help researchers to explore the observed variables "without imposing an a priori structure on the dimensions…" (Widaman, 2012, p. 362). Once certain items were identified as unsuitable, those items were dropped. For example, nine measured items (indicators) were available that appeared to represent math behavior intention (i.e., intention), but five items were eliminated due to either loadings that were close to zero or loadings onto multiple factors, thus only four items were kept for confirmatory factor analyses. Having identified potentially good items, CFAs were conducted to assess whether the selected items (indicators) were well representing the designated latent factors. Detailed discussions on the differences between CFAs and EFAs are available (Widaman, 2012).

According to the results of CFAs, all factor loadings were above .45 and all of them were statistically significant. The fit indexes also indicated that the model was a reasonably good fit; χ^2 (209) = 10,729.42, $p < .001$, *RMSEA* = .049 (90% CI, .048 - .049), *CFI* = .937, and *TLI* = .924, which suggests that each latent factor (i.e., construct) consisted of well-corresponding measured items. Simply put, Items 29q01, 29q03, 29q04, and 29q06 had strong loadings on the Attitude factor. Items 43q01, 43q02, and 43q05 were good indicators of the Perceived Control factor, and three items (43q03, 43q04, and 43q06) did not show high loadings on the Perceived Control factor, and they were dropped. The three deleted items were 43q03 (*Family demands or other prevent me from*

putting a lot of time into my mathematics), 43q04 (*If I had different teachers, I would try harder in mathematics*), and 43q06 (*I do badly in mathematics whether or not I study for my exam*). Table 8 shows the results of the CFA for the entire sample from 4 countries based on the reduced number of items for each factor.

Table 8. *Standardized Factor Ladings of 23 Indicators for the Entire Sample*

Latent factors	**Measured Items**	**Standardized Factor loadings**	
Attitude	29q01	0.722	(0.005) ***
	29q03	0.852	(0.003) ***
	29q04	0.906	(0.002) ***
	29q06	0.843	(0.003) ***
SN-Friend	35q01	0.736	(0.008) ***
	35q02	0.647	(0.008) ***
	35q03	0.487	(0.009) ***
SN-Parent	35q04	0.834	(0.005) ***
	35q05	0.860	(0.005) ***
	35q06	0.488	(0.007) ***
SN-Teacher	77q02	0.697	(0.006) ***
	77q04	0.736	(0.005) ***
	83q02	0.808	(0.004) ***
	83q03	0.823	(0.004) ***
control	43q01	0.841	(0.006) ***
	43q02	0.616	(0.007) ***
	43q05	0.655	(0.007) ***
intention	46q01	0.691	(0.006) ***
	46q02	0.827	(0.005) ***
	46q03	0.657	(0.006) ***
	46q04	0.648	(0.006) ***
math	PV1	0.975	(0.005) ***
	PV2	0.962	(0.005) ***

Note. *Standard error in parentheses. *p < .05, **p < .01, ***p < .001.*

Although all 23 indicators had high loadings and all were statistically significant, I decided to reduce the number of indicators for each latent factor in further data analyses so these subsequent analyses would be more efficient. To reduce the number of indicators, I formed parcels of items for each factor for the practical measurement purpose. "A parcel can be defined as an aggregate-level indicator comprised of the sum (or average) of two or more items, response, or behaviors" (Little, Cunningham, Shahar, & Widaman, 2002, p. 152).

For example, for the Attitude factor, 29q01 and 29q04 together were summed to form an *att1* manifest variable, while 29q03 and 29q06 were summed to form an *att2* manifest variable. Thus, the Attitude factor now had two indicators (*att1* and *att2)* instead of four indicators (29q01, 29q03, 29q04, and 29q06). I did this for each latent factor, so that each factor had only two manifest indicators. Thus, instead of 92 observed variables in total (23 x 4 countries), now 56 observed variables (14 x 4 countries) were used for further analyses.

To assess whether analyses using 14 manifest variables would fit as well as or better than the analyses using 23 manifest variables, CFAs were conducted. Table 9 shows the results of the CFA analysis for the entire sample ($N = 21363$). All factor loadings were high from .59 to .98, and they were all statistically significant. The fit indexes also indicated that the 14-indicator model had excellent fit to the data, χ^2 (56) = 1015, $p < .001$, *RMSEA* = .028 (90% CI, .027 - .030), *CFI* = .991, and *TLI* = .986.

Table 9. *Standardized Factor Loadings for 14 Indicators of the Entire Sample*

Constructs	Indicators	Standardized factor loadings
Attitude	Att1	.877 (.004) ***
	Att2	.955 (.004) ***
SN-Friend	Frnd1	.700 (.012) ***
	Frnd2	.691 (.012) ***
SN-Parent	Prnt1	.875 (.007) ***
	Prnt2	.773 (.007) ***
SN-Teacher	Tea1	.788 (.012) ***
	Tea2	.793 (.012) ***
Control	Con1	.898 (.010) ***
	Con2	.590 (.008) ***
Intention	Int1	.785 (.006) ***
	Int2	.939 (.006) ***
Math	pv1	.975 (.006) ***
	pv2	.962 (.006) ***

Note. *Standard error in parentheses.* $*p < .05$, $**p < .01$, $***p < .001$. χ^2 (56) = 1015, $p < .001$, *RMSEA* = .028 (.027 - .030), *CFI* = .991, and *TLI* = .986.

The CFA showing that all 14 indicators had high loadings was based on the correlation among the 14 indicators, which are presented in Table 10 for the entire sample from four countries (N = 21363).

Certain correlational patterns in Table 10 should be noted. For example, Friend item #1 (r = -.02), teacher item #1 (r = .03), teacher item #2 (r = .05), and intention item #1 (r = .05 and .04) were weakly correlated with both plausible values (pv1 and pv2) - the outcome variables (i.e., math achievement).

Table 10. *Estimated Correlations among 14 Indicators for the Entire Sample*

	1	2	3	4	5	6	7
1. att1	1						
2. att2	.84	1					
3. frnd1	.17	.19	1				
4. frnd2	.18	.18	.48	1			
5. prnt1	.32	.35	.24	.23	1		
6. prnt2	.26	.30	.22	.20	.68	1	
7. tea1	.21	.27	.18	.21	.15	.12	1
8. tea2	.22	.28	.19	.19	.15	.12	.63
9. con1	.38	.41	.15	.11	.32	.30	.19
10. con2	.23	.27	.08	.09	.17	.18	.19
11. int1	.33	.37	.16	.15	.25	.20	.20
12. int2	.40	.45	.13	.16	.25	.19	.24
13. Pv1	.25	.21	**-.02**	.12	.11	.09	**.03**
14. Pv2	.24	.20	**-.02**	.11	.10	.08	**.03**

	8	9	10	11	12	13	14
8. tea2	1						
9. con1	.21	1					
10. con2	.19	.53	1				
11. int1	.20	.22	.15	1			
12. int2	.24	.28	.21	.74	1		
13. Pv1	**.05**	.19	.15	**.05**	.13	1	
14. Pv2	**.05**	.19	.15	**.04**	.12	.94	1

Note. Bolded values are close to zero. att = attitude, frnd = subjective norm friend, prnt = subjective norm parent, tea = subjective norm teacher, con = perceived control, int = intentional math behavior, pv = plausible values in mathematics. All are statistically significant.

Testing Measurement Invariance of the TPB Model

Having identified that the 14-indicator model provided very close fit to data from the entire sample of four countries, the next step was to assess whether the measurement

model for the 14 indicators was equivalent across countries (i.e., USA, Germany, Japan, and Korea). As mentioned earlier, the sample used for the present research was large (N = 21363). Thus, before evaluating measurement invariance/equivalence across groups, the initial data from the four nations (N = 21363) was divided into 2 sets of data. The first half sample (N = 10681) was used for testing measurement invariance in multi-group analyses and the second half sample (N = 10682) was employed to replicate results across a new, independent sample. As discussed below, the attempted replication yielded similar results to those from the first half sample. The first half sample from each country was USA (N = 2489), Germany (N = 2500), Japan (N = 3176), and Korea (N = 2516). Using the first half sample from four countries with the 14-indicator model (N = 10681), a series of factorial invariance tests were conducted, based on the work of Widaman and Reise (1997).

Table 11 presents the results of the sequential measurement invariance testing in order to establish the most optimal model. The first step, configural invariance testing, evaluated if the pattern of the 14-indicator factor model was equivalent across the four national groups; the pattern was fixed, whereas the parameter estimates were allowed to vary from country to country. The configural invariance model was used as a baseline for model comparisons. Model 1 (configural invariance) had very good fit to data as the fit indexes indicated; χ^2 (224) = 749.92, p < .001, *RMSEA* = .030 (90% CI, .027 - .032), *CFI* = .990, and *TLI* = .984, which means that the pattern of fixed and free loadings in the factor structure of the TPB model components was invariant across groups. Stated differently, constructs (latent factors) had been measured by the same items/indicators

across four national groups (i.e., patterns of indicator-latent factor relations were equivalent).

Table 11. *Summary of Measurement Invariance Testing for the Derivation Sample*

Model	χ^2	*df*	RMSEA (CI)	CFI	TLI
Model 1	749.92	224	.030 (.027 - .032)	.990	.984
Model 2	864.74	245	.031 (.029 - .033)	.989	.983
Model 3	2310.36	266	.054 (.052 - .056)	.962	.948
Model 3a	1869.16	265	.048 (.046 - .050)	.970	.959
Model 3b	1672.12	264	.045 (.043 - .047)	.974	.964
Model 3c	1489.78	263	.042 (.040 - .044)	.977	.969
Model 4	1092.40	262	.034 (.032 - .037)	.985	.979

Note. Model 1 = Configural invariance (same pattern across groups); Model 2 = Weak invariance (same factor loadings across groups); Model 3 = Strong invariance (same factor loadings and intercepts across groups); Model 3a = partial strong invariance (relaxing one intercept, con2, on Japan); Model 3b = partial strong invariance (relaxing two intercepts - att1on Korea and con2 on Japan); Model 3c = partial strong invariance (relaxing three intercepts - int2 on Japan, att1 on Korea, and con2 on Japan); Model 4 = partial strong invariance (relaxing four intercepts - frnd1on USA and Germany, int2 on Japan, att1 on Korea, and con2 on Japan). RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index.

When factor loadings were constrained to be equal across groups, the result for Model 2 (Weak invariance) did not differ much from the baseline model (Configural invariance) in terms of model fit significance. In other words, the fit of Model 2 was also very good, χ^2 (245) = 864.74, $p < .001$, *RMSEA* = .031 (90% CI, .029 - .033), *CFI* = .989, and *TLI* = .983. Although the difference in fit between Model 2 and Model 1 was statistically significant, $\Delta\chi^2$ (21) = 114.82, $p < .001$, all practical fit indices were virtually unchanged, supporting acceptance of the more restricted weak factorial invariance over Model1. Having attained Weak (Metric) invariance across groups means that the

constructs of the TPB model components (i.e., Attitude, SN-Friend, SN-Parent, SN-Teacher, Perceived Control, and Intention) carried reasonably the same meaning to 15-year old students in all four different countries - USA, Germany, Japan, and Korea.

However, when indicator (i.e., item) intercepts were constrained to be equal across groups, the fit of the resulting Model 3 (strong invariance) to the data was not good, χ^2 (266) = 2310.36, $p < .001$, *RMSEA* = .054 (90% CI, .052 - .056), *CFI* = .962, and *TLI* = .948. The chi-square increased significantly from Model 2 to Model 3, $\Delta\chi^2$ (21) = 1445.62, which suggests that some of the item intercepts were substantially different from country to country. Thus, several steps of partial strong invariance testing were conducted. The four models moving from Model 3a to Model 4 are the result of the sequential partial strong invariance testing. Having relaxed invariance constraints on 4 indicators (sn-friend#1 for USA and Germany, attitude#1 for Korea, and control#2 and intention#2 for Japan), Model 4 achieved relatively good fit to the data as the fit indexes indicated, χ^2 (262) = 1092, $p < .001$, *RMSEA* = .034 (90% CI, .032 - .037), *CFI* = .985, and *TLI* = .979.

Table 12 shows the results of the overall sequential multiple group analyses of the TPB measurement model for both the first half and second half samples.

Table 12. *Summary of an Overall Sequential TPB Measurement Model*

	The First Half Sample (Derivation, $N = 10681$)				
Model	χ^2	*df*	RMSEA (CI)	CFI	TLI
Model 1	749.92	224	.030 (.027 - .032)	.990	.984
Model 2	864.74	245	.031 (.029 - .033)	.989	.983
Model 3	2310.36	266	.054 (.052 - .056)	.962	.948
Model 4	1092.40	262	.034 (.032 - .037)	.985	.979
	The Second Half Sample (Cross-Validation, $N = 10682$)				
Model 1	761.57	224	.030 (.028 - .032)	.990	.984
Model 2	889.62	245	.031 (.029 - .034)	.988	.982
Model 3	2371.88	266	.054 (.052 - .056)	.961	.947
Model 4	1188.45	262	.036 (.034 - .039)	.983	.976

Note. Model 1 = Configural invariance (same pattern across groups); Model 2 = Weak invariance (same factor loadings across groups); Model 3 = Strong invariance (same factor loadings and intercepts across groups); Model 3a = partial strong invariance (relaxing one intercept, con2, on Japan); Model 3b = partial strong invariance (relaxing two intercepts - att1on Korea and con2 on Japan); Model 3c = partial strong invariance (relaxing three intercepts - int2 on Japan, att1 on Korea, and con2 on Japan); Model 4 = partial strong invariance (relaxing four intercepts - frnd1on USA and Germany, int2 on Japan, att1 on Korea, and con2 on Japan). RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index.

From both the derivation (the first half) and cross-validation (the second half) samples, the difference in fit between Model 2 and Model 1 was statistically significant, $\Delta\chi^2$ (21) = 114.82, and $p < .001$, all practical fit indices were very similar, supporting acceptance of the more restricted weak factorial invariance over Model1. the chi-square increased significantly from Model 1 (configural invariance) to Model 3 (strong invariance), $\Delta\chi^2$ (42) = 1560.44, and $\Delta\chi^2$ (42) = 1610.31, ps < .001, respectively, and the results of other fit indices indicated the fit of model 3 became worsened. Although the

change of the chi-square was statistically significant when moving from Model 1 (configural invariance) to Model 4 (partial strong invariance) in the two samples, $\Delta\chi^2$ (18) = 342.48, and $\Delta\chi^2$ (18) = 426.88, respectively, other fit indices were still good when moving from Model 1 to Model 4. Also, note that the chi-square decreased substantially and significantly for both the derivation and cross-validation samples, when moving from strong invariance model (Model 3) to the partial strong invariance - final optimal model (Model 4), $\Delta\chi^2$ (4) = 1217.96 and $\Delta\chi^2$ (4) = 1183.43, respectively. Lastly, improvements in the statistical index of fit moving from Model 3 to Model 4 were accompanied by improvements in all of the practical indices of fit.

Table 13 contains detailed results of the measurement invariance testing from Model 4 from the first half (derivation) sample, which shows that all item loadings were equal across national groups. The yellow highlighted estimates are item intercepts for which invariance constraints were relaxed (Friend#1 for USA and Germany [these two were constrained equal, but different than Japan and Korea], Control#2 and Intention#2 for Japan, and Attitude#1 for Korea).

Table 13. *Estimated Intercepts and Unstandardized Factor Loadings for Derivation Sample*

	Inter	Att	SNfnd	SNpnt	SNtch	PerCon	Int	Math
USA								
att1	2.39	1.42(.03)	0	0	0	0	0	0
att2	2.96	1.55(.03)	0	0	0	0	0	0
frnd1	1.82	0	0.39(.01)	0	0	0	0	0
frnd2	1.76	0	0.56(.02)	0	0	0	0	0
prnt1	3.93	0	0	0.92(.02)	0	0	0	0
prnt2	2.19	0	0	0.56(.02)	0	0	0	0
tea1	4.58	0	0	0	1.01(.03)	0	0	0
tea2	4.65	0	0	0	1.10(.04)	0	0	0
con1	4.70	0	0	0	0	1.12(.03)	0	0
con2	2.33	0	0	0	0	0.41(.01)	0	0
int1	3.70	0	0	0	0	0	0.99(.02)	0
int2	3.92	0	0	0	0	0	1.24(.03)	0
Pv1	481.30	0	0	0	0	0	0	87.77(1.35)
Pv2	481.55	0	0	0	0	0	0	86.32(1.32)
Germany								
att1	2.39	1.42(.03)	0	0	0	0	0	0
att2	2.96	1.55(.03)	0	0	0	0	0	0
frnd1	1.82	0	0.39(.01)	0	0	0	0	0
frnd2	1.76	0	0.56(.02)	0	0	0	0	0
prnt1	3.93	0	0	0.92(.02)	0	0	0	0
prnt2	2.19	0	0	0.56(.02)	0	0	0	0
tea1	4.58	0	0	0	1.01(.03)	0	0	0
tea2	4.65	0	0	0	1.10(.04)	0	0	0
con1	4.70	0	0	0	0	1.12(.03)	0	0
con2	2.33	0	0	0	0	0.41(.01)	0	0
int1	3.70	0	0	0	0	0	0.99(.02)	0
int2	3.92	0	0	0	0	0	1.24(.03)	0
Pv1	481.30	0	0	0	0	0	0	87.77(1.35)
Pv2	481.55	0	0	0	0	0	0	86.32(1.32)

	Inter	Att	SNfnd	SNpnt	SNtch	PerCon	Int	Math
Japan								
att1	2.39	1.42(.03)	0	0	0	0	0	0
att2	2.96	1.55(.03)	0	0	0	0	0	0
frnd1	1.52	0	0.39(.01)	0	0	0	0	0
frnd2	1.76	0	0.56(.02)	0	0	0	0	0
prnt1	3.93	0	0	0.92(.02)	0	0	0	0
prnt2	2.19	0	0	0.56(.02)	0	0	0	0
tea1	4.58	0	0	0	1.01(.03)	0	0	0
tea2	4.65	0	0	0	1.10(.04)	0	0	0
con1	4.70	0	0	0	0	1.12(.03)	0	0
con2	2.65	0	0	0	0	0.41(.01)	0	0
int1	3.70	0	0	0	0	0	0.99(.02)	0
int2	4.36	0	0	0	0	0	1.24(.03)	0
Pv1	481.30	0	0	0	0	0	0	87.77(1.35)
Pv2	481.55	0	0	0	0	0	0	86.32(1.32)
Korea								
att1	2.75	1.42(.03)	0	0	0	0	0	0
att2	2.96	1.55(.03)	0	0	0	0	0	0
frnd1	1.52	0	0.39(.01)	0	0	0	0	0
frnd2	1.76	0	0.56(.02)	0	0	0	0	0
prnt1	3.93	0	0	0.92(.02)	0	0	0	0
prnt2	2.19	0	0	0.56(.02)	0	0	0	0
tea1	4.58	0	0	0	1.01(.03)	0	0	0
tea2	4.65	0	0	0	1.10(.04)	0	0	0
con1	4.70	0	0	0	0	1.12(.03)	0	0
con2	2.33	0	0	0	0	0.41(.01)	0	0
int1	3.70	0	0	0	0	0	0.99(.02)	0
int2	3.92	0	0	0	0	0	1.24(.03)	0
Pv1	481.30	0	0	0	0	0	0	87.77(1.35)
Pv2	481.55	0	0	0	0	0	0	86.32(1.32)

Note. Inter = intercepts, Att = attitude, SNfnd = subjective norm friend, SNpnt = parent, SNtch = teacher, PerCon = perceived control, Int = intention, and Math = plausible values in mathematics.

Table 14 shows means and variances for latent factors from the final model of the measurement invariance testing when comparing the four countries. Means were set to "0" and variances were set to "1" for the USA sample, so that means and variances on latent variables for the remaining countries could be estimated.

Table 14. *Means and Variances on Latent Variables for the Most Optimal Model from the Partial Strong Invariance (Model 4) using the first half sample by Each Country*

	USA		**Germany**		**Japan**		**Korea**	
Factors	mean	Var	mean	Var	mean	Var	mean	Var
Attitude	0	1	-0.21	1.15	-0.34	1.07	-0.40	0.90
SN-Friend	0	1	-0.70	0.53	-0.50	1.14	0.10	1.18
SN-Parent	0	1	0.12	1.16	-1.05	1.45	-0.50	1.23
SN-Teacher	0	1	-0.99	1.87	-0.45	1.35	-0.58	0.79
Control	0	1	-0.12	1.28	-0.57	1.16	-0.35	1.15
Intention	0	1	-0.31	1.36	-0.59	1.13	-0.86	1.22
Math	0	1	0.38	1.16	0.63	1.08	0.83	1.18

Note. Math = Plausible Values in Mathematics in PISA; Var = Variance. All were statistically significant ($p < .05$).

For Germany, the scale mean scores of SN-Friend (-0.70) and SN-Teacher (-0.99) were noticeably different from USA, while SN-Parent (0.12) and control (-0.12) were similar to USA. For Japan, the scale mean score of SN-Parent (-1.05) was significantly different from USA. For Korea, the scale mean scores of SN-Teacher (-0.58) and intention (-0.86) were significantly different from USA, while SN-Friend (0.10) was close to USA. In terms of math achievement, the scale mean score in Korea (0.83) was significantly different USA, while the scale mean score of Germany (0.38) was rather closer to USA.

Table 15. *Correlations among Latent Factors from Final Strong Partial Invariance Model for the First Half (Derivation) Sample by Each Country*

	1	2	3	4	5	6	7
USA							
1. Att	1						
2. SN-F	.19	1					
3. SN-P	.53	.27	1				
4. SN-T	.33	.27	.29	1			
5. Con	.44	.17	.44	.31	1		
6. Int	.48	.22	.33	.34	.43	1	
7. Math	.17	**.07^{+}**	.15	.17	.27	.18	1
Germany							
1. Att	1						
2. SN-F	.32	1					
3. SN-P	.33	.27	1				
4. SN-T	.48	.29	.27	1			
5. Con	.58	.15	.39	.31	1		
6. Int	.48	.26	.21	.31	.15	1	
7. Math	.26	**.0^{+++}**	-.13	**-.12^{++}**	.27	**.09^{+}**	1
Japan							
1. Att	1						
2. SN-F	.22	1					
3. SN-P	.44	.52	1				
4. SN-T	.37	.33	.27	1			
5. Con	.46	**.11^{++}**	.41	.43	1		
6. Int	.43	.22	.36	.37	.30	1	
7. Math	.37	.32	.31	.17	.23	.18	1
Korea							
1. Att	1						
2. SN-F	.24	1					
3. SN-P	.43	.39	1				
4. SN-T	.24	.33	.28	1			
5. Con	.51	.28	.52	.30	1		
6. Int	.72	.30	.47	.39	.53	1	
7. Math	.44	.33	.49	.24	.44	.47	1

Note. $p < .05 = {}^{+}$; $p < .01 = {}^{++}$; $p > .05 = {}^{+++}$

TPB Structural Models

TPB Structural relations of the TPB model comparisons

The primary goal of the present research was to investigate the relations between non-cognitive attributes and academic achievement in mathematics across national groups (USA, Germany, Japan, and Korea) by utilizing Ajzen's (1991) TPB framework. As aforementioned, the TPB model structure (i.e., the conceptual model for the present study) includes seven constructs: attitude, subjective norm friend, subjective norm parent, subjective norm teacher, perceived control, intention, and behavior (i.e., math achievement). Thus, once measurement invariance for factor loadings and intercepts across groups was established (see Model 4 from Table 12), the next step was to examine the structural relations of the TPB model components across the four national groups simultaneously in order to answer the present study's research questions. The multi-group analyses allowed me to examine the similarities and differences between Western (USA and Germany) and Asian (Japan and Korea) countries concerning the influences of non-cognitive factors on math achievement.

First, based on the final partial strong invariance model (Model 4), several sequential structural model analyses were conducted. Figure10 provides a visual representation regarding the changes of the Chi-Square and the degrees of freedom for each model from no restrictions on regression weights across groups (Model A) to full constraints – same regression weights across groups (Model E). Table 16 provides the TPB structural model comparisons that were shown in Figure 10.

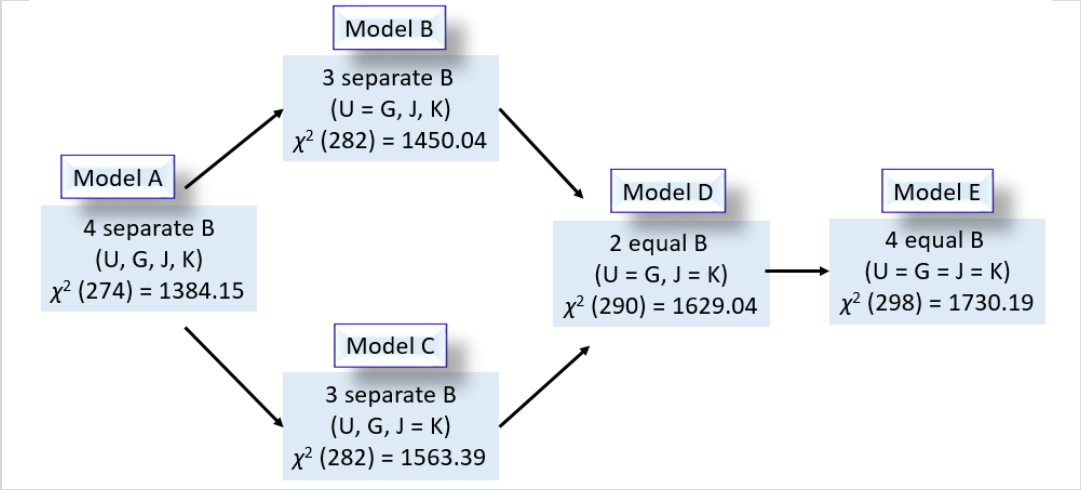


Figure 10. Structural model comparisons from all free (Model A) to all constrained (Model E)

Table 16. *Structural Model Comparisons from the First Half Sample (Derivation)*

Model	χ^2	*df*	RMSEA (CI)	CFI	TLI
Model A	1384.15	274	.039 (.037 - .041)	.979	.973
Model B	1450.04	282	.039 (.037 - .041)	.978	.972
Model C	1563.39	282	.041 (.039 - .043)	.976	.969
Model D	1629.04	290	.042 (.040 - .044)	.975	.969
Model E	1730.86	298	.042 (.041 - .044)	.973	.969

Note. Model A = Strong partial regression (regression weights were freely estimated across groups), Model B = regression weights were constrained in USA and Germany, Model C = regression weights were constrained only Japan and Korea, Model D = regression weights were constrained Western vs. Asian countries, Model E = regression weights were constrained across all four countries

Model A had no constraints on regression coefficients across nations and indicated good fit to the data, χ^2 (274) = 1384.15, $p < .001$, *RMSEA* = .039 (90% CI, .037 - .041), *CFI* = .979, and *TLI* = .973. Model B invoked equality constraints on all eight path coefficients across USA and Germany (but no restrictions on Japan and Korea), and the change in fit when moving from Model A to Model B was significant and fairly large, $\Delta\chi^2$ (8) = 65.89, $p < .001$. This suggests that relaxing invariance constraints on some of the regression weights for USA and Germany might improve model fit. Model C made

equality constraints on all eight path coefficients for Japan and Korea (but no restrictions on USA and Germany), and the change in model fit was significant and even larger, $\Delta\chi^2$ (8) = 179.24, $p < .001$). This also suggests that relaxing some equality constraints on regression weights for Japan and Korea might be needed to improve model fit. Model D included simultaneously the constrains employed in Models B and C, and the fit of model deteriorated substantially as expected, $\Delta\chi^2$ (16) = 244.89, $p < .001$. The most highly constrained model, Model E, showed that constraining to equality all eight regression coefficients across all four countries led to much poorer model fit. Specifically, the worsening in fit moving from Model A to Model E was substantial, $\Delta\chi^2$ (24) = 346. 71, $p < .001$.

Next, using Model A as a base model (with no restrictions on regression paths across countries), further sequential structural model analyses were conducted by applying systematically varying forms of constraint on regression coefficients across groups. I tested each regression weight separately for all eight paths (e.g., Attitude → Intention, SN-Friend → Intention). For the Attitude → Intention path, I compared the change in model fit under full restriction (Model A1, USA = Germany = Japan = Korea) and relaxing a specific constraint in Korea (Model A1a, USA = Germany = Japan, Korea). When the fit of those two models (Model A1 and Model A1a) is examined (see Table 16), model fit improved dramatically moving from Model A1 to Model A1a, $\Delta\chi^2$ (1) = 108.43, $p < .001$, which indicated that Model A1a exhibited much better fit to the data. In addition, Model A1a now did not differ significantly from the fit of Model A1,

$\Delta\chi^2$ (2) = 2.86, *ns*, even with the equality constraint across three of the countries (USA = Germany = Japan).

On the other hand, for the SN-Friend → Intention path, the change in model fit when moving from Model A to Model A2, was nonsignificant, $\Delta\chi^2$ (3) = 7.62, *ns*. This indicated that Model A2 was preferable to Model A because of its lack of substantial difference in fit and greater parsimony, with an equality constraint on the SN-Friend → Intention path coefficients across all four countries.

I followed the same procedure in testing the remaining six path coefficients, and results of these analyses are shown in Table 17. That is, I first invoked equality constraints on a given path coefficient across all four countries. If this constraint led to negligible worsening of model fit, I accepted the constraint across all four countries. But, if model fit worsened significantly, I relaxed the constraint that would most improve model fit.

In brief, the findings for the remaining six path coefficients were as follows: For both the SN-Parent → Intention and the SN-Teacher → Intention paths (see Models A3 and A4), equality constraints across all four countries led to such small differences in fit that the a priori equality constraints across all four countries could be accepted. For the Control → Intention path (see Model A5), model fit worsened significantly relative to Model A. After relaxing a single constraint (allowing the USA coefficient to differ from Germany = Japan = Korea), the resulting Model A5a showed much improved fit and returned model fit to approximately the level of Model A. Similar "single constraint relaxations" were needed for the remaining three paths. For the Attitude → Math path,

allowing the constrained coefficients for Western countries (USA = Germany) to differ from the constrained coefficients for East Asian countries (Japan = Korea) enabled fit of the resulting Model A6a to lessen the misfit associated with full constraints across all four countries in Model A6. For the Intention → Math path, allowing the path coefficient for Korea to differ from the constrained path across the other three countries (USA = Germany = Japan) in Model A7a improved model fit over Model A7. Finally, for the Control → Math path, allowing the path coefficient for Japan to differ from the constrained estimate across the other three countries (USA = Germany = Korea) in Model A8a showed much improved fit over that for Model A8, which had full constraints across all four countries.

After testing separately each of the eight regression paths of the TPB structural relations across groups, I formed a composite model, termed Model F. Model F contained the sets of constraints that were found to be optimal in Models A1 through A8, as shown in Table 16. That is, if full constraints across all four countries were associated with adequate fit for a given path, these full constraints were invoked in Model F (such as for Models A2, A3, and A4). Or, if certain constraints had to be relaxed (as in Models A1a, A5a, A6a, A7a, and A8a), then these specific constraint relaxations were made in Model F. Although the change of Chi-Square was statistically significant when moving from Model A to Model F, $\Delta\chi^2$ (19) = 44.73, $p < .001$, other fit indices were virtually the same or even slightly better for Model F, supporting acceptance of the more restricted and therefore more parsimonious model (Model F) over Model A:

Model A (32 estimates): *RMSEA* = .039 (90% CI, .037 - .041), *CFI* = .979, *TLI* = .973

Model F (13 estimates): *RMSEA* = .038 (90% CI, .036 - .040), *CFI* = .979, *TLI* = .974. Furthermore, Model F relaxed only five constraints in Model E, yet had much improved model fit, $\Delta\chi^2$ (5) = 301.98, $p < .001$, and Model E had noticeably worse fit than did either Model A or Model F:

Model E (8 estimates): *RMSEA* = .042 (90% CI, .041 - .044), *CFI* = .973, *TLI* = .969.

Table 17 provides the detailed sequential comparison analyses from Model A to Model A8a. Table 17 provides an overall summary of the model comparisons -- Model A with no restriction, Model E with complete restrictions across all four countries, and Model F (which relaxed five equality constraints), which turned out to be the optimal model to evaluate the multiple group analyses for the TPB structural relations across all four countries.

Table 17. *Sequential Analyses of Each Regression Path of the TPB Structural Models*

Model	description	χ^2	*df*	RMSEA (CI)	CFI	TLI
Model A (32 estimates)	Baseline: all free (U, G, J, K)	1384.15	274	.039 (.037 - .041)	.979	.973
Model A1 (29 estimates)	Att → Int (U=G=J=K)	1495.44	277	.041 (.039 - .043)	.977	.970
Model A1a (30 estimates)	Att → Int (U=G=J, K)	1387.01	276	.038 (.036 - .040)	.979	.973
Model A2 (29 estimates)	SNf → Int (U=G=J=K)	1391.77	277	.039 (.037 - .041)	.979	.973
Model A3 (29 estimates)	SNp → Int (U=G=J=K)	1391.51	277	.039 (.037 - .041)	.979	.973
Model A4 (29 estimates)	SNt → Int (U=G=J=K)	1388.72	277	.039 (.037 - .041)	.979	.973
Model A5 (29 estimates)	Con→ Int (U=G=J=K)	1420.98	277	.039 (.037 - .041)	.979	.972
Model A5a (30 estimates)	Con→ Int (U, G=J=K)	1391.91	276	.039 (.037 - .041)	.979	.973
Model A6 (29 estimates)	Att → Math (U=G=J=K)	1424.95	277	.039 (.037 - .041)	.979	.972
Model A6a (30 estimates)	Att→ Math (U=G, J=K)	1391.82	276	.039 (.037 - .041)	.979	.973
Model A7 (29 estimates)	Int → Math (U= G=J=K)	1400.67	277	.039 (.037 - .041)	.979	.973
Model A7a (30 estimates)	Int → Math (U=G=J, K)	1388.36	276	.039 (.037 - .041)	.979	.973
Model A8 (29 estimates)	Con →Math (U=G=J=K)	1409.47	277	.039 (.037 - .041)	.979	.972
Model A8a (30 estimates)	Con → Math (U=G=K, J)	1392.32	276	.039 (.037 - .041)	.979	.973
Model F (13estimates)	Combined all significant models	1428.88	293	.038 (.036 - .040)	.979	.974

Note. Model F (final optimal model) = resulted from accumulative analyses by comparing significant testing differences ($\Delta\chi^2/df$) between models with restriction and relaxing a relevant constraint on each regression path.

Table 18. *Overall Sequential Structural Model Comparisons from both Derivation and Cross-Validation Samples*

	The First Half Sample (Derivation, $N = 10681$)				
Model	χ^2	*df*	RMSEA (CI)	CFI	TLI
Model A	1384.15	274	.039 (.037 - .041)	.979	.973
Model E	1730.86	298	.042 (.041 - .044)	.973	.969
Model F	1428.88	293	.038 (.036 - .040)	.979	.974
	The Second Half Sample (Cross-Validation, $N = 10682$)				
Model A	1422.12	274	.040 (.038 - .042)	.979	.972
Model E	1846.40	298	.044 (.042 - .046)	.972	.965
Model F	1529.88	293	.040 (.038 - .042)	.979	.972

Note. Model A = Base model, regression weights were freely estimated in all countries, Model E = regression weights were constrained across all four countries, Model F = final optimal model, five constraints on regression path are Att → Int, Con → Int, Att → Math, Int → Math, Con → Math.

For both the derivation (the first half) and cross-validation (the second half) samples, the Chi-Square increased significantly from Model A to Model E, $\Delta\chi^2$ (24) = 346.71, and $\Delta\chi^2$ (24) = 424.28, *p*s < .001, respectively. The change of the Chi-Square index of fit was statistically significant when moving from Model A to Model F in the two samples, $\Delta\chi^2$ (19) = 44.73, and $\Delta\chi^2$ (19) = 107.76, respectively, but other fit indices were virtually the same or slightly improved when moving from Model A to Model F. These results support the rather highly constrained Model F over the unconstrained Model A. Also, note that the Chi-Square decreased substantially and significantly for both the derivation and cross-validation samples, when moving from the fully restricted model (Model E) to the final optimal model (Model F), $\Delta\chi^2$ (5) = 301.98 and $\Delta\chi^2$ (5) = 316.52, respectively. Finally, improvements in the statistical index of fit moving from Model E to F were accompanied by improvements in all of the practical indices of fit.

Multiple-Group Analyses of the TPB Model Components

According to the TPB model, the Intention factor should be predicted by Attitude, Subjective Norm (friend, parent, and teacher), and Perceived Control factors, and Math Achievement should be predicted by Attitude, Intention, and Perceived Control factors. Based on the final optimal model (Model F), all regression weights by each country were reported in Table 18, which showed the results of the TPB Structural model analyses. Although the values of the regression coefficients from cross-validation analysis were not identical to the results from the first half sample's structural model analysis, the pattern of each latent factor's effects on Intention and on Math Achievement were quite consistent. In other words, the overall findings from the cross-validation analysis were very similar to those of the first half sample's TPB structural model analysis.

In terms of the estimated magnitude of effects on Math Intention, a substantial amount of the variation was explained by the TPB model components, although these amounts of explained variance varied considerably across countries from both the derivation and cross-validation samples: 33% and 37% in USA, 17% and 21% in Germany, 20% and 22% in Japan, and 51% and 53% in Korea. However, a significant cultural difference was found in Math Achievement explained variance. That is, predictors of math achievement accounted for only 6% and 5% of the variance in USA and in Germany, respectively, whereas the comparable estimates were 12% and 12% in Japan and 24% and 28% in Korea, respectively. In other words, a significant amount of the Math Achievement variation was explained by Attitude, Intention, and Perceived

Control factors in Asian countries (Japan and Korea), but much less in Western countries (USA and Germany).

Table 19. *Regression Analysis for the TPB Structural Relations and the Estimated Effect Size for Both the Derivation and the Cross-Validation Samples*

	Unstandardized Regression Coefficients			
	USA	**Germany**	**Japan**	**Korea**
Derivation Sample (the first half, *N* = 10681)				
Att → Int	.36 (.02)***	.36 (.02)***	.36 (.02)***	**.83** (.04)***
SN-Frnd → Int	.08 (.02)***	.08 (.02)***	.08 (.02)***	.08 (.02)***
SN-Prnt → Int	.07 (.02)***	.07 (.02)***	.07 (.02)***	.07 (.02)***
SN-Tech → Int	.14 (.03)***	.14 (.03)***	.14 (.03)***	.14 (.02)***
Con → Int	**.30** (.03)***	.09 (.02)***	.09 (.02)***	.09 (.02)***
Att → Math	.10 (.02)***	.10 (.02)***	.31 (.02)***	.31 (.02)***
Int → Math	.02 (.01)+	.02 (.01)+	.02 (.01)+	**.11** (.02)***
Con → Math	.18 (.02)***	.18 (.02)***	**.06** (.02)*	.18 (.02)***
	Estimated Effect Size			
Int R^2	.33	.17	.20	.51
Math R^2	.06	.06	.12	.24
Cross-Validation Sample (the Second half, *N* = 10682)				
Att → Int	.38 (.03)***	.38 (.03)***	.38 (.03)***	**.78** (.04)***
SN-Frnd → Int	.09 (.02)***	.09 (.02)***	.09 (.02)***	.09 (.02)***
SN-Prnt → Int	.07 (.02)***	.07 (.02)***	.07 (.02)***	.07 (.02)***
SN-Tech → Int	.14 (.03)***	.14 (.03)***	.14 (.03)***	.14 (.03)***
Con → Int	**.32** (.04)***	.14 (.02)***	.14 (.02)***	.14 (.02)***
Att → Math	.03 (.02)+	.03 (.02)+	.34 (.03)***	.34 (.03)***
Int → Math	.01 (.01)+	.01 (.01)+	.01 (.01)+	**.11** (.03)***
Con → Math	.21 (.02)***	.21 (.02)***	**.07** (.02)**	.21 (.02)***
	Estimated Effect Size			
Int R^2	.37	.21	.22	.53
Math R^2	.05	.05	.12	.28

Note. Standard error in parentheses. Att = Attitude, SN-Frnd = Subjective Norm Friend, SN-Prnt = Subjective Norm Parent, SN-Tech = Subjective Norm Teacher, Con = Perceived Control, Int = Intention. *$p < .05$, **$p < .01$, ***$p < .001$, +$p > .05$

In addition, the effect of the Attitude factor on Intention was significant in all countries from both the derivation and cross-validation samples, but the regression coefficient of Attitude → Intention in Korea was the strongest (β = 0.83 and 0.78), and smaller (β = 0.36 and 0.38), but invariant across the remaining three countries. In terms of the direct effect of Attitude on Math Achievement, cultural differences were found. That is, results of the derivation sample indicated that the regression weight of Attitude → Math Achievement was statistically significant across all four countries, but much greater in Asian countries (Japan and Korea, β = 0.31) than in Western Countries (USA and Germany, β = 0.10). Results of the cross-validation sample also yielded cultural differences between Western and Asian countries, meaning that the regression weight of Attitude → Math Achievement was much stronger in Japan and Korea (β = 0.34, $p <$ .001) than in USA and Germany (β = 0.03, p > .05, ns).

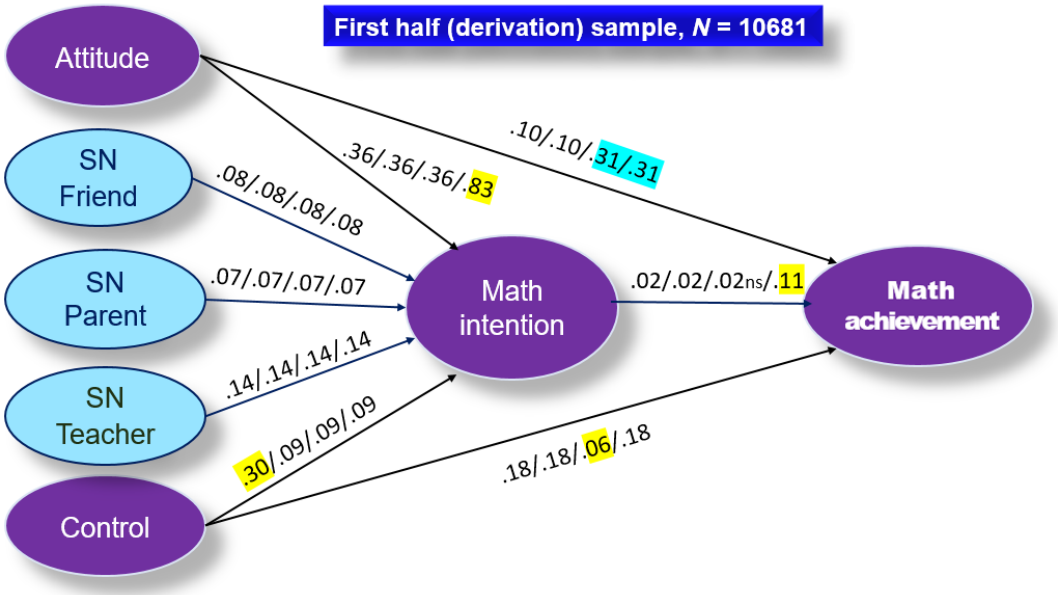
Next, the effect of the Perceived Control factor on Intention was significant in all countries, but the regression weight of Perceived Control → Intention in USA was the largest (β = 0.30 and 0.32), and smaller (β = 0.09 and 0.14), but invariant across the remaining three countries. However, the direct effect of Perceived Control on Math Achievement was statistically significant in USA, Germany, and Korea from both the derivation and cross-validation samples (β = 0.18 and 0.21, p < .001), but not in Japan (β = 0.06 and 0.07, p > .05).

On the other hand, the effect of Intention on Math Achievement was very small and not statistically significant in USA, Germany, and Japan except in Korea, from both the derivation and the cross-validation samples (β = 0.02 and 0.01, respectively, *p*s >

.05). In other words, the regression coefficient of Intention → Math Achievement in Korea from both the derivation and cross-validation samples was stronger than the other all three countries (two βs = 0.11, $p < .001$).

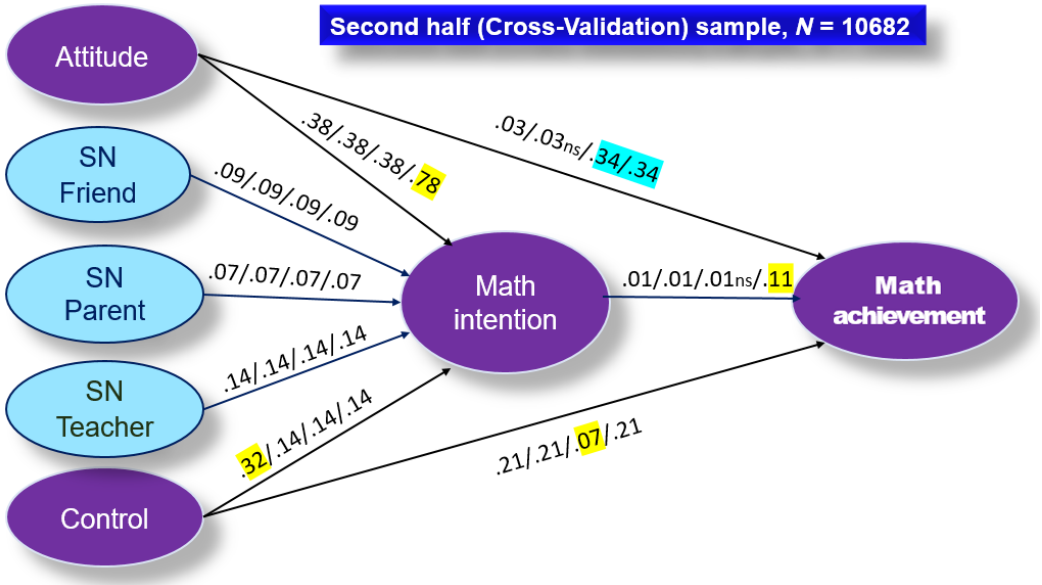
In terms of social influences on Math Intention, no cultural and national differences were found. The effect of the Subjective-Norm Friend factor on Intention was consistent across four countries. The regression weights of the SN-Friend factor → Intention from both the derivation and cross-validation samples were invariant across four countries (β = 0.08 and 0.09, respectively). The regression weights of the SN-Parent → Intention from both samples were also invariant across all four countries (four βs = 0.07). Although consistent across four countries, the effect of SN-Teacher on Intention was larger than the effect of the other two Subjective Norm factors (i.e., Friend and Parent factors) on Intention. That is, the regression weights of the SN-Teacher → Intention from both the derivation and cross-validation samples were consistent across four countries (four βs = 0.14).

Figure 11 shows a visual representation of the TPB structural model analyses conducted across the four national groups, which was based on the findings from Table 19. For graphical simplicity, correlations among predictor variables of the TPB model were not included in Figure 11. Instead, Table 20 shows the correlations among predictor variables of the TPB model components from the derivation and cross-validation samples by each country.



All regression coefficients were statistically significant ($p < .05$) except the Intention → Math achievement in USA, Germany, and Japan

---



All regression coefficients were statistically significant ($p < .05$) except the Attitude → Math Achievement in Western countries (USA and Germany) and the Intention → Math achievement in USA, Germany, and Japan

Figure 11. Estimated regression coefficients - USA, Germany, Japan, and Korea, respectively.

Table 20. *Correlations among Predictor Variables from both the First Half (derivation, N = 10681) and the Second Half (Cross-validation, N = 10682) Samples*

Correlated with		USA		Germany	
		First half Sample	Second half Sample	First half Sample	Second half Sample
SN-F	& Att	.20	.32	.33	.30
SN-P	& Att	.53	.47	.33	.41
	& SN-F	.28	.41	.27	.30
SN-T	& Att	.34	.39	.46	.53
	& SN-F	.28	.33	.32	.20
	& SN-P	.29	.31	.25	.28
Con	& Att	.42	.48	.58	.56
	& SN-F	.19	.29	.19	.16
	& SN-P	.42	.45	.39	.45
	& SN-T	.32	.51	.32	.31

Correlated with		Japan		Korea	
		First half Sample	Second half Sample	First half Sample	Second half Sample
SN-F	& Att	.23	.19	.26	.29
SN-P	& Att	.44	.52	.44	.43
	& SN-F	.53	.46	.40	.40
SN-T	& Att	.37	.35	.25	.34
	& SN-F	.35	.27	.34	.35
	& SN-P	.28	.32	.29	.33
Con	& Att	.48	.49	.57	.63
	& SN-F	.07	.09	.32	.34
	& SN-P	.42	.42	.56	.63
	& SN-T	.49	.53	.36	.44

Note. SN-F = Subjective Norm Friend, SN-P = Subjective Norm Parent, SN-T = Subjective Norm Teacher, Att = Attitude, Con = Perceived Control.

Although not identical, the patter of correlations among the predictor variables from both the derivation and cross-validation samples were similar. One noticeable correlation was found between Control and Subjective Norm Friend factor in Japan.

Research Questions

RQ #1: Do non-cognitive factors (i.e., the TPB model components) convey approximately the same meaning in four different countries – USA, Germany, Japan, and Korea?

According to the results of the multi-group measurement invariance testing (see Model 4 in Table 12), it seems that non-cognitive factors conveyed reasonably the same meaning in all countries for both the first half (derivation), χ^2 (262) = 1092.40, $p < .001$, *RMSEA* = .034 (90% CI, .032 - .037), *CFI* = .985, and *TLI* = .979, and the second half (cross-validation) samples, χ^2 (262) = 1188.45, $p < .001$, *RMSEA* = .036 (90% CI, .034 - .039), *CFI* = .983, and *TLI* = .976.

RQ #2: Would the structure of the extended TPB model of the present study be supported by the 2012 PISA data across the four national groups?

According to the results of the multi-group analyses, the extended TPB structural model was well supported by the 2012 PISA data across all four countries. The results of the first half sample (derivation) indicated the model had very close fit to the data; χ^2 (293) = 1428.88, $p < .001$, *RMSEA* = .038 (90% CI, .036 - .040), *CFI* = .979, and *TLI* = .974. Additionally, the results of the second half sample (cross-validation) of four countries from the 2012 PISA data revealed that the TPB model also had close fit to the data, χ^2 (293) = 1529.88, $p < .001$, *RMSEA* = .040 (90% CI, .038 - .042), *CFI* = .979, and *TLI* = .972. Meanwhile, it is important to note that the results of the present study showed the significant direct effect of Attitude on Math Achievement, especially in Japan and Korea. This finding is inconsistent with the study by Lipnevich et al. (2011) in terms of

the original TPB structural model applicability (i.e., the direct effect of Attitude on Math achievement was not tested in the Lipnevich et al.'s research).

RQ #3: Would a substantial amount of variance in math behavior Intentions be explained by non-cognitive factors (i.e., TPB model components) across national groups?

Based on Ajzen's (1991) TPB model framework, the conceptual model of the current study consisted of Attitude, Perceived Control, and 3 distinctive Subjective Norm factors (i.e., friend, parent, and teacher) that should predict math behavior Intentions (i.e., intention). According to the results of the multi-group analyses on the TPB structural relations ($1^{st}/2^{nd}$ sample), variance in Intention accounted for by Attitude, SN-Friend, SN-Parent, SN-Teacher, and Control factors was substantial in all countries, 33% and 37% in USA, 17% and 21% in Germany, 20% and 22% in Japan, and 51% and 53% in Korea; but were notably greater in Korea than in the rest of the three countries.

RQ #4: Are cultural differences evident in the magnitude of effects of each Subjective Norm factor on Intention to work hard on mathematics between Western and Asian countries?

No cultural differences were found in terms of the magnitude of effects of each Subjective Norm factor on Intention. The effect of Subjective Norm Friend factor on Intention was consistent across all four countries. The TPB structural model analyses revealed that the regression coefficients ($1^{st}/2^{nd}$ sample) of Subjective Norm Friend → Intention were statistically significant for USA, Germany, Japan, and Korea, β = 0.08/0.09, *ps* < .05. Although not large, the Subjective Norm Parent factor also had a positive effect on Intention across all four nations. From the multi-group structural model

analyses, the regression coefficients ($1^{st}/2^{nd}$ sample) of Subjective Norm Parent → Intention were statistically significant and invariant across all four countries, β = 0.07/0.07, *ps* < .05. Furthermore, the effect of the Subjective Norm Teacher factor also had a positive effect on Intention across all four countries as the regression weights ($1^{st}/2^{nd}$ sample) of Subjective Norm Teacher → Intention showed, β = 0.14/0.14, *p*s < .05.

RQ #5: Would a substantial amount of the variance in Math Achievement be explained by Attitude, Intention, and Perceived Control across national groups?

The aim of the fifth research question was to test the direct effects of Attitude and Perceived Control on Math Achievement along with an effect of Intention on Math Achievement. According to the results of the multi-group TPB structural analyses ($1^{st}/2^{nd}$ sample), variance in Math Achievement explained by the three predictors was 6% and 5% both in USA and Germany, 12% and 12% in Japan, and 24% and 28% in Korea. That is, the amount of variability of Math Achievement (which was explained by Attitude, Intention, and Perceived Control) was much weaker in Western countries (5% - 6%) than in Asian countries (12% - 28%).

CHAPTER FIVE

DISCUSSION

Overview

Considerable empirical evidence has supported the applicability of the TPB model in various behavioral domains and disciplines (Armitage & Conner, 2001; Lee et al., 2010; Sheeran, 2001). However, there is limited research on utilizing the TPB model in educational contexts aside from one cross-cultural study by Lipnevich et al. (2011). To address the need for investigating the applicability of the TPB model in academic domains, the current cross-national comparison study focused on investigating how non-cognitive attributes, the TPB model components, affect students' math intention and math achievement in the USA and their peers in three other countries, Germany, Japan, and Korea.

One of the key findings of the current research was that the extended TPB model was successfully applied to the internationally representative large-scale samples from the 2012 PISA data across all four countries: USA, Germany, Japan, and Korea. That is, the results of the multiple-group analyses supported the conclusion that the measurement model for the TPB constructs was invariant/equivalent across the four national groups. Fit indices indicated that the partial-strong factorial invariance for the TPB model provided close fit to the data for the first half sample (derivation, N = 10681), χ^2 (262) = 1092.40, $p < .001$, *RMSEA* = .034 (90% CI, .032 - .037), *CFI* = .985, and *TLI* = .979, and also for the second half sample (cross-validation, N = 10682), χ^2 (262) = 1188.45, p

< .001, *RMSEA* = .036 (90% CI, .034 - .039), *CFI* = .983, and *TLI* = .976.

Another major finding of the present cross-national study was that a substantial amount of the variation in Math Intentions was explained by the TPB model components, although these amounts of explained variance varied considerably across countries from both the derivation and cross-validation samples: 33% and 37% in USA, 17% and 21% in Germany, 20% and 22% in Japan, and 51% and 53% in Korea. However, a significant cultural difference was found in Math Achievement explained variance. That is, predictors of math achievement accounted for only 6% and 5% of the variance in the USA and in Germany, respectively, whereas the comparable estimates were 12% and 12% in Japan and 24% and 28% in Korea, respectively. In what follows, the relations of each non-cognitive factor of the TPB model components with Math Intention and Math Achievement across the four countries are discussed.

Cross-national Comparisons of Non-cognitive Factors

Attitude Factor

The Attitude factor affected Math Intention significantly in all four countries, USA, Germany, Japan, and Korea, but had a much stronger effect in Korea than the other three countries. The regression coefficients of Attitude → Math Intention from both the derivation and cross-validation samples were β = 0.36 and 0.38 (*ps* < .05), respectively, for USA, Germany, and Japan, and β = 0.83 and 0.78, respectively, in Korea. The findings of strong effects of Attitude on Intention across all four countries were consistent with prior TPB research findings (Armitage & Conner, 2001; Lipnevich et al., 2011). However, an explanation for the very large and much different estimate for the

Korean sample when compared with the much smaller, but still moderate to strong estimates for the remaining countries should be the topic of future research.

In most Attitude-Behavior relation studies, the direct effect of Attitude on Behavior outcomes was not tested because the proponents of the TPB model (e.g., Ajzen, 1991) have long argued that the relation of Attitude to Behavior is only indirect, through the effect of Attitude on Intention. For example, a recent cross-cultural study that examined the relations between math attitude and math achievement utilizing the TPB model, the effect of Attitude on Math achievement was not tested (Lipnevich et al., 2011).

However, in one study of attitude-behavior relations based on the Theory of Reasoned Action (TRA) model – a precursor model of TPB, Bentler and Speckart (1979) tested the direct effect of Attitude factor on Behavior outcomes. They reported that the regression weights of Attitude → Alcohol ($\beta = 0.47$), Attitude → Marijuana ($\beta = 0.57$), and Attitude → Hard drug consumption ($\beta = 0.48$) were all statistically significant (*ps* < .05). Therefore, the present study also examined how Attitude directly affected Math Achievement. The results of the TPB structural model analyses revealed that there were cultural differences between Western and Asian countries in terms of the direct effect of Attitude factor on Math Achievement. More specifically, the regression coefficients of Attitude → Math Achievement from both the derivation and cross-validation samples were $\beta = 0.10$ ($p < .001$) and $\beta = 0.03$ ($p > .05$, *ns*), respectively, in Western countries (USA and Germany) and $\beta = 0.33$ ($p < .001$) and $\beta = 0.34$ ($p < .001$), respectively, in the two samples in the Asian countries (Japan and Korea).

Subjective Norm Factors

No cultural differences were found in the magnitudes of effects of the three Subjective Norm factors (Friend, Parent, and Teacher factors) on Math Intention. First, the effect of SN-Friend on Math Intention was invariant across the four countries. Although not large, but statistically significant, the regression weights of SN-Friend → Math Intention from both the derivation and cross-validation samples were β = 0.08 and 0.09 (*ps* < .001), respectively, in USA, Germany, Japan, and Korea. Second, the effect of SN-Parent on Math Intention was small but also statistically significant across all four countries. The regression weights of SN-Parent → Math Intention from both the derivation and cross-cultural samples were βs = 0.07 and 0.07 (*ps* < .001), respectively, in USA, Germany, Japan, and Korea. Third, the effect of SN-Teacher on Math Intention was strongest among the three SN factors. That is, the regression weights of SN-Teacher → Math Intention from both the first half (derivation) and the second half (cross-validation) samples were βs = 0.14 and 0.14, respectively, (*ps* < .001) in USA, Germany, Japan, and Korea.

Perceived Control Factor

No cultural differences were found between Western and Asian countries in effects of the Perceived Control factor, but significant national differences were found in terms of the effects of Control factor on Math Intention and on Math Achievement. First, the regression coefficients of Control → Math Intention were significant in all four countries from both the derivation and cross-validation samples, β = 0.30 and 0.32,

respectively, in USA and β = 0.09 and 0.14, respectively, in Germany, Japan, and Korea (all *ps* < .001). Notably, the regression weights of Control on Math Intention were different and much stronger for the USA relative to estimates for the other three countries from both samples.

Second, the direct effects of Control on Math Achievement were also invariant across groups from both the derivation and cross-validation samples. Specifically, the regression weights of Control → Math Achievement were all statistically significant from both samples in USA, Germany, and Korea, β = 0.18 and 0.21 (*p*s < .001), respectively, but in Japan were much smaller from both samples; the first half sample, β = 0.06 (*p* < .05) and the second half sample, β = 0.07 (*p* < .01).

Summary

To summarize, the non-cognitive factors associated with the TPB model tended to have significant influences on both Math Intention and Math Achievement, but these effects varied in certain instances from country to country. Succinctly stated, the present cross-national investigation highlights as follow:

1. Two findings were consistent with "West vs. Asia" contrast; (a) the influence of Attitude on Math Achievement was much stronger for Japan and Korea than USA and Germany; (b) the estimated effect size of Math Achievement was much greater in Japan and Korea than USA and Germany.
2. Most other findings were contrasts showing that one Western country or one Asian country differed from the other three countries, so were inconsistent with a generic "West vs. Asia" cultural comparison model; (a) the influence of Control

on Math Intention in USA was distinctively greater from the other three countries; (b) the direct effect of Control on Math Achievement was strong in USA, Germany, and Korea, but weak in Japan; (c) lastly, the influence of Attitude on Math Intention and Math Intention on Math achievement in Korea were much stronger compared to the other three countries.

3. The reasons for the lack of full invariance of regression coefficients across countries are beyond the scope of the present study – and deserve future research to uncover why certain constructs had much stronger effects in certain countries than in other countries.
4. One key strength of the current investigation was the use of two random samples from each country (the derivation and cross-validation samples). Consistent with this strength, the differences in regression weights found across countries (such as the very large weight of Attitude to Math Intention for Korea relative to the other three countries) based on analyses of data from the derivation sample were replicated and therefore confirmed in analyses of data from the cross-validation sample. This gives considerable confidence that the findings represent true substantial differences across countries in effects of various model components, even though the basis for these differences cannot be uncovered from the current data.

Limitations

Because of the cross-sectional nature of the current samples, no causal claims should be made from the current research findings, although the term ‘effect’ was used

throughout the present study as a part of SEM terminology. It is quite possible that high math achievement may influence students' positive attitude toward math, so the direction of effect may, in fact, be the reverse of the direction suggested by TPB theory. For example, students in Japan and in Korea have much higher mean PISA scores in 2012 (536 and 554, respectively) than in USA and Germany (481 and 514, respectively).

Another limitation of the current research was the fact that more school-related indices of student achievement, such as students' math grades, were not included. Perhaps the Attitude factor and the SN-Teacher factor may have long-term positive effects on students' math learning outcomes with more contextual outcomes such as math course grades. Another limitation was that other non-cognitive factors (e.g., grit, perseverance, motivation, etc.) were not included because the present study was based on the Theory of Planned Behavior framework. It is quite possible that other non-cognitive factors might affect students' Math Intention more significantly.

What's Next?

The Attitude factor of the TPB model components was the strongest predictor of Intention in all four countries, which is consistent with prior findings from other behavioral studies that applied the TPB model. Also, the direct effect of Attitude on Math Achievement was much stronger in the high-performing Asian countries – Japan and Korea – than in the two Western countries. One implication of this finding concerns efforts to implement instructional strategies that foster students' self-efficacy and motivation to increase their positive attitudes toward math, which, in turn, could influence math engagement and achievement. Unfortunately, the smaller effects that

Attitude had on Math Achievement in the Western countries suggests that such work may lead to less impact on Math Achievement than would occur in Asian countries. On the other hand, implementing such strategies might lead to stronger relations between math attitudes and math achievement in Western countries, so the implementation of strategies might have effects on the estimated relations among the TPB model components.

Teacher education programs also should rigorously recruit teacher-candidates who are keenly aware of the important role that non-cognitive factors (e.g., attitude, self-efficacy, etc.) can play in students' math outcomes.

References

Aiken, L. R. (1970). Attitudes toward mathematics. *Review of Educational Research, 40*(4), 551-596.

Aiken, L. R. (1974). Two scales of attitude toward mathematics. *Journal for Research in Mathematics Education, 5*(2), 67-71.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes. 50*(2), 179–211.

Alloway, T. P. (2011). *Improving working memory: Supporting students' learning.* Thousand Oaks, CA: Sage Publications, Inc.

Allport, G.W. (1935). Attitudes. In C.A. Murchison (Ed.) *A handbook of social psychology*. Worcester, Mass: Clark University Press.

Ames, C. (1990). Motivation: What teachers need to know. *Teachers College Record, 91*(3), 409-421.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*(4), 471-499.

Asparouhov, T., & Muthe´n, B. (2010). *Plausible values for latent variables using Mplus: Technical report*. Los Angeles, CA: Muthèn and Muthèn.

Baddeley, A.D. (1992). Working memory. *Science Magazine, 255*, 556-559.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Becker, B. E., & Luthar, S. S. (2002). Social-emotional factors affecting achievement outcomes among disadvantaged students: Closing the achievement gap. *Educational Psychologist, 37*(4), 197-214.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246.

Bentler, P. M., & Speckart, G. (1979). Models of attitude–behavior relations. *Psychological Review, 86*(5), 452-464.

Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23*, 329-345.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.). *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

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.

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

Chun, K. T. (1980). The myth of Asian American success and its educational ramifications. *IRCD Bulletin*, *15*(1), 1-13.

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

Coleman, J. S., Campell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office.

Cragg, L., & Gilmore, C. (2013). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education, 3*(2), 63-68.

Crowley, S. L., & Fan, X. (1997). Structural equation modeling: Basic concepts and applications in personality assessment research. *Journal of Personality Assessment, 68*(3), 508-531.

Dewey, J. (1934). The need for a philosophy of education. From The new era in home and school. In R.D. Archambault (Ed.), *John Dewey on education: Selected writings* (pp. 3-14). NY: The Modern Library.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology, 92*(6), 1087 - 1101.

Dweck, C. (2002). Messages that motivate: How praise molds students' beliefs, motivation, and performance (in surprising ways). In J. Aronson (Ed.), *Improving*

academic achievement: *Impact of psychological factors on education* (pp. 37–60). San Diego, CA: Academic Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.

Eccles, J. S., & Jacobs, J. E. (1986). Social forces shape math attitudes and performance. *Signs, 11*(2), 367-380.

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

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*(2), 117-142.

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology, 82*, 221–234.

Fishbein, M. (1967). Attitude and the prediction of behavior. In M. Fishbein (Ed), *Readings in attitude theory and measurement* (pp. 477 - 492). New York: John Wiley & Sons, Inc.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Geary, D. C. (1996). International differences in mathematical achievement: Their nature, causes, and consequences. *Current Directions in Psychological Science, 5*, 133–137.

Goodenow, C. (1992). Strengthening the links between educational psychology and the study of social contexts. *Educational Psychologist, 27*(2), 177-196.

Hallinan, M. T. (2008). Teacher influences on students' attachment to school. *Sociology of Education, 81*(3), 271-283.

Hamre, B. K., & Pianta, R. C. (2001). Early teacher–child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development, 72*(2), 625-638.

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

Hernández, M. M., Robins, R. W., Widaman, K. F., & Conger, R. D. (2014). School belonging, generational status, and socioeconomic effects on Mexican-origin

children's later academic competence and expectations. Journal of Research on Adolescence, ***(*), 1-16.

Horn, J. L., McArdle, J. J., & Mason, R. (1983). When is invariance not invariant: A practical scientist's look at the ethereal concept of factor invariance. *Southern Psychologist, 1*, 179–188.

Hu, L. T., & 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.

Jerrim, J., & Choi, Á. (2014). The mathematics skills of school children: how does England compare to the high-performing East Asian jurisdictions?. *Journal of Education Policy, 29*(3), 349-376.

Jones, S. M., Bailey, R., & Jacob, R. (2014). Social-emotional learning is essential to classroom management. *Phi Delta Kappan, 96*(2), 19-24.

Joreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*, 409–426.

Kim, J. I., Schallert, D. L., & Kim, M. (2010). An integrative cultural view of achievement motivation: Parental and classroom predictors of children's goal orientations when learning mathematics in Korea. *Journal of Educational Psychology, 102*(2), 418 - 437.

Khine, M. S., Al-Mutawah, M., & Afari, E. (2015). Determinants of Affective Factors in Mathematics Achievement: Structural Equation Modeling Approach. *Journal of Studies in Education, 5*(2), 199-211.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford.

Lee, J., Cerreto, F. A., & Lee, J. (2010). Theory of planned behavior and teachers' decisions regarding use of educational technology. *Educational Technology & Society, 13*, 152-164.

Leung, F. K. (2002). Behind the high achievement of East Asian students. *Educational Research and Evaluation, 8*(1), 87-108.

Lindgren, H. C., Silva, I., Faraco, I., & Rocha, N. S. D. (1964). Attitudes toward problem solving as a function of success in arithmetic in Brazilian elementary schools. *The Journal of Educational Research, 58*(1), 44-45.

Linnenbrink, E. A., & Pintrich, P. R. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading and Writing Quarterly, 19*(2), 119-137.

Lipnevich, A. A., MacCann, C., Krumm, S., Burrus, J., & Roberts, R. D. (2011). Mathematics attitudes and mathematics outcomes of US and Belarusian middle school students. *Journal of Educational Psychology, 103*(1), 105-118.

Lipnevich, A. A., & Roberts, R. D. (2012). Noncognitive skills in education: Emerging research and applications in a variety of international contexts. *Learning and Individual Differences, 22*(2), 173-177.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151-173.

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.

Ma, X. & Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for Research in Mathematics Education, 28* (1), 26-47.

Madden, T. J., Ellen, P. S., & Ajzen, I. (1992). A comparison of the theory of planned behavior and the theory of reasoned action. *Personality and Social Psychology Bulletin, 18*(1), 3-9.

Majeed, A.A., Darmawan, G.N., & Lynch, P. (2013). A confirmatory factor analysis of the Attitudes Toward Mathematics Inventory (ATMI). *The Mathematics Educator, 15*, 121-135.

Maker, C. J. (1982). *Teaching models in education of the gifted*. Rockville, MD: Aspen system corp.

Mau, W. C. (1997). Parental influences on the high school students' academic achievement: A comparison of Asian immigrants, Asian Americans, and White Americans. *Psychology in the Schools, 34*(3), 267-277.

McLeod, D. (1992). Research on affect in mathematics education: a reconceptualization. In D. Grows (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp.575-596). New York: McMillan Publishing Company.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525-43.

Minato, S., & Yanase, S. (1984). On the relationship between students attitudes towards school mathematics and their levels of intelligence. *Educational Studies in Mathematics, 15*(3), 313-320.

Mislevy, R. J., & Sheehan, K. M. (1989). Information matrices in latent-variable models. *Journal of Educational Statistics, 14*(4), 335-350.

Monseur, C., & Adams, R. (2009). Plausible values: How to deal with their limitations. *Journal of Applied Measurement, 10*(3), 320-334.

Motlagh, S. E., Amrai, K., Yazdani, M. J., Abderahim, A. H., & Souri, H. (2011). The relationship between self-efficacy and academic achievement in high school Students. *Procedia-Social and Behavioral Sciences, 15*, 765-768.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30-38.

Muthén, L.K. and Muthén, B.O. (1998-2012). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Nachtigall, C., Kroehne, U., Funke, F., & Steyer, R. (2003). (Why) should we use SEM? Pros and cons of structural equation modeling. *Methods of Psychological Research Online, 8*(2), 1-22.

National Center for Education Statistics (2011). Mathematics achievement of fourth- and eighth-graders in 2011.
Retrieved from http://nces.ed.gov/timss/results11_math11.asp.

National Commission on Excellence in Education (NCEE). (1983). *A Nation at Risk: The imperative for educational reform: A report to the nation and the Secretary of Education, United States Department of Education.* Washington, DC: The Commission. Retrieved from https://www2.ed.gov/pubs/NatAtRisk/risk.html.

National Council of Teachers of Mathematics (NCTM). (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA.

National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.

Neale, D. (1969). The role of attitudes in learning mathematics. *The Arithmetic Teacher*, 631-641.

Noftle, E. E., Robins, R. W. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*(1), 116-130.

Organization for Economic Co-operation and Development (OECD) (2009). *PISA Data Analysis Manual: SAS 2nd ed*. Paris: OECD.

OECD (2013). *Ready to Learn*: *(Volume III): Students' Engagement, Drive and Self-Beliefs.* OECD Publishing.

OECD (2014a). *PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know*. OECD publications; Paris.

OECD (2014b). *PISA 2012 Results: Creative Problem Solving: Students' Skills in Tackling Real-Life Problems (Volume V*), PISA, OECD publishing. http://dx.doi.org/10.1787/9789264208070-en

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66(*4), 543-578.

Peng, S. S., & Wright, D. (1994). Explanation of academic achievement of Asian American students. *The Journal of Educational Research, 87*(6), 346-352.

Phelan, P., Davidson, A. L., & Cao, H. T. (1992). Speaking up: Students' perspectives on school. *Phi Delta Kappan, 73*(9), 695-704.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322-338.

Raykov, T., Marcoulides, G. A., & Li, C. H. (2012). Measurement invariance for latent constructs in multiple populations: A critical view and refocus. *Educational and Psychological Measurement, 72*(6), 954-974.

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

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology, 104*(3), 700-712.

Riebl, S. K., Estabrooks, P. A., Dunsmore, J. C., Savla, J., Frisard, M. I., Dietrich, A. M., ... & Davy, B. M. (2015). A systematic literature review and meta-analysis: The Theory of Planned Behavior's application to understand and predict nutrition-related behaviors in youth. *Eating behaviors, 18,* 160-178.

Rosenthal, R., & Rubin, D. B. (1982). A note on percent variance explained as a measure of the importance of effects. *Journal of Applied Social Psychology, 9,* 395-396.

Rubin, D.B. (1987). *Multiple imputations for non-response in surveys*. New York: Wiley.

Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18*(4), 210-222.

Schneider, B., & Lee, Y. (1990). A model for academic success: The school and home environment of East Asian students. *Anthropology & Education Quarterly, 21*(4), 358-377.

Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools,* 22(2), 208-223.

Schunk, D. H. (2012). *Learning theories: An educational perspective* (6th Ed.). Boston, MA: Pearson Education.

Sheppard, B., Hartwick, J., & Warshaw, P. (1988). The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research. *Journal of Consumer Research, 15*(3), 325-343.

Smith, T. M. (2004). Curricular reform in mathematics and science since a nation at risk. *Peabody Journal of Education, 79*, 105-129.

Sun, Y. (1998). The academic success of East-Asian–American students—An investment model. *Social Science Research, 27*(4), 432-456.

Suydam, M. N., & Weaver, J. F. (1975). Research on mathematics learning. In J. N. Payne (Ed.), *Mathematics learning in early childhood: Thirty-seventh yearbook* (pp. 44-67). Reston, VA: National Council of Teachers of Mathematics.

Tapia, M. (1996). *The attitudes toward mathematics Instrument*. Paper presented at the annual meeting of the Mid-south Educational Research Association, Tuscaloosa, AL (ERIC Reproduction Service No. ED 404165).

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

Teo, T., & Tan, L. (2012). The theory of planned behavior (TPB) and pre-service teachers' technology acceptance: A validation study using structural equation modeling. *Journal of Technology and Teacher Education, 20,* 89-104.

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review, 41*, 1-32.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1–10.

Ullman, J. B. (2006). Structural equation modeling: Reviewing the basics and moving forward. *Journal of Personality Assessment, 87*(1), 35-50.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.) Cambridge, MA: Harvard University Press.

Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90,* 202-209.

Widaman, K. F. (2006). III. Missing data: What to do with or without them. *Monographs of the Society for Research in Child Development, 71*, 42–64.

Widaman, K. F. (2012). Exploratory factor analysis and confirmatory factor analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 3: Data analysis and research publication* (pp. 361-389). Washington, DC: American Psychological Association.

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*, 10–18.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: methodological advances from alcohol and substance abuse research* (pp. 281-324). Washington, DC: American Psychological Association.

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods, 8*, 16-37.

Windle, M., Iwawaki, S., & Lerner, R. M. (1988). Cross-cultural comparability of temperament among Japanese and American preschool children. *International Journal of Psychology, 23*, 547-567.

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation, 31*, 114-128.

Zan, R., & Di Martino, P. (2007). Attitude toward mathematics: Overcoming the positive/negative dichotomy. *The Montana Mathematics Enthusiast, 3*, 157-168.

Zimmerman, B. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*, 82-91.

Zint, M. (2002). Comparing three attitude-behavior theories for predicting science teachers' intentions. *Journal of Research in Science Teaching, 39*(9), 819-844.