editorial UNIVERSIDAD DE ALMERÍA

Electronic Journal of Research in Educational Psychology
ejrep

UNIVERSIDAD DE ALMERÍA

Ilustre Colegio Oficial de Psicología de Andalucía Oriental

Exploring Latent Profiles of Stereotype Threat Susceptibility in U.S. and Colombian Students

Katherine Picho[1], Tatiana Rojas-Ospina[2], Adriana Maria Caicedo-Tamayo[2]

[1] Human Development and Psychoeducational Studies, Howard University, Washington DC
[2] Departamento de Ciencias Sociales, Pontificia Universidad Javeriana, Cali

United States of America

Colombia

Correspondence: Katherine Picho, Human Development and Psychoeducational Studies, Howard University, 2441 4th St NW, Washington DC 20059. United States of America. E-mail: katherine.picho@howard.edu

Abstract

Introduction: The present study investigated the theoretical Stereotyping Threat-susceptibility groups proposed by Steele (1997) by using a latent class analysis.

Method: 413 undergraduate students from the U.S and Colombia, majoring in various Science Technology Engineering and Math (STEM) and non-STEM disciplines completed a stereotype threat susceptibility measure-- the Social Identities and Attitudes Scale, SIAS (Picho & Brown, 2011).

Results: For U.S. women in STEM results indicated the presence of three ST susceptibility profiles (i.e., low and high ST susceptibility classes and a disengaged class) and two variations of an un-identified class in the non-STEM sample. High and low susceptibility to ST classes were found for Colombian women in STEM, while the non-STEM sample yielded disengaged and un-identified classes. In both countries, over 70% of the women in STEM were classified as highly susceptible to ST.

Discussion: This is the first study investigating latent profiles of susceptibility to ST (SST) so additional replication with samples from different populations is strongly recommended. Extensive investigation into latent profiles of ST susceptibility could provide the insight required to develop differentiated ST reduction strategies for students in STEM and non-STEM fields of study.

Keywords: Stereotype threat, latent class analysis, mathematics education, latent profile analysis, females, STEM

Resumen

Introducción: El presente estudio investigó los grupos teóricos de susceptibilidad a la amenzada de estereotipo propuesta por Steele (1997) por medio de un análisis de clases latentes.

Método: 413 estudiantes de pregrado de Estados Unidos y Colombia, en áreas de Ciencias, Tecnología, Ingeniería y Matemática (STEM) y disciplinas no-STEM diligenciaron un cuestionario que mide la susceptibilidad a la amenaza de estereotipo-- the Social Identities and Attitudes Scale, SIAS (Picho & Brown, 2011).

Resultados: Para las mujeres de Estados Unidos de áreas STEM los resultados indicaron la presencia de tres perfiles de susceptibilidad a la amenaza de estereotipo (Por ejemplo, las clases Baja y Alta susceptibilidad y una clase No Involucrada) y dos variaciones de una clase No Identificada en la muestra no-STEM. En el caso de las mujeres de Colombia en áreas STEM, fueron encontradas las clases Baja y Alta susceptibilidad a la amenaza de estereotipo, mientras que las mujeres de la muestra no-STEM se encontraron las clases No Involucrada y No Identificada. En ambos países, cerca del 70% de las mujeres en áreas de STEM fueron clasificadas como altamente susceptibles a la a la amenaza de estereotipo.

Discusión o **Conclusión**: Este es el primer estudio investigando perfiles latentes de susceptibilidad a la amenzada de estereotipo (SST), por lo tanto, estudios adicionales con diferentes poblaciones son recomendados. Extensa investigación acerca de los perfiles latentes de susceptibilidad a la amenza de estereotipo podría proveer nueva información requerida para el desarrollo de estrategias deiferenciadas para la reducción de amenaza al estereotipo para estudiantes en áreas de estudio STEM y no-STEM.

Palabras Clave: Amenza de estereotipo, análisis de clases latentes, educación en matemática, mujeres, CTIM

Introduction

Recent analyses of national US data reveal that the gender gap in mathematics starts much earlier, between kindergarten and the fifth grade (Robinson & Lubienski, 2011), prompting a call for tailored interventions for the content domains where gaps persist. The gender gap in mathematics has been shown to vary based on socio-cultural factors (Else-Quest et al., 2010). Implicated in these factors has been gender stereotypes regarding women's ability in stereotypically masculine domains (Casad et al., 2018; Dennehy et al., 2018; Nosek et. al, 2009; Steele, 1997; Seyranian et. al, 2018).

Research on gender differences in mathematics, studied extensively in North America and Europe, has only begun to gain traction in Latin America in the past few decades (Baldeón-Padilla et al., 2020; Del Río et al., 2016; González-Gutiérrez et al., 2018; Espinosa, 2010). The role of gender and gender related factors as a potential contributor of the performance gap in mathematics at all levels of education has spurred research from Mexico, Chile, Colombia, Brazil, Argentina and Peru. Studies from these countries indicate a persistent gender gap favoring males in the mathematics performance (Aguiar et al., 2011; Cervini & Dari, 2009; Espinosa, 2004; Fernández & Hauri, 2016; Figueroa & Ortega, 2010; Gonzalez-Pineda et al., 2012; March, 2009; Mella, 2006; Reali et al., 2016; Salazar et al., 2010; Valentova et al., 2017). The gender gap in language and mathematics performance has been linked to social norms and expectations that create distinct gender roles, which are often stereotypical in nature. These stereotypes trickle down to academics and thus impact women's attitudes towards Science Technology Engineering and Mathematics (STEM) subjects. Studies from Peru, Mexico, and Brazil indicate that females strongly endorse gender stereotypes related to mathematics (Gonzalez-Pineda et al., 2012; Valentova et al., 2017; Winkler, 2004), and teachers and parents seem to contribute to those stereotypes since early in education (Correa, 2015; Del Río et al., 2016).

Gonzalez (2005) and Gonzalez-Pineda et al. (2012), found that women endorsed gender stereotypes more strongly than men, and that women who endorsed beliefs of mathematics being a male subject tended to have low math performance scores. Similarly, research indicates that stereotype endorsement with regards to mathematics and language abilities tend to be more pronounced among students from low socio-economic backgrounds (Espinosa, 2004). In Mexico, Chile and Brazil, the percentage of women who enroll in STEM disciplines is significantly lower than that of men. Data provided by the National Ministry of Education of Colombia (2018), through SPADIES (2018) indicate that the highest proportion of students who are enrolled in engineering are men and that a significant proportion of women pursue careers in stereotypically feminine domains like the humanities (Espinosa, 2004; Figueroa & Ortega, 2010; Gonzalez, 2009; March, 2009). In sum, previous studies show that the cultural and educational practices seem to affect women's choice of careers and performance in mathematics and the physical sciences.

Despite the link between gender stereotypes, stereotype endorsement, and choice of career paths, as well as the under-representation of women in STEM, there remains a deficit of research investigating these relationships in Latin America. Research from other parts of the globe indicate that the negative gender stereotypes questioning the ability of women in quantitative domains can negatively affect their performance in these subjects (Armenta, 2010; Eriksson & Lindholm, 2007; Huguet & Regner, 2007; Picho & Stephens, 2012). The phenomenon, commonly called stereotype threat (ST), is social-psychological. It arises when one performs difficult tasks in a context or domain where negative stereotypes about one's group is made salient (Steele, 1997). ST undermines academic achievement by inducing anxiety which affects performance. ST is not restricted to women or to mathematics; it negatively affects the performance of individuals from any social group whose abilities in a specific domain are questioned. For instance, negative societal stereotypes about elderly individuals' memory have been shown to actually affect memory of members of this group when the stereotype is activated (Kit et al., 2008). Several studies conducted in North America have also demonstrated negative ST effects on the performance of ethnic minorities on tests of academic ability (Gonzales et al., 2002; Steele & Aronson, 1995). Even White males seem to perform poorly when confronted with the stereotype that Asians are superior in mathematics (Aronson et al., 1999). ST is pervasive in domains where negative stereotypes about a particular group exists. However, the present study focuses solely on ST as it relates to the performance of women in quantitative domains, specifically mathematics.

Literature Review

In their seminal article, Steele and Aronson (1995) demonstrated the negative impact of gender- ability stereotypes on the test performance of bright women on the GRE quantitative test. Since the inception of stereotype threat theory (STT, hereafter) hundreds of articles and dissertations investigating the threat of negative stereotypes (or stereotype threat) on academic performance have reported similar effects on the performance of females in quantitative domains (Brown & Pinel, 2003; Keller & Molix, 2008; Schmader, 2002).

For women, math-related ST is mediated by one's identification to the domain (i.e. math identification), awareness of but not necessarily belief in negative stereotypes about women's ability in mathematics (gender stigma consciousness), and negative emotions like self-doubt and anxiety, experienced while performing difficult mathematics tasks (Steele, 1997). Domain identification is the key mediator of ST, since non-identification to mathematics renders the negative gender ability-stereotype irrelevant to the individual's performance, which buffers the individual against the negative ST effects. For math-identified women who are aware of the negative gender stereotype related to math ability, ST is further mediated by negative emotions experienced while performing difficult quantitative tasks. Negative emotions like anxiety and task-related worries degrade limited working memory required for successful problem solving (Schmader et al., 2008), which lead to poor cognitive processing (Beilock et al., 2007; Schmader & Johns, 2003), and hence lower task performance, which confirms the stereotype. Over time, repeated underperformance due to ST leads to 'disidentification'-- a situation where the members of the stigmatized group disengage from and lose interest in the stereotyped domain (Aronson et al., 2001; Casad et al., 2018; Dennehy et al., 2018).

Theoretical ST susceptibility profiles. Steele (1997) proposed three qualitatively different profiles of ST susceptibility to stereotype threat (hereafter, SST), largely differentiated by domain identification, and stigma consciousness. There is the domain unidentified group which consists of individuals unsusceptible to ST because their non-identification with mathematics makes gender stereotypes in this domain irrelevant, which buffers against the activation of ST. Second is the domain-identified group. Individuals belonging to this group are theorized to be highly susceptible to ST (hereafter high SST) because they strongly identify with the domain and have a vested interest in succeeding in it but also identify with the stereotyped group (i.e. gender) which makes ST more likely. Identification to both the stereotyped domain (mathematics) and their gender creates cognitive dissonance, which creates anxiety in the individual about confirming the gender stereotype (Schmader et al., 2008). Thus, it is the interplay between group and domain identification, mediated by anxiety, that co-opts limited working memory and leads to suboptimal performance, which inadvertently confirms the stereotype (Schmader et al., 2008). Finally, there is the disidentified or disengaged group, which comprises individuals who have experienced ST over extended periods and as a result, no longer identify with the domain.

Domain disidentification serves the primary function of protecting one's self-esteem (Steele, 1997) by diminishing the relevance of the stereotyped domain to one's sense of self (i.e. I am not doing well but I do not like math anyway). In so doing, the self- esteem of the disidentified student is less impacted by poor performance on the intellectual task in the stereotyped domain, than it otherwise would, had he or she been domain-identified (Steele, 1997). Because ST is not universal to all members of stigmatized groups (Steele, 1997), members of stereotyped groups who strongly identify with the stereotyped domain but who also remain unaffected by ST, would naturally be classified as a low SST or normative (reference) group.

Although a substantive amount of ST research has been conducted in recent decades, inquiry into the SST profiles proposed by Steele (1997) has been largely ignored. The present study sought to fill this gap in the literature. Specifically, the purpose of the current investigation was to assess the tenability of Steele's theoretical SST profiles among college students, using latent class analysis. To our knowledge, this is the first study to validate these profiles using this approach, with important implications for ST research which we discuss here. SST profiles are an essential component of ST theory; they help us differentiate sub-groups of individuals within a marginalized group which are more likely to experience ST from those which are not. Ergo, they help set some of the boundary conditions for ST theory—at least, as far as which types of individuals are predisposed to experiencing ST. Boundary conditions are a set of conditions that set limits on our generalizability of when a phenomenon occurs; they also facilitate predictions about how the relationship between a phenomenon with other variables should change under different situations (Busse et al., 2016). Boundary conditions are crucial for theory development (Aguinis et al., 2013; Mathieu et al., 2012) because they guide the interpretation of empirical findings and further empirical work to extend or revise existing theory (Klein, 2014). Therefore, being the first study to validate these SST profiles, the present investigation makes a significant contribution to the ST literature. More precisely, findings from this study will provide an empirical basis for future studies to test the aforementioned ST boundary conditions more directly. This could initiate new directions for ST research where the differential impact of ST within these distinct SST groups is examined and, over time, facilitate a more comprehensive understanding of ST.

An overview of latent class analysis is presented, followed by a report and discussion of the investigation of these latent classes using samples from both the U.S.A and Colombia.

Latent Profile Analysis

Latent class analysis (LCA) is a statistical method used to identify unobserved class membership based on categorical or continuous data (Samuelson & Dayton, 2010). LCA classifies individuals into substantively meaningful groups (latent classes) based on similarity in response patterns assessing an array of observations. Class membership in LCA is probabilistic, not deterministic (Samuelson & Dayton, 2010). For example, one might seek to categorize respondents based on their smoking behaviors (observations) into different types of smokers (latent classes). This could lead to finding categories such as non-smokers, occasional smokers, and chain smokers. When the observations are continuous (as is the case with ST susceptibility measure used in this study), the method is called latent profile analysis, LPA (Lubke & Neal, 2006). In LPA, two key parameters are estimated: (1) unconditional, and (2) conditional class probabilities. The former estimates the probabilities that respondents belong to a given latent class, k, and provides information on the size of the latent class. The latter allows one to obtain various probabilities conditional on class membership, i.e. the probability of endorsing an item based on class membership.

LCA has been used to investigate categories of substance use, and other mental disorders. In Psychiatry and clinical psychology for example, it is used to analyze ADHD classification based on the DSM IV's diagnostic criteria for the disorder (Hudziak et al., 1998; Neuman et al., 1999; Rasmussen et al., 2002), as well as alcohol dependence and abuse based on the DSM IV's diagnostic criteria of the same (Muthén, 2006). To the best of our knowledge, this is the first time that LPA has been used to investigate typologies of ST-susceptibility.

Research aims and hypotheses

The primary aim of the current study was to examine whether the theoretical profiles proposed by Steele (1997) could be reproduced in a sample of American and Latin-American students in STEM and non-STEM domains.

The literature on gender-related ST is well established and most of these studies have been conducted in North America. Here, studies have shown that negative gender stereotypes about quantitative ability make some women particularly vulnerable to ST when they perform quantitative tasks (Spencer et al., 1999; Galdi et al., 2014; Picho & Schmader, 2018). According to ST theory, women predisposed to high levels of domain and gender identification are more likely to experience ST (Steele, 1997). Therefore, we hypothesized that for American women, evidence would be found to support all four classes of ST susceptibility proposed by Steele.

ST theory presents domain identification as a precursor to ST. Hypothetically, SST profiles should vary considerably among participants who self-selected into STEM versus those in non-STEM domains. The reason for this is that, more than likely, individuals who self-select into STEM do so because they identify with the domain. For this reason, as part of the validation process, SST profiles for individuals in both STEM and non-STEM were examined and compared. According to ST theory, domain identification is crucial to activating ST because strong domain identification signals investment and a desire to be successful in the domain, which is threatened when the negative stereotype is made salient (Steele, 1997). On the other hand, non-identification with the domain buffers against ST because even though negative stereotypes might be present, the fact that one does not care for or is not invested in the stereotyped domain makes the stereotype irrelevant to one's performance (Steele, 1997). Therefore, we predicted that the SST profiles found among women in STEM would differ from that found among women in non-STEM domains. More precisely, we hypothesized low, high and disidentified SST profiles for women in STEM. We expected to find low and high SST profiles in the STEM sample because both of these profiles require moderate to strong identification with the stereotyped domain, which is mathematics intensive. Since not all women are expected to experience ST, a low SST group was therefore postulated. Based on prior studies that have shown ST to be a burden for women who identify with math (Steinberg et al., 2012; Keller, 2010; Keller & Molix, 2008) and their gender (Schmader, 2002), a high SST group was also hypothesized. Finally, because the long-term effect of ST is disidentification (Steele, 1997), a disidentified group for this sample was also hypothesized. By contrast, there is likely to be more variability in mathematics identification among non-STEM students, as some majors within fields like Psychology (e.g. quantitative psychology) can be heavily quantitative, while other majors like History, or Communications, for example, might hardly have any quantitative components. To that end, we hypothesized that depending on individual level characteristics assessed by the SST measure, any of the theorized SST classes, i.e. low, high, unidentified and disidentified class, was possible. As reviewed earlier, gender stereotypes are prevalent in Latin American countries. Therefore, we expected SST profiles based on samples from Colombia, to be relatively consistent with what would be found for the U.S. samples.

Method

The present study used latent profile analysis to investigate the above-mentioned SST groups described by Steele (1997).

Participants

After IRB approval, students from two private Universities in the United States and Colombia were solicited to complete an attitudinal survey- the Social Identities and Attitudes Scale (SIAS) (Picho & Brown, 2011). Information sheets delineating the purpose of the study, which was to assess student attitudes towards mathematics, and student perceptions of their social identities relative to their academics, were given to potential participants. To ensure that responses to indicators on the ST susceptibility subscale assessing mathematics identification were relevant to respondents in the non-STEM domains, the selection of non-STEM participants was restricted to students taking a quantitative course (i.e., calculus or statistics) during the semester that the SIAS was administered. A total of 413 undergraduate women from both countries participated in the study. A full description of the sample is available in table 1.

Table 1. *Student demographic information in percents- US and Colombian sample*

	US		Colombia	
	STEM (*n* = 211)	Non-STEM (*n* = 176)	STEM (*n* = 104)	Non-STEM (*n* = 61)
Ethnicity				
Asian	4.1	5.7		
Black	16.3	19.9	10.9	3.6
Hispanic	4.5	9.1	46.1	29.7
White	69.2	59.7		
Indigenous			1.2	1.2
Year in College				
Freshman	1.7	4	7.3	7.3
Sophomore	32.9	4	24.8	7.9
Junior	16	42.7	20.6	9.7
Senior	21.1	29.3	9.1	9.7
Last year (Colombia)			1.2	1.8

Note: In some cases, percents don't equate to 100 because of missing data.

Instrument

The SIAS (Picho & Brown, 2011) was used to evaluate stereotype threat susceptibility. The SIAS is a six- factor scale assessing key constructs that distinguish ST susceptibility classes proposed by Steele (1997): domain (math) identification, negative emotions experienced during test taking (negative affect), group identification (gender and ethnicity) and stigma consciousness of the same. Indicators for the latent factors assessed by the SIAS subscales are continuous and measured on a seven-point Likert scale where 1 = strongly disagree, 2= disagree 3= somewhat disagree 4 = neither agree nor disagree, 5= somewhat agree 6= agree, and 7 = strongly agree. The SIAS, previously validated with college (Picho & Brown, 2011) and high school samples (Picho, 2011), has thus far demonstrated strong psychometric properties. For this sample, subscale reliabilities were strong, ranging from .76 -.94 (see table 2).

Table 2. *Reliability estimates for SIAS subscales for the entire sample*

		USA		**Colombia**	
Subscale	No. Items	*Alpha*	CI_{95}	*Alpha*	CI_{95}
MI	6	.94	.92-.95	.91	.89-.92
NA	6	.94	.92-.95	.87	.85-.89
GI	3	.76	.70-.81	.71	.64-.76
GS	5	.81	.78-.84	.81	.77-.84

Procedure

Participants completed the SIAS in the second half of the fall semester-- after they had all taken at least one exam in the quantitative subject. The latter was an attempt to facilitate the accesibility of one's direct experience in the current quantitative class as a basis for more accurate reports on group stigma consciousness, identification, and negative affect relative to mathematics. Composite scores for the SIAS subscale factors were created and used to conduct latent profile analyses in Mplus 7.

Because the U.S. and Colombian samples were predominantly White and White Hispanic, respectively, scales assessing group identification and stigma consciousness related to ethnicity were not used in the study. Therefore, four instead of six subscales of the SIAS were used to assess gender related SST: math identification, gender identification, gender stigma consciousness and negative affect. Gender identification was assessed because of the link between gender stereotypes and math-related ST (which was the focus of the investigation), and also because previous studies have identified gender identification as a moderator of ST (Armenta, 2010; Erikson & Lindholm, 2007; Keller & Molix, 2008; Schmader, 2002). Based on STT and on the aforementioned theoretical ST profiles delineated in Steele's (1997) original article, ST susceptibility groups were easily defined a priori by subscales on the SIAS as follows: (1) students highly susceptible to ST (high SST) would rank high on all factors measured by the SIAS, (2) the normative group (low susceptibility to ST or low SST) would rank high on math identification and low on all other factors, (3) disengaged students, by virtue of previously experiencing ST over extended periods prior to domain disidentification (per STT), were expected to rank low on math identification and high on other SIAS factors, and finally (4) unidentified students were expected to rank low on math identification and all other SIAS factors.

Model building

Using STT as a guide, a model-building approach was taken to building latent profile models (LPMs). First, a two- class model was fitted (high vs. low SST), followed by a three-class model (high SST, low SST, and disengaged), and finally, a four- class model (high SST, low SST, disengaged, and unidentified). The models were fitted separately for samples from each country therefore, results presented below are presented separately for the U.S and Colombian samples.

Results

Latent profile models were evaluated using entropy values, class separation, Bayesian information Criteria (BIC), the adjusted likelihood ratio test (aLRT) of model fit and where warranted, the bootstrap likelihood ratio test (bLRT). Entropy is a measure of how well latent classes can be distinguished (i.e. classification accuracy). Entropy values for a given model can range from 0-1, with values at or near 1 indicating a high degree of certainty of classification, and values at or close to 0 indicating the opposite (Muthén, 2006). An entropy value of 0.8, for instance, indicates 90% correct classification rates. Good models are denoted by low BIC values (Muthén, 2006), as well as low p values ($p < .05$) for the aLRT and BLRT. The latter provide tests of improved model fit in moving from a model with one latent class less than the model being estimated so a low p value indicates a preference for the estimated model (k) over the simpler model (k-1) and vice versa, in both cases.

Interpretation of SST latent profile models. Both STT and the meaning of the points on the Likert scale guided the interpretability of ST susceptibility profiles for each of the factors in both American and Colombian samples. As described in the measures section, the items assessing the factors under investigation were anchored on a seven-point scale (1= strongly disagree, 4 = neutral and 7= strongly agree). Therefore mean scores hovering above .5 points on each point of any SIAS sub-scale (e.g., 3.5, and 4.5) indicated average responses for the participants leaning toward the next higher point on the scale (e.g., 4 for neutral and 5 for somewhat agree, respectively) and vice versa.

Latent Profile Analysis for the U.S. Sample

Table 3. *USA: Latent Profile Models of Susceptibility to Stereotype-Threat for STEM and non-STEM women*

		STEM		
Classes	*Entropy*	*BIC*	*aLRT*	*bLRT*
2	0.97	2049.83	.049	<. 001
3	**0.88**	**2043.99**	**.122**	**<.001**
4	0.81	2051.48	. 535	<.001
		Non-STEM		
2	**0.90**	**2772.08**	**<. 001**	<. 001
3	0.84	2754.05	.121	<.001
4	.78	2758.66	.370	<.001

Table 3 shows model fit characteristics for the fitted latent profile models of U.S. women in both domains. The entropy of all fitted models was above .80 (ranging from .81 - .97), indicating adequate to very good classification rates. Based on the aforementioned model fit criteria, the three-class model was the best fitting model for women in STEM. The BIC and aLRT did not agree on the number of classes for the best fitting model. The BIC favored the three-class model while the aLRT favored the two-class model. The bLRT, which is usually recommended to decide between classes when model fit indices do not agree (Muthen, 2006), failed to discriminate between the number of classes for all fitted models, so reliance on theory, and an inspection of graphs were used to aid in the selection of number of classes. Although the two-class model had the higher entropy and the aLRT suggested a preference for the two- class model over the three- class model (p =. 05), the three-class model had the lower BIC, and a close graphical inspection of its latent profiles showed a substantively distinct third class congruent with STT that was not captured in the two class latent profile model. Therefore, the three-class model was retained.

The three-class model for U.S. women in STEM

This model had strong classification accuracy (entropy= .88), and strong class separation, as indicated by average latent class probabilities for the most likely class membership (see table 4). The predicted probabilities (diagonal values) were very high for the most likely class membership and were either zero or extremely low (< .10) for the other classes. For example, for individuals in the high SST group (class 3), the probability of being in that class given their responses to the SIAS was .95, but their probability of being in the first and second classes were zero and .05, respectively.

Table 4. *U.S.: Average latent class probabilities for most likely class membership for STEM vs. non-STEM Women*

Conditional probabilities- STEM			
Class	1	2	3
1 (Disengaged, 7.2%)	**0.98**	0.00	0.02
2 (Low SST, 21.6%)	0.01	**0.89**	0.10
3 (High SST, 71.2%)	0.00	0.05	**0.95**
Conditional probabilities- non-STEM			
Class	1	2	
1 (Unidentified, 18.5%)	**0.95**	0.05	--
2 (Non-committal, 81.5%)	**0.02**	**0.98**	--

STEM latent profiles of SST. Latent profiles for the best fitting model for the STEM sample are depicted in fig 1. Results found three profiles consistent with STT for STEM students: low SST, high SST and disengaged classes.

The low SST class. The normative group- represented as the second class in figure 1, constituted 21.6% of the STEM sample. Congruent with STT's definition of domain-identified individuals not susceptible to ST, this group strongly identified with mathematics (with average scores over 6 on the 7-point Likert scale) and had low ratings (disagree) on all other ST moderators.

The high SST class. Students fitting this profile constituted the third class depicted in the figure. 71.2% of STEM women sampled fell into this category. Respondents belonging to this class strongly endorsed math identification items, and ranked moderately high on all other ST moderators, which was also consistent with ST theory's definition of individuals highly susceptible to ST.

The disengaged class. This was the smallest class (7.2%), but responses indicated a class substantively and qualitatively different from previous classes. In this class, students did not identify with mathematics but ranked moderately high on all other ST moderators. This group also ranked higher than other groups on negative affect experienced during test taking. This profile was also consistent with STT's definition of disengaged groups.

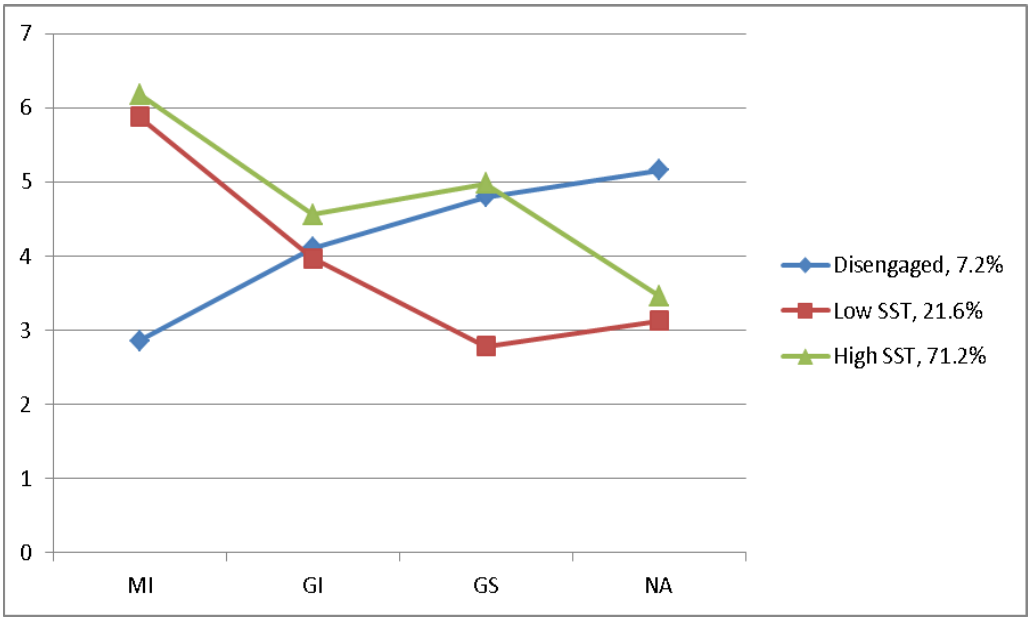

Figure 1. Latent Stereotype Threat Susceptibility Profiles for STEM Women—US

The two- class model for U.S. non-STEM women

Similar to the STEM female sample, the model fit indices for the non-STEM sample did not reach consensus on the best fitting model; the aLRT supported a two class model, but the three class model had the lower BIC, and the bLRT did not discriminate between either models (p <.001 for all tested models). Therefore, as in the STEM female sample, plotted latent profiles of the two and three class models were inspected and STT was used to guide model selection. A comparison of the latent profiles revealed no distinct differences between the first and the third latent profiles in the three-class LPM; except for lower means on all subscales for the third class, there were no qualitative differences between the two profiles. Therefore, the two-class model was retained for the non-STEM sample. The classification accuracy for the two-class model was very high (entropy= .90); it also had strong class separation with very high predicted probabilities for the most likely class membership, as shown in table 4.

Non-STEM latent profiles of SST. The latent profiles for the two-class model (see figure 2) revealed two variations of un-identified SST latent profiles: a non-committal class (class 2) and an unidentified class (class 1). Given the anchors of the SIAS scale, with a score of 4 representing ‘neutral’, average responses for math identification leaned towards weak identification for non-committal (4.5), and no identification to mathematics for the unidentified class.

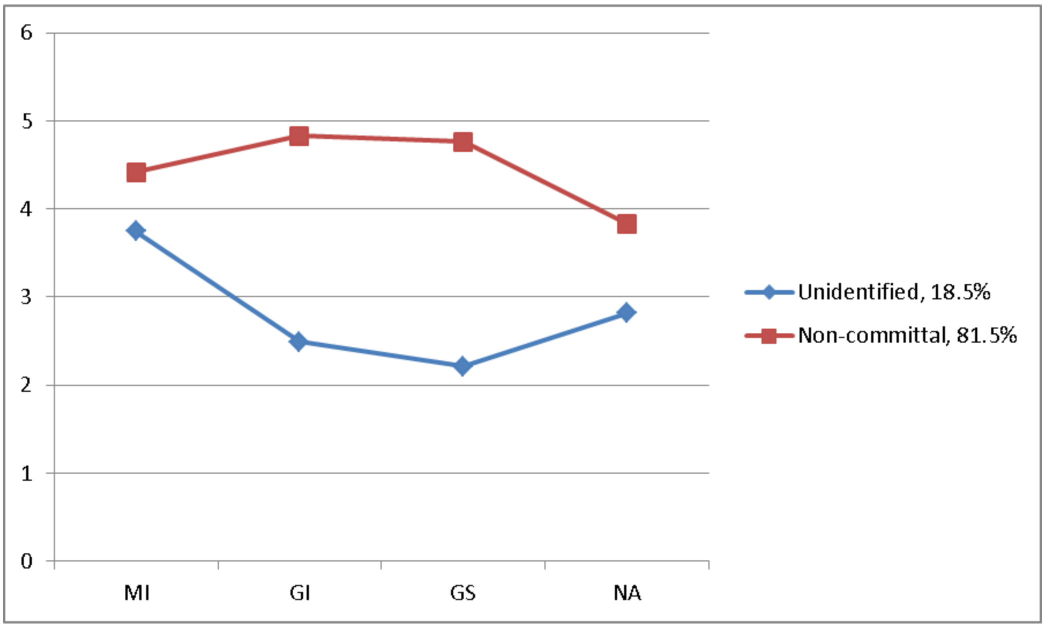

Figure 2. Latent Stereotype Threat Susceptibility Profiles for non-STEM Women- USA

The unidentified class. Based on mean responses on the SIAS subscales, this group did not identify with math; they also ranked very low on stigma consciousness related to gender and negative affect. This profile was consistent with STT's proposition regarding unidentified groups, discussed earlier in this article, and it constituted 18.5% of the non-STEM women sampled.

The non-committal class. Women belonging to this class could be perceived as 'sitting on the fence' because their average responses on the SIAS subscales were non-committal. 81.5% of non-STEM women sampled fell under this category, which was denoted by weak math identification and similar ratings on the remaining ST moderators. Of note is that this group also had much higher mean scores on gender identification and stigma consciousness factors compared to individuals in the unidentified class.

Latent Profile Analysis for the Colombian Sample

The entropy of all fitted models ranged from .73 -.89. Based on the model fit criteria stated earlier, the two-class model was the best fitting model for Colombian women in STEM. The BIC favored the four-class model, but the three-class model had the highest entropy and the aLRT favored the two -class model. Close graphical inspection of its latent profile graphs showed that in the three-class model, the third class was quantitatively but not qualitatively different from the second class. However, the extreme low values on all SIAS factors in the third class indicated an unidentified class, whereas the second class indicated moderate math identification, and very low scores on other ST susceptibility variables, indicating a low susceptibility to ST class. Therefore, the three-class model was retained.

The three-class model for Colombian women in STEM

Table 5. *Colombia: Latent Profile Models of Susceptibility to Stereotype-Threat for STEM and non-STEM women*

		STEM		
Classes	Entropy	BIC	aLRT	bLRT
2	.83	1499	.06	<. 001
3	.88	1498	.14	<.001
4	0.87	1495.80	<.001	<.001
		Non-STEM		
2	**.73**	**894.77**	**.03**	**--**
3	.89	903.82	.17	.25
4	.85	915.41	.84	--

This model had high entropy values (i.e. strong classification accuracy). It also had strong class separation as indicated by average latent class probabilities for the most likely class membership (see table 5), and high predicted probabilities for the most likely class membership.

STEM latent profiles of SST. Figure 3 shows latent profiles for Colombian women majoring in STEM. With the exception of lower mean scores on mathematics identification, the susceptibility to ST group was similar to that found in the U.S. sample: a high SST class (65.1%), a low SST class (31%) and an unidentified class (3.8%).

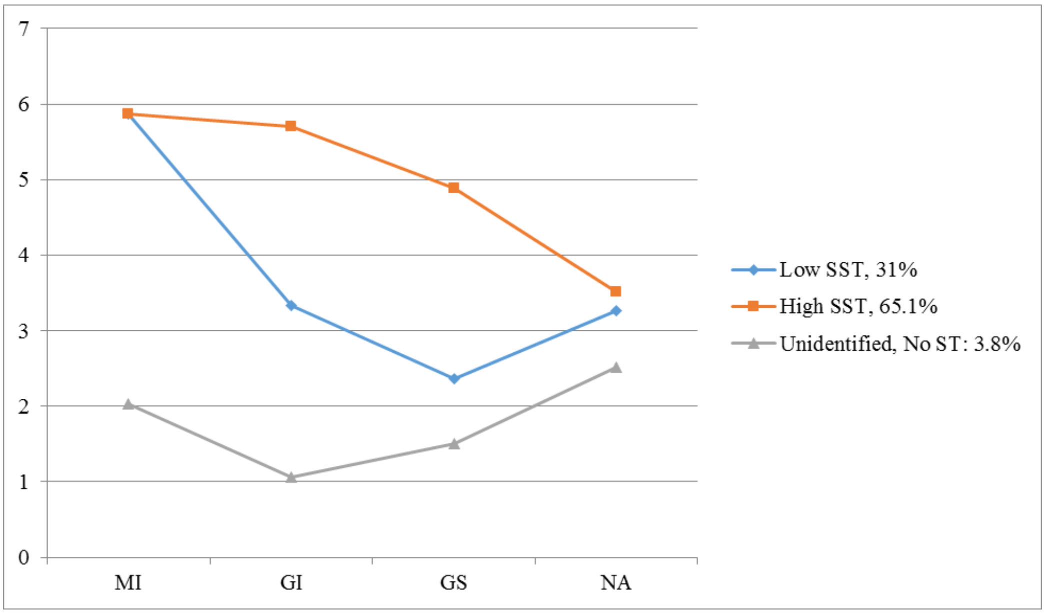

Figure 3. Latent Profiles for STEM women- Colombia

The two-class model for Colombian non-STEM women

For the non-STEM sample, the low BIC and aLRT (p =.03) favored the two-class model as the best fitting model (see table 5). This model, characterized by an entropy value of .73, also had strong class separation with very high predicted probabilities for the most likely class membership (see table 6).

Non-STEM latent profiles of SST. Latent profiles for non-STEM women revealed a class that did not identify with mathematics but was relatively higher on the rest of the SST factors. Following STT guidelines, this class mimicked a disengaged class. 60.2% of women in this sub-sample were classified as being disengaged. A low SST class was also found in this sub-sample, with women exhibiting moderate identification with mathematics and ranking low on all the other SST factors. 39.8% of the sample of non-STEM students were classified as low susceptibility to ST.

Table 6. *Colombia: Average latent class probabilities for most likely class membership for STEM vs. non-STEM Women*

Conditional probabilities- STEM			
Class	1	2	3
1(Low SST, 31%)	**.91**	.01	.09
2 (High SST, 65.1%)	.04	**.96**	.00
3 (Unidentified, 3.8%)	.04	.00	**.96**
Conditional probabilities- non-STEM			
Class	1	2	
1 (Disengaged, 39.8%)	**0.91**	0.09	
2 (Low SST, 60.2%)	**0.06**	**0.94**	

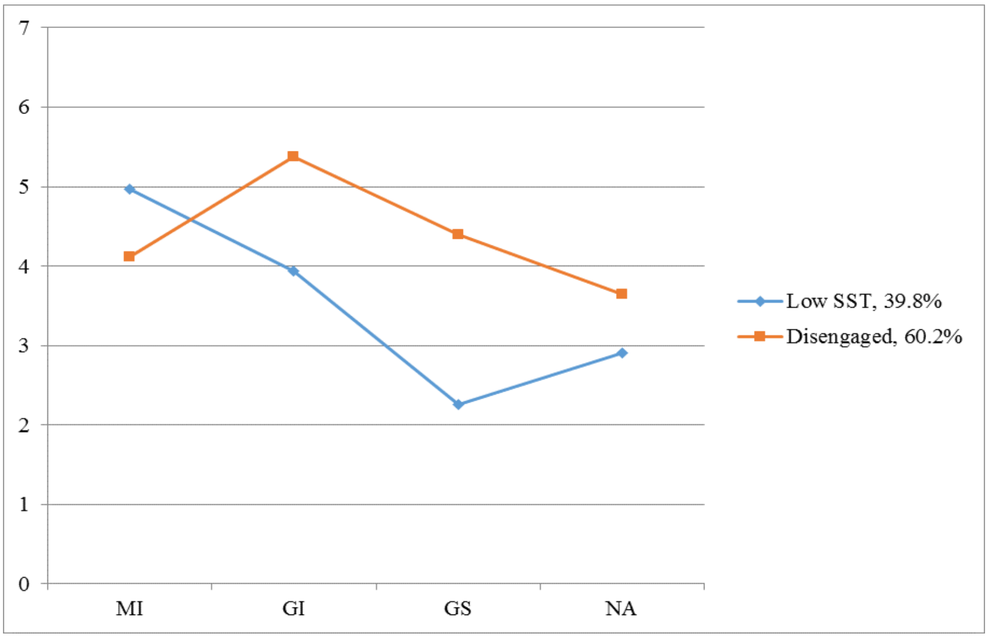

Figure 4. *Latent Profiles for non-STEM women- Colombia*

Discussion

Overall, results from the latent profile analyses revealed the presence of three out of the four ST susceptibility profiles proposed by STT (Steele, 1997) in both Colombian and US students. In the U.S sample three SST classes for women in STEM (low and high SST, plus a disengaged class), and two variations of unidentified groups among women from non-STEM domains were identified. Similar profiles were In the Colombian sample, a high and low SST class was found for women undertaking STEM under-graduate studies, while low SST and disengaged classes were found for those in non-STEM domains.

The difference in SST latent profiles by academic domain appear tenable and in line with STT. The latent profiles of non-STEM women in both samples, marked by average to below average levels of math identification were reasonable since undergraduate disciplines outside STEM tend not to be as math-intensive, and students generally tend to self- select into disciplines that they identify with. Conceptually it would appear reasonable not to find an unidentified SST class among STEM majors because most students self-select into majors of their choice in college and the probability of math unidentified majoring in STEM would be highly unlikely. For non-STEM undergraduate students on the other hand, an unidentified class could well be expected as a potential SST class. Therefore, the presence of these classes in the respective domains is plausible.

However, the latent profiles of U.S. women from non-STEM domains found variations of an unidentified class, which was not expected because it deviated from empirical research finding ST effects on women from non-STEM domains. Based on these previous studies, the presence of at least a high SST class would have been expected in the current sample. It is possible that, as with other research using LCA for classification purposes (Hudziak et al., 1998; Neuman et al., 1999; Rasmussen et al., 2002), the current results might have been largely a function of sample characteristics. The university from which the sample was drawn (a teaching university) tends not to offer an array of quantitative courses in for majors outside STEM; for example the Psychology department, from which the majority of non-STEM students in the study were drawn, lacks a quantitative psychology program and offers only one statistics and two research methods courses for its graduate and undergraduate curriculum. The departmental offerings at this specific university most likely reflect interests of the student demography. Therefore, replication studies in other schools offering programs outside of STEM with strong, quantitative components (like quantitative psychology), might yield the expected theoretical profiles that were lacking in the current sample under study.

The latent profiles also revealed a difference in STEM and non-STEM profiles with regards to the disengaged class. This class was found among STEM U.S. students and in non-STEM Colombian students. The presence of a disengaged class in the STEM sample is tenable, given the long term effects of ST. The presence of such a group in the non-STEM South American sample possibly speaks to the possible idea underlying the choice of careers: women in non-STEM fields selected these domains because they did not identify with STEM prior to enrolling at the university. This conjecture is based on empirical research presented earlier indicating that women endorse gender stereotypes more strongly than men and that few enroll in STEM disciplines. However, the paucity of research with Latin American samples in this field precludes inference of a causal relationship between the aforementioned.

ST susceptibility for women in STEM. 71.2% and 74.03% of the STEM-sample in the U.S and Colombia, respectively, were classified as being susceptible to ST, which, at first blush, appears to be quite high. But this result is tempered by the restrictive assumption of local independence under which LPA operates. Generally, LPA (LCA) assumes no heterogeneity within class i.e., all observations in the same class have the same susceptibility to ST. Although STT provides distinctions between qualitatively different types of ST susceptibility, it is still unclear from the literature whether susceptibility to the phenomenon is dimensional or strictly categorical. The former would place ST as being a matter of degree, while the latter would limit SST to qualitatively different classes. Given that the psychological constructs that constitute the key parameters of STT are themselves dimensional (i.e. domain identification, and varying levels of group identification and stigma consciousness), the idea that the phenomenon could also be dimensional is plausible as well as possible. Modeling latent classes that allowed for degree of ST susceptibility within each class might not only be a happy medium between the two extreme camps defining the nature of the phenomenon, but also a more realistic representation of SST. Latent variable approaches to modeling group membership, like factor mixture analysis (FMA), might be useful in exploring ST susceptibility further.

FMA relaxes the assumption of local independence and allows one to model both class (i.e. different classes of ST susceptibility) and within class heterogeneity (varying degrees of ST susceptibility within each class). Modeling dimensionality within each class probably could provide more insight into ST susceptibility, and inform the development of ST alleviation strategies for stigmatized groups based on level and degree of ST susceptibility. Unfortunately, the sample size in this study was not large enough to facilitate FMA, which would have allowed for the investigation of heterogeneity within ST susceptible classes. The current study was limited this way and future researchers are encouraged to explore the dimensionality of ST susceptibility further using FMA.

General discussion

Findings from this study have significant implications for the broader ST literature. So far, hundreds of experimental studies have been conducted to examine the impact of ST on the performance of women in quantitative subjects like mathematics. For the most part, these studies have demonstrated that ST degrades the performance of women in subjects like mathematics where negative gender stereotypes are prevalent (Spencer et al., 2001; Keller & Molix, 2008; Tomasetto et al., 2011; Tomasetto & Appoloni, 2013). However, these and most other ST studies did not gauge participants' susceptibility to ST prior to the study so it is possible that these study samples had a mixture of individuals who were susceptible and not susceptible to ST. As a result, and because ST does not affect all individuals the same way (Steele, 1997), the ST effects observed and reported in existing ST studies might not reflect true ST effects that might be obtained from studies conducted exclusively with participants who are actually susceptible to the phenomenon. It follows then, that being able to discriminate between ST susceptible and non-susceptible individuals could improve current estimates of ST effects. Thus, empirical validation of these SST profiles makes it possible for researchers to discriminate more easily between ST susceptible and non-susceptible individuals during selection of study participants. Consequently, and as a result of this investigation, researchers should now be able to conduct experiments with appropriate samples i.e. participants who are susceptible to the phenomenon, leading to an overall improvement in estimates of ST effects. The latter is very important because findings from ST research have provided a basis for numerous interventions. Therefore, where decisions have to be made regarding the allocation of limited efforts and resources, it is vital that these decisions and corresponding initiatives to narrow gender disparities in STEM are based on the magnitude of impact of the phenomenon. Studies based on ST susceptible participants would offer more accurate estimates on the impact.

Findings from this study also have practical implications. Specifically, the empirical validation of these SST profiles could be used as a basis to enhance ST interventions and corresponding research. In the United States, the literature on ST interventions is fairly well established. Research in this area suggests that teaching about ST (Johns et al., 2005), self- affirmation (Bowen et al., 2012; Miyake et al., 2010), and promoting an incremental view of intelligence (Aronson et al., 2001; Good et al., 2003) could effectively reduce ST effects on the performance of at-risk members of marginalized groups within a given domain. Currently, these interventions are generic and catering to all students regardless of their susceptibility to ST. However, having a sound, theoretically and empirically validated SST classification system would propel further research into the psychology of individuals belonging to the different SST profiles. Results from this study make it possible to develop more effective ST interventions—at least in countries where the phenomenon has been extensively investigated. In particular, adequate replication of the classes empirically validated in this study should help corroborate mathematics/ quantitative SST classes, which could eventually be used to tailor interventions for different SST groups. Findings from these studies would make it possible for researchers to tailor interventions to meet the unique needs and challenges of students belonging to the different classes of susceptibility to ST (SST). Such focused interventions are likely to be more effective in attenuating ST among the vulnerable. Therefore, findings from this study serve as a bridge towards a deeper understanding of ST, and to the development of more specialized, and potentially highly effective ST reduction interventions.

Lastly, this study has not only demonstrated that the psychometric properties of the SIAS remain consistently strong, but it has also provided further evidence of the validity of the scale in appropriately discriminating latent SST profiles by gender and academic domain, as discussed previously. The properties of the SIAS (i.e., strong reliability, and discriminant validity) speak to its potential as an effective SST classification tool that intervention researchers can use to develop differentiated ST reduction strategies for students. For instance, the SIAS could be used in mixed methods research, where students belonging to different SST profiles are identified and interviewed. This information could then be used to tailor interventions for each of these SST groups accordingly. Because ST affects students from stereotyped groups who actually have the desire, skills and capacity to be successful in the stereotyped domain (Steele, 1997), specialized (as opposed to generic) ST interventions are more likely to benefit a broader range of 'at-risk' students, and as a result seal areas in the leaking STEM pipeline resulting from stereotype threat related factors.

Limitations and recommendations for future research

It is worth mentioning that although a large proportion of women in STEM were classified as high SST, notifying them about their status and offering remediation was not considered because the SIAS is stereotype threat (ST) susceptibility and not a stereotype threat experience measure. Susceptibility to ST does not automatically mean experience of ST. Studies show that whether individuals susceptible to ST or not they actually experience the phenomenon is moderated by several other contextual and intra-personal psychological factors like locus of control (Cadinu et al., 2006), performance expectancies (Cadinu et al., 2003), and stereotype endorsement (Erikson & Lindholm, 2007), to name a few. Therefore, it is possible for one to be highly susceptible to ST but not experience the phenomenon depending on one's ranking on these and other ST moderators. As such, highlighting one's susceptibility to a deleterious phenomenon based (1) solely on an instrument (the SIAS) in its nascent stages of development, and (2) using a classification system (latent profile analysis) that is itself not deterministic but probabilistic, would have been both premature, and unethical since it would have increased the likelihood of making wrong diagnoses, and creating problems where there might have been none, to begin with. To misconstrue ST susceptibility for ST experience based solely on the SIAS and to act accordingly could have had damaging ramifications for high SST classified individuals. Potential problems might have included, self-fulfilling prophecies for students who might not have otherwise been affected by the phenomenon, for example. Researchers are therefore cautioned against using the SIAS for diagnostic purposes at this stage in its development.

Although the ultimate goal would be for the SIAS to be used as a classification tool, considerable work is yet to be done in this area. The current study reported the first step towards this goal-- making sure that the SST susceptibility profiles assessed by the SIAS aligned with Steele's (1997) SST classifications. Accordingly, findings from this study provided support of the SIAS in this regard. The next steps involve establishing criterion validity, which would involve not only replicating these findings across samples, but also conducting qualitative research (e.g. phenomenology) with students from different susceptibility classes to understand the experience better and using this information to potentially refine the instrument further if necessary. Once these steps have been taken, the SIAS could be used in tandem with other external measures (e.g. measures of academic achievement and motivation) to tailor interventions members of marginalized groups at risk of experiencing ST. And even then, the SIAS would first have had to strongly demonstrate criterion validity—which is the next step in the development of the scale.

Being the first study of its kind to investigate typologies of ST susceptibility, caution should also be exercised in interpreting the results reported here. Specifically, the results presented here are not immutable but rather offer a basis for further exploration and replication of ST susceptibility in different domains. Certainly, more replication in this area is strongly encouraged before drawing strong conclusions about the differential nature of ST susceptibility in STEM versus non-STEM domains. It is recommended that future replication studies be conducted using samples from different populations e.g., high schools, and samples from populations or nations culturally distinct from the United States, where ST research has been conducted. To the best of my knowledge, this is the first study to investigate latent profiles of ST susceptibility. The present study showed that ST susceptibility classes differed as a function of academic domain and gender. It would have been useful to add ethnicity as a latent factor in examining ST susceptibility profiles of students but there weren't enough ethnic minorities satisfying the categories of interest (i.e. minority males and females in STEM and non-STEM) to facilitate the analyses. Future studies should examine how SST profiles vary among students from marginalized groups when ethnic group identification and stigma consciousness of the same are added as latent SST factors.

Also, although the data provide preliminary support for the possible existence of math-related stereotype threat among Colombian women majoring in STEM, this dearth of experimental research directly investigating ST in Latin America bounds the interpretation of results from the Colombian sample in this study. That is, further empirical research directly testing ST effects, and the causal relationship between ST and disengagement in Latin America, is required to corroborate the results reported in this study.

Conclusion.

This study provides preliminary support for Steele's theoretical SST profiles. As such, it opens up new directions for ST research, especially where interventions are concerned. Based on this and other studies (e.g. Smith & Cokely, 2016), the SST measure, SIAS, had demonstrated strong psychometric properties. It is therefore hoped that researchers can take advantage of the SST measure and use it to select appropriate study samples (i.e. participants for whom ST is likely to apply). This simple measure of improving sample selection in ST studies could make a difference in producing more accurate estimates of ST effects.

References

Aguiar, M., Gutiérrez, H., Lara-Barragán, A. y Villalpando, J. (2011). El rendimiento académico de las mujeres en matemáticas: análisis bibliográfico y un estudio de caso en educación superior en México. *Revista Electrónica Actualidades Investigativas en Educación*, *11*(2), 1-24.

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management, 39*, 1490-1528.

Armenta, B. E (2010). Stereotype Boost and Stereotype Threat Effects: the moderating role of ethnic identification. *Cultural Diversity and Ethnic Minority Psychology, 16* (1), 94-98 doi:10.1037/a0017564

Aronson, J., Fried, C. B & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology*, *38,* 113–125 doi:10.1006/jesp.2001.1491

Aronson, J., Lustina, M. J., Good, C., Keough, K., Steele, C. M., & Brown, J. (1999). When white men can't do math: Necessary and sufficient factors in stereotype threat. *Journal of Experimental Social Psychology, 35(1),* 29-46.

Baldeón-Padilla, D. S., Valencia-Serrano, M. & Alvarado-Bueno, J. I. (2020). Amenaza de estereotipo, género y desempeño académico en matemáticas. *Magis, Revista Internacional de Investigación en Educación*, *13*, 1-22. doi: 10.11144/ Javeriana.m13.aegd

Beilock, S. L., Rydell, R. J. & McConnell, A. R., (2007). Stereotype threat and working memory: Mechanisms, alleviation & spillover. *Journal of experimental psychology, 136* (2) 256-276 doi:10.1037/0096-3445.136.2.256

Bowen, N. K., Wegmann, K. M., & Webber, K. C. (2012). Enhancing a briefwriting intervention to combat stereotype threat among middle-school students. *Journal of Educational Psychology*. Advance online publication. doi: 10.1037/a0031177

Brown, R.P., & Pinel, E.C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology, 39,* 626-633 doi:10.1016/S0022-1031(03)00039-8

Busse, C., Kach, A. P., & Wagner, S. M. (2016). Boundary conditions: What they are, how to explore them, why we need them, and when to consider them. *Organizational Research Methods*. Advance online publication. doi:10.1177/1094428116641191

Cadinu, M., Maas, A., Rosabianca, A., Figerio, S., & Latinotti, S. (2003). Stereotype Threat: the effect of expectancy performance. *European Journal of Social Psychology, 33,* 267-285. doi: 10.1002/ejsp.145

Cadinu, M., Maas, A., Rosabianca, A., Lombardo, M., & Figerio, S. (2006). Stereotype Threat: The moderating role of locus of control beliefs. *European Journal of Social Psychology, 36*, 183-197. doi:10.1002/ejsp.303

Casad, B., Petzel, Z., & Ingalls, E. (2018). A Model of Threatening Academic Environments Predicts Women STEM Majors' self-esteem and engagement in STEM. *Sex Roles, 80,* 469-488, https://doi.org/10.1007/s11199-018-0942-4

Cervini, R. & Dari, N. (2009). Género, escuela y logro escolar en matemática y lengua de la educación media. *Revista Mexicana de Investigación Educativa, 14*(42), 1051-1078.

Correa, J. B. (2015). Desempeño académico y diferencias de género en Colombia: un análisis con base en las pruebas TIMSS 2007. *Sociedad y Economía, 30,* 15-42.

Del Río, M. F., Strasser, K. y Susperreguy, M. I. (2016). ¿Son las habilidades matemáticas un asunto de género? Los estereotipos de género acerca de las matemáticas en niños y niñas de kínder, sus familias y educadoras. *Calidad en la educación, 45,* 20-53.

Dennehy, T. C., Smith, J. S., Moore, C., & Dasgupta, N. (2018). Stereotype threat and stereotype inoculation for underrepresented students in the first year of college. In R. S. Feldman (Ed.), *The first year of college: Research, theory, and practice on improving* *the student experience and increasing retention* (p. 309–344). Cambridge University Press.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: *A meta-analysis. Psychological Bulletin*, *136*(1), 101–127

Eriksson, K., & Lindholm, T. (2007). Making gender matter: The role of gender-based expectancies and gender identification on women and men's math performance in Sweden. *Scandinavian Journal of Psychology, 48*, 329–338 doi:10.1111/j.1467-9450.2007.00588.x

Espinosa, C. (2010). Diferencias entre hombres y mujeres en educación matemática: ¿Qué pasa en México? *Investigación y Ciencia, 18*(46), 28-35.

Espinosa, G. (2004). Currículo y equidad de género en la primaria: Una mirada desde el aula. Estudio realizado en tres escuelas estatales de la ciudad Lima. En M. Benavides (Ed.), *Educación procesos pedagógicos y equidad. Cuatro informes de investigación* (pp. 69-129). Lima, Perú. Retrieved from http://bibliotecavirtual.clacso.org.ar/ar/libros/peru/grade/educa/doc2.pdf

Fernández, M. C. y Hauri, S. (2016). Resultados de aprendizaje en la araucanía. La brecha de género en el Simce y el androcentrismo en el discurso de docentes de lenguaje y matemática. *Calidad en la Educación, 45,* 54-89.

Figueroa, A. & Ortega, M. (2010). Condición de género y elección profesional. El área de físico-matemático en las mujeres. *Investigación y Ciencia*, *18*(46), 18-27.

Galdi, S., Cadinu, M., & Tomasetto, C. (2014). The roots of stereotype threat: when automatic associations disrupt girls' math performance. *Child development, 85*(1), 250–263. https://doi.org/10.1111/cdev.12128

Gonzales, P. M., Blanton, H., & Williams, K. J. (2002). The effects of stereotype threat and double-minority status on the test performance of Latino women. *Personality and Social Psychology Bulletin, 28,* 659–670. doi:10.1177/0146167202288010

González, D. E. (2009). Factores individuales que afectan la demanda de educación superior en ingenierías: caso de la Pontificia Universidad Javeriana de Cali. *Cuadernos de administración, 22*(39), 307-333.

González-Gutiérrez, N., Sepúlveda-Delgado, O., & Espejo-Lozano, R. (2018). Formación matemática en Colombia: una mirada desde una perspectiva de género. *Revista de Investigación, Desarrollo e Innovación, 8*(2), 251-264. https://doi.org/10.19053/20278306.v8.n2.2018.7519

González, R. M. (2005). Un modelo explicativo del interés hacia las matemáticas de las y los estudiantes de secundaria. *Educación Matemática, 17*(1), 107-128.

González-Pineda, J., Fernández-Cueli, M., García, T., Suárez, N., Fernández, E., Tuero-Herrero, E. y da Silva, E. (2012). Diferencias de género en actitudes hacia las matemáticas en la enseñanza obligatoria. *Revista Iberoamericana de Psicología y Salud, 3*(1), 55-73.

Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Applied Developmental Psychology, 24*, 645-662. doi:10.1016/jappdev.2003.09.002

Huguet, P., & Regner, I. (2007). Stereotype Threat among school girls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology, 99*, 545-560

Hudziak, J.J., Heath, A.C., Madden, P.F., Reich, W., Bucholz, K.K., Slutske, W., Bierut, L.J., Neuman, R.J., & Todd, R.D. (1998). Latent class and factor analysis of DSM-IV ADHD: A twin study of female adolescents. *Journal of American Academy of Child and Adolescent Psychiatry, 37*, 848–857 doi:10.1097/00004583-199808000-00015

Johns, M., Schmader, T., & Martens, A. (2005). Knowing is half the battle: Teaching stereotype threat as a means of improving women's math performance. *Psychological Science 16*,175-179. doi:10.1111/j.0956-7976.2005.00799.x

Keller, J., & Molix, L. (2008). When women can't do math: the interplay of self-construal, gender identification, and stereotypic performance standards. *Journal of Experimental Social Psychology, 44*, 437-444 doi:10.1016/j.jesp.2007.01.007

Keller, J. (2010). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British Journal of Educational Psychology*, *77*(2), doi: 10.1348/000709906X113662

Kit, K. A., Tuokko, H. A., & Mateer, C. A. (2008). A review of the stereotype threat literature and its application in a neurological population. *Neuropsychology Review*, *18*(2), 132-148.

Klein, S. B. (2014). What can recent replication failures tell us about the theoretical commitments of Psychology? *Theory and Psychology, 24*(3), 326-338.

Lubke, G., & Neal, M. C. (2006). Distinguishing between latent classes and continuous factors: Resolution by maximum likelihood? *Multivariate Behavioral Research, 41*(4), 499–532 doi: 10.1207/s15327906mbr4104_4.

March, D. A. (2009). *Diferencias de género en rendimiento académico: Efectos en la evaluación a nivel escolar* (Memoria). Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Santiago, Chile. Retrieved from http://www.ideaseneducacion.cl/wp-content/uploads/2009/06/memoria-daniela-march.pdf

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, *97*, 951-966.

Mella, O. (2006). Factores que afectan los resultados de la escuela pública chilena. Revista *Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación (REICE)*,*4*(1), 29-37. Retrieved from http://redalyc.uaemex.mx/pdf/551/55140104.pdf

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330,* 1234-1237. doi: 10.1126/science.1195996

Ministerio de Educación Nacional (2018). *Matriculados en educación superior - Información 2018* – Colombia 2018. [Date file]. Retrieved from https://www.mineducacion.gov.co/sistemasinfo/Informacion-a-la mano/212400:Estadisticas

Muthén, B. (2006). Should substance use disorders be considered as categorical or dimensional? *Addiction, 101* (Suppl. 1), 6–16 doi:10.1111/j.1360-0443.2006.01583.x.

Neuman, R.J., Todd, R.D., Heath, A.C., Reich, W., Hudziak, J.J., Bucholz, K.K., Madden, P.A.F., Begleiter, H., Porjesz, B., Kuperman, S., Hesselbrock, V., & Reich, T. (1999). The evaluation of ADHD typology in three contrasting samples: A latent class approach. *Journal of American Academy of Child and Adolescent Psychiatry, 38*, 25–33 doi:10.1097/00004583-199901000-00016

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., et al. (2009). *National differences in gender-science stereotypes predict national sex differences in science and mathematics achievement.* Retrieved on 10/10/2013 at www.pnas.org_cgi_doi_10.1073_pnas.0809921106

Picho, K. (2011). *Stereotype Threat in Context: The psychosocial experience of Stereotype Threat susceptible urban high school girls in honors mathematics classes.* (unpublished dissertation). University of Connecticut, Storrs.

Picho, K., & Brown, S.W. (2011). Can stereotype threat be measured? A validation of the social identities and attitudes scale (SIAS). *Journal of Advanced Academics*, *22* (3), 374 – 411

Picho, K. & Schmader, T. (2018). When do stereotypes impair math performance? A study of stereotype threat among Ugandan adolescents. *Sex Roles*, *78*(3), 295-306. doi: 10.1007/s11199-017-0780-9

Picho, K. & Stephens, J. M. (2012). Culture, context and stereotype threat: A comparative analysis of single-sex and coed schools in Uganda. *Journal of Educational Research, 105*, 52-63. doi:10.1080/00220671.2010.517576

Rasmussen E. R., Neuman R. J., Heath A. C., Levy F., Hay D.A, Todd R.D. (2002). Replication of the latent class structure of attention-deficit/hyperactivity disorder (ADHD) subtypes in a sample of Australian twins. *Journal of Child Psychology and Psychiatry, 43,* 1018–28

Reali, F., Jiménez-Leal, W., Maldonado-Carreño, C., Devine, A., & Szücs, D. (2016). Examining the link between math anxiety and math performance in Colombian students. *Revista Colombiana de Psicología, 25*(2), 369-379. doi: 10.15446/rcp.v25n2.54532

Robinson, J. P., & Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: Examining direct cognitive assessments and teacher ratings. *American Educational Research Journal, 48* (2), 268-302. doi: 10.3102/0002831210372249

Salazar, L., Hidalgo, V. y Blanco, H. (2010). Estudio sobre diferencias de género en el aula de matemáticas. *Revista Latinoamericana de Etnomatemática, 3*(2), 4-13.

Samuelson, M. K. & Dayton, C. M. (2010). Latent class analysis. In G.R. Hancock, & R.O. Mueller (Eds): The reviewer's guide to quantitative methods in the social sciences, p 173-184Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38,*194-201 doi:10.1006/jesp.2001.1500

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38,*194-201

Schmader, T. & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 88,* 934-947 doi:10.1037/0022-3514.85.3.440

Schmader, T., Johns, M & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115* (2), 336-356 doi: 10.1037/0033-295X.115.2.336.

Seyranian, V., Madva, A., Duong, N., Abramzon, N., Tibbetts, Y. & Harackiewicz, J. (2018). The longitudinal effects of STEM identity and gender on flourishing and achievement in college physics. International Journal of STEM *Education, 5,* 40. https://doi.org/10.1186/s40594-018-0137-0

Smith, L. V., & Cokley, K. (2016). Stereotype Threat Vulnerability: A Psychometric Investigation of the Social Identities and Attitudes Scale. *Measurement and Evaluation in Counseling and Development, 49* (2), 145-162. doi:10.1177/0748175615625752

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4–28. https://doi.org/10.1006/jesp.1998.1373

Steele, C. M. (1997) A threat in the air: how stereotypes shape intellectual identity and performance. *American Psychologist, 52*, 613-629 doi:10.1037/0003-066X.52.6.613

Steele, C.M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African-Americans. *Journal of Personality and Social Psychology, 68*, 797-811 doi:10.1037/0022-3514.69.5.797

Steinberg, J. R., Okun, M. A., & Aiken, L. S. (2012). Calculus GPA and math identification as moderators of stereotype threat in highly persistent women. *Basic and Applied Social Psychology, 34*(6), 534–543. https://doi.org/10.1080/01973533.2012.727319

Tomasetto, C. & Appolini, S. (2013). A lesson not to be learned? Understanding stereotype threat does not protect women from stereotype threat. *Social Psychology of Education, 16*, 199-213. doi: 10.1007/s11218-012-9210-6

Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology, 47*(4), 943–949. https://doi.org/10.1037/a0024047

Valentova, J. V., Otta, E., Silva, M. L. y McElligott, A. G. (2017). Underrepresentation of women in the senior levels of Brazilian science. PeerJ 5:e4000. https://doi.org/10.7717/peerj.4000

Winkler, D. (2004). Investigaciones sobre etnicidad, raza, género y educación en las américas. In D. Winkler y S. Cueto (Eds.), *Etnicidad, raza, género y educación en América Latina* (pp. 1-32). Perú: PREAL. Retrieved from http://thedialogue.org/PublicationFiles/EtnicidadRazayGenero.pdf

Received: 17-09-2019
Accepted: 23-06-2020