DO (OUTGROUP) TEACHERS ALWAYS HEIGHTEN ANXIETY AND THREATEN STEM STUDENTS' CLASSROOM OUTCOMES?

JAYA KUMAR KARUNAGHARAN

Thesis submitted to the University of Nottingham Malaysia

For the degree of Doctor of Philosophy in Psychology

January 2023

Abstract

Malaysia, along with the rest of the world, needs more Science, Technology, Engineering and Mathematics (STEM) literate professionals. However, the persistent gender gaps in STEM make it difficult to meet the demand and stand in the way of scientific and economic progress. Although minority group participation in education has improved over the years, generally, fewer ethnic minority and female students tend to pursue STEM-related fields. One prominent explanation is that there aren't enough ingroup role models in STEM education to boost confidence in pursuing this profession, given the anxiety faced by underrepresented groups due to the stereotype that STEM is not for them. However, despite the positive effects of role models, these effects are disproportionate which suggests that there could be other factors that counter the inoculating effect of role models.

For this PhD thesis, five studies were conducted. Two surveys were conducted to confirm the assumptions regarding the sources of anxiety within the classroom and the effect of anxiety on students' confidence in their classroom outcomes, as well as to identify meta-stereotypes that students expect their teachers to have of them. Subsequently, three field experiments were used to evidence the impact that exposure to ingroup (versus outgroup) role models could have on stereotyped underrepresented STEM students' educational outcomes as a function of *ingroup spotlighting* (being given heightened attention by the increased exposure to exemplary counter-stereotypical ingroup role models) and stereotype-related anxieties.

The surveys were conducted using established tools of self-reported measures that showed anxiety being a reality in the classroom setting especially in relation to outgroup teachers with meta-stereotypes being typically negative and the concerns being more pronounced in the context of the competence related traits especially in relation to the outgroup teacher. The field studies were conducted in actual classroom experiments, using a

novel measure of voice jitter as a proxy for anxiety. The hypothesis was that ingroup spotlighting would exacerbate the pressure to likewise excel in STEM. The results corroborated the assumptions that role models could have positive effects on underrepresented students' academic outcomes but differs from these assumptions when the students suffer from an imposter syndrome and feel "spotlighted" by the counter-stereotypical ingroup role model.

The present findings could challenge the current thinking around the persistent gender and racial gap in STEM and will help to inform a robust/inclusive government policy around STEM education in Malaysia.

Keywords. STEM; anxiety; meta-stereotypes; classroom outcomes; teacher identity; role model; stereotype inoculation; ingroup spotlighting

List of Publications

Karunagharan, J. K., & Owuamalam, C. K. (March, 2022). The impact of role models on "spotlighted" underrepresented students in STEM. Talk presented at the Malaysian Pscyhology Student Assembly (MAPSA), University of Nottingham Malaysia, Semenyih. (Winner of Best Presenter Award)

Karunagharan, J. K., & Owuamalam, C. K. (July, 2017). Do outgroup teachers enhance or diminish academic outcomes of stereotyped students relative to ingroup teachers? In Editorial Universidad de Granada (Ed.), 18th General Meeting of The European Association of Social Psychology: Programme and abstract book (pp. 130). Granada, Spain: Universidad de Granada. ISBN: 978-84-338-6096-5.

Karunagharan, J. K., & Owuamalam, C. K. (May, 2017). Do outgroup teachers enhance or diminish academic outcomes of stereotyped students relative to ingroup teachers? Talk presented at the Malaysian International Psychology Conference 2017, Shah Alam, Selangor, Malaysia.

Karunagharan, J. K., & Owuamalam, C. (September, 2015). Shedding light on the Stereotype Inoculation Model by infusing predictions from the Stereotype Content Model. Talk presented at the British Psychological Society Development Section and Social Section Annual Conference 2015, Manchester, England.

Karunagharan, J. K., & Owuamalam, C. (April, 2015). West is best? Effects of race-based status and stereotypicality on anxiety and classroom participation of Malaysian science students. Talk presented at the Society of Australasian Social Psychologist Conference 2015, Newcastle, NSW, Australia. (Winner of the Margaret Foddy Postgraduate Travel Award)

Acknowledgements

My utmost appreciation to my God and Saviour, Jesus Christ, who is a Keeper of promises. Throughout this journey, His promise in Isaiah 41:10 was a source of strength, “Fear not, for I am with you; be not dismayed, for I am your God; I will strengthen you, I will help you, I will uphold you with my righteous right hand.” The submission of this thesis is evidence of the fulfilment of that promise!

I have also had the privilege of working with an amazing supervision team. I must mention Dr Chuma Owuamalam who is the primer and pillar of the work reported here. The late Prof Ganakumaran Subramaniam was always a source of inspiration and special thanks goes to Dr Hazel Ramos and Dr Michael Mustafa for stepping in to help push this through to the finish line.

To make this possible, some of the biggest sacrifices came from my wife, Su Mei. She has been on the receiving end of both extremes: diarrhoea of emotions and coldness of distance. She has been a champion as well as selfless and patient through it all.

I am grateful for the great host of helpers who surrounded me and cheered me on by their love, support, and prayers. My mothers, Amma and Mum, who constantly cover me with love and prayers. My loving family who sticks by me no matter what and my spiritual family in BTGC, KGC and many other places who prayerfully cheer me on. My helpful colleagues who were generous with their time to assist where needed. And the host of people who have had a hand in making this possible: Prof Andrew Morris, Prof Kuan Hon Lim, Prof Festo Massawe, Dr Matthew Ashfold, Coey Wong, Dr Mishaari Weerabangsa, Dr Treshi-Marie Perera, Prof Steve Janssen, Dr Marieke de Vries, Prof Martijn Meeter, Daniel Lutscher, Lianne de Vries, and the School of Psychology in UNM.

Finally, I dedicate this work to the three arrows in my quiver: Asya, Dhanya and Trisha. May you live out your God given potential. Papa loves you to the moon and back!

Table of Contents

List of Tables and Figures

Tables

Figures

Chapter One

Introduction

"*The need for a vast, talented workforce in STEM-related fields has never been more necessary…*" – Bridget Long, Dean of the Harvard Graduate School of Education (O'Rourke, 2021)

The world needs more Science, Technology, Engineering and Mathematics (STEM) literate professionals and leaders as it becomes more digitalised and demand for healthcare-related skills increase amidst the pandemic threat. *STEM fields represent a group of academic and professional disciplines that form the basis for jobs such as (and not limited to) biologists, software developers, engineers, data scientists, chemists, mathematicians, and doctors.* Across the globe, growth in STEM fields can be seen in various ways. The Edge Malaysia cited the Malaysian Digital Economy Corporation (MDEC) which found that the number of digital job vacancies in Malaysia almost tripled from 19,000 in June 2020 to more than 56,000 in April 2021 (Nair, 2021). Many of the top 20 fastest growing occupations in some developed economies, like the U.S., also require knowledge and skills in STEM (U.S. Bureau of Labor Statistics, 2022). In 2018, the Pew Research Center reported that the STEM employment had grown 79% since 1990 and computer jobs increased 338% over the same period in the U.S. (Funk & Parker, 2018).

It has been projected that STEM jobs in the U.S. would grow 11% from 2020 to 2030 (O'Rourke, 2021). The Australian Government's Department of Education also predicts that workers in the future would spend more than double the amount of time spent today on work tasks requiring STEM skills such as science, maths, and critical thinking (Department of Education, 2022). The United Nations' 2030 Agenda for Sustainable Development serves as an impetus for further growth in STEM as its policies established 17 Sustainable Development Goals (SDGs) to address major global issues such as climate change and food

shortage. To achieve the SDGs, STEM literacy would play a vital role as it would be the conduit for innovative solutions contributed by a multidisciplinary scientific workforce that has the know-how and skill set to develop new technology (Ng, 2019).

Considering the need for STEM literacy, the Ministry of Education in Malaysia gave attention to STEM education from as early as 2010, and as part of the process, STEM education initiatives have been incorporated into the Malaysian Education Blueprint 2013-2025 (Halim et al., 2021). The aims of these initiatives are to equip students with the necessary skills to face a world that is being transformed by STEM and to sufficiently lay a solid foundation at the school level to ensure that Malaysia has enough STEM graduates for the industries. The plan is to do this by raising interest among students, improving delivery by enhancing the teaching skills and by increasing STEM awareness among students and parents (Ministry of Education Malaysia, 2013). Indirectly affirming the Malaysian initiatives as steps in the right direction, the Australian Government's Chief Scientist reported in 2014 that high performing countries would consistently produce STEM graduates with skills that are highly sought after by employers and a STEM-literate population that values discovery and entrepreneurship (Office of the Chief Scientist, 2014).

Although the growth and prospects of STEM related education and occupation has been promising, the persistent racial and gender gaps in STEM make it difficult to meet the demand and stand in the way of scientific and economic progress (Beede, et al., 2011; UNESCO, 2012). Reports indicate that women and ethnic minorities are underrepresented in STEM fields relative to their representation in the overall population (Goy, et al., 2018; National Science Board, 2015). Although women made up about half of the employed U.S. workforce in general in 2019, they made up only 34% of the people employed in the STEM workforce specifically and only made up a small share of scientists and engineers. This overall underrepresentation of women in STEM is contributed by the significant gender gap

in occupations related to engineering (16% women) and computer science (26% women) (National Science Foundation, 2022). As for leadership, the information technology industries globally had less than a quarter of its top leadership positions being women while the representation of women among board of directors was only 28% in 2021 (Catalyst, 2020). Turning the focus to education, UNESCO recently reported that only 30% of STEM students in higher education globally are women (UNESCO, 2017).

In Malaysia, although female participation in education has improved over the years, lesser female students tend to pursue STEM-related fields even though female students, in total, outnumber their male counterparts (Goy, et al., 2018; Ministry of Education Malaysia, 2018; Sultana Alam, et al., 2021). The irony is that an annual Graduate Tracer Study (GTS) by the Ministry of Higher Education (MOHE) found that women made up a larger proportion of STEM graduates from the Malaysian higher education institutions (53.2% in 2021) and yet, it is the men STEM graduates that are being hired at a higher rate than women (Gong, 2023). *Zooming in on specific STEM related skill sets, Gong (2023) also reported that in Malaysia, there are larger gender gaps when it comes to possessing more advance skills in the area of information and communications technology (ICT) such as configuring software, with men surpassing women. The skills gap was attributed to the opportunities of on-the-job experience afforded to men. In Malaysia's general labour force, the consistent participation over the last five years has been skewed towards men with about 80% working or looking for work as compared to women with about 55%* (Department of Statistics Malaysia, 2023). *The GTS by MOHE also suggests that though Malaysia is better compared to the global trends when it comes to STEM education, it is still showing a large gap when it comes to employment* (Gong, 2023). *Women are hired at lower rates than men even though they make up the larger percentage of STEM graduates. The gender gap problem in Malaysia and the leaky pipeline persists and is more acute in fields like engineering. The Board of Engineers*

Malaysia reported in 2019 that a mere 29% of registered graduate engineers and 8% of registered professional engineers in Malaysia were women (Board of Engineers Malaysia, 2019).

In general, the gender gap in STEM is well documented as evidenced by a survey by Dos Santos et al. (2022) of the bibliography of papers published and included in the Web of Science. In their search of papers published between 2015 and 2019, they managed to identify 130 papers that addressed the gender gap in STEM. *They also suggested that the gender gap in both STEM education and careers is due to cultural factors or social context, more specifically the low number of role models for women, instead of biological or genetic factors.*

Similar patterns are found with regards to ethnic minorities where the underrepresentation extends to leadership positions across STEM-related industries, academia, and the federal workforce (Wu & Jing, 2013). The National Science Foundation in the U.S. reports that Blacks or African Americans, Hispanics or Latinos, and American Indians or Alaska Natives racial and ethnic groups are still underrepresented in science and engineering degrees relative to their representation in the overall population although the share has increased over the past decade (National Center for Science and Engineering Statistics, 2021). While these ethnic minorities made up 30% of the overall employed population in the U.S., it only represented 23% of the STEM workforce in 2019 which also stems from an underrepresentation at the higher education level (National Science Foundation, 2022). Among underrepresented minorities (URM), a steady decline in percentage at the higher education level has been observed, reducing from 14.7% obtaining STEM Bachelor's degrees to 12.6% STEM Masters degrees to only 8.3% STEM Doctorates. The decline continued as those with STEM Doctorates who worked as academics were down to a mere 7.3% (Estrada, et al., 2016).

In Malaysia, there is a pervasive racial gap in education in general which affects the STEM fields as well. Despite the perception of being an equal opportunity platform, the Malaysian education system is criticised for perpetuating racial disparities. Anne (2022) in an article from the Institute for Democracy and Economic Affairs (IDEAS), highlighted a study that found nine out of ten ethnic Indian respondents felt racially discriminated against in school because of their ethnicity, skin colour, and religious beliefs. Admissions of Chinese and Indian minorities are side-lined which affect number of these groups within STEM education (Anne, 2022). However, there is a lack of race-based figures within STEM in the Southeast Asian region and yet, considering the systemic issues related to race, it is important to include the racial context within the studies in this thesis.

Women from ethnic minority groups have an even lower representation in STEM fields (Kricorian et al., 2020; Stockard et al., 2021). Pietri et al. (2019) reported how because of both their gender and ethnicity, Latinas face *double jeopardy* in STEM. Some scholars have termed this the *double bind* (Malcom et al., 1976) and the editors of the Harvard Educational Review highlighted that the meeting of both race and gender in STEM magnifies the disparity in achievement levels of women of colour as compared to white women or men of colour. The disparity is especially acute at the most advanced levels (Unraveling the double bind: Women of color in STEM, 2011). This double bind phenomenon which leads to the underrepresentation of ethnic minority women in STEM has been partly attributed to social identity threat (Steele et al., 2002) which in turn could result in concerns about belonging or lack of interest, as well as low performance in STEM (Ireland, et al., 2018). Therefore, it is becoming more and more apparent that the need to understand and address this gender and racial gap is urgent, especially when its impact on academic self-concept, subject interests and, ultimately, career aspirations and choices is great (Makarova et al., 2019).

Problem Statement

In response to the underrepresentation in STEM of groups like ethnic minorities and women, there is a call to empower students from underrepresented groups to feel more confident about their place in STEM. One way forward being proposed by the *stereotype inoculation model* for example, is increasing the number of ingroup experts as role models, which could provide identity-safe cues that boost confidence and neutralise the stereotype related anxieties faced by underrepresented groups in STEM (e.g., more teachers who are minorities; see Olsson & Martiny, 2018, for a review; Dasgupta, 2011; Pietri et al., 2019; Stout et al., 2011). However, the problem is despite the positive effects of role models, these effects have been and are projected to be disproportionate which suggests that there could be other factors that counter the inoculating effect of role models. Not enough is known about these other factors that could counter the ingroup role model effects and whether outgroup teachers *always* threaten STEM students' classroom outcomes. So, this PhD research aims to establish if teachers, specifically those that are of a different social identity from the students (i.e., the outgroup), *always* threaten stereotyped underrepresented STEM students' academic related outcomes *when social identity is salient* and it also explores the question of the possibility that the academic outcomes of underrepresented groups could be worsened by the fact that they are "spotlighted" by the increased exposure to exemplary counter-stereotypical ingroup role models. It is crucial to fill in these gaps as they suggest a more nuanced approach in intervention to increase the representation of ethnic minorities and women in STEM as compared to the broad-brush approach of increasing ingroup role models currently taken.

Significant Contributions

Building upon previous research on intergroup anxiety, meta-stereotypes, stereotype threat and role model effects, the present research programme attempts to address gaps in three main ways and in doing so, makes important contributions to knowledge in the field.

First: *Empirically*, this research **provides evidence for the sources of anxiety** within the classroom and the effect of anxiety on students' confidence in their classroom outcomes. It then **identifies meta-stereotypes** that students expect their teachers to have of them. Finally, **evidence of the impact an ingroup (versus outgroup) expert/role model** has on STEM performance in racial groups and women are presented.

Second: *Methodologically*, there has not been any previous study on intergroup anxiety, to the best of the author's knowledge and through search in peer-reviewed databases, that employed a **direct physiological measure of anxiety like vocal jitter**. This novel measure of anxiety has been applied in linguistic research, but its use has not extended into social psychology studies. Vocal jitter was used here as an involuntary index of anxiety to provide a less biased and more conclusive test of stereotype threat effects on anxiety and performance.

Third: *Theoretically*, this study proposes the **ingroup spotlighting thesis**, which is a form of spotlighting that stems from individuals' own group (ingroup) but that can negatively affect underrepresented students' outcomes in STEM *because exposure to an ingroup expert might put pressure on the students to also perform in STEM fields (a more detailed discussion on how ingroup spotlighting negatively affects underrepresented students' outcomes in STEM provided in Chapter 2)*. As far as the author knows, there has not been any systematic exploration of this potential side effect of exposing students from underrepresented backgrounds to ingroup STEM experts.

Research Questions

Specifically, and systematically, this thesis sought to answer the following research questions in a series of five studies:

1. Is there empirical evidence to validate the assumptions regarding the sources of anxiety within the classroom? Does such anxiety have an impact on students' confidence in classroom outcomes? (As presented in Chapter 4)
2. What is the specific content of the meta-stereotypes in the student versus teacher context? Do students' stereotype concerns differ depending on the social identity of the teacher? (As presented in Chapter 5)
3. Are there positive or negative impacts of an ingroup (versus outgroup) expert/role model on STEM performance in underrepresented groups as a function of stereotype threat and anxiety? (As presented in Chapters 6 and 7)
4. Do the *ingroup spotlighting thesis* or the *stereotype inoculation model* provide better explanations for the poorer outcomes of underrepresented groups in the STEM context of higher education? (As presented in Chapters 6 and 7)

In the process of answering these empirical and theoretical research questions, an important methodological question was also addressed: does vocal jitter provide an accurate, less biased, and more conclusive test of stereotype threat effects on anxiety and performance? (See Chapter 6)

Summary of Thesis Chapters

Ultimately, this PhD thesis aims to establish if teachers or professors (the experts/role models that students are exposed to in classrooms; Maylor, 2009), specifically outgroup teachers, *always* threaten stereotyped underrepresented STEM students' academic related outcomes *when social identity is salient* and investigate the question of the possibility that the academic outcomes of underrepresented groups could be worsened by the fact that they are

"spotlighted" by the increased exposure to exemplary counter-stereotypical ingroup role models.

The thesis begins by discussing the important theoretical underpinnings by way of a literature review in the next chapter (Chapter 2). The chapter begins by discussing the factors that contribute towards the representation gaps and underperformance of underrepresented groups in STEM. It then scans the literature on important areas that have a bearing on this thesis such as, anxiety in the classroom, meta-stereotyping, stereotype threat, stereotype inoculation model (or ingroup role model effect) and provides a preamble on the novel ingroup spotlighting thesis being posited here. The chapter ends by summarising the hypotheses for each of the five empirical studies conducted.

Considering that it was important to explain the approach taken across the empirical chapters, Chapter 3 then goes on to outline the rationale behind some of the significant choices made in the research methodology for this thesis. The chapter summarises the ethical considerations within the research framework and then discusses the approach taken to address data missingness, the reasons behind choosing to do field classroom experiments, the theory that informed the development and use of the novel anxiety measure, and the justifications for the stereotype salience manipulations used in the empirical studies. The chapter ends by also outlining the rationale that informed some of the choices made regarding the manipulation of significant variables that needed to be separated out for more accurate conclusions.

The empirical chapters in this thesis begin with Chapter 4 (Study 1) validating the assumptions regarding the sources of anxiety within the classroom setting and how such anxiety could have important consequences for students' confidence in their classroom outcomes. Malaysian students reported shame and embarrassment, a concern over their social image/identity, as the largest source of anxiety which was highest in the context of their

teachers. The results showed that the increased anxiety caused by teachers also lowered students' self-reported confidence in participating in classroom activities.

Chapter 5 (Study 2) then proceeded to identify specific contents of the views that students expect their teachers to have of them that could also later be used to manipulate meta-stereotypes in the following studies. The study aimed to identify some of the specific contents of the stereotypes that students think their teachers hold of them, expecting that competence-related meta-stereotypic concerns would be more an issue when students are taught by an outgroup teacher compared to an ingroup teacher. Results showed that Malaysian students felt that their teachers (especially outgroup teachers) generally viewed them more negatively and this was especially true within the context of competence-related traits.

Subsequently, Chapters 6 and 7 (Studies 3-5) examined the positive and potentially negative impacts of an ingroup (versus outgroup) expert on STEM performance in racial groups and women as a function of stereotype threat and anxiety. These three empirical studies (Studies 3-5) revealed that stereotype-related anxieties can harm classroom outcomes of vulnerable students who also come from stigmatised (or underrepresented) social groups, especially when they are exposed to ingroup (rather than outgroup) teachers, which contradicts SIM but supports the newer ingroup spotlighting thesis that is being advanced in this thesis. While answering these more empirical and theoretical questions, Chapter 6 (Study 3) also addresses an important question methodologically regarding whether vocal jitter could provide a suitable test of stereotype threat effects on anxiety and performance.

Finally, general discussions and conclusions are presented in Chapter 8 especially regarding the boundary conditions of the stereotype inoculation model and the comparison of the findings in the light of the stereotype inoculation model versus the proposed ingroup

spotlighting thesis. The chapter ends with an outline of policy implications that may revolutionise the typical and blanket approach taken so far.

Chapter Two

Theoretical Underpinnings

Factors Influencing the STEM Representation Gap

In the literature, the underrepresentation of women and other minority groups within STEM fields was originally attributed to natural differences. For example, men's brains were thought to be better adapted for science and math relative to women's (see Stewart-Williams & Halsey, 2021, for a review) and intrinsic psychological factors were thought to affect choices in STEM for minority groups (Kricorian et al., 2020). Stewart-Williams and Halsey (2021) wrote that "STEM gender gaps are due in part to average sex differences in a small subset of STEM-relevant cognitive capacities, which result in somewhat more men than women having a suitable profile of aptitudes for working in some STEM fields" (p. 7). This nativist explanation has persisted across time and multiple cultures (Lippa, 2010). It finds its basis in evolution (Buss, 1995) and genetics (Iervolino et al., 2005). However, it fails to explain several other findings.

First, girls in grade school excel in math and science, and often outperform boys (Halpern, et al., 2007; Stoet & Geary, 2015). Hill et al. (2010) also highlighted significant early interest in STEM among girls. However, around puberty—a time of identity development—the advantage for girls begins to shrink and reverse (Hyde et al., 1990a). Previous studies have also shown that eventually, female students were less inclined to pursue STEM related qualifications although they outperformed their male counterparts academically (Abd Rahim et al., 2019). In semi-structured interviews conducted among women in STEM from both academia and industry, Abd Rahim et al. (2019) goes on to report that the gender disparity in STEM reaches even the top management in Malaysia with fewer women being in leadership roles within their organisations and this is propagated throughout all levels (i.e., recruitment, retention, and rehiring of women talents). *Although the semi-*

structured interviews in the Abd Rahim paper could be useful, the reported experience in the paper would be limited in its generalisability as there were only four participants from both academia and industry. Combining the qualitative interviews with a quantitative survey might have provided a more comprehensive and generalisable result.

Looking at the change over time, Hyde and team published two meta-analysis papers in the same year that examined the gender difference in mathematics attitudes and affect as well as mathematics performance *which was typically measured using tests of varying cognitive levels (e.g., computation, concepts, and problem solving) and content (e.g., arithmetic, algebra, geometry, and calculus). The assumption in these meta-analyses is that combining various studies would provide a reliable overview of gender differences in mathematics attitudes and performance even though differences in study design, cultural contexts, and educational systems in the studies also need to be taken into consideration.* In Hyde et al. (1990a), a sampling of studies from seven different sources yielding 254 independent effect sizes from 100 usable sources which represented the testing of 3,175,188 subjects, differences in mathematical problem solving that favoured men were seen emerging in high school (d = 0.29) and in college (d = 0.32), even though there were no differences in elementary and middle school. In another meta-analysis study, Hyde et al., (1990b) found that there were gender differences in attitudes and self-confidence toward mathematics among high school and college students which were lower among younger students (Hyde et al., 1990b). A subsequent meta-analysis study in 2020, looking into gender and mathematics performance from 242 studies involving 1,286,350 people published between 1990 and 2007, revealed almost similar findings such that although the gender difference neared zero for elementary and middle school students, there were small effects favouring male high school and college students (ds = +0.23 and +0.18) (Lindberg et al., 2010). These patterns were also present among later aged groups of women. For example, women expressed much less

interest in STEM majors by the time they reach first year in college and by graduation, men tend to outnumber women in STEM with a further decline in the number of women at the graduate level and the workplace (Hill et al., 2010).

These gender differences in STEM interest also permeates race. African American girls expressed positive dispositions toward STEM compared to their White counterparts but lose confidence and interest over time (Young, Young, et al., 2017). Young, Ero-Tolliver, et al. (2017) affirmed that African American girls start off having a strong affinity for STEM professions, however, the extant literature is filled with evidence for a decline over time and African American girls begin to lose interest in STEM in middle school and by the time they reach high school and college, they choose not to take STEM related courses (Hanson, 2009; King & Pringle, 2019; Pringle et al., 2012).

Taking into consideration these two general findings, several studies have suggested factors that influence this STEM representation gap. According to Robnett (2015), women in math-intensive courses encountered gender bias especially from their male peers. *Many studies, however, assume that gender gaps or disparities stem from implicit biases in education, recruitment, and workplace environments but it is important to also acknowledge that other factors such as cultural influences can also play a part.* Girls are sometimes not rewarded or encouraged by peers, and even teachers and parents, to succeed in STEM classes the same way that boys are (e.g., Lavy & Sand, 2015). Sultana Alam et al. (2021) reported that teacher stereotypes (i.e., teachers being more likely to encourage boys to take STEM subjects) negatively predicted the self-efficacy of female students to pursue STEM education (β = -.0157, p = .021) while Šimunović and Babarović (2021) reported that parental STEM-supportive behaviour predicted increase in longer term STEM interest but boys received significantly higher STEM-related support compared to girls. These effects are particulary significant in Malaysia, where the patriarchal societal norms generally expect women to

focus on the family. This ostensibly steers women away from pursuing further education, including STEM-related education (Sultana Alam et al., 2021). *Malaysia being a multi-racial society, this perspective becomes more complicated as there are also differences in the levels of patriarchy among the races. The culture of Malaysian Chinese and Malaysian Indian communities tend to be more patriachal due to closer ties with the traditional cultures of East and South Asia as compare to the Malay majority population* (Hirschman, 2016). *In Malaysia, the partriachal worldview is also closely associated with the assumption of an interest in childbearing for the women or "the motherhood penalty" which could lead to married women not being hired for jobs that are perceived as less women friendly, such as the tech industry or engineering* (Abd Wahab, 2022).

Stereotypes about the nature of STEM occupations and about the people working in STEM are also reported to push women away from STEM related fields (Gonzalez-Perez et al., 2020). In fact, Cheryan et al. (2017) described the STEM culture as being hostile to women on a regular basis. In a study of girls and women in STEM related fields, 61% reported experiencing gender bias which led to lower STEM self-concept (Robnett, 2015). Witteman et al. (2019) reported in a comprehensive study of 23,918 grant applications and 7093 principal investigators between 2011 and 2016, instead of the quality of their proposed research, it is the assessment of women as principal investigators that are less favourable that ultimately contribute towards the gender gaps in grant funding. This further evidences how hostile an environment STEM fields are for women, which could also lead to achievement gaps, *however, it does not take into account the possibility of a change in attitudes and policies about gender diversity and equity over time*. Cross-cultural research (Halpern, 2012; Reynolds et al., 2015), however, suggests that an achievement gap may result from lack of experience due to less exposure of women to tasks requiring STEM related skills stemming from the lack of opportunity.

Scanning the literature then, several intrinsic psychological factors that are relevant to this thesis may influence choices in STEM fields for underrepresented groups. These factors either interact with or underlie the factors mentioned so far. The sense of ***identity and belonging*** *in STEM is one important factor which can help an individual feel connected to the field and perceive it as a place they belong (Hurtado et al., 2009).* ***Self-efficacy*** *or the belief in one's ability to do well in STEM fields is also important as it affects the likelihood of the individual pursuing STEM related education and persisting on to STEM careers (Lent et al., 1994). Apart from the two mentioned so far, the fear of confirming negative stereotypes about an individual's social group or* ***stereotype threat*** (Steele & Aronson, 1995)*. This latter one can have a negative impact on the confidence and aspirations of underrepresented groups. Finally,* ***role models and mentoring*** *play a crucial role in impacting choices in STEM for underrepresented groups as it can positively affect self-perception within STEM (Stout et al., 2011). These psychological factors that impact the choices in STEM fields for underrepresented groups are important to this thesis and will be discussed further in the following sections and chapters.*

In summary, the representation gap in STEM fields continue to persist due to social forces—such as socialization patterns and stereotyping—that (a) discourage underrepresented groups from pursuing or persisting in STEM, (b) provide fewer opportunities for them to develop the needed competencies, and (c) cause them to underperform in STEM. It is precisely for these reasons that advocacy groups are calling for conditions that empower students from underrepresented groups to feel more confident about their place in STEM. One such initiative concerns boosting the number of individuals from underrepresented groups who could act as role models to students with similar backgrounds and provide identity-safe cues (e.g., more teachers, mentors and supervisors who are women or minorities; see Olsson & Martiny, 2018, for a review; Pietri et al., 2019). The argument is

that role models or identity-safe cues could boost confidence in pursuing professions in STEM, given the anxiety underrepresented groups face in STEM fields due to stereotypes (Dasgupta, 2011; Stout et al., 2011). The assumption made in this argument is there aren't enough role models to help inoculate the underrepresented against the stereotype-related anxiety that they do not belong in the STEM fields. A longitudinal study of 499 universities in the US over a 16-year period, however, revealed that even if the role models in science and engineering departments increases by 10%, the percentage of female majors in biological sciences, physical sciences and engineering can only be expected to increase by 1.2% (Sonnert et al., 2007). Hence, despite the positive effects of role models, these effects are disproportionate which suggests that there could be other factors that counter the inoculating effect of role models.

Anxiety in the Classroom

Student fear has been referred to as the *elephant in the classroom* (Bledsoe & Baskin, 2014) and student anxiety can be triggered by various classroom related experiences (England et al., 2017). The American Psychological Association defines anxiety as "an emotion characterized by feelings of tension, worried thoughts, and physical changes like increased blood pressure" (American Psychological Association, 2022, para. 1). Anxiety can be experienced in various situations even if a person has not been diagnosed with anxiety disorder. Examples include riding an airplane, public speaking, talking to a stranger, attending a job interview, asking for a raise, facing an exam and even being in the presence of a dog. Symptoms also vary from one individual to another, but the common symptoms can include blushing, sweating, trembling, dizziness, rapid heartbeat, feelings of self-consciousness, difficulty to make eye contact or speaking with a soft voice (American Psychological Association, 2022; National Institute of Mental Health, 2022). As such,

anxiety is generally an uncomfortable and unwelcomed emotion although generally experienced in situations that are not life threatening (Morin, 2017).

Stress and anxiety, however, are known to be key contributors to optimal performance. According to Clark (2018), there is sufficient evidence through the years to establish a strong correlation between moderate stress and peak performance. There are at least four ways in which moderate anxiety can be a useful tool towards performing better: a) *It helps with focus*. Ohman et al. (2001) showed the ability of anxiety to redirect or capture the attention when fear-relevant stimuli were detected more quickly than fear-irrelevant stimuli, even in the presence of distractors, $F(1,24) = 22.29$, $p < .0001$ (reaction time measured in ms). When moderate levels of anxiety is present, focus is optimised and that helps boost performance. b) *It increases motivation*. A review paper in Neuron reported that dopamine which has been shown to be released as a result of pleasure, is also produced as a result of stress, pain and loss. Salamone and Correa (2012) reported that dopamine is instrumental in maintaining the level of activity that is required to achieve what is intended. In other words, it is not only released when we are involved in an activity that brings pleasure but studies show that it is produced before the activity to motivate one to act. *Anxiety motivation* has also been reported to mediate the relationship between anxiety and outcomes such as academic achievement, persistence and job satisfaction (Strack et al., 2017). c) *It keeps you alert*. A study in France using computational modeling of human behaviour and electrical brain activity suggests that anxiety leads to better ability to detect threats which in turn would result in prompt action (El Zein et al., 2015). And d) *It improves attention to detail*. Due to the tendency to imagine the worst when anxious, stemming from negativity bias (Vaish et al., 2008), potential risks can be exaggerated, and extra caution would be exercised.

Although some anxiety could create a degree of alertness needed to do well in the classroom, excessive levels of this emotion could have negative consequences for students, especially for those students whose group identities are negatively stereotyped in the subjects they are studying (e.g., women in STEM). This agrees with the Yerkes-Dodson law (Yerkes & Dodson, 1908) which posits a bell-shaped curve for the relationship between anxiety and performance. According to the law, although anxiety at the midrange could heighten performance, very low and very high anxiety can lead to the impediment of performance. For example, a study of 1,414 graduate students found that female students were more affected by higher levels of test anxiety than their male counterparts which led to a decrease in their grade point average (Chapell, et al., 2005). Generally, there is an agreement in the literature that anxiety during evaluations is associated with decreased academic performance (Harris, et al., 2019; Hembree, 1988; Hill & Wigfield, 1984; Schwarzer, 1990; Seipp, 1991). One study reported that 38.5% (30.0% males and 46.3% females) of undergraduates reported suffering from test anxiety (Gerwing et al., 2015) and another paper suggested that even minor low-stakes pop quizzes could lead to text anxiety (Khanna, 2015). Apart from test anxiety, another form of anxiety that has been well documented within the context of the classroom is math anxiety. Van Mier et al. (2019) investigated math anxiety and performance in 124 second- and fourth-graders and found that math anxiety significantly correlated with math performance only in girls. They also found only in girls that math anxiety negatively moderated math performance. In the literature, math anxiety has been suggested to be the primary reason women and ethnic minority students stay away from courses related to advanced mathematics specifically, and STEM generally (Samuel et al., 2022).

Another significant and relevant form of anxiety that could lead to detrimental effects within the classroom is intergroup anxiety. Stephan and Stephan (1985) first used the term intergroup anxiety as referring to the anxiety that stems from interacting with outgroup

members and traced its origins to the fear of negative evaluations by ingroup or outgroup members. In more recent times, intergroup anxiety has been defined as a type of anxiety that is experienced in anticipation of or engaging in intergroup interaction (Stephan, 2014, p. 240). Kanamori et al. (2022) highlighted that it is the expectation of negative consequences to the self during contact that could lead to intergroup anxiety. In a review of studies on intergroup anxiety, Stephan (2014) proposed a number of potential causes and consequences of intergroup anxiety revolving around three interrelated components: affective, cognitive and physiological. Among other things, one of the proposed causes and consequences of intergroup anxiety is negative expectations and stereotypes (Stephan, 2014; Van Zomeren et al., 2007). In other words, intergroup anxiety can be a result of anticipating negative evaluations by outgroup members or negative meta-stereotype activation (Finchilescu, 2010; Laher & Finchilescu, 2010; Ma et al., 2022). Finchilescu (2010) also found that negative meta-stereotype strength was a significant indicator of intergroup anxiety whereby higher levels of intergroup anxiety were indicated by higher score of meta-stereotypes.

Hence, intergroup anxiety could also have detrimental effects. As mentioned earlier, consequences include negative expectations and stereotype (Stephan, 2014). But apart from that, intergroup anxiety could reduce cognitive resources (Birtel & Crisp, 2012), interfere with executive functioning (Richeson & Shelton, 2003), and ultimately decrease performance when it comes to more complex cognitive tasks and evaluative conditions (Abrams et al., 2006; Stephan, 2014).

Considering the causes and possibly negative effects of anxiety within the context of the classroom, it was important to this PhD research to begin by understanding not only whether anxiety manifests in the classroom settings of interest, but crucially whether the anxiety could be linked to a specific source (e.g., teachers).

Meta-stereotyping as a Source of Anxiety

Meta-stereotypes have been defined as the stereotypes that individuals think that those outside of their group hold of a social group to which they belong (Fowler & Gasiorek, 2020). In other words, a meta-stereotype is a belief or thinking or even awareness that an outgroup that is relevant has a certain opinion of the ingroup (Owuamalam & Zagefka, 2014). These stereotypes or opinions could either be positive or negative. Examples include African Americans expecting their White counterparts to stereotype them as being athletic (Sigelman & Tuch, 1997), or Moroccan teenagers in the Netherlands believing that the general Dutch population stereotype them as being involved in criminal activities (Kamans et al., 2009), or White Canadians thinking that Aboriginal Canadians perceive that they are well-groomed or wealthy (Vorauer et al., 1998), or even women in STEM having concerns about others perceiving them as lacking competence or warmth (Gilrane et al., 2019). The following paragraphs will describe these examples.

Sigelman and Tuch (1997) cited a *Time*/CNN survey of 504 African Americans adults conducted in the U.S. which included questions about perceptions that most White Americans hold of African Americans. The 12-part question on perception included both positive and negative traits, namely, "lazy, religious, more likely to commit violent crimes, would rather live off welfare than work, less intelligent than whites, are better athletes than whites, have no self-discipline, patriotic, low moral standards, good parents, always whining about racism, and more likely to abuse drugs and alcohol." Most of the African American adults surveyed thought that Whites viewed them negatively as "violent, unintelligent, immoral, lazy, undisciplined whiners who abuse drugs and alcohol and would rather live off welfare than work" (Sigelman & Tuch, 1997, p. 89). But at the same time, African Americans also perceived that most Whites see them as being religious and athletic which were more positive traits. Sigelman and Tuch also go on to present findings that suggest that the meta-stereotypes

held by the African American adults surveyed were generally accurate, which means that these were indeed general views that Whites had of African Americans.

In the Netherlands, Moroccan teenagers were asked in interviews how they thought Dutch people perceived Moroccans (Kamans et al., 2009). The findings showed that four negative meta-stereotypical characteristics were evident in the sample population, namely criminality, aggression, Islamic extremism, and loitering (a term often used to describe problems with Moroccan teenagers from areas with low socioeconomic status), suggesting a generally negative meta-stereotype amongst Moroccans in the Netherlands. These negative meta-stereotypes are especially salient when Dutch Moroccan teenagers are relatively negative about the Dutch, and they believe that they are personally viewed via the meta-stereotypes of their group.

On the other hand, Vorauer et al. (1998) investigated meta-stereotypes among White Canadian students by asking them to indicate their perceptions about Aboriginal Canadians' stereotypes of Whites based on 76 traits that were presented. Participants were asked to fill in the blanks for target and baseline ratings which stated, "According to the stereotype that exists in Native Indian society, about ___% of White Canadians possess this trait." (target) and "According to the stereotype that exists in Native Indian society, about ___% of Native Indians possess this trait." (baseline) (p. 921). The diagnostic ratio was then calculated based on dividing the target ratio by the baseline and each diagnostic ratio was then analysed using a 1-sample *t* test to determine which traits were significantly different from 1 which indicated a meta-stereotype. As predicted, there were a considerable number of negative traits although there were also positive characteristics that made up the list of significant meta-stereotypes. The positive traits included ambitious, independent, wealthy, well-groomed, well-mannered, not rebellious and sociable. Their study showed that White Canadians (the dominant group)

believed that individuals from the Aboriginal group (lower status group) held extremely negative expectations about them due to their racial group.

Within the STEM field, Gilrane et al. (2019) surveyed a total of 174 women who were recruited from applicants to STEM conferences from 2006 to 2010. They employed a different method to measure the meta-stereotype content compared to Vorauer et al. (1998). To assess how women in STEM think others stereotype them, Gilrane et al. (2019) adapted items from the stereotype content model (Fiske et al., 2002) which included both the competence (intelligent, competitive and independent) and the warmth (warm, tolerant and family oriented) dimensions. Participants were asked to rate these items on a Likert scale from 1 (extremely uncharacteristic) to 5 (extremely characteristic) as they considered how most males view women in STEM professions. They found that women in STEM had negative meta-stereotypes in both the competence and warmth dimensions. *Gilrane et al. (2019) also found that women within STEM fields engage in impression management behaviours that affect their ratings of supervisors. It is the initial encounters with others that prompt individuals to engage in impression management. These impression management concerns would heighten the pressure to disconfirm stereotypes (i.e., stereotype disconfirmation anxiety) especially when meeting a new professor.*

The meta-stereotype literature, therefore, suggests that it is the valence of the characteristics that is more important compared to the stereotype content (Finchilescu, 2010; Gordijn et al., 2008). Although there have been previous work that report positive traits in meta-stereotypes of various groups (Sigelman & Tuch, 1997; Vorauer et al., 1998), meta-stereotypes in general continue to be mainly negative (Vorauer et al., 1998). According to Fowler and Gasiorek (2020) and Vorauer (2006), this could be due to people naturally understanding that ingroup favouritism may lead individuals of an outgroup to perceive them more negatively than their fellow ingroup members.

Before the influential paper by Vorauer et al. (1998), Joachim Krueger from Brown University tested the *projected-favouritism hypothesis* (which in essence was meta-stereotypes) by surveying 85 African Americans and White undergraduate students. Participants were presented with twenty trait adjectives that included both positive and negative characteristics which were either stereotypic of African Americans or Whites. The traits were as follows: *aggressive, ambitious, arrogant, athletic, copying, family-oriented, friendly, hard-working, intelligent, lazy, materialistic, morally loose, musical, practical, prone to violence, self-confident, selfish, tolerant, unmotivated,* and *unreliable*. In the first stage, participants were first asked to indicate their personal beliefs about their ingroup (personal stereotype) by responding to the following instructions:

> "We would like you to think about a number of attributes, and how they pertain to Blacks (African-American)/Whites (Euro-Americans). Rate the extent to which each adjective is *actually descriptive of Blacks/Whites*…" (Krueger, 1996, p. 539)

Participants then provided their ratings of "cultural stereotypes" by responding to:

> "think about the cultural stereotype of Blacks (African-American)/Whites (Euro-Americans). According to the cultural stereotypes of Blacks/Whites, how characteristic is each of the following attributes?" (Krueger, 1996, p. 539)

In the subsequent stage, participants were asked to anticipate the outgroup's responses by responding to the following instructions:

> "Again, rate the cultural stereotypes and the descriptive characteristics of your own groups. Now, however, instead of giving *your* thoughts and beliefs, you will be asked to estimate the responses that were given by members of the opposite group. *Specifically, we would like you to guess the average of the ratings made my members of the other racial group that participated in this study*…In other words,

> when completing this section, do your best to "think" as if you were a representative member of the opposite group" (Krueger, 1996, p. 539).

The study concluded that participants perceived that their respective outgroups rated members of their ingroup less favourably than was the case, again showing that metasterotypes usually have negative valence. Another study that provides evidence for the predominantly negative valence of meta-stereotypes is that of Finkelstein et al. (2013) as they investigated age-based meta-stereotypes in the workplace. For their paper, they surveyed 247 participants consisting of younger psychology students at a particular university (M_{age} = 21.00; SD = 2.07) as well as middle-aged (M_{age} = 46.38; SD = 4.95) and older (M_{age} = 57.52; SD = 6.82) acquaintances of psychology students at another university. The survey consisted of two major divisions and each division had two parts. In the first part of the first division which focused on stereotypes, the participants responded to the following instructions, which was an adaptation of the instructions given in Vorauer et al. (1998):

> "When people first meet, they often have expectations about each other based on group membership (e.g., gender, ethnic group, student versus nonstudent). In this question we ask you about expectations that most members of *your* age group might have about a younger/middle-aged/older worker on the basis of his/her age group. *In the space provided below, please describe the kinds of qualities and traits that most members of your age group might imagine younger/middle-aged/older workers to possess because of their age group (i.e., before they get to know the person)*."
>
> (Finkelstein et al., 2013, p. 638)

Then in the second part of the first division, participants were told:

> "In this section we would like you to think about stereotypes that *most* people in your age group would have about members of the younger/middle-aged/older worker age

> group. We are not interested in your personal beliefs, but in how you think *most* members of your age group view this other group." (Finkelstein et al., 2013, p. 638)

Participants then rated 180 trait words which included positive, negative, and neutral adjectives on a scale of 1 = "extremely uncharacteristic" to 7 = "extremely characteristic". The second division focused on meta-stereotypes and the instructions were tweaked as follows:

> "In this question, we ask you about the expectations you think that *most younger/middleaged/older* workers would have about *members of your age group. In the space provided below, please describe the kinds of qualities and traits that most younger/middleaged/older workers might imagine that members of your age group are likely to possess (i.e., before they get to know you as a person)*." (Finkelstein et al., 2013, p. 639)

In the second part of the second division, the traits were the same as the first division, but the instructions were as follows:

> "In this section we would like you to think of how *most younger/middleaged/older* workers view *members of your age group*. We are not interested in your personal beliefs, but in how you think members of *your* age group are viewed by another age group." (Finkelstein et al., 2013, p. 639).

The results of the study revealed that although younger participants attributed a few positive characteristics to the older group such as "responsible" and "hard working", the older participants assumed negative traits attributed to them by the younger which were not even in the younger's list such as "grumpy" and "boring". In fact, the older group did not include the positive stereotypes in their meta-stereotypes. The younger group on the other hand, perceived that the older group stereotyped them as being "lazy" and "irresponsible" which were not actually attributed to them by the older group. They also did not include

characteristics such as “enthusiastic” in their meta-stereotypes which was attributed to them by the older group.

Fowler and Gasiorek (2020) said that there can be important consequences stemming from what people think others believe about them based on their social identity. Although generally negative, meta-stereotypes can be useful for people to better understand how they are viewed by outgroup members especially in the context when the opinion of the outgroup matters (Finkelstein et al., 2013; Fowler & Gasiorek, 2020; Frey & Tropp, 2006; Yzerbyt et al., 2009). Positive meta-stereotypes, especially, have been reported to have beneficial effects for stereotyped groups as well as among high-status members (Matera et al., 2015; Owuamalam et al., 2013; Vezzali, 2017). Matera et al. (2015) reported that non-deaf persons had more favourable attitudes towards deaf individuals when exposed to positive meta-stereotypes as opposed to negative meta-stereotypes. Similarly, Vezalli (2017) reported that when exposed to positive meta-stereotypes, participants indicated more positive attitudes and feelings towards an intergroup interaction. According to Owuamalam et al. (2013), when positive meta-stereotypes rather than negative meta-stereotypes were activated, levels of anger among individuals in the low status ingroup reduced significantly, which was followed by less unfavourable perception of the outgroup.

Yet, despite these promising effects, meta-stereotypes could still have adverse effects on the behaviour, thoughts and feelings in intergroup situations due to the fact that meta-stereotype valence is generally negative (Gordijn & Boven, 2009; Vorauer et al., 2000). Vorauer et al. (1998) showed that reduced likelihood to approach others (behaviour), reduced self-esteem (thoughts), and negative feelings can be the result of negative meta-stereotypes (see Gordijn, 2010, for reduced self-esteem). On the level of behaviour, activating negative meta-stereotypes results in avoidance of intergroup contact and perception of distance (Fowler & Gasiorek, 2020; Gomez, 2003) which in turn could lead to loneliness (Gordijn &

Boven, 2009), disidentification with the ingroup (Owuamalam & Zagefka, 2011), impression management behaviours (Klein & Azzi, 2001), selective interaction or hostile reaction (Gomez, 2003), and conforming to the negative stereotypes at times to the detriment of the ingroup and society in general (Kamans et al., 2009). As for the thought or cognitive processes, negative meta-stereotypes can lead to social identity threat (Shelton et al., 2006), increased negativity towards the outgroup (Kamans et al., 2009; Shelton et al., 2005), lowered perceptions of societal fairness (Owuamalam & Zagefka, 2012), and undermined employability beliefs of individuals in disadvantaged groups (Owuamalam & Zagefka, 2014) among other things.

Finally, in the area of feelings, Lammers et al. (2008) put it succinctly when they said that an individual's interactions with people can become loaded with negative emotions if they believe that that person holds negative thoughts about them. Gordijn et al. (2017) reported negative effects on the personal well-being of the person(s) holding on to negative meta-stereotypes. A couple of studies indicate why this would be the case. Owuamalam et al. (2013) showed arousal of anger towards the outgroup while Crocker et al. (1998) posited that when a disadvantaged group activates meta-stereotypes about the dominant group, it could lead to guilt. Apart from that, in a laboratory experiment involving 62 female college students, there was a reduction in cooperative behaviour, measured using the ultimate game, when negative meta-stereotypes were activated and this behavioural change was mediated by feelings of intergroup anxiety (Ma et al., 2022). Finchilescu (2010) found that the degree of negative metasterotypes predicted intergroup anxiety in studies conducted over a 5-year period involving university students. In that study, negative racial meta-stereotypes that people called to mind aroused anxiety about the possibility of intergroup interaction. Shelton et al. (2006) also showed that a person's concern about how those outside of their group view their own group could cause anxiety and this was especially true of underrepresented or

minority groups who are negatively stereotyped which agrees with Lammers et al. (2008) who said that powerless groups in particular engage in meta-stereotypes. These feelings of intergroup anxiety is of particular interest in this research programme especially in the context of the classroom and the effect it has on students' academic outcomes.

Students in a classroom who presume the stereotypes their teachers might hold about their social identity or social group could experience academically crippling intergroup anxiety that has dire consequences on their academic performance (Stephan, 2014; Stephan & Stephan, 1985). Classrooms tend to create an environment in which concerns about being judged according to the salient negative meta-stereotypes within that context could be evoked (Steele & Aronson, 1995). These concerns could be directed towards the teacher or instructor especially if the teacher or instructor is a member of the outgroup since the student-teacher relationship tends to be hierarchical (Brown & Dobbins, 2004). In their investigation, Brown and Dobbins (2004) studied 86 university psychology students that included both people of colour and European Americans. Participants were exposed to a teaching assistant (TA) that was either matched or not matched to their ethnicity, gender being controlled. They were also instructed to imagine either a non-evaluative or evaluative interaction with the TA. After the exposure, participants' response to a survey that included a measure of their perception of the TA in the form of 10 pairs of descriptors: *relaxed/intense, easygoing/competitive, lenient/demanding, reassuring/intimidating, insecure/confident, insincere/sincere, comforting/overbearing, competent/incompetent, qualified/unqualified*, and *good person to learn from/bad person to learn from*. The study concluded that the evaluative role that teachers inherently have could transect with the meta-stereotype that people of colour have of European Americans, resulting in relatively negative expectations when such teachers evaluate them (Brown & Dobbins, 2004). Their results suggests that stigmatized or underrepresented students could have negative expectations for their teacher or instructor

when they expected an evaluative interaction with a teacher or instructor who was an outgroup member (see also Brown, 1998). Gender and race are of particular interest to this research, but other studies have also shown that meta-stereotypes depend on the specific social identities being considered which are not limited to only gender (Gilrane et al., 2019; Ma et al., 2022; Owuamalam & Zagefka, 2014) and race (Vorauer et al., 1998; Kamans et al., 2009). Some of the social groups that have been reported on include weight (Gordijn, 2010), occupation (Xu et al., 2021), and age (Finkelstein et al., 2013).

Considering the social identities mentioned so far, the social identity approach is discussed in the next section. But to conclude this section and move to the next, a point from Fowler and Gasiorek (2020) would be appropriate as it especially highlights the connection between meta-stereotypes, and stereotype threat which is discussed in a following section. They pointed out that the phenomenon of stereotype threat centres on the activation of meta-stereotypes. In other words, the activation of a belief or cognition that other groups perceive our group in a certain way *(i.e., meta-stereotypes)* is what causes deficits in task-related performance.

Social Identity Approach

The social identity approach, made up of the social identity theory and self-categorisation theory, has revolutionised the thinking around many group-mediated phenomena and has become a prominent influential theory not only in the group processes and intergroup relations literature but also beyond social psychology (Hornsey, 2008). The Social Identity Theory (SIT) was developed by Henri Tajfel and John Turner in the 1970s (Tajfel & Turner, 1979) by arbitrarily assigning participants into meaningless groups and then measuring the number of points participants would allocate to their own group members (ingroup) and to the members of the other group (outgroup). According to the SIT, people belong to an array of social groups and these group memberships shape the way that

individuals view themselves, as well as those who share similar or different social attributes to them and influences the way individuals behave (Tajfel & Turner, 1986). These shared or unshared attributes can be based either on social categories (e.g., race and occupation), or interests and abilities (Marx & Ko, 2012) and leads to the formation of an "ingroup" and an "outgroup". In fact, the self-categorisation theory (SCT) differentiates between social and personal identity, the former being dependent on group memberships and the latter being independent of it. SCT suggests that the individual's behaviour is influenced by either social or personal identity processes depending on which is salient or important in any given situation. When both are salient at the same time, then behaviour can be motivated by a interaction of both (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987).

This kind of categorisation can eventually lead to social comparison and the development of positive biases towards the ingroup and negative biases towards the outgroup which could result in stereotyping, prejudice, and discrimination. Importantly, SIT assumes that people are normally motivated to maintain a positive social identity or try to enhance it when threatened in some way given its consequences for members' self-esteem (Brown R. , 2000; Owuamalam, Tarrant, Farrow, & Zagefka, 2013; Tajfel & Turner, 1986). People's self-esteem is boosted by a positively valued social identity (Snyder, Lassegard, & Ford, 1986) and, likewise, people experience deficits in self-esteem when their identities are compromised in some way, e.g. by activating negative meta-stereotypes (Branscombe, 1998; Gordijn & Boven, 2009; Owuamalam & Zagefka, 2011; Vorauer, Main, & O'Connell, 1998).

Within the classroom context, people can experience the disruptive negative emotion of anxiety when a stereotyped dimension on which their ingroup compares unfavourably with outgroups is salient, leading to deficits in academic performance (Steele & Aronson, 1995). Thus, although people can, at times experience performance boost in a stereotyped dimension (Crisp, Bache, & Maitner, 2009; Walton & Cohen, 2003), evidence exist within the literature

on stereotype threat suggesting the opposite detrimental effect of unflattering ingroup stereotypes on academic performance, particularly before an outgroup audience or expert (Marx & Goff, 2005), and presumably because such context heightens stereotyped group members' evaluation anxieties (Vorauer, Hunter , Main, & Roy, 2000). Hence the suggested use of ingroup role models to inoculate stereotyped group members from the often harmful effects of the anxieties arising from the outgroup's negative stereotypes of the ingroup (Dasgupta N. , 2011; Marx & Ko, 2011).

Based on the social identity approach, underrepresented students in STEM may feel a particularly higher need to maintain a positive social identity of their group when the lecturer teaching them is a member of the outgroup (e.g., male) compared to when the lecturer is an ingroup member (e.g., female). They would want to project a positive image of their ingroup because they are concerned that their behaviour (e.g., incorrect negative responses in class) might reinforce the outgroup's negative opinions of them. In that way, perceived negative meta-stereotypes could influence student participation in class.

Stereotype Threat

The origin of the role model interventions mentioned in the first section of this chapter as well as the phenomenon of intergroup anxiety and meta-stereotypes are all rooted in or extended from the impact of stereotype-related anxieties, otherwise known as *stereotype threat* (ST; see e.g., Steele & Aronson, 1995; Spencer et al., 1999). To define it simply, "ST is being at risk of confirming, as self-characteristic, a negative stereotype about one's group" (Steele & Aronson, 1995, p. 797). For example, a woman seeking to perform in a STEM setting may feel that others could see her performance as confirming a negative stereotype about women and STEM if she performs badly, and this could be psychologically threatening and anxiety inducing. There's a significant body of literature that shows that stereotype threat can lead to people performing worse than they can perform if the conditions were evaluative.

In other words, the person believes ability is being evaluated (Inzlicht & Schmader, 2012; Steele, 1997; Steele, 2010; Walton & Spencer, 2009). Stereotype threat is to be distinguished from social identity threat in that the latter encompasses broader cues pointing to the possibility that an individual's identity may be devalued in a particular situation (Purdie-Vaughns et al., 2008; Steele et al., 2002), while the former is about the anxiety that an individual feels due to be possibility of confirming a negative stereotype of the individual's group.

Steele (1997, 2010) also suggested the conditions that could induce stereotype threat: a) the awareness by group members of the existence of a stereotype of the group (e.g., stereotypes of women's ability in STEM), b) a task being perceived as diagnostic of ability or evaluative of skills that are negatively stereotyped (e.g., ability in maths), c) the task is pushing the limits of the individual's skills, and d) the person identifies with the group being stereotyped and is invested in doing well. As such, stereotype threat could affect a wide range of people groups and arise within a wide range of contexts. Some examples include when the memory of the elderly is undermined (Hess et al., 2003), when the individuals from lower socioeconomic backgrounds are expected to take a verbal test (Croizet & Claire, 1998), when Whites or Blacks underperform in golf when the task is said to assess athletic ability or intelligence (Stone et al., 1999), and when White men are given a math test that is said to be used to test why Asians are so good at math (Aronson, et al., 1999). And yet, the past research on stereotype threat has predominantly focused on women in STEM specifically and ethnic minorities in academics generally (Walton et al., 2015) due to its importance.

Stereotype threat has been shown to have a wide range of effects or consequences on stereotyped individuals, mostly negative in nature. Some of the major consequences reported in the extant literature so far include internal attributions for failure (Koch et al., 2008), reactance (Kray et al., 2001), ironic effects (Goff et al., 2008), self-handicapping (Keller,

2002), task discounting (Forbes et al., 2008), distancing the self from the stereotyped group (Goff et al., 2008), disengagement and disidentification (Woodcock et al., 2012), and altered professional identities and aspirations (Casad & Bryant, 2016). Apart from the consequences mentioned, decreased performance is an effect that has been most studied. According to Walton et al. (2015), the body of literature contains more than 400 studies at their time of publishing that evidence the effect stereotype threat has in on performance. Performance in literature has been measured by way of tasks that include adaptations of standardised test, intelligence tests, subject-specific tests, memory tests, emotional-sensitivity tests, and physical tasks (Steele, 2010; Steele et al., 2002). A meta-analysis of 17 years of research on ST, involving 103 papers and 5,588 females, revealed that females in ST conditions underperformed on *mathematics tests that were reportedly diagnostic of ability* compared to their counterparts in the control conditions (d = |0.24|) (Picho, Roriguez, & Finnie, 2013). Agreeing to this, another meta-analytical report examining 76 papers containing 116 studies published between 1995 to April 2006 with 8,277 data points from stereotyped groups showed that both women and ethnic minorities suffer from performance deficits when *mathematics tests* were difficult although women experienced smaller performance differences (Nguyen & Ryan, 2008). The underachievement could stem from stereotype confirmation (or even disconfirmation) anxieties which could impair their cognitive function (Ajilchi & Nejati, 2017) by diverting attentional resources away from the task at hand (Cheryan & Bodenhausen, 2000; Mrazek, et al., 2011).

On the flipside, when stereotype threat is reduced, individuals from stereotyped groups performed better (Steele, 2010). In two meta-analyses, one for stereotype threat experiments and one for intervention field experiments, data was combined from 18,976 students in five countries (Canada, France, Germany, Sweden, and the United States). There were four inclusion criteria in the first analysis: the negatively stereotyped group in the

domain tested and the group not stereotyped had to be included in the studies, stereotype threat had to be manipulated, the *subsequent performance of participants on a relevant test had to be measured*, and performance in that same domain had to be tested in a real-world setting. The inclusion criteria for the second analysis were identical to the first except the manipulation had to be an intervention to reduce stereotype threat instead and the subsequent performance measure had to assess real-world classroom performance. Both meta-analyses showed that the performance of stereotyped students was better than their non-stereotyped counterparts (Walton & Spencer, 2009). Nadler and Clark (2011) also conducted a meta-analysis that examined the effects of removing stereotype threat but on ethnic minorities, namely African Americans and Hispanic Americans. The analysis found moderate levels of increment in *test scores* for both groups when stereotype threat was removed or nullified (d = 0.52).

Considering these established consequences of stereotype threat, the next question is, who would then be vulnerable to it? Some of the factors relevant to this thesis that have been identified as having the potential to increase the susceptibility of an individual to stereotype threat are group membership, group identification, and stigma consciousness. Group membership is when an individual identifies with certain social groups and is perceived to belong to that group. It is treated as a matter of reality that can be more or less salient depending on the context (Roth et al., 2018). A study by Harvard University psychologists tested the implicit activation of group membership to see if it would facilitate or impede performance. Shih et al. (1999) made salient the group membership of their participants by presenting them a short questionnaire with an item regarding their identity as women (e.g., *Do you prefer single sex or coed college dormitories?*), Asian Americans (e.g., *How many generations of your family have lived in America?*) or members of a social category that is not relevant to the domain (cable TV subscribers). They found that Asian American women

who activated their ethnic group membership performed better on a mathematics test but performed worst of the three groups when they activated their gender identity, and this difference was due to stereotype attached to the identity within that the domain of mathematics and not the identity *per se*. Along the same lines, McGlone and Aronson (2006) primed 90 liberal arts undergraduates with different group memberships before administering the Vandenberg Mental Rotation Test, which is a standardised visual-spatial test (there is a performance gap in spatial reasoning which is the largest recorded sex difference in any cognitive ability, with women scoring lower than men). The priming involved probe questions that made salient participants' sex or gender (e.g., *List three reasons why one might prefer living on a coed floor in a dormitory*) or participants' status as a student at a private school (e.g., *List three reasons why one might attend a private liberal arts college*) or participants' status as a resident of the area which was a control condition (e.g., *List three reasons why one might prefer living in the Northeast to the other parts of the U.S.*). Results showed that women primed to think about their sex, performed worse than those who focused on their group membership as students.

Group identification is defined as "the act or process of associating oneself with a group and its members, such that one imitates and internalizes the group's actions, beliefs, standards, objectives, and so forth" (American Psychological Association, n.d. a, para. 1). In a study on age group identification, the memory of old adults was tested as they were asked to recall a prose passage either under stereotype threat conditions or under normal conditions (Kang & Chasteen, 2009). Participants who strongly identified with the age group experienced more performance deficit than those who were weak identifiers, showing that group identification (in this case age group) was an important factor that contributed to the susceptibility to stereotype threat. Armenta (2010) reported similar findings in a study that examined ethnic identification as a moderator of stereotype threat effects on the math

performance of Latino undergraduates who are negatively stereotyped within the domain. Participants were asked to complete a 7-item ethnic identification measure with items such as "I have a strong sense of belonging to my ethnic group" and then proceeded to take a difficult math exam one to two weeks later, having been either assigned to an experimental condition where ethnic stereotype cues were salient or a control condition where it was not. The study confirmed that group identification increases susceptibility to stereotype threat for Latino undergraduates as those who were high group identifiers performed worse under stereotype threat conditions. Armenta (2010) summarised it well by stating that when someone belongs to a stereotyped social group, performance could be affected when in stereotype-relevant situations and the performance often follows the direction of the stereotype.

Apart from group membership and identification, stigma consciousness is also an important factor contributing to the susceptibility to stereotype threat. By way of defining it, stigma consciousness refers to the extent to which targets think that their stereotyped status permeates their interactions with outgroup individuals. Basically, it is in reference to the focus on one's stereotyped status (Pinel, 2004). In a study involving 49 female psychology undergraduates who at least moderately identified with math and scored either high or low in gender-based stigma consciousness (measured using the 10-item stigma consciousness questionnaire. E.g., "When interacting with men, I aways feel as though they interpret my behavior in terms of the fact that I am a woman" and "Stereotypes about women have never affected me personally" – reverse scored), read cover stories that divided them into the low threat and high threat conditions. Participants in the *low threat* condition read that the study was about factors that predict performance on mathematical ability tests and that the test had been shown to be *completely free of gender bias* that men and women tend to perform equally well on it. In the *high threat* condition, participants read that the study was about factors that explain why there was a big difference in performance on standardised math tests between

men and women. After that, they were given a test consisting of 20 math problems. The results showed that when stereotype threat is salient, women scored worse on the math test when stigma consciousness was high (Brown & Pinel, 2003).

Putting the possible consequences of stereotype threat and the risk factors that lead to susceptibility to stereotype threat together, it becomes clearer that students from stereotyped underrepresented backgrounds could underachieve in STEM fields. The experience of stereotype threat could, at times, undermine the performance of students from groups that are stereotyped in the relevant subjects (Krendl et al., 2008). This classic stereotype threat effect has been used to explain the achievement gap between students from stigmatised backgrounds, especially in STEM fields (Rice et al., 2013; Shapiro & Williams, 2012). One example is a Spanish study in which women's performance in mathematics was reliably worse-off in a condition that heightened the salience of negative stereotypes about women in math, especially for participants who reported being anxious (Delgado & Prieto, 2008). These ST effects also manifest in the context of underrepresented ethnic groups (see Steele, 2010, for a review). In a meta-analysis, for example, Nguyen and Ryan (2008) found that ST-activating cues produced a large negative effect on STEM-related performance (*measured by mathematics test scores*) amongst students from ethnic minority backgrounds (*d* Cohen = |.64|). The evidence does show that ST can (at least partly) explain the achievement gap in some STEM fields in a diverse range of stigmatised groups (Davies et al., 2002; Johnson et al., 2012; Marchand & Taasoobshirazi, 2013). ST is shown to affect women's confidence in engineering (Cadaret et al., 2017), performance in engineering tests (Bell et al., 2003), and performance in math tests (Nguyen & Ryan, 2008).

The main tenet of the ST theory and literature is that levels of anxiety would increase as an effect of ST conditions, causing individuals to underperform. Osborne (2007) tested this assumption within the STEM setting using real-time physiological measures of arousal (i.e.,

heart rate, skin conductance, skin temperature and blood pressure). A total of 43 undergraduate students were subject to the experimental manipulations which manipulated the perception of whether the test they were about to take produced gender differences (high ST condition, 10 males and 14 females) or not (low ST condition, 11 males and 8 females). Participants were assigned to their condition based on the flip of a coin. Results were consistent with the general hypothesis of ST and showed evidence high psychological reactance in girls in the high ST condition and the reactance is consistent with an anxiety reaction. While in a separate study within the STEM field involving a sample size of 28,240 high school students from 1,015 schools in the U.S., anxiety partially explained racial and sex differences in academic performance in mathematics (Osborne, 2001). Conversely, situations that ease stereotype-related anxieties tend to enable people from stereotyped groups to overcome ST effects and bring their level of performance up to those of their counterparts from non-stigmatised groups (Spencer et al., 1999). In other words, ST-attenuating strategies can be effective in reducing ST effects (Nguyen & Ryan, 2008) and strategies to overcome ST has a role to play in reducing the achievement gap in STEM.

Interestingly, however, not all the evidence supports ST especially in more recent studies (e.g., Flore et al., 2018; Flore & Wicherts, 2015; Stoet & Geary, 2012). *In fact, Jussim and Honeycutt (2023) go even further to point out that the initial ST studies were politically biased by claiming that they vindicated some political narrative, the results were misinterpreted or misrepresented in ways that advanced a politicised narrative and ideology-affirming conclusions were made based on weak data. According to Jussim & Honeycutt (2023), every pre-registered study that attempted to replicate the effect of stereotype threat on women's performance in math has failed to do so and there is no pre-registered published study on the effects of stereotype threat on racial/ethnic differences in performance.*

In one of the pre-registered large-scale studies, Flore et al. (2018) studied 2,064 high school students in the Netherlands, who were 13 to 14 years old, to examine the generalisability of the stereotype threat effect on female students in the domain of mathematics. These were students who were average to high achieving, recruited from 21 schools that were chosen via convenience sampling. Participants were given a mathematics test that included an introduction that differed according to the condition they were randomly assigned. The two conditions were an explicit ST condition where ST was heightened by adding "The most recent study carried out four years ago showed that boys and girls do not perform equally well on this mathematics test. There was a difference in the average grade on the test between boys and girls" to the introduction, and a control condition where the threat was nullified by adding "The most recent study carried out four years ago showed that boys and girls performed equally well on this mathematics test. There was no difference in the average grade on the test between boys and girls" (Flore et al., 2018, p. 147). Despite implementing an optimal experimental design to increase the chances of finding an effect in line with the ST theory, they found that Dutch school girls within the age group studied did not perform significantly worse relative to boys in maths.

An earlier pre-registered study that claimed that replications of ST studies were relatively uncommon tested the effect in a large and more diverse sample. Finnigan and Corker (2016) studied 590 American female participants and found that there was no significant main or interaction effects for ST. In their study, they measured mathematics performance using 10 multiple choice questions from the GRE quantitative reasoning section. They chose these questions after a pilot study with three tests of different levels of difficulty showed that this was the most difficult and so, it was the most suitable to detect ST. Much of the methodology of the paper is sound, however, participants were recruited via Mechanical Turk (MTurk). Although MTurk has become increasingly popular in social psychology

studies due to its cost-effective and accessible features, there are several limitations associated with it that needs to be taken into consideration. Participants are usually paid a fee (in this study it was $1.00) for taking part in studies which would attract individuals motivated more by financial incentives rather than an interest in the research (Mason & Suri, 2012). *Because of the fee, "professional participants" who make a living by participating in numerous studies may not pay full attention to the study which could lead to unreliable responses* (Chandler & Shapiro, 2016). *Finally, the limited control over experimental conditions is the issue most pertinent to this thesis. The control that the researcher has over the experimental conditions is limited on MTurk compared to traditional lab studies or field experiments. There could be a lot of noise in the data due to things like participant settings, distractions around the participants, and technological glitches* (Paolacci, Chandler, & Ipeirotis, 2010).

Recent meta-analyses about stereotype-related achievement gaps in STEM also suggest that the induction of ST probably only causes a small negative effect. Flore and Wicherts (2015) investigated the effects of stereotype threat on *performance in tests related to math, science and spatial skills (MSSS) for girls* with a mean age of under 18 years. Their inclusion criteria were the studies that included school going girls and gender stereotype threat was manipulated, the average age of the participants were below 18, the participants were randomly assigned to streotype threat and contol conditions, the stereotype threat manipulation was treated as a between-subjects factor, and the dependent variable was the score for a MSSS test. This yielded 26 papers or unpublished reports containing 47 independent effect sizes from a total sample size of 3,760 girls. They found that the estimated mean effect size was only -0.22 and that there were indications of the presence of publication bias which could have affected the reports of stereotype threat effects among school girls in the literature. In fact, when Stoet and Geary (2012) reviewed replication attempts in the

stereotype threat literature, they found that only 30% of unconfounded replicated studies that presented their experimental design actually replicated the original results. Their meta-analysis of the effects in these studies showed that only studies with *adjusted math test scores* showed the stereotype threat effect.

The generally small effect size in the meta-analyses by Flore and Wicherts (2015) is also seen to be practically wiped-out when testing is conducted in operational real-world scenarios where motivational inducements (such as the need to excel) are more apparent (Shewach et al., 2019; Stafford, 2018; but see Smerdon et al., 2020, for a rebuttal). *What might be the cause of this weak and conflicting evidence for ST? The reason could be methodological and/or theoretical. The possible methodological reason would be discussed in the next chapter when the anxiety measure used in this thesis is discussed. From the more theoretical aspect, firstly, there seems to be a misinterpretation of the original findings of Steele and Aronson (1995). Sackett et al. (2004) showed that 90.9% of journal articles and 55.6% of textbooks interpreted the findings or the original study by Steele and Aronson incorrectly. Fifteen years later, Tomeh and Sackett (2022) found that the misinterpretation rate dropped significantly in journal articles but not in textbooks. This misinterpretation could affect methodology of the experiments and also the interpretation and conclusion of the findings. Secondly, if we take the identity-in-context model (Easterbrook, Hadden, & Nieuwenhuis, 2019) into consideration, identity is said to be malleable and influenced by the context. As experiencing ST is dependent on the identity that one activates in any particular context, it becomes more complex to control the identity that participants are activating while undergoing any ST study. Finally, evidence suggests that ST plays less a role in the achievement gap of stigmatised students when experimental tasks were administered by ingroup experimenters who could act as role models for these students.* Over a decade of research on role model interventions and *stereotype inoculation* show that such “ingroup

ambassadors" can act as social "vaccines" that immunise stigmatised students against the experience of anxiety that causes ST effects (Dasgupta, 2011) which might also explain why Flore et al. (2018) found that girls' math performance was not affected by ST in their study. Thus, the present research will focus on the positive, and potentially negative, effects of an ingroup role model or expert.

Counter-stereotypical Role Models and Stereotype Inoculation Model

Several studies in the past have shown that ingroup role models that are counter-stereotypical to their stereotyped or marginalised group do produce better outcomes for women or ethnic minorities in STEM (Herrmann, et al., 2016; Lockwood & Kunda, 1997; Shin et al., 2016; Stout et al., 2011). In fact, supportive evidence for the idea that ingroup role models can be used as means to combat ST effects has been found in numerous contexts including, gender (Dasgupta & Asgari, 2004; Marx & Roman, 2002; McIntyre et al., 2003; Stout et al., 2011), ethnicity (Marx et al., 2009), among middle and high school students (Lin-Siegler et al., 2016; O'Brien, et al., 2016), and even university students (Latu et al., 2013; Young et al., 2013).

Considering the amount of written work using the term "role model" (Morgenroth et al., 2015, reported 400,000 scholarly articles at the time of writing their paper and a Google Scholar search at the time of writing this thesis produced 8,310,000 hits), it is important to first define what is meant by the term "role model". The term has been defined in various ways in the literature. Gladstone and Cimpian (2021) defined role models in general as persons who, by acting as a successful exemplar, can positively shape a student's motivation and made a contextual distinction by focusing on the effects on motivation from exposure to STEM experts with whom students have no prior connection with. However, Morgenroth et al. (2015) have produced the most comprehensive write up on the definition of the term and, by piecing together all angles, they defined role models as "individuals who

influence…achievements, motivation, and goals by acting as behavioral models, representations of the possible, and/or inspirations" (p. 468). *Role models that come from diverse backgrounds can foster inclusivity and broaden the range of aspirational figures especially for underrepresented groups. Exposure to diverse role models helps to challenge stereotypes and promotes self-perception (Eagly & Karau, 2002) as well as enhances a sense of belonging and encourages underrepresented groups to pursue atypical aspirations (Murphy & Taylor, 2012).* A *counter-stereotypical* role model, on the other hand, would be someone who assumes a role that is antithetical to the group stereotypes (e.g., a male nurse, a female Nobel Prize winner for mathematics, or a female CEO) (Olsson & Martiny, 2018).

There is also a huge body of literature that has reported on the various improved outcomes when individuals from stigmatised or stereotyped groups are exposed to counter-stereotypical ingroup role models. Among other things, there is evidence that ingroup role models increase girls' *interest* in STEM (Steele et al., 2002). Stout et al. (2011) examined the effect of brief interactions with a female or male expert on the self-conception, effort, and performance in mathematics for 73 undergraduate female students majoring in STEM undergraduate programmes. In the same paper but a separate study, Stout et al., also found that female students who were exposed to female professors reported greater *self-efficacy* and generally more *positive orientation* towards science, technology, engineering, and math (STEM) than those exposed to outgroup (male) professors. Female students' performance on a difficult mathematics test did not differ as a function of exposure to an ingroup expert but *effort*, operationalised as total number of questions attempted, did significantly increase. In a related study, Marx and Goff (2005) had Black and White Harvard undergraduates undergo a verbal test after being exposed to either a Black or White experimenter who was portrayed as an expert. Results showed that the *performance* of Black participants exposed to an ingroup expert rose to match those of their White counterparts. This trend was reversed when the

expert was an outgroup member, so that Black students were outperformed by their White counterparts. In general, role model interventions have been reported to inoculate from stereotype threat (Dasgupta, 2011; Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2003; Stout et al., 2011), and reduce implicit self-stereotyping (Asgari et al., 2012).

Considering these positive effects on stereotyped underrepresented groups, it is not surprising then that Casad et al. (2018) lists role model intervention as a "wise" intervention to improve gender and racial equality in STEM. They provide three reasons for doing so: 1) role model interventions target a specific psychological need (e.g., belonging, identity), 2) it has an effect and produces results downstream (e.g., academic performance), and 3) it has been rigorously tested in both the lab and classroom settings. To add to these reasons, the role model interventions exist in various formats other than in person exposure. Examples include exposure to online ingroup role models that led to higher course grades and lower withdrawal rates for female students in STEM (Herrmann, et al., 2016), and reading biographies of successful ingroup exemplars that resulted in more positive implicit attitudes and higher identification, confidence, self-efficacy, and career goals in for women in engineering (Shin et al., 2016; Stout et al., 2011). Consequently, for role models to be effective they need to be perceived as competent (Marx & Ko, 2012), they need to be of the same group membership (Marx & Goff, 2005; Marx & Roman, 2002; McIntyre et al., 2003), and they need to be seen as successful in a mutual area of interest (Marx & Roman, 2002). However, direct contact is not a necessity (Herrmann, et al., 2016; Stout et al., 2011). In a meta-analytic review of the effectiveness of ST interventions, Liu et al. (2021) states that the encouraging results of role model interventions are made possible because upward social comparison could help in removing the stigma, disconfirming negative ingroup stereotypes and preventing evaluative concerns as it is self-enhancing especially when the focus is on the similar and successful

other. *Performance outcomes on the review was measured based on task performance, test scores or grade point average among other things.*

Probing the parameters of the role model intervention, or more specifically, the *stereotype inoculation model* (SIM; Dasgupta, 2011), the assumption is that the performance of stigmatised underrepresented students on STEM-related tasks is undermined by them feeling out of place or having a sense that they do not belong in STEM classes (i.e., as imposters). The *imposter syndrome*, or in other terms imposter phenomenon, fraud syndrome, perceived fraudulence, or imposter experience (Bravata, et al., 2020), is a term coined by Clance and Imes (1978) when they studied high-achieving women and found that women who experienced imposter syndrome persisted in believing that they were not intelligent and that they had fooled anyone who might think otherwise despite their objectively excellent performance both academically and professionally. In other words, the imposter syndrome is the fear of being exposed as an imposter and an inability to internalize success (Jensen & Deemer, 2020). According to Dasgupta (2011), individuals who feel like imposters are more likely to make external attributions (e.g., luck, effort, or personal charm) for their success than internal attributions (e.g., ability), more likely to be anxious and less likely to be sure of their ability. Ample examples exist of individuals who are particularly vulnerable to imposter-related anxiety for both the gender context, where high-achieving women are more susceptible compared to their male counterparts (King & Cooley, 1995; Kumar & Jagacinski, 2006), and the racial context, where African American students' imposter anxiety is correlated with lower academic self-efficacy (Ewing et al., 1996).

The sense of imposterism is especially true in situations that align with negative stereotypes about one's group and associated anxieties. Although the presence of an outgroup (a group that a person does not identify with e.g., men in the case of women) can accentuate evaluative anxieties (Vorauer et al., 2000, see also Chapter 5 in this thesis), such concerns

tend to be less severe in the presence of fellow ingroup members. Hence, the anxiety over being negatively judged should be weaker when people are with an ingroup rather than with an outgroup audience (Hopkins, et al., 2007; Owuamalam & Rubin, 2014). Ingroup experts (e.g., female STEM professors) directly challenge negative stereotypes that underrepresented students do not have what it takes to succeed in the field, which can boost students' subject-specific confidence/efficacy (Stout et al., 2011). Hence, based on SIM, anxious students from stereotyped backgrounds who feel out of place in STEM are especially likely to benefit from the buffering effect of ingroup experts because the latter can calm anxiety and improve STEM-related outcomes by allowing these students to devote attentional/cognitive resources to the task at hand (Dasgupta et al., 2015).

In short, it is entirely possible that the inconsistent effects of stereotype induction in ST research could be attributed to exposure to identity-matching versus identity-mismatching experts or role models. For example, the absence of the ST effect in the gender context that was reported by Flore et al. (2018), could have been due to the stereotype-disconfirming presence of an expert ingroup member in the form of a female "experiment leader" who gave the instructions to participants. They cited a meta-analysis by Doyle and Voyer (2016) which showed that differences in effect sizes between studies run by female and male "experiment leaders" were negligible as reason to continue with a female "experiment leader", however, the presence of the female experimenter could have diluted (even eliminated) stereotype-related anxiety – the critical mechanism that is theorised to underlie ST effects. *There is sufficient evidence for this assumption especially when considering that Marx and Roman (2002) and Stout et al. (2015) highlights the importance of the presence of an ingroup experimenter.*

The Ingroup Spotlighting Thesis in the STEM Context

Despite the promising effects of ingroup experts, they could also create a different type of stress for some students from underrepresented groups. Casad et al. (2018) highlights some of the limitations of expert or role model interventions in stating that it needs to be varied depending on the student population. Exposing anxious minorities to ingroup experts could undermine their performance (Flore et al., 2018) instead of *inoculating* them against anxiety as proposed by Dasgupta (2011) and many others (see previous section). The notion that exposing students to STEM experts from their ingroup could *worsen* STEM-related outcomes of anxious women or ethnic minorities could be rooted in several processes. For example, studies have shown that people are especially mindful of the opinions of those they care about (e.g., ingroup members) and are dismissive of the opinions of those they do not care so much about (e.g., outsiders, Hornsey et al., 2002). Thus, a female teacher's evaluation, for example, might be particularly important (and so anxiety producing) to girls compared to boys, especially if those girls have a lower sense of self-efficacy to begin with (Hoyt, 2013). Furthermore, a female expert also removes the excuse that STEM is not for women (Marx & Roman, 2002), which could generate additional demands for girls to excel in the domain. Anxious women may feel especially *spotlighted* in this context.

To be clear, spotlighting has also been used elsewhere in the literature in several ways. A general understanding of the research on the *spotlight effect* is that people commonly believe that they come under the scrutiny and attention of others more than they actually do in reality (Gilovich et al., 2000). In other words, people tend to overestimate how much their actions and appearance are given attention or noted by others. Spotlighting is also used to refer to behaviours that single women out due to their gender in ways that make them uncomfortable (McLaughlin, 2005). The term in McLaughlin's paper was further subdivided into three different types of spotlighting. Type I spotlighting referred to *overt sexism* which is

defined according to the general definition already given, Type II spotlighting or tacit sexism refers to behaviours that single out women but without any intention to harm or help (neutral intentions), and Type III spotlighting which is behaviours that single out women by gender with the intention to help. The exploration of the spotlighting phenomenon by McLaughlin (2005) ostensibly focused on behaviours perceived as overtly or tacitly sexist, or those that are presumptive and demeaning (e.g., unsolicited help in STEM based solely on gender), and often *emanates from the outgroup*. Lastly, Crosby et al. (2014) reported on the *minority spotlight effect* when they found that underrepresented participants felt most spotlighted when they were the minority in a class where the *outgroup* drew attention to their group and because they could be more vigilant to the cues connected to social identity threat.

The present thesis, however, proposes a different form of spotlighting originating from within one's own group, which can negatively impact underrepresented students' outcomes in STEM: The pressure that exposure to an ingroup expert or role model might put on underrepresented students to likewise excel in STEM fields. In the presence of such an expert, there is attributional ambiguity that could subsequently harm their self-esteem (Crocker et al., 1991). That is, when women or minorities are exposed to (or are being evaluated by) an outgroup professor, underachievement in STEM could be attributed to structural barriers like sexism, lack of motivation/interest, or lack of aptitude/experience and so on (Stewart-Williams & Halsey, 2021). However, in the presence of an ingroup member who is successful in STEM, attributing the cause of potential failure in STEM to bias and/or sexism may be hard to do because an ingroup expert provides a glaring example that underrepresented groups can overcome these barriers (Marx et al., 2009; Marx et al., 2005). To the extent that externally directed attributions for potential setbacks in STEM become limited, the search for excuses for potential failure could deplete cognitive resources and increase the burden of having to prove oneself in this type of situation, particularly for those

who are less confident in the subject, and/or are anxious about their place in STEM to begin with. In short, according to this *ingroup spotlighting thesis*, highly anxious women and/or ethnic minorities should achieve *worse* (not better) STEM outcomes and should be even more motivated to *search for excuses* for this underachievement when exposed to ingroup expert (contrary to SIM's inoculation thesis; Dasgupta, 2011). Despite the intuitive appeal of this ingroup spotlighting thesis, there is, as yet no systematic exploration of this potential side effect of exposing students from underrepresented backgrounds to ingroup STEM experts or role models.

Hypotheses

The research in this thesis ultimately aimed to explore the possibility that the academic outcomes of underrepresented groups could be worsened by the fact that they are "spotlighted" by the increased exposure to exemplary counter-stereotypical role models that might exacerbate the pressure to likewise excel in the field. Therefore, it was first important to examine the general assumption that students experience anxiety in classrooms, the source of this anxiety and how this anxiety might impact students' classroom outcomes. Prior to Study 1 (Chapter 4), it was hypothesised that (a) the most common source of anxiety in the classroom would relate to concerns over social image/identity, (b) teachers would be the most likely key source of this classroom anxiety, and (c) such anxiety could reduce students' confidence and classroom outcomes in terms of classroom participation and/or engagement.

Study 2 (Chapter 5) aimed to determine the meta-stereotype content in the students versus teacher context and whether the social identity of the teacher would influence these meta-stereotypic concerns. The prediction was that the meta-stereotype content of students regarding their teachers would be generally negative, and this would be particularly salient in the outgroup teacher condition compared to the ingroup. Also, competence-related meta-stereotypic concerns would be more an issue when students are faced with outgroup teachers

rather than ingroup teachers because this is a context in which people are especially mindful of their social image.

The main goal of the subsequent studies (Studies 3 to 5) is to understand the fundamental processes that shape the STEM outcomes of students from underrepresented populations, using insights from the new *ingroup spotlighting thesis* that is proposed here, while accommodating predictions that could be deduced from a more entrenched (and competing) theoretical framework (the *stereotype inoculation model*; Dasgupta, 2011). At its core, the stereotype inoculation model (SIM) assumes that the inoculative benefits of ingroup experts on the outcomes of underrepresented STEM students should be most visible when they are exposed to an expert ingroup (versus outgroup) teacher, especially amongst those students with strong imposter sentiments (i.e., students who feel they do not have what it takes to succeed in a relevant field; Dasgupta, 2011). This pattern of results under SIM was expected because an ingroup role model should help to dilute (or inoculate against) stereotype-related anxieties that could prompt underrepresented students to question their place in STEM (i.e., whether they are good enough). The reduction of stereotype-related anxieties by ingroup role models then frees up the cognitive space to concentrate on the task at hand, and this should lead to improved outcomes in STEM, according to SIM. Specifically, the inoculation thesis would predict that anxiety (i.e., the jitter in students' voices) should be accentuated when negative ingroup stereotypes are salient versus non-salient, especially in classes that are taught by outgroup (but not ingroup) teachers (Study 3 in Chapter 6). This anxiety should predict increases in performance errors (Studies 4 and 5 in Chapter 7). Overall, it is expected that students will perform better on a STEM test with an ingroup teacher compared to an outgroup teacher. However, for anxious students with low subject efficacy (i.e., students with imposter feelings), increased performance errors (the ingroup spotlighting hypothesis) rather than reduced performance errors (as per the SIM) is predicted

especially for those who are exposed to an ingroup and not to an outgroup teacher (Studies 4 and 5). Apart from these predictions, both the stereotype inoculation and ingroup spotlighting theses is expected to have longitudinal effects (Study 4) and post-lecture confidence in mathematics as well as post-quiz attributions were also investigated in Study 5 (hypotheses summarised in Table 1).

Study 3 (Chapter 6) was also used to validate vocal jitter as an index of anxiety. If vocal jitter is indeed a valid measure of anxiety, then it is expected that there would be a significant correlation between the ratings of anxiety from students' recorded responses and the actual vocal jitter extracted from recordings, and also that there would be an effect of the experimental treatments on judges' ratings of selected emotions (especially anxiety) that also correspond to the effect of the experimental treatments on vocal jitter. Table 1 presents a summary of the research hypotheses for each study.

Table 1

Summary of Research Hypotheses

Number	Hypotheses
	Study 1
H1a	Concerns related to social image/identity would be the most common source of anxiety in the classroom.
H1b	Concerns related to social image/identity would be the most common source of anxiety in the classroom within the teacher context compared to friends or family.
H1c	Concerns related to social image/identity within the teacher context would negatively affect confidence in the classroom.
	Study 2
H2a	Meta-stereotype content among Malaysian students regarding their teachers would be generally negative.
H2b	Meta-stereotype content among Malaysian students would be more negative in the Caucasian teacher (outgroup) condition than Malaysian teacher (ingroup)
H2c	More competence-related meta-stereotypes would be reported in the outgroup condition.
	Study 3
H3ai	The actual vocal jitter extracted from the recordings would be significantly correlated to judges' ratings of anxiety from students' recorded responses.
H3aii	*The experimental conditions would affect the judges' ratings of anxiety, corresponding to its effect on vocal jitter.*
H3b	Students' anxiety (vocal jitter) would increase when negative ingroup stereotypes are salient compared to non-salient.
H3c	Students' anxiety (vocal jitter) would increase when negative ingroup stereotypes are salient only when taught by outgroup (not ingroup) teachers.
H3d	Students' anxiety would be affected at the onset of the class but not at the end of the class.
	Study 4
H4ai	Anxious students will perform worse on a STEM test in the presence of an outgroup teacher compared to an ingroup teacher (SIM).
H4aii	Anxious students with imposter syndrome will perform worse on a STEM test in the presence of an ingroup teacher compared to an outgroup teacher (ingroup spotlighting).
H4b	Both stereotype inoculation and ingroup spotlighting would have longitudinal effects on performance.
	Study 5
H5ai	Anxious students would have lower post-lecture confidence in mathematics when exposed to an ingroup teacher compared to an outgroup teacher (ingroup spotlighting).
H5aii	Anxious students would have higher post-lecture confidence in mathematics when exposed to an ingroup teacher compared to an outgroup teacher (SIM).
H5b	Anxious students would endorse more self-esteem preserving attributions when exposed to an ingroup teacher compared to an outgroup teacher.

Summary

To summarise, among the many factors that influence the representation gap in STEM, social factors such as stereotyping, contribute significantly towards these gaps being maintained by discouraging stigmatised underrepresented groups from pursuing or persisting in STEM. As advocacy efforts for the underrepresented increases, one significant initiative being strongly argued for is the increase of role models with similar backgrounds to the underrepresented groups that could provide identity-safe cues and boost the confidence and ultimately performance of underrepresented individuals. However, the literature suggests that there could be other factors that counter the inoculating effect of role models who are generally reported to have the ability to reduce the stereotype threat related anxieties in groups that are stigmatised. Although these anxieties could have a positive effect, generally and at excessive levels, these emotions have been reported to have negative consequences for those whose group identities are negatively stereotyped. For these groups, concerns about the stereotypes that outgroup members have towards their ingroup have been shown to compound the problem. Meta-stereotypes, as they are called, could even lead to performance deficits stemming from the activation of a belief or cognition that the outgroup perceives the ingroup in a negative way. The stereotype inoculation model literature suggests that this stereotype threat effect could be countered by exposing negatively stereotyped groups to ingroup role models that act as social “vaccines” that immunise the stigmatised groups against the experience of anxiety. However, there is reason to believe that exposing anxious underrepresented groups to ingroup role models could have detrimental effects instead of positive results. The ingroup spotlighting effect could eventually lead to more pressure on underrepresented groups since the exposure to ingroup experts or role models could lead to these groups feeling that they need to likewise excel in the field where they are the minority.

Chapter Three

Introduction to the Research Methodology

Having discussed the theoretical underpinnings of this doctoral work in Chapter 2, the approach adopted across the studies presented in the empirical chapters (Chapters 4 to 7) will first be discussed in this chapter. This chapter will outline the rationale that informed the significant decisions made in the research methodology, beginning with the ethical considerations and permissions sought in line with best practice guidelines. This is then followed by a thorough discussion of key elements of the data collection across the studies undertaken as part of this thesis including the missing data for Study 2 (Chapter 5), and the reasons for choosing to do field classroom experiments as well as the anxiety measures and stereotype salience manipulations used in Studies 3 to 5 (presented in Chapters 6 and 7). This chapter will then conclude with the rationale around the manipulation of other significant variables that have previously been conflated in the ST literature.

Ethical Considerations

Participants for the studies presented in the subsequent chapters were informed that the study they were participating in was about social attitudes and perceptions in the classroom. They were informed clearly in the participant information sheet and verbally that they were free to withdraw at any point in the study and free to refuse the usage of any partial data that was given. Extra care was taken to ensure that the studies excluded participation from those who were below the age of 18. The Information Sheet and consent forms were distributed prior to the study, and stated, “If you are not 18 years old, please do not proceed with this study.” Each participant was also asked to double confirm their age by asking them to state their age in the demography section of the survey.

Participants were also assured of their confidentiality and anonymity prior to giving their consent to take part. Consent forms with participant details were provided separately

from the questionnaires especially for studies in Chapters 4 and 5. For all studies, any data on the questionnaires that could lead to the identification of a participant were removed and the data anonymised. Information on the consent forms were only accessed by the researchers involved in the studies.

The studies in Chapters 6 and 7 involved some mild deception within the stereotype salience manipulations with information of context being withheld from the participants. The Scientific and Ethical Review Board (VCWE) at VU University Amsterdam were particularly concerned about this but the concern was addressed by including a detailed and thorough debrief since the deception involved was mild and was not expected to have any adverse effect on the participants wellbeing.

In the recruitment of human subjects for this doctoral work, all ethical protocols recommended by The British Psychological Society (The British Psychological Society, 2014) were observed. All studies presented in the subsequent chapters obtained relevant ethics approvals as follows:

1. Study 1 (presented in Chapter 4): from the School of Education Research Ethics dated 30 September 2013 (Application identification number: None given).
2. Studies 2 and 3 (presented in Chapters 5 and 6): from the University of Nottingham Malaysia Research Ethics Committee dated 3 December 2013 (Application identification number: JKK271113).
3. Study 3 (presented in Chapter 6; Validating Vocal Jitter as an Index of Anxiety): from the Science & Engineering Research Ethics Committee (SEREC) in the University of Nottingham Malaysia dated 16 December 2014 (Application identification number: JKK240714).
4. Study 4 and 5 (presented in Chapter 7): from SEREC in the University of Nottingham Malaysia Campus dated 2 October 2017 (SEREC Reference: JKK110917) and the

Scientific and Ethical Review Board (VCWE) of the Faculty of Behavior and Movement Sciences, VU University Amsterdam dated 27 September 2017 (Manuscript number: VCWE-2017-136).

Missing Data

The problem of data attrition is common in almost all research, especially in paper-based surveys. In educational and psychological studies, a missing rate of 15% to 20% was found to be common (Enders, 2003). A survey of quantitative studies published from 1998 to 2004 in 11 education and psychology journals found that 48% had missing data (Peng et al., 2006). This prevalence of missing data in research poses a number of problems for researchers such as reducing the statistical power of analyses, causing bias in the estimation of parameters, reducing the representativeness of the samples and complicating the analysis (Kang, 2013).

Understanding the mechanisms that led to the missing data was important in deciding how to handle the missing data encountered. The classification system Donald B. Rubin theorised based on the reasons for the missing data is still in use today and includes, missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR; Rubin, 1976; Little & Rubin, 2020). MCAR is when the missingness is totally unsystematic and the probability of the missing data on a particular variable is unrelated to other measured variables and to the values of the particular variable itself (Baraldi & Enders, 2010; Kang 2013). Many find MCAR a strict and unreasonable assumption to be satisfied in practice (Raghunathan, 2004; Muthén et al., 1987) but MCAR can be ideal because the analysis and estimated parameters remain unbiased by the absence of the data (Kang, 2013). MAR, on the other hand, is a less strict and more realistic assumption for missing data because the data missingness is related to other variables in the dataset but not to missing values (McNeish, 2017). Being ignorable response mechanisms, the reasons for the missing

data could be ignored when data are MCAR or MAR, thus simplifying the model-based methods used for missing data analysis (Pigott, 2001). Finally, MNAR is when the probability of missing data for a particular variable is related to the values of that variable itself (Enders, 2010), which makes dealing with MNAR data generally problematic.

Researchers have employed a plethora of techniques to deal with missing values. Among the studies with missing data surveyed by Peng et al. (2006), 97% used the listwise deletion (LD) or the pairwise deletion (PD) method to deal with the data missingness. LD (also known as complete-case analysis or casewise deletion) is the most basic and common approach to missing data. In LD, analysis is limited to cases that have complete data while discarding cases with missing values. LD, therefore, produces a complete data set that lends itself nicely to standard analysis techniques especially when the data is MCAR (Baraldi & Enders, 2010; McNeish, 2017). However, the LD method reduces the sample size and is prone to bias, and is therefore largely discouraged (Schafer & Olsen, 1998). Kang (2013) also suggests that LD may not be suitable if the sample size is not large enough. PD (also known as available-case analysis) on the other hand, excludes information only when the particular data-point needed to test a particular assumption is missing (Kang, 2013). That way, more cases can be preserved for analysis which is an improvement to the LD method. PD, however, shares the same limitation as LD which is the data should be MCAR or it will produce biased estimates (Baraldi & Enders, 2010). Pigott (2001) also suggests that a major reason why PD is not a useful general choice of method is the fact that one cannot predict when PD will provide adequate results.

Newer and more nuanced methods such as the maximum likelihood (ML) method and the multiple imputation (MI) method provide better estimates for parameters than either LD or PD. The huge advantages of these methods to the more traditional methods are: 1) no data gets thrown out, and 2) the estimates are unbiased with both MCAR and MAR data.

Although they could produce biased estimates for MNAR data, it is still at a lower level compared to LD or PD (Baraldi & Enders, 2010). In contrast to MI, the ML (also known as full information maximum likelihood, FIML) method does not impute or predict any values but uses available data, complete or incomplete, from each case to work around the missing values by computing the maximum likelihood estimates that have the highest probability of producing the sample data. As such, the ML estimate of a parameter is the value of the parameter that is most likely to have resulted in the observed data or the estimation is based on the likelihood of the observed data (Baraldi & Enders, 2010; McNeish, 2017; Pigott, 2001). ML is known to work better with large samples because it does not require the infinite number of imputations that MI does to achieve a high level of efficiency when the samples are large (von Hippel, 2016). Although von Hippel proposes conditions in which ML could be more efficient for smaller sample sizes as well, Jakobsen et al. (2017) make it very clear when MI could be used to address data missingness. Jakobsen et al. points out that multiple imputation could be used to handle missing data if the missingness is above 5% and cannot be ignored, the missing data proportion is not too large (less than 40%) and the MCAR and MNAR assumptions are not plausible. As the data missingness in the Chapter 5 study was close to 15% cases with missing values on multiple items in the questionnaire and MAR assumptions most plausible, *the robust MI technique was used* (Sinharay et al., 2001).

MI gives parameter estimates and standard errors that take into account the uncertainty caused by missing values (Rubin, 1987). *To verify that the estimates generated from the MI analysis here more closely approximates the real value, this operation was performed five times using the predictive mean matching approach because in most applications, just three to five imputations have been shown to be sufficient to provide good results (Sinharay et al., 2001)*. Plus, according to calculations, for 40% missing information, five imputations give 93% efficiency (Rubin, 1987). *Results of 5 imputed (predicted) values*

were averaged to give a single estimate for each missing data point. Additionally, given the present focus on ratios, all zero scores for all items were replaced with 1 to allow divisibility between the numerator and denominator (Vorauer et al., 1998).

Field Classroom Experiments

Previous ST studies have mainly used artificial classroom experiments or lab settings. *Laboratory experiments* are not necessarily held in a laboratory *per se,* but they are studies conducted under very controlled conditions, where highly accurate measurements are assumed to be possible. The researchers would usually determine the parameters of the experiments such as the place of the experiment, the timing, the type of participants and the conditions of the experiment, plus they usually use standardised protocols (McLeod, 2012). Shewach et al. (2019) looked at 212 studies in their recent meta-analysis paper, which is the largest meta-analysis on stereotype threat to date. *They meta-analysed the effect of ST on cognitive ability tests* and found that 208 studies were within lab settings and have not looked at the question of generalisability into the real-world situations. That is a whopping 98% of all studies which included samples that had a mean age of 18 years old or older but not the elderly, that used an experimental design with a stereotype threat manipulation and included a test of cognitive ability as a dependent variable and that included statistics that could be converted to a weighted effect size of Cohen's *d* (Shewach et al., 2019). Keeping in mind that the sample was from mainly lab experiments, Shewach et al. found that the overall ST effect in adults (which was their focus) was comparable with the estimates of ST in children and adolescents as evidenced by Flore and Wicherts (2015), and a little stronger than the meta-analysis of ST effects in adult and student samples by Nguyen and Ryan (2008). The sheer number and preponderance of lab based ST experiments does indicate that they do have advantages (Shewach et al., 2019).

A notable advantage of lab-based ST experiments is having the ability to better control and deliberate manipulation of the contexts and variables, as well as the possibility of randomly allocating participants to treatments and control groups, all of which allow for the inferences of cause and effect relationships (Heng et al., 2018; Wilson et al., 2010). Another advantage of lab-based experiments is that it is easier to replicate due to its standardised protocols, making it easier to assess the reliability of the research or results obtained (Falk & Heckman, 2009). Wei (2012) highlighted that labs also have low stake settings that do not result in major consequences on the participants outside of the experiment, nor do they involve large amounts of money (Wei, 2012). This could be an advantage of lab experiments in general but when it comes to testing stereotype threat, this feature could be a disadvantage.

A major criticism often directed at ST research is the question of the generalisability of its mostly lab-based findings to real-world settings (Heng et al., 2018; Sackett & Ryan, 2012). In other words, it has low ecological validity due to how artificial the settings are, producing behaviour that does not necessarily reflect real life and lacks mundane realism (Wilson et al., 2010). The mainly lab-based experiments are usually conducted in contrived environments that are often alien to the participants, who are also usually asked to perform tasks that are not typical of everyday living. In their review of ST research conducted in the field, specifically in high stakes or operational settings, which included experimental, quasi-experimental, and observational studies, Sackett and Ryan (2012) found that there's not enough research providing evidence of consistent and replicable threat effects in high-stakes settings. They suggest that more studies in operational testing settings is needed. For this reason, the studies in Chapters 6 and 7 diverted from previous approaches to go directly into the classroom to examine the processes of interest as it helps address the issues related to artificial environments (Sackett & Ryan, 2012). The approach chosen here is by no means to devalue the contribution of lab-based studies and their findings, especially since lab

experiments are very useful to test predictions and establish theoretical frameworks (Aronson & Dee, 2011; Mook, 1983). Instead, it is to emphasise the need for and importance of testing ST and stereotype priming effects in a more diverse and realistic setting.

Field classroom experiments used in this thesis included studies conducted in a real-life classroom setting where the experimenter still controlled the independent variables but not the extraneous variables. The control, however, is within the everyday environment of the participants *and reflected the real-world classroom situations that students face*. This contrasts with a *natural experiment* that does not allow for the experimenter to have any control over the independent variables, and which is allowed to take its natural course or occurs naturally in the real world (McLeod, 2012). A key benefit of field experiments is the higher ecological validity compared to lab experiments. In other words, the behavioural measures in a field experiment would more likely reflect real life because of its natural setting (Reis & Gosling, 2010). A further benefit of the field approach used, is the reduced likelihood of demand characteristics affecting the results as the study is conducted within a more natural environment and the participants are not aware of the aims of the study. Participants would not, therefore, behave in a way that they think the researchers would want them to but instead, they would be responding to the independent variable (McCambridge et al., 2012). These strengths make the field classroom experiment an ideal approach for the studies in Chapters 6 and 7, tying in with the purpose of using a physiological anxiety measure, as the attempt is to obtain measures closer to the natural responses of the participants.

Moreover, with regards to the aspect of role model intervention, understanding when role models are helpful in the field presents even greater practical implications (Lawner et al., 2019). In their meta-analysis, Lawner et al. (2019) examined the literature on ingroup role model intervention research *for improvement of the performance and interest among*

individuals from underrepresented groups within the STEM field. They synthesised 45 separate studies that met the inclusion criteria (21 lab studies and 24 field studies) and covered an N size of 9,930. They found that ingroup role models had a significant positive overall effect ($d = 0.20$) in field studies but no significant overall effect in lab studies with level of interaction being the only significant moderator. Their results do indicate that ingroup role models could be effective especially in field studies.

Field classroom experiments, however, can be handicapped by low samples (Shewach et al., 2019), that is, it depends on the actual size of the classroom. So, to maximise sample size across the relevant studies, the approach taken was augmented by using bootstrap simulations for the analysis of key results to be more sure of the robustness of emergent trends. *Bootstrapping is particularly advantageous when dealing with small sample sizes as it is a process by which statistics are generated over a very large number of replications and involves repeatedly sampling with replacement from the observed data to create multiple simulated datasets* (Tabachnick & Fidell, 2001). Preacher and Hayes (2004) highlights that bootstrapping is a nonparametric approach to effect-size estimation and hypothesis testing. The approach does not make any assumption regarding the shape of the distributions of the variables or the sampling distribution of the statistic, *which allows the researcher to make inferences without assuming a specific distribution for the data and which makes it a versatile resampling technique. Others such as Bollen and Stine (1990), together with Preacher and Hayes, agree that bootstrapping could be used as a way to circumvent the lack of power and it can be applied with confidence to small samples. Therefore, bootstrapping was employed in this thesis to address the smaller sample size, although it is also well noted that Preacher and Hayes (2004) suggest that bootstrapping could help in low N-size scenarios, but not by a lot.*

Anxiety Measure

In the previous chapter, the question was asked about the reason behind the weak and conflicting evidence for ST. The theoretical explanation was offered in Chapter 2 itself, while the possible methodological reason is further discussed here. The presumed mechanism of anxiety in the ST literature has mostly been measured with self-reports (e.g., in Delgado & Prieto, 2008; Flore et al., 2018), which is not surprising as participants' self-reports can be useful. Paulhus and Vazire (2007) discuss some of the reasons why many researchers take for granted that self-reported measures are the best and why it is a popular choice. Two of the advantages discussed that is particularly relevant to the choices made in this thesis are, information richness, and practical considerations. *Information richness* is based on the notion that nobody has access to the quantity, breadth, and quality of information that the people being studied themselves have, especially more private information related to very personal thoughts and emotions such as anxiety (Robins et al., 1999). An advantage like this is particularly important in the study reported in Chapter 4 and 5 where the finer nuances of the anxiety experienced by students were examined. Self-reports were also important for its practicality at those earlier stages of this thesis since they are very efficient to execute and extremely inexpensive (Paulhus & Vazire, 2007). Also, at later parts of this thesis, self-reports were advantageous to measure constructs where self-reports were the only appropriate method to use (Ozer & Reise, 1994). Paulhus and Vazire (2007) gives self-efficacy as an example of a construct that must be measured by self-reports as it is naturally a self-perception.

Although self-report measures can be useful, they do not always accurately assess the intended phenomenon (Davidson et al., 2022; Littleford et al., 2005). Feldman and Saletsky (1986), for example, reported non-verbal behaviour conveying more negative attitudes than self-reported attitudes. In a study that compared non-verbal and self-reported measures used

to capture ST-related anxiety, Bosson, Haymovitz, and Pinel (2004) investigated 37 heterosexual and 28 gay men aged between 18 and 22 years in their interaction with preschool children and quality of childcare skills under ST or control conditions. They found for gay men who were stereotype threatened, more non-verbal anxiety was demonstrated but not more self-reported anxiety. They also found that non-verbal anxiety mediated the effects of ST on the participants' quality of childcase skills while suggesting that if they had depended on self-reports alone, they would have obtained little evidence that ST produces anxiety.

Hence, the age-old question has always been regarding the credibility of self-reports for constructs like anxiety, as respondents are often influenced by consistency seeking, self-enhancement, and self-presentation in their self-perceptions. Self-reports are also plagued with inaccuracies due to self-deception and memory related weaknesses (Paulhus & Vazire, 2007). As self-reports are subject to self-regulation and control, they are also sensitive to socially desirable responding. This is important considering prevailing egalitarian and feminist norms that could pressure women to minimise their negative emotions (e.g., anxiety) – since displaying such emotions could reflect poorly on their competence and question their equality with men (Lewis, 2000; also see Brescoll, 2016, for a review). Thus, a non-reactive measure of anxiety was used in the present research.

Non-reactive measures in essence refer to measures of behaviour that is obtained without the participant being aware of it or without unduly alerting the participant that a measurement is being made or that they are being observed (American Psychological Association, n.d. b; Oxford Reference, n.d.). Other terms most commonly used to describe this approach are *unobtrusive measures* or *concealed measures*. Participants are usually unaware because data collection in non-reactive measures do not usually involve direct gathering of data from participants via a survey or questionnaire (Connelly, 2017), which is a

major advantage of non-reactive measures. Research subjects' behaviour is not influenced by the measure since they are not aware of the measurement, so there is a lack of reactivity. Non-reactive measures can be based on tangible, at times physical, evidence rather than depending on perceptions, either of the researcher or participant, as questionnaires, surveys and interviews do (Marrelli, 2007). In Marrelli's list of advantages of non-reactive measures, it is also mentioned that the results from such measures, therefore, could provide more credibility and it can be used to corroborate other results which have been obtained via other methods without introducing another new source of reactivity. Marrelli (2007) also went on to list their disadvantages, but they mainly revolved around observer error due to lack of vigilance and incorrect inferences, both of which are avoided with the method chosen here to measure anxiety.

Taking the pros and cons of self-reported (reactive) verses physiological (non-reactive) measures, the novel measure of anxiety that has been applied in linguistic research was used in this thesis, namely vocal jitter. Humans engage nearly a hundred muscles when speaking, whose movements are often sensitive to the emotions that people are feeling (Planalp et al., 1996), which is why vocal cues (e.g., jitter) are the most common direct means of determining others' emotional state (Planalp et al., 1996). In a chapter reviewing the nature of psychophysiological processes and the use of psychophysiological measures as state-of-the-art indexes of psychological constructs, Blascovich (2000) also supported the use of physiological measures to assess arousal and reactivity. *Emotions may affect short-term fundamental frequency fluctuations of voice (i.e., vocal jitter), caused by poor laryngeal musculature in response to emotional stress (Ozdas et al., 2004; Titze, 1994). Vocal fold vibrations have been predicted to become more irregular with emotional stress induced laryngeal tension. Irregular vocal fold vibrations result in higher voice jitter* (Scherer, 1986). *In fact, higher levels of jitter have been correlated with increased signs of weakness with*

muscles connected to voice (Shao et al., 2010). Emotionally, jitter, or voice tremors, has been associated with social anxiety disorder, especially when answering questions, as well as other emotions (Silber-Varod et al., 2016). Research has also shown that elevated levels of jitter in speech denote higher levels of anxiety and it was found not to differ across coping styles (Fuller et al., 1992). This suggests that jitter may be a more useful indicator of anxiety, whether clinically or in the research lab (Ozdas et al., 2004).

Hence, vocal jitter was capitalized on as an involuntary index of anxiety to provide a less biased and more conclusive test of ST effects on anxiety and performance (unlike reactive self-reported measurements that can be consciously/deliberately distorted). However, although vocal jitter allows the possibility of measuring the "experience" of anxiety, it falls short of tapping into the different nuances of that anxiety (e.g., embarrassment, fear of making mistakes, etc.). So vocal jitter did not suit the initial investigations in this thesis whose aims were simply to detect the sort of anxiety that are relevant in the classroom context to better interpret readings from the vocal jitter that were used in the later studies.

Stereotype Salience

The ST literature presents several methods to manipulate the stereotype salience of among participants. Previous experimental manipulations to heighten ST have included tests that are diagnostic versus non-diagnostic in nature for ethnic and intelligence based stereotypes (Croizet, et al., 2004; McKown & Weinstein, 2003; Steele & Aronson, 1995), making the gender identity salient before a math test (Ambady et al., 2001; Neuville & Croizet, 2007), letting females know that a task given measured ability in geometry versus memory strength (Huguet & Regner, 2007), using vignettes about famous mathematicians (mostly male) before a math test among female participants (Muzzatti & Agnoli, 2007), providing mathematical material that is said to produce gender difference in test performance

versus difference (Sekaquaptewa & Thompson, 2003), and priming females to think that their opponent in chess was male instead of female (Rothgerber & Wolsiefer, 2014).

Although these methods have their advantages, *the approach taken instead in this doctoral thesis was to directly manipulate ST by highlighting competence and/or warmth related stereotypes generated from the study on meta-stereotypes or the greater body of literature on relevant stereotypes.* This approach was used rather than the typical ST induction that focuses on fabricated stories of gender difference in test performance (Spencer et al., 1999) for three reasons: (a) different operationalisations of a construct should yield the same results, especially in light of (b) scepticisms over the efficacy of the performance-related induction of ST in recent times (Sackett et al., 2004), and (c) to focus on the impact of negative stereotypes on anxiety, the exact mechanism proposed in ST theory. It is entirely possible that prior research using information about test performance of women may not be triggering anxiety associated with a stereotype but a different source of anxiety (e.g., test anxiety).

Control of Other Significant Variables

This research also aimed to address empirical holes in ST research by specifically manipulating some of the variables that have been previously conflated with ST (teacher's group membership) as well as other measures that could introduce stereotype-irrelevant sources of anxiety that should be teased out: formality of the teacher; and timing of the measurements (at the start of class versus before an assessment).

Regarding formality, research using the classic Trier Social Stress Test has shown that placing individuals in formal settings where they are likely to be evaluated (such as classrooms) is a robust predictor of stress and anxiety (Kirschbaum et al., 1993). In other words, the formality of a panel of experts can be used as a stressor that introduces the element of social-evaluative threat (Allen, et al., 2017; Frisch et al., 2015; Vors et al., 2018).

The test is widely used to induce acute psychosocial stress in participants by exposing them to a simulated social-evaluative environment (which is exactly what a classroom is). Factors that increase anxiety could combine with stereotype-related anxiety, thus making the unique effects of stereotype-driven anxiety difficult to discern. One might expect that teachers who adopt a strictly formal demeanour in the classroom could also heighten anxiety in students and, this might confound the specific stereotype-related anxiety that this thesis is interested in. Thus, the present study will control for professor's formality by manipulating it along with our key ST variables.

With regards to timing, this thesis aimed to isolate stereotype-induced anxiety from anxiety caused by academic evaluation expectancy (i.e., 'exam fever'; Tempel & Neumann, 2014). It was expected that only anxieties tied to stereotype induction should undermine women's performance (based on ST/SIM). For this reason, participants' anxiety was measured at the onset of class when impression management concerns should heighten the pressure to disconfirm stereotypes, and after class but prior to taking a STEM test when impression management stress should be low, but exam anxiety should be high.

Chapter Four

Study 1: Understanding Sources of Students' Anxiety in the Classroom

Anxiety is the crucial mechanism that underlie the prominent theories that informed this thesis (i.e., stereotype threat, and stereotype inoculation model). In a nationally representative sampling of high school students, Osborne (2001) reported that anxiety explained 38.8% to 41.4% of the racial gap in performance. When discussing possible explanations for the performance deficit stemming from ST, several earlier (Aronson et al., 1998; Spencer et al., 1999; Steele, 1997) and later (Osborne, 2007) papers argue that anxiety caused by ST explains (mediates) the impairment of test performance, diminishing the cognitive processing. On the other hand, and in line with the stereotype inoculation model, exposure to ingroup role models prevented women's anxiety from increasing over the course of the academic year, while being exposed to outgroup mentors or not having mentors at all led to an increase in anxiety (Dennehy & Dasgupta, 2017; Dennehy et al., 2018).

Anxiety within the classroom is of particular interest in this study, considering the negative effects it can have on academic outcomes and performance especially under evaluative conditions and particularly for underrepresented stigmatised students. However, the extant literature on the topic shows *several* different types of anxiety that *could* affect students in the classroom. The common anxieties experienced by students that have been reported include test anxiety (von der Embse et al., 2018), math anxiety (Van Mier et al., 2019), foreign language classroom anxiety (Botes et al., 2020), social anxiety which have to do with social image/identity concerns (Stephan & Stephan, 1985), and even anticipatory anxiety which is the fear or worry that bad things could happen in the future (Grupe & Nitschke, 2013).

Having identified the most potent source of anxiety, it would be important to identify the condition under which these concerns were most visible amongst students (i.e., teacher

context versus friends and parents/guardians). According to Hoferichter et al. (2021), friends and parents constitute a category of audience before whom students would not normally be anxious due to evaluative concerns. However, within the classroom setting, there is good reason for students to be particularly concerned about the way their teachers perceive them as opposed to their friends (peers) who form their support group and parents/guardians who are not present. Teachers have been shown as a major contributing factor in identity-based anxiety within the language learning context (Stroud & Wee, 2006) *and they have been identified as a potential source of anxiety especially within the aspect of instructor-learner interactions (Young, 1991). Students within the STEM field in particular, could feel anxious when thinking of their teachers because STEM faculty members have been described as "unapproachable" by students* (Seymour & Hewitt, 1997)*. Thus, social image/identity-based concerns or anxiety would be highly expected within an evaluative environment where the teachers often are the evaluators. When Shapiro and Williams (2012) wrote about "other-as-source stereotype threat", they mentioned that these threats "emerge as a function of perceptions of how others might assess one's performance" (p. 179). Students would be especially concerned in an evaluative classroom setting that they are embarrassing themselves by confirming what the teacher already thinks of them. Interestingly, Young suggested three potential sources of anxiety (i.e., the learner, the teacher, and the instructional practice) but none of them were friends or parents/guardians (Alrabai, 2015; Young, 1991).*

One of the central aims of the current thesis, therefore, was to first identify specific source of anxiety that are relevant in the classroom context and to then locate the source most likely to cause them. Hence, the first study in this thesis seeks to empirically validate the critical assumptions that (a) anxiety is common within the classroom setting, (b) the most common anxiety relates to concerns over social image/identity, (c) teachers are likely the key

sources of this classroom anxiety, and (d) such anxiety can have important consequences for students' confidence and classroom outcomes (e.g., classroom participation and/or engagement). A summary of the research questions and hypotheses are as follows:

RQ1 - Is there empirical evidence to validate the assumptions regarding the sources of anxiety within the classroom? Does such anxiety have an impact on students' confidence in classroom outcomes?

H1a - Concerns related to social image/identity would be the most common source of anxiety in the classroom.

H1b - Concerns related to social image/identity would be the most common source of anxiety in the classroom within the teacher context compared to friends or family.

H1c - Concerns related to social image/identity within the teacher context would negatively affect confidence in the classroom.

Method

Participants

A total of 198 students (140 females, 41 males and 17 unknown) aged from 18 to 21 years old (M_{age} = 18.08, SD_{age} = 0.43) enrolled in the Foundation in Science programme *from the Faculty of Science and Engineering* at a *private* higher education institution in *Selangor*, Malaysia took part in this study. Given the current focus on Malaysian students, the responses of five participants who reported nationalities other than Malaysian were excluded from analysis which brought the final N size to 193. *An a priori power analysis was conducted using G*Power* (Faul et al., 2007) *to determine the minimum sample size required to test the study hypotheses. The output indicated that the minimum required sample size to achieve 95% power for detecting a medium effect at a significance of α = .05, was N = 176. Therefore, the obtained sample size of N = 193 is adequate to test the hypotheses.*

Participants were informed that the study was about social anxieties and of their right

to withdraw from the study without any penalty to them prior to signing a consent form to take part.

Design

An independent measures or between-subjects design was used in which participants were assigned to one of three anxiety source conditions (teachers versus friends versus parents/guardians) *as the independent variable*. *The between-subjects design was used to avoid participants altering their natural responses to the questions posed to them, potentially arising from deciphering what the experimental hypothesis is which is often a problem when using a within-subjects design.* The dependent measures were anxiety and confidence in classroom participation.

Procedure

To establish the context in which classroom anxieties are most apparent, three versions of the questionnaire *described in the next paragraph* were created: in one condition, participants reported the anxieties they experienced in relation to their teachers; while in a second condition, the anxieties that they reported were in connection to their friends. In a third condition, students reported the anxieties they experienced in relation to their parents or guardians. If the expectation was that teachers would be the primary source of classroom anxiety amongst students, then, to confirm this assumption, self-reports of anxiety should be greatest amongst participants assigned to the teacher condition relative to those assigned to the potentially less evaluative audiences (i.e., friends and parents/guardians).

Anxiety was measured using an adapted version of the Performance Failure Appraisal Inventory (PFAI), a multidimensional measure of cognitive-motivational-relational appraisals, developed by Conroy et al. (2002). Conroy et al. (2002) indicated good construct validity for the PFAI, showing that it is internally consistent and reliable with the α range of the sub-scales being between .74 and .81 (see also Conroy, 2001; Conroy & Elliot, 2004;

Conroy et al., 2003; for validity, reliability, and stability of the PFAI). As the PFAI has been used in multiple classroom related studies (Bartholomew et al., 2018; De Castella et al., 2013), it was adapted to suit the classroom context for this study especially since this scale provides an insight into the reasons for the anxiety that students feel within the classroom setting as opposed to the common run-of-the-mill anxiety scales. Plus, the PFAI does show evidence of being correlated to performance anxiety or anxiety in evaluative situations (Conroy et al., 2002; Conroy et al., 2003). As such, to further focus the scale on the evaluative concerns of students within the classroom, the items were phrased to prime participants to think about making mistakes and how they would feel or react. The items began with the words, "When I am making mistakes…" or "When I am not getting things right…"

The modified scale used for this study contained 27-items consisting of five different factors with (a) six items tapping the experience of *shame and embarrassment*, $\alpha = .74$; (b) seven items tapping anxieties over *devaluing one's self-estimate*, $\alpha = .80$; (c) four items measuring anxieties over an *uncertain future*, $\alpha = .69$; (d) five items assessing anxieties over *important others losing interest*, $\alpha = .87$; and (e) five items tapping anxieties over *upsetting important others*, $\alpha = .75$ (see Appendix A for a full list of items). Participants also completed a 2-item measure of confidence in classroom participation, $r(191) = .39$, $p < .001$. *The measure was limited to 2 items to reduce the participant burden when answering the questionnaire considering the risk of participants skipping over items in an already long survey. A balance between a measure that is brief enough but also reliable and valid was aimed for and so a 1-item measure was avoided.* All the scales used in this study were measured on a 6-point Likert scale except the confidence in classroom participation measure that was measured on a 7-point scale (see Appendix A).

Results

Examining the Potent Anxieties Within the Classroom Context

To examine the anxiety that is most salient for students, their scores were compared across the five anxiety subscales to a point on the scale that indicates ambivalence (i.e., the scale's midpoint of 3.5). *Since there is a single group of observations and the aim was to test the hypothesis regarding the mean for the anxiety subscale significantly differing from the midpoint, a one-sample t test was used for analysis*. One-sample *t* test revealed, as expected, that the strongest endorsement of sources of anxiety was over shame and embarrassment (social image; Conroy et al., 2002) – being the only subscale having the highest absolute positive difference from the scale's midpoint (see Table 2 and Figure 1).

Table 2

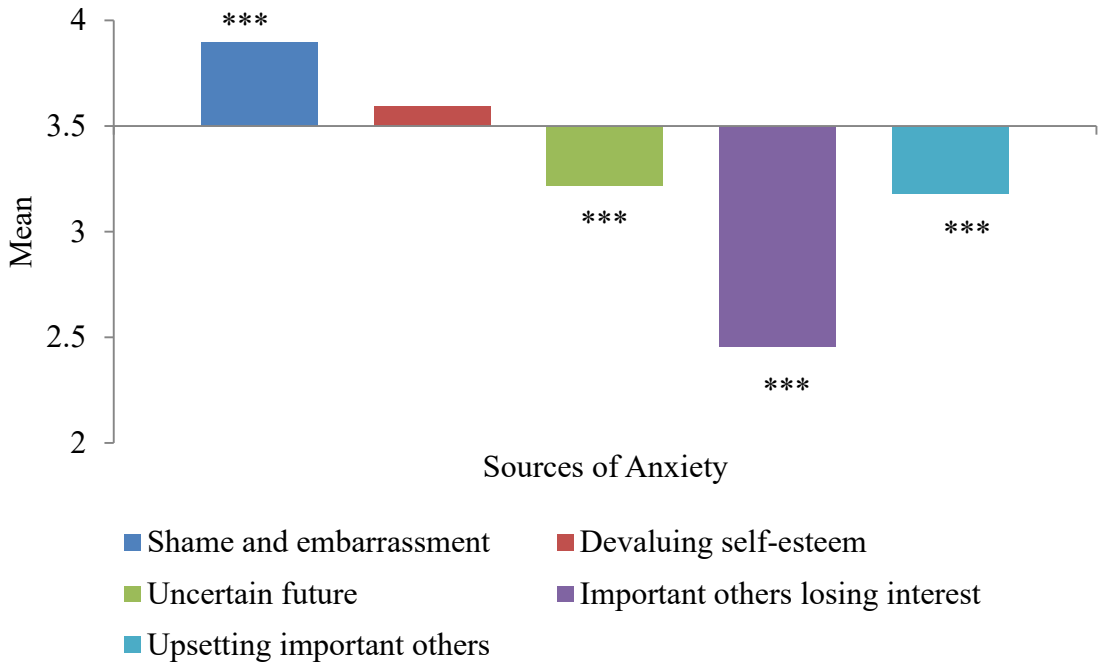
Results of the One-Sample T Test for Sources of Anxiety

Sources of Anxiety	*M*	*SD*	*df*	*t*	*p*	Cohen's *d*
Shame & Embarrassment	**3.89**	**0.98**	**191**	**5.59**	**<.001**	**0.40**
Devaluing Self-Estimate	3.59	0.99	190	1.29	.198	0.09
Uncertain Future	3.22	1.10	188	-3.53	.001	-0.26
Important Others Losing Interest	2.46	1.11	189	-12.91	<.001	-0.94
Upsetting Important Others	3.18	0.98	186	-4.44	<.001	-0.32

Note. Sources of anxiety with the highest endorsement are highlighted in bold.

Figure 1

Mean Difference from the Midpoint (3.5) for the Six Sources of Anxiety



Note. *** $p \leq .001$

Main Analysis

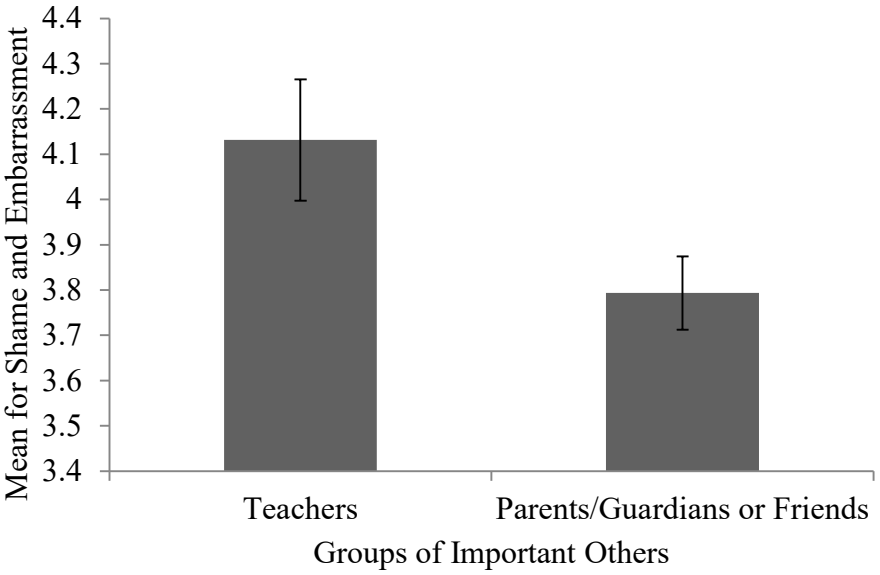
Having identified the most potent source of anxiety as "losing face" (i.e., shame and embarrassment), it becomes imperative to further explore what conditions within the classroom that these concerns manifest themselves more prominently especially addressing the question of who elicits such anxiety the most. Therefore, the next analysis aimed to identify the condition under which these concerns were most visible amongst students (i.e., teacher context versus friends and parents/guardians). As friends and parents constitute a category of audience before whom students would not normally be anxious due to evaluative concerns (Hoferichter et al., 2021), the scores across these two categories were collapsed and grouped together into a single "comfort zone" group, especially since an independent-samples *t* test showed no significant difference between friends ($M = 3.83$, $SD = 1.01$) and

parents/guardians (M = 3.76, SD = 0.89) for the shame and embarrassment subscale, $t(133)$ = 0.390, p = .697. *Also, to create a binary variable that will lend itself to a mediation analysis the data, the data needed to be organised in a way that made the most theoretical sense. It made sense pragmatically and for theoretical reasons where the groups were two conditions that were expected to be low in one and high on the other for comparison.*

The score for shame and embarrassment in this "comfort zone" condition was then compared to those in the teacher condition using an independent-samples t test to confirm whether the concerns for each source were heightened in the teacher condition or the friends and parents/guardians condition. Results from this analysis revealed, as expected, that shame and embarrassment scores were higher in the teacher condition (M = 4.13, SD = 1.01) compared to the "comfort zone" condition (M = 3.79, SD = 0.95), $t(190)$ = 2.213, p = .028 (see Figure 2).

Figure 2

Mean Difference between Teachers' and Parents'/Guardians' or Friends' Contribution to Mean of Shame and Embarrassment



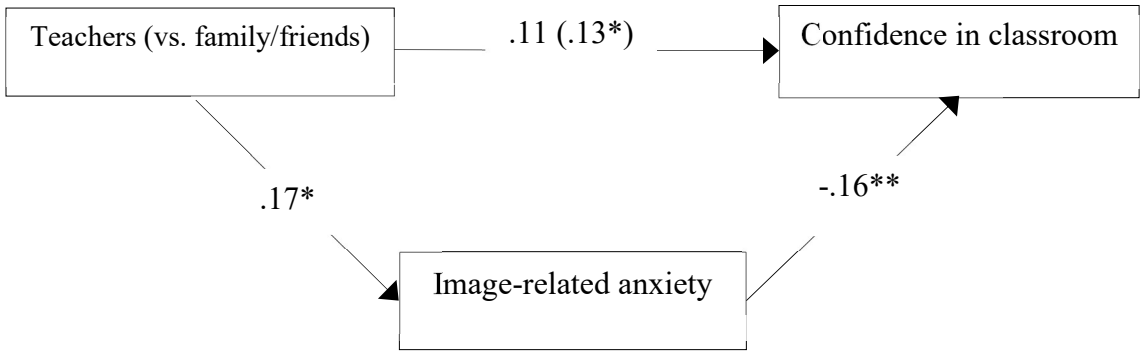
Note. Error bars represent standard errors of the means.

Does Teacher-related Anxiety Explain Students' Classroom Confidence?

So far, two things have been identified: (a) that concerns about social image (in the shape of shame and embarrassment) is the most salient form of anxiety within the classroom context, and (b) that the anxiety in the form of social image concerns is most apparent when students contemplate their relations with teachers but not people in their "comfort zone" (e.g., friends and parents/guardians). To understand the role of this anxiety in explaining classroom outcomes (in terms of students' confidence in class participation), a mediated regression analysis was performed using the bootstrap empirical simulation approach recommended by Preacher and Hayes (2004) *with a bootstrap sample size of 5000*. According to this approach, a mediated effect is established if zero lies outside the 95% confidence intervals (CI) for the estimate of the indirect effect (Hayes, 2013). The SPSS macro developed by Hayes (2013, Model 4) was used for this purpose. *Considering that the friends/parents/guardians conditions were merged to form a combined group, a weighted effect coding was used to adjust for the uneven N-sizes across the now two conditions. This is typically done using a weighted effect approach that expresses the relative ratio across both conditions in ways that equals zero* (Grotenhuis, et al., 2017). The weighted coding was used in which the teacher condition was coded 1, while the friends/parents/guardians was coded -.414 given the uneven cases across these two conditions. Consistent with the theorising in this study, teachers (versus friends/family) caused anxiety (i.e., shame and embarrassment) to increase, which in turn lowered students' self-reported confidence in participating in classroom activities although there was a positive main effect of being exposed to teachers on that classroom confidence (Figure 3). Results from a bias corrected and accelerated bootstrapped 95% CI formalised this conclusion and revealed a negative indirect effect of teachers on students' classroom confidence via anxiety over their social image, B_{IE} = -.03, *SE* = .02, 95% *CI* = [-.073, -.002].

Figure 3

The Indirect Effect of Source on Confidence in Classroom Participation via Social Image Anxiety



Discussion

In reference to the anxieties that affect students, social anxiety is said to impact an estimate of 13% of people (England et al., 2017). Jefferson (2001) defines social anxiety as a "marked and persistent fear of social or performance situations in which embarrassment may occur". Along the same lines, Chekroun and Nugier (2011) and Tangney (1991) also report that shame and embarrassment are linked to social situations. As such, the classroom becomes a relevant example of such social situations and the students surveyed would be susceptible to feeling shame and embarrassment in that context. When compared with the other sources of anxiety tested in this study, shame and embarrassment in particular, are emotions that are highly connected to self-image and self-presentation, and they are responses to how individuals are perceived by others (Chekroun & Nugier, 2011). The literature reports data that suggest that shame and embarrassment could lead to individuals responding by wanting to hide away or to avoid the social situation altogether (Tangney, 1991; McGregor & Elliot, 2005), and could also lead to feelings of incompetence or self-deficiency (Sagar & Stoeber, 2009).

In a Southeast Asian culture especially, shame and embarrassment are associated with the face culture. Face culture emphasizes maintaining social harmony, avoiding public embarrassment, and preserving one's social image or "face" (Aslani, Ramirez-Marin, Semnani-Azad, Brett, & Tinsley, 2013; Ting-Toomey, 2015). In face cultures, the fear of losing face may influence academic attitudes and behaviour. Students would tend to be cautious about public failures and feel pressured to strive for success to maintain a good social standing. Interestingly, face cultures tend to prioritize group harmony and maintaining a positive social image (Ting-Toomey, 2015).

Regarding the other sources of anxiety examined, although showing a positive difference, devaluing self-estimate did not differ significantly from the midpoint. This would agree with the findings of Henry et al. (2021) who reported that the devaluing self-estimate dimension did not emerge among responses from STEM undergraduates which may mean that students within the STEM context are not concerned about harm to their self-estimate within the classroom. It could be argued that since concerns about devaluing one's self-estimate are particularly associated with a lack of purposeful engagement in a specific activity (Conroy et al., 2007), students who chose to be in a Foundation in Science programme would generally be invested and purposefully engaged in the programme *for two main reasons, (a) to pursue the Foundation in Science programme is a choice made as it is not part of the compulsory education programme, and (b) the students enrolled are mostly high performing students as the entry requirements are higher than usual*. Hence, these particular students may not necessarily be significantly concerned about devaluing their self-estimate. Having an uncertain future, important others losing interest and the fear of upsetting important others, on the other hand, were significantly lower than the midpoint, indicating that these sources were not a concern for the participants surveyed (see Table 2 and Figure 1). Considering that the participants were focused on the effects of making mistakes instead of

the more far-reaching failure, it is unsurprising that the participants were not too concerned about an uncertain future or affecting important others in a negative way.

As shame and embarrassment is a response to how others perceive the students (Chekroun & Nugier, 2011), within the classroom setting, students would be particularly concerned about the way their teachers perceive them as opposed to their friends (peers) who form their support group and parents/guardians who are not present. Teachers have also been shown as a major contributing factor in identity-based anxiety within the language learning context (Stroud & Wee, 2006) and they have been identified as a potential source of anxiety especially within the aspect of instructor-learner interactions (Young, 1991). STEM faculty members have been described as "unapproachable" by students (Seymour & Hewitt, 1997). Therefore, there is reason for students within the STEM field to feel anxious when thinking of their teachers because. If that's the case, as mentioned in the introduction, social image/identity-based concerns or anxiety would be highly expected within an evaluative environment where the teachers often are the evaluators.

There is sufficient scholarly work to evidence that teachers can be very influential in the self-efficacy and persistence in STEM of students (Hand, Rice, & Greenlee, 2017). Hand et al. (2017) also suggest that teachers do play an important role in providing the encouragement that students need to pursue their STEM related goals. This is particularly true of students in their transitioning points from school to higher education when they must make decisions regarding STEM disciplines. Within the same paper, they go on to posit that teacher support in STEM-related subjects correlated with increase in self-efficacy and positive perceptions. Exposure to teachers should normally boost students' confidence because they provide the information or expertise needed for students to overcome their insecurities regarding incompetence, however, the results imply that the thoughts that students have over the ways in which their teachers might view them (or meta-stereotypes –

Vorauer et al., 1998; Owuamalam & Zagefka, 2011) may be particularly relevant in understanding why students might feel image related anxieties in the first place. This matter was investigated further in the following study.

Summary and Conclusions

This study aimed to validate the assumptions regarding the sources of anxiety for STEM-related students within their classroom setting and do ascertain whether such anxiety has an impact on students' confidence in classroom outcomes (RQ1). The results in this study so far demonstrated that social image-related anxieties were the most prominent form of classroom anxiety for Foundation in Science students at the higher education institution (H1a), and that teachers contributed to the social image concerns more so than friends and parents combined (H1b). Importantly, the results showed that exposure to teachers should ordinarily boost students' confidence, however, when thoughts of their teachers (relative to friends and parents/guardians) raise social image related anxieties, there is a negative impact on students' confidence in classroom participation (H1c). Given the negative impact that image-related anxieties can have on students' confidence concerning participation in classroom activities, the next aim was to identify some of the specific contents of the stereotypes that students might think their teachers hold of them.

Chapter Five

Study 2: Students' Meta-stereotypes – The Content and the Effect of Teacher Identity

Results from Study 1 (Chapter 4) showed that the classroom anxiety experienced by students at a higher education institution was related to social image, or more specifically, shame and embarrassment. As shame and embarrassment are responses to how others, in this case teachers, perceive the students (Chekroun & Nugier, 2011), Study 2 in this chapter aimed to identify some of the specific contents of the stereotypes that students might think their teachers hold of them. This chapter sought to establish the content of meta-stereotypes in the student versus teacher context, while exploring, in addition, whether students' stereotype concerns were different depending on the social identity of the teacher. Such a study is possible because evidence indicates that people such as females (Crosby, 1982), African Americans (Sigelman & Tuch, 1997), and even people with mental illnesses (Gibbons, 1981), are aware of how the ingroup is perceived by the outgroup even if the extent of that awareness can vary from one member of the group to another (Pinel, 1999). Study 2 also aimed to generate relevant meta-stereotype triggers that can be used to manipulate stereotype threat in the subsequent studies as the stereotype threat literature indicates that the awareness of the ingroup being negatively stereotyped on a given domain negatively affects the individual's performance on that domain (Spencer et al., 1999; Steele & Aronson, 1995).

Previous studies (see detailed discussion of meta-stereotypes as a source of anxiety in Chapter 2) have shown that meta-stereotypes can depend also on the specific social identities being considered (Brown & Dobbins, 2004; Finkelstein et al., 2013; Gordijn, 2010; Owuamalam & Zagefka, 2014; Vorauer et al., 1998; Xu et al., 2021), and the context of the current study is such that teachers/lecturers come from varied social groups (most prominently involving Caucasian experts and local Malaysian lecturers). In trying to understand the effect of meta-stereotypes on students in a classroom, it is not only important

to understand the social identity settings but also the content of the meta-stereotypes. *The stereotype content model (SCM) posited by Fiske et al. (2002) suggests that people assess others along two main dimensions, namely warmth and competence. This stems from the need to know the intent (i.e., warmth) and capability (i.e., competence) of others and these evaluations of warmth and competence matter because they reflect the dimensions along which people tend to evaluate and categorise social groups. The evaluations along these two dimensions are important because they influence social judgments and lead to intergroup emotions such as anxiety (Ryan et al., 2015). Studies suggest that warmth is normally attributed to the ingroup and competence to the outgroup (Cuddy et al., 2008). Examples of the different levels of warmth and competence perceptions include elderly individuals who might be seen as warm but less competent in specific domains (high warmth, low competence), or a successful professional from a different cultural background could be perceived as more competence but less approachable (low warmth, high competence), or a homeless person who might be perceived as being neither warm not competent (low warmth, low competence; Fiske et al., 2002). So, the question in this chapter would be, might students in the classroom think that their lecturers stereotype or perceive them as being aloof or incompetent, or maybe friendly and capable? Considering the social identity concerns and the meta-stereotype content, it would be expected that students would react differently depending on how they are perceived on these dimensions.*

The hypotheses for this study then is that the meta-stereotype content of students regarding their teachers would be generally negative and competence-related meta-stereotypic concerns would be more an issue when students are faced with outgroup teachers (i.e., Caucasian lecturers) rather than the ingroup teachers (i.e., Malaysian lecturers) because this is a context in which people are especially mindful of their social image (Klein & Azzi, 2001; Owuamalam & Rubin, 2014). After all, Caucasians are accorded a position of prestige

and status in Malaysian society (Owuamalam & Rubin, 2017; Owuamalam & Matos, 2019) and given that status is linked to competence (Fasel et al., 2021), it seems reasonable to also expect students to also be exceptionally weary of competence-related meta-stereotypes in the context of Caucasian (versus Malaysian) teachers. This study was also the first of its kind considering the context of the participants and the outgroup teachers.

A summary of the research questions and hypotheses for this study are as follows:

RQ2 - What is the specific content of the meta-stereotypes in the student versus teacher context? Do students' stereotype concerns differ depending on the social identity of the teacher?

H2a – Meta-stereotype content among Malaysian students regarding their teachers would be generally negative.

H2b – Meta-stereotype content among Malaysian students would be more negative in the Caucasian teacher (outgroup) condition than Malaysian teacher (ingroup)

H2c - More competence-related meta-stereotypes would be reported in the outgroup condition.

Method

Participants

One hundred and eleven Malaysian students (72 females and 39 males) aged 18 to 20 years old (M_{age} = 18.04, SD_{age} = 0.27) enrolled in the Foundation in Science programme *in the Faculty of Science and Engineering* at a *private* higher education institution in *Selangor*, Malaysia were recruited via convenience sampling for the study. *An a priori power analysis was conducted using G*Power* (Faul et al., 2007) *to determine the minimum sample size required to test the study hypotheses. The output indicated that the minimum required sample size to achieve 95% power for detecting a medium effect at a significance of α = .05, was N = 45. Therefore, the obtained sample size of N = 111 is more than adequate to test the*

hypotheses. The ethnicity composition of participants was mainly Chinese (76.6%) followed by Malay (19.8%), Indian (0.9%) and others (2.7%). Participants were informed that the study was about social attitudes and perceptions and were assured of their confidentiality and anonymity prior to giving their consent to take part.

Design

A *repeated measures* or within-subject design was used in which participants provided ratings for both Caucasian and Malaysian teachers *as the independent variables*. The dependent measures were the diagnostic ratios which indicated the existence of meta-stereotypes, the traits constituting the meta-stereotypes and the meta-stereotype valence.

Materials and Procedure

A paper and pencil survey that closely followed the diagnostic ratio approach developed by Vorauer et al., (1998) was administered to identify salient meta-stereotypes among students. Participants were provided 20 positive and 20 negative traits from Anderson's 555 personality-trait list (Anderson N. H., 1968) on which their ratings were based. Anderson's paper is a much-cited paper both in the social sciences literature and in general. Dumas et al. (2002) reported 713 published reports citing Anderson's work since 1977 to the time of writing their paper according to *Social Science Citation Index* data. At the time of writing this thesis, Google Scholar indicated that Anderson's article was cited 2539 times. Anderson's work basically provides personality-trait words that are mostly adjectives describing personality, disposition, or behaviour. Considering how meta-stereotypes tend to be person or even group descriptive, Anderson's list of traits became an invaluable resource.

The diagnostic ratio method by Vorauer et al. (1998) of establishing meta-stereotypes involves 3 stages: (a) participants' estimates of the proportion of Malaysian students that either their Caucasian (outgroup) or Malaysian (ingroup) teacher would describe along the traits provided – i.e., *target rating*; (b) participants' estimates of the proportion of the

Caucasian and Malaysian teachers that the Caucasian and Malaysian teachers themselves would describe along the same traits – i.e., *baseline rating*. The rationale being that we judge others based on how we view our own selves on the same evaluative dimension and, consequently in this case, students' expectations of their teachers' impressions of them should be based on how they themselves might view their ingroup on the relevant dimensions. And finally (c) dividing the target ratings by the baseline ratings to generate the *diagnostic ratio*. Diagnostic ratios that are significantly greater than 1 are deemed to be meta-stereotypes – i.e., such traits are assumed to be more descriptive of the ingroup relative to the reference outgroup. Specifically, for the target ratings, participants were instructed as follows:

> "*According to the stereotypes that exist among Caucasian/Malaysian university lecturers, what percentage of Malaysian foundation year students do you think could be described using the following traits? Write the percentage in the space next to each trait (0-100%).*"

In the baseline condition, the instruction read:

> "*According to the stereotypes that exist among Caucasian/Malaysian university lecturers, what percentage of Caucasian/Malaysian university lecturers do you think could be described using the following traits? Write the percentage in the space next to each trait (0-100%).*"

Results

Main Analysis

The data missingness in the current study was close to 15%. So, the more robust multiple imputation (MI) technique was used to address the issue (See Chapter 3 for a detailed justification). The operation was performed five times using the predictive mean

matching approach (Sinharay et al., 2001) to ensure that the estimates generated from the MI analysis more closely approximates the real value. Thus, results of five imputed (predicted) values were averaged to give a single estimate for each missing data point. A diagnostic ratio was then calculated for each of the participants and for each of the forty traits by dividing the target rating by the baseline rating. The value of one was used to replace zero for the calculation of the diagnostic ratio as it allows divisibility between the numerator and denominator (Vorauer et al., 1998). The ratios were then averaged across participants to produce a mean diagnostic ratio for each trait. The significance of the diagnostic ratios was tested by comparing the resulting value for each trait against 1 using a one-sample *t* test. Results from this analysis are presented in Table 3 (negative traits) and Table 4 (positive traits).

In line with previous studies showing that the content of meta-stereotypes is generally negative (Vorauer et al., 1998, 2000), all 20 negative traits were selected as being part of Malaysian students' meta-stereotypes in relation to both Caucasian and Malaysian teachers *and the diagnostic ratios were generally higher for the negative traits than the positive ones* (Table 3). There were only 2 positive meta-stereotype traits, meticulous and inquisitive, that students endorsed in relation to the two teacher groups (traits highlighted in italics in Table 4). These 22 traits (20 negative and 2 positive) are characteristics that Malaysian students perceive that both their Caucasian and Malaysian lecturers think of them.

Table 3

Diagnostic Ratios of Malaysian Students' Negative Meta-stereotypes in Relation to Caucasian and Malaysian Teachers

Traits	Caucasian Teachers		Malaysian Teachers	
	Diagnostic Ratios	*t-values*	*Diagnostic Ratios*	*t-values*
Immature	6.70***	4.77	7.79***	4.04
Shy	17.11***	6.58	10.29***	5.51
Unoriginal	4.90***	4.31	3.51***	3.34
Mediocre	3.42**	3.05	2.31**	2.77
Dependent	4.13**	3.03	3.10**	2.66
Unimaginative	5.84***	4.05	2.82***	3.56
Unenthusiastic	5.70***	4.55	2.39***	3.86
Inhibited	5.85***	3.71	2.72***	3.81
Passive	7.20***	4.76	3.88***	3.49
Aimless	6.08***	4.81	5.25***	3.80
Shallow	4.91***	4.83	3.34***	3.73
Childish	7.09***	5.19	5.71***	4.90
Irresponsible	6.01***	4.32	4.64***	3.68
Prejudiced	5.08***	3.89	4.07***	3.67
Unreliable	4.07***	4.10	5.95***	3.39
Impolite	4.65**	3.57	3.90***	3.94
Boring	4.49**	3.15	3.06*	2.43
Disrespectful	3.85***	3.10	4.00**	2.78
Narrow-minded	6.55***	3.94	3.32*	2.41
Whiny	5.84***	3.61	3.27***	3.93

Note. *$p \leq .05$, **$p \leq .01$ & ***$p \leq .001$. Note that according to Vorauer et al. (1998) diagnostic ratios more than 1 are meta-stereotypes.

Table 4

Diagnostic Ratios of Malaysian Students' Positive Meta-stereotypes in Relation to Caucasian and Malaysian Teachers

Traits	Caucasian Teachers		Malaysian Teachers	
	Diagnostic Ratios	*t-values*	*Diagnostic Ratios*	*t-values*
Intelligent	0.91***	-3.40	0.92*	-2.60
Dependable	1.48	1.43	2.58	1.88
Resourceful	0.87***	-3.57	0.85*	-2.27
Talented	1.05	0.86	0.98	-0.49
Respectful	1.28**	2.84	1.13	1.19
Capable	1.13	0.58	1.11	0.52
Productive	0.96	-0.99	0.87***	-4.08
Responsible	0.98	-0.44	0.92	-1.69
Creative	0.86***	-4.62	1.07	0.98
Independent	1.42	1.08	2.03	1.23
Self-motivated	0.92	-1.79	0.87***	-3.55
Disciplined	0.97	-0.60	0.86***	-5.26
Studious	3.50**	2.67	1.55	1.76
Ambitious	1.66	1.26	1.03	0.62
Confident	0.75***	-7.74	0.84***	-5.43
Meticulous	2.53*	2.51	2.83*	2.57
Inquisitive	2.28**	2.90	2.10**	3.23
Outspoken	1.39	1.13	1.47	1.60
Hardworking	1.04	0.77	0.92*	-2.19
Rational	0.97	-0.70	2.03	1.35

Note. *$p \leq .05$, **$p \leq .01$ & ***$p \leq .001$. Note that according to Vorauer et al. (1998) diagnostic ratios less than 1 are not part of the meta-stereotypes and the significant ratios in the table are therefore traits that participants believed are reserved for the outgroup rather than for the ingroup (indicated in **bold**). Traits in *italics* are positive meta-stereotypes for the ingroup.

Apart from that, traits with diagnostic ratios that were significantly lesser than one, intelligent, resourceful, and confident, are not part of the meta-stereotypes but are traits that students believed are reserved for the outgroup rather than the ingroup. In other words, the students believed the positive traits were more descriptive of especially their Caucasian teachers rather than themselves. At the same time, the size of the diagnostic ratios for negative meta-stereotypic endorsements in relation to Caucasian teachers seemed greater than those in relation to Malaysian teachers. Indeed, an independent-samples *t* test of the mean of the diagnostic ratios for all negative traits in relation to the Caucasian teacher context and those in relation to the Malaysian teacher context revealed a stronger endorsement of negative meta-stereotypes overall in relation to Caucasian teachers ($M = 5.97$, $SD = 2.84$) compared to Malaysian teachers ($M = 4.27$, $SD = 1.96$), $t(38) = 2.21$, $p = .03$.

The results also showed that there was a stronger endorsement of competence-related negative meta-stereotypes specifically in relation to the Caucasian teachers as compared to the Malaysian teachers. The list of negative traits could be divided to 10 traits related to competence (i.e., immature, unoriginal, mediocre, unimaginative, unenthusiastic, aimless, shallow, childish, irresponsible, and narrow-minded) and 10 traits related to warmth (i.e., shy, dependent, inhibited, passive, prejudiced, unreliable, impolite, boring, disrespectful, and whiny) according to the trait adjectives found in the multidimensional scaling solution of social (warmth) and intellectual (competence) dimensions reported in Cuddy et al. (2008). An independent-samples *t* test of the mean of the diagnostic ratios for all competence-related negative traits in relation to the Caucasian teacher context and those in relation to the Malaysian teacher context revealed a stronger endorsement of competence-related negative meta-stereotypes overall in relation to Caucasian teachers ($M = 5.72$, $SD = 1.07$) compared to Malaysian teachers ($M = 4.11$, $SD = 1.74$), $t(18) = 2.50$, $p = .02$.

Discussion

As suggested by Fowler and Gasiorek (2020) as well as Vorauer (2006), the perception of mainly negative meta-stereotypes for the students and the positive traits for the teachers could stem from the phenomenon whereby people generally think that ingroup favouritism could lead outgroup individuals to perceive them more negatively than their fellow ingroup members. That is because negative outgroup stereotypes provide positive inputs to social identity as they positively differentiate the ingroup (Klein & Azzi, 2001). However, as often is the case, these negative meta-stereotypes may not necessarily have a basis as it may not be the actual stereotypes their teachers think of regarding these students (Finkelstein et al., 2013). Investigating the actual stereotypes that teachers of varying group memberships have of Malaysian students would make a good future follow up to this study.

More important to the present study, as students report more negative meta-stereotypes, the ST effect they experience (Shelton et al., 2006) could lead to underperformance in the domain they are in (Stephan, 2014) due to the anxiety that it induces (Ma et al., 2022). The literature also highlights the relation of meta-stereotypes to self-esteem whereby negative meta-stereotypes can undermine the self-esteem of individuals from stereotyped groups (Gordijn, 2010; Vorauer et al., 1998; Vorauer et al., 2000). As such, individuals from stereotyped groups typically display both cognitive and affective reactions when negative meta-stereotypes are activated (Finkelstein et al., 2015). What's worse is that there are studies that suggest that even when primed with positive information about their group, people could struggle to believe that they could successfully counter the prevalent negative meta-stereotype (Finkelstein et al., 2020). Therefore, stigmatised underrepresented students who have predominantly negative meta-stereotypes of their teachers, could experience anxiety, and perform badly within classroom settings which consequently would lead to further disadvantage academically.

The size of the diagnostic ratios for negative meta-stereotypic endorsements in relation to Caucasian teachers was greater than those in relation to Malaysian teachers suggesting that although these meta-stereotypes were generally apparent even in the Malaysian teacher context, the degree to which the current sample of Malaysian students endorsed them in relation to their Caucasian teachers were noticeably stronger. Although within a different context, Gordijn (2010) similarly reported that women were likely to apply more negative meta-stereotypes about being overweight when they believed they were being evaluated by an attractive man (i.e., the higher status outgroup).

As the student-teacher relationship tends to be hierarchical, the concerns about being evaluated according to the negative meta-stereotypes by the outgroup teacher would be especially salient (Brown & Dobbins, 2004). Therefore, the findings in Study 2 so far are consistent with the idea that outgroup teachers (i.e., Caucasian teachers) increase the salience of negative meta-stereotypes and possibly arouse greater anxiety amongst Malaysian students (Steele & Aronson, 1995). This is especially true as the inherently evaluative role of the teacher is coupled with the meta-stereotypes that Malaysian students have of their outgroup teachers which lead to negative expectations when such teachers are anticipated to evaluate them (Brown & Dobbins, 2004). According to Vorauer et al. (1998), there is ample amount of research showing that if individuals believe that someone is holding negative preconceptions about them, their interactions with that person will be adversely affected. More specifically, if students believe that their Caucasian teachers have negative views of them, they are more likely to feel anxious about contacts with their Caucasian teachers as compared to their Malaysian teachers because of worries about negative responses (Kim & Oe, 2009; Stephan & Stephan, 1985; Vorauer et al., 2000).

As mentioned in the introduction, the competence dimension generally reflects a group's status in society (Fasel et al., 2021). As such, the lower status group members' (i.e.,

Malaysian students') meta-stereotypes of how they are viewed by the higher status group (i.e., Caucasian teachers) has centred around the negative perceptions of their group's competence. A perception of negative evaluations on the competence dimension could be harmful as status also reflects access to resources, influence, and respect, and as it has implications on performance and learning outcomes (Fasel et al., 2021). Although the results in this study agree with the initial predictions based on the literature, the findings are still novel in that there have been no known presentations of such findings among Malaysian students in relation to their ingroup (Malaysian) and outgroup (Caucasian) teachers especially in relation to the competence and warmth domains.

Summary and Conclusions

Study 2 aimed to investigate the specific content (both valence and SCM domain) of the meta-stereotypes in the student versus teacher context and whether these meta-stereotypes differ depending on the social identity (outgroup verses ingroup) of the teacher (RQ2). If students were particularly concerned about social image/identity (i.e., shame and embarrassment) in the context of their teachers as shown in Study 1, then it would naturally be expected that they would also be concerned about the way their teachers viewed them. The results of Study 2 showed that Malaysian students generally felt their teachers perceived them more negatively especially in the context of Caucasian teachers as compared to their Malaysian teachers (H2a and H2b). This concern seemed more pronounced in the context of competence-related traits (H2c). Coupling the results reported in this study together with the findings in the body of literature concerning meta-stereotypes, students who believe that their outgroup teachers have negative views of them, especially in the competence domain, are more likely to feel anxious about contacts with those outgroup teachers.

Fasel et al. (2021) do make some interesting recommendations on how to address the problem by suggesting that training provided for teachers could include "help to avoid

patronising speech, reduce behaviour that reinforces incompetence meta-stereotypes, and encourage respectful and empowering communication styles" (p. 2787). However, there is still need for the approach to be more nuanced considering the theoretical implications of the stereotype inoculation model and the ingroup spotlighting thesis. The findings in Study 2 informed the focus in Study 3 which is reported in the next chapter on the effect of activating the negative meta-stereotypes that are held by outgroup (compared to ingroup) teachers on student's classroom participation especially since negative meta-stereotypes are known to have a negative impact on stereotyped students.

Chapter Six

Study 3: Investigating the Impact of the Salience of Negative Ingroup Stereotype on Students' Classroom Outcomes

Study 2 in Chapter 5 showed that the meta-stereotypic beliefs about the students' incompetence were especially strong when the teachers under consideration belonged to the outgroup rather the ingroup. In this chapter, the broader concern over the incompetence expectation that students assume their teachers would hold of them was tapped into to investigate how anxieties linked to such beliefs might affect students' classroom outcomes. Traditional research on stereotype threat (ST) conceives incompetence, not in broad terms, but narrowly linked to a specific dimension (e.g., intellectual inferiority and poor math prowess as in the case of African Americans and women respectively, Steele & Aronson, 1995). Hence, the study reported here aims to verify key assumptions that ST makes in terms of the undermining impact of negative ingroup stereotypes (e.g., intellectual inferiority and incompetence) on the classroom outcomes of stereotyped underrepresented students.

A key factor that has been reported in the literature as a source of anxiety is the social identity of the teacher and whether the teacher is perceived as an ingroup or outgroup member (Stephan, 2014). As seen in Study 2, concerns exist in students' perceptions regarding what outgroup teachers might think about them especially when it comes to their competence. One can imagine that within the classroom setting, this could have adverse effects on the students' emotional state and, subsequently, performance, especially if the perception is believed to be negative (Fowler & Gasiorek, 2020). For example, Vorauer et al. (2000) reported an increase in evaluative anxieties in the presence of the outgroup. Following this argument, the expectation is that teachers who are perceived as an outgroup member would cause an increase in the levels of anxiety in students. In this study, Malaysian women were exposed either to a White (outgroup) or to a Malaysian (ingroup) male professor to

establish whether the salience of "incompetence" stereotypes induced threat/anxiety depending on the ethnic background (social identity) of the professor. *As mentioned in Chapter 3, vocal jitter was the measurement taken as a proxy for anxiety and it was extracted from recordings of participants' response to a short interview question (for details refer to the "Method" section.*

Southeast Asians, of which Malaysians are a part, have often been stereotyped negatively in academic domains and are often associated with underachievement compared to their East Asian counterparts (Ngo, 2006; Reyes, 2017). According to Reyes (2017), Southeast Asians have been subjected to similar stereotypes traditionally associated with African Americans. As such, it was expected that the exposure to the outgroup professor would induce identity threat concerns in these Malaysian students. Basically, this pits the two competing theories, *stereotype inoculation model* (SIM) and the alternative *ingroup spotlighting thesis*, against each other. *The assumption in SIM is that the performance of stigmatised underrepresented students on STEM-related tasks is undermined by them feeling out of place or having a sense that they do not belong in STEM classes, however, ingroup experts or role models can boost students' performance and efficacy. Specifically, the SIM would predict that the jitter in students' voices should be accentuated when negative ingroup stereotypes are salient versus non-salient, especially in classes that are taught by outgroup (but not ingroup) professors (Dasgupta, 2011). The ingroup spotlighting thesis suggests that highly anxious underrepresented individuals would achieve worse STEM outcomes when exposed to an ingroup expert.*

Although the experience of stereotype-related anxieties is a key mechanism underlying ST effects, the reality in the classroom is that there may be other sources of this largely negative emotional state. Hence, to be sure that emergent effects are purely due to the operation of stereotype threat, there is the need to also hold constant (or at least control for)

other potential sources of anxiety within the classroom context. One obvious source of anxiety (potentially unrelated to stereotypes) is the perceived formality of a teacher. The classic Trier Social Stress Test paradigm (Kirschbaum et al., 1993), has shown that the formality of a panel of experts can be used as a stressor that introduces the element of social-evaluative threat (Allen, et al., 2017; Frisch et al., 2015; Vors et al., 2018). For this reason, one might expect that teachers who adopt a strictly formal demeanour in the classroom could also heighten anxiety in students and, this might confound the specific stereotype-related anxiety that this thesis is interested in. To address this potential confound—i.e., decoupling formality-related anxiety from stereotype-related anxiety—students' perceptions of the formality of their teacher/professor were manipulated following recommendations by Chesebro (2003). Assuming the anticipated undermining effect of stereotype-related anxieties on classroom outcomes is real (as per ST framework), then such process should be visible, especially when the anxiety linked to teacher formality is controlled for.

A summary of the research questions and hypotheses for this study are as follows:

RQ3 - Are there positive or negative impacts of an ingroup (versus outgroup) expert/role model on STEM performance in underrepresented groups as a function of stereotype threat and anxiety?

RQ4 - Do the *ingroup spotlighting thesis* or the *stereotype inoculation model* provide better explanations for the poorer outcomes of underrepresented groups in the STEM context of higher education?

RQ5 - Does vocal jitter provide an accurate, less biased, and more conclusive test of stereotype threat effects on anxiety and performance?

H3ai - The actual vocal jitter extracted from the recordings would be significantly correlated to judges' ratings of anxiety from students' recorded responses.

H3aii - *The experimental conditions would affect the judges' ratings of anxiety, corresponding to its effect on vocal jitter.*

H3b - Students' anxiety (vocal jitter) would increase when negative ingroup stereotypes are salient compared to non-salient.

H3c - Students' anxiety (vocal jitter) would increase when negative ingroup stereotypes are salient only when taught by outgroup (not ingroup) teachers.

H3d - Students' anxiety (vocal jitter) would be affected at the onset of the class but not at the end of the class.

Method

Participants

One hundred and thirty-eight women undergoing their third semester in the Foundation in Science program *in the Faculty of Science and Engineering* at a *private* higher education institution in *Selangor*, Malaysia took part in this field (classroom) experiment (M_{age} = 18.87, SD_{age} = 0.60). *These were different participants from the previous studies. An a priori power analysis was conducted using G*Power* (Faul et al., 2007) *to determine the minimum sample size required to test the study hypotheses. The output indicated that the minimum required sample size to achieve 95% power for detecting a medium effect at a significance of α = .05, was N = 210. Therefore, the obtained sample size of N = 138 is lower than the required to test the hypotheses.* Participants were informed that the study was about social attitudes and perceptions and were assured of their confidentiality and anonymity prior to giving their consent to take part.

Design

A *2 x 2 x 2 between-subjects design* was used for this study where participants were randomly assigned to different conditions of a stereotype salience manipulation (salient versus non-salient); professor formality conditioning (formal-acting versus non-formal-

acting) and professor ethnic identity treatment (*Malaysian ingroup versus British White outgroup*).

The primary dependent measure was vocal jitter, which was used as a direct behavioural/physiological index of anxiety. As discussed in Chapter 3, although self-reported measures for anxiety were used in the first empirical study and was useful to address the research questions there, a novel and non-reactive measure of anxiety was used here to measure the experience of anxiety more accurately within the given conditions. The measure of vocal jitter has been used in linguistic research and is a common direct means of determining others' emotional state (Planalp et al., 1996). Vocal jitter (in percentage) was extracted using Praat version 6 (Boersma & Weenink, 2014) from brief recordings of students' response to the question, "How much *do you anticipate enjoying* this workshop?" This question was framed to mirror items commonly used to tap subject aspirations in the stereotype inoculation literature (e.g., "How likely are you to pursue a professional job in engineering"; Dasgupta et al., 2015, supplementary material p. 1). They were asked this question at the outset of the class when the convening professor was about to begin the class. Recordings were made at the recommended optimal distance for speech sound capture of 1ft (Titze, 1994), using an Olympus digital voice recorder (VN-6800PC, set at HQ recording mode with a sampling frequency of 44.1kHz/32kbps). This early recording was taken because it was a period in the session that an impression management concern (i.e., stereotype disconfirmation anxiety) with regards to meeting their new professor should be most apparent.

Procedure

After a session on biology that lasted about 30 minutes, the vocal jitter was also extracted from a subsequent recording of students' response to the question, "How much *did you enjoy* this workshop?" This was done to test whether the ST induction in this study

continued to influence students' anxiety beyond the initial meeting with their new professor when impression management pressures should have eased somewhat. This was to provide an initial gauge to determine whether the distinction being made with regards to stereotype-relevant anxiety versus exam fever following Fuller et al.'s (1992) example was even possible using this novel approach. That is, stereotype-related impression management concerns should be rife at the onset of the class but should be less influential after such initial encounters.

Professor identity. This variable was manipulated by presenting the female students with two professors of roughly similar years of teaching experience (> 5years) *which they have not been exposed to prior to this study*. Roughly half the students were exposed to a British White male professor (outgroup condition, $n = 74$), and the other half were exposed to a Malaysian male professor (ingroup condition, $n = 64$). Hence, the gender identity of the professor was constant, with racial identity being the distinguishing factor across the two conditions *and the professors were male instead of female due to the availability of candidates to fulfil this role for the study*. For this study, the outgroup condition was limited to only a British White professor due to the perceived status in Malaysian society given to them (Owuamalam & Rubin, 2017; Owuamalam & Matos, 2019) and the link to competence due to that perceived status as also mentioned in Study 2 (Fasel et al., 2021). The key inoculating ingredient in SIM is a "shared social identity" and this was met by exposing Malaysian women to a male professor who shared a racial identity in common with them (versus a male professor who obviously did not; Allen & Collisson, 2020; Marx et al., 2009). Hence, in this context, an inoculative benefit should manifest for Malaysian women in the Malaysian male professor condition. Importantly, the lecture notes across the two professor identity conditions were the same, so that resulting effects cannot be easily attributed to lecture quality.

Stereotype salience. Participants were led to believe that they would undergo a memory task involving word fragments. Prior to the class starting, half the students were exposed to incompetence-related stereotypes on PowerPoint slides on the computer screen that was mounted on each desk in the lab (i.e., salient stereotype condition, $n = 72$). In a non-salient stereotype condition, another half were exposed to pictures of beautiful scenery, also presented on PowerPoint slides ($n = 66$). In the salient stereotype condition, two incompetence stereotypes ("incapable" and "dumb") were presented on-screen via a desktop assigned to each participant, along with other stereotypes often associated with Malaysian students ("shy"; "timid", and "fearful"; see also findings from Study 2). *In this salient condition, the instructions given to the participants read as follows:*

> *"Before you proceed with today's workshop, we wanted to show you 5 statements based on a previous study conducted to identify stereotypes Caucasian lecturers have towards Malaysian Foundation Year students. Your task is to remember these statements. At the end of the 5 pictures, you will be given a word fragment task to complete."*

Some might argue that the latter list of stereotypes used here could have caused a self-fulfilling effect in the sense that knowing the ingroup is stereotyped to be anxious could cause anxiety to increase beyond any specific effect of the incompetence stereotypes. This would be in line with what others like Kamans et al. (2009) reported regarding Dutch Moroccans acting in line with the negative meta-stereotypes as well as Guyll et al. (2010) regarding Latinas confirming the stereotypes of their ingroup. However, it was precisely because of the mechanism of self-fulfilling prophecy, potentially caused by the anxiety-related stereotype, that these attributes were paired with the incompetence stereotypes, so that participants' anxious reactions can be interpreted from the lens of how they might feel in the current context (Madon et al., 2011).

To determine the efficacy of this novel stereotype-threat induction, participants were afterwards presented with a word-fragment completion task designed to check that the intended stereotypes were indeed salient across the two conditions. *The instructions given to participants were:*

> *"Complete the following word fragments as quickly as possible by filling in your answer in the space provided. If you are unable to think of a word immediately, please skip the question and move on to the next one."*

The five word-fragments which were t_m_d, _ea_ful, s_y, d_m_, and in_ _ _ able could be completed with at least one other word that were unrelated to the stereotypes that participants in the salient stereotype condition were exposed to (e.g., t**i**m**i**d/t**a**m**e**d). Hence, a stereotype was salient if the word-fragment completions corresponded to the stereotypes that participants in the salient stereotype condition viewed.

Professor formality and other controls. The reasoning here was that the professors' *formality* could contribute to feelings of anxiety in a classroom (Kirschbaum et al., 1993). To control for this possibility, the perceived formality of the professor's behaviour in the classroom was manipulated, using an adaptation of a procedure developed by Chesebro (2003). Specifically, half the participating students were exposed to a professor that acted formal: he wore a formal looking buttoned-up shirt with a tie and properly rolled down sleeves. He started the class with a set of rules and instructions that must be adhered to. Facial expressiveness was also relatively emotionally neutral (e.g., he did not smile) and he maintained limited eye contact with participants (n = 78). The other half (n = 60) were exposed to a professor who was non-formal-acting: he wore a casual shirt with rolled up sleeves and no tie, started the class with some humour, was generous with facial expressions (e.g., he smiled a lot) and maintained eye-contact with participants. The reasoning was that a formal-acting professor should elicit more teacher-related anxiety amongst students relative

to a less formal-acting one. The validity of this assumption was assessed using an 8-item measure of professor-related class anxiety adapted from the Questionnaire on Teacher Interaction (QTI) scale developed by Fisher et al. (1995): "we are afraid of this lecturer"; "this lecturer's standards are very high"; "this lecturer can take a joke" (reverse scored); "this lecturer is strict"; "this lecturer has a sense of humour" (reverse scored); "this lecturer's class is pleasant" (reverse scored): "this lecturer is friendly" (reverse scored); "we have to be silent in this lecturer's class" (1 = *strongly disagree*, 7 = *strongly agree*; Cronbach's α = .73).

Moreover, social interactions also tend to be rife with interpersonal/group anxieties (Shelton et al., 2006; Stephan, 2014). One argument against the use of vocal jitter as an index of anxiety, therefore, might be that participants' vocal fluctuations could result from mimicking the vocal modulations of the teaching assistant (TA) posing the question that students responded to (Rueff-Lopes et al., 2015). This is especially true given past evidence of emotional contagion (Hatfield et al., 1993) and mimicry of vocal properties (Eriksson & Wretling, 1997; Rueff-Lopes et al., 2015). To counter this argument, the vocal jitter from the TA was extracted as well and factored in into the models here as a covariate, so that its effects are statistically held constant across participants/conditions.

After the experiment, participants were thoroughly debriefed about the true purpose of the study and thanked (with chocolates) for taking part. The workshop was as part of the Genetics and Living Systems module (or course), designed to link the understanding from the module to real life, although the contents of the workshop was not part of the module's assessment.

Validating Vocal Jitter as an Index of Anxiety

Research on spontaneous trait inference shows that people can make accurate judgements of other's emotions via thin slices of behavior that are typically a few seconds long (Ambady et al., 2000; Wu et al., 2016). This accuracy in emotional detection seems

much more evident in voice-only scenarios than in visual-only or even mixed scenarios (Krause, 2017). This human capability was used to validate the connection made between anxiety and vocal jitter across this investigation. If the sounds of peoples' voices convey emotions to others, then there must be an acoustic property that perceivers use as heuristic to detect anxiety in others (e.g., vocal jitter). Hence, five independent judges who were blind to the experimental treatments were recruited to rate the anxiety in the slices of voice recordings that were collected from participants (on a 7-point Likert scale: 1 = *not at all present*, 7 = *very much present*), along with five other emotions carefully selected to be relevant in the classroom context. These emotion detection ratings were done for the onset and after class recordings of participants'/students' vocal responses, reasoning that people who are not privy to the experimental treatments ought to be able to connect the vocal jitter in the recordings to the students' experience of anxiety at a level well above chance.

Assuming vocal jitter is a valid measure of anxiety then one should expect confirmation of the following predictions:

Hypothesis H3ai: a significant correlation between the ratings of anxiety from students' recorded responses and the actual vocal jitter extracted from those recordings.

Hypothesis H3aii: there should be an effect of the experimental treatments on the judges' ratings of these emotions (especially anxiety) that also correspond to the effect of the experimental treatments on vocal jitter.

For construct validity especially in relation to the testing of Hypothesis H3ai, convergent and discriminant validations were used, as originally proposed by Campbell and Fiske (1959), to be able to justify the novel measure of anxiety via vocal jitter. Convergent validity tests the constructs to see if the constructs that are hypothesised to be related are indeed related in reality, whereas discriminant (or divergent) validity tests the constructs to

see if the constructs that should not have a relationship indeed do not have a relationship in reality (Taherdoost, 2016).

Convergent Validity. One of the five emotions was the "depressed" affect, and this was chosen because it is not only prevalent in schools (Lewinsohn et al., 1993), but also highly comorbid with anxiety (Kessler et al., 2005). Therefore, it allows the test for convergent validity with the patterns expected for anxiety ratings (Campbell & Fiske, 1959; Raykov, 2011). The judges were expected to also be able to connect their ratings of depressed affect to the students' vocal jitter, as with their ratings for anxiety.

Discriminant Validity. For the sake of discriminant validity, two things were done. Firstly, another negative emotion (restless) that was also deemed relevant in the classroom situation (Rafidi, 2020) was included. Assuming vocal jitter signals from students' recorded speech were simply down to the experience of negative emotions in general, rather than the anticipated feelings of anxiety, then judges should also be able to consistently detect this negative emotion in their ratings of students' recordings at a level that is well above chance.

Secondly, the judges were asked to rate 3 positive emotions ("confident," "inspired," and "excited") that were deemed relevant in a classroom context. Students who are in awe of their teachers (i.e., especially those who are inspired or excited) ought to show breaks in their speech patterns that might contribute to vocal jitter (Kamiloğlu et al., 2020). Hence, to be sure that vocal jitter is not symptomatic of other positive emotions that students might experience in the classroom, especially before a teacher that they have met for the first time, then the judges should be less likely to connect the vocal jitter in the recordings of students' responses in class to either of these emotions at a level well above chance, or not at all (i.e., a null association).

Results and Discussion

Assumption and Manipulation Checks

Validating Vocal Jitter as an Index of Anxiety. ***Preliminary Analysis.*** The main question in this section was, did the judges converge in their ratings of the emotional reality in the students' voice recordings? To answer this question, a confirmatory factor analysis was performed on the judges' ratings for each emotion type to be sure that the same latent emotion was captured in their ratings of the onset and after class voice-only recordings. This analysis was performed in *jamovi* version 1.0.4 (The jamovi project, 2019), using its model modification indices to select-out ill-performing ratings that either led to singularity or that did not permit the model to converge. Results from this analysis are presented in Table 5, showing that the judges largely converged in the emotional reality that they rated.

Testing Hypothesis H3ai. To establish whether the judges could accurately detect each of the 6 emotions from the vocal jitter in the students' voices, a series of bivariate correlations was then performed, testing the association between emotionality ratings and the students' actual vocal jitter. Confirming the assumption that vocal jitter indexed anxiety, results revealed that the independent judges were able to associate the vocal jitter in the students' voiced recordings to the experience of anxiety at a level well above chance (i.e., *p*s < .050 for the onset and after class ratings, see Table 6). These findings agree with Fuller et al. (1992) and provides further evidence that vocal jitter is a valid and reliable indicator of anxiety. Also, the judges' ratings of anxiety and depressed affect were moderate-to-highly positively correlated (*r*s = .39 - .56; *p*s < .001; Kessler et al., 2005) and, unsurprisingly, they were also able to connect the vocal jitter in the students' voice recordings to this comorbid emotion at a level well above chance as highlighted earlier from Ambady et al. (2000) and Wu et al. (2016) (*convergent validity*; see Table 6). Evans et al. (2021) reported good convergent validity for anxiety and depression on the self-reported anxiety and depression

subscales of the 21-item Depression, Anxiety, and Stress Scales which is used frequently to measure symptoms of depression, anxiety, and stress in adults. Their findings along with several others who used self-reported measures (Gotlib, 1984; Reidy & Keogh, 1997) are now here further supported by the physiological measure of vocal jitter which, as argued earlier in this thesis, has the potential of being a less biased and more conclusive test of anxiety (see Chapter 3 for a detailed discussion).

Interestingly, however, and reinforcing the *discriminant validity* of the approach taken here, a corresponding association of the vocal jitter in the students' recordings to the other negative emotion of restlessness was not observed (see Table 6). As mentioned earlier, if the judges were perceiving negative emotions in general rather than the feelings of anxiety specifically, they would be able to consistently detect these other negative emotions in their ratings. However, the absence of a significant correlation between the vocal jitter and the other negative emotion provides further evidence supporting the suggestion that the judges were able to specifically detect levels of anxiety in vocal jitter at a level above chance.

Finally, the judges' ratings of the vocal jitter in the students' recorded voices were not indicative of the experience of the three positive emotions of confidence, inspiration, and excitement, as shown by their reliably negative correlations with this acoustic property. In other words, the judges did not connect the vocal jitter in the recordings of students' responses in class to any of these emotions at a level well above chance. Although Kamiloğlu et al. (2020) proposed that vocal jitter could be a "likely candidate" to index positive emotions like happiness, they did admit that the empirical evidence was limited and that they could not draw a clear conclusion. The evidence put together here demonstrates that the independent judges who were blind to the experimental thesis, could detect anxiety specifically from the vocal jitter in the recorded responses for the average student, accumulating the support for the use of vocal jitter as a physiological index for anxiety.

Table 5

Did the Judges Converge in Their Ratings of the Emotional Reality in Participants' Voice Recording? A Confirmatory Factor Analysis

	Ratings of Class Onset Recordings					
	Anxiety	**Depressed**	**Restless**	**Excited**	**Confidence**	**Inspired**
X^2 (*df*)	9.97 (5) {1.78 (2)}	xx {1.91 (2)}	xx {6.31 (2)}	8.12 (5)	16.9 (5) {3.18 (2)}	18.8 (5) {1.33 (2)}
***p*-value**	0.08 {0.41}	xx {0.38}	xx {0.04}	0.15	0.01 {0.20}	0.00 {0.52}
CFI	0.79 {1.00}	xx {1.00}	xx {0.68}	0.97	0.91 {0.99}	0.88 {1.00}
SRMR	0.06 {0.03}	xx {0.03}	xx {0.05}	0.04	0.05 {0.02}	0.05 {0.02}
RMSEA	0.09 {0.00}	xx {0.00}	xx {0.13}	0.07	0.13 {0.07}	0.14 {0.00}
AIC	2630 {2075}	xx {1943}	xx {1897}	2331	2039	2411 {1972}
	Ratings of After Class Recordings					
X^2 (*df*)	xx {1.36 (2)}	xx {*<0.00 (0)*}	xx {3.33 (2)}	3.06 (5)	5.25 (5)	6.83 (5)
***p*-value**	xx {0.51}	xx {*NaN*}	xx {0.19}	0.69	0.39	0.23
CFI	xx {1.00}	xx {*1.00*}	xx {0.96}	1.00	0.99	0.96
SRMR	xx {0.03}	xx {*3.10e-9*}	xx {0.04}	0.03	0.04	0.04
RMSEA	xx {0.00}	xx {*0.00*}	xx {0.07}	0.00	0.02	0.06
AIC	xx {1838}	xx {*1251*}	xx {1618}	2020	2243	2073

Note. Numbers outside parenthesis reflect fit indices when all 5 ratings were included in the model. Numbers in square brackets represent fit indices when a single non-performing rating was excluded from the model (i.e., a 4-item model): A 3-item model is also presented in the square bracket, but with italicized statistics. The "xx" sign means that the relevant model either did not achieve convergence or that the information matrix was singular. *NaN* = not available due to zero *df*.

Table 6

Bivariate Correlations between Vocal Jitter and Judges Ratings of Participants' Emotionality

	Anxious		**Depressed**		**Restless**		**Excited**		**Confidence**		**Inspired**	
	Onset	After	Onset	After	Onset	After	Onset	After	Onset	After	Onset	After
Unadjusted ratings	.20*	.23*	.39**	.18*	.08	.15	-.39**	-.27**	-.32**	-.14	-.34**	-.28**
Adjusted ratings	.19*	.19*	.34**	.10	.11	.16	--	--	-.34**	--	-.32**	--

Note. Statistics presented in the unadjusted row depict the correlation between vocal jitter (onset and after class) and the unadjusted average of judges' emotional ratings, while those in the adjusted row depict the association between vocal jitter (onset and after class) and the transformed average of judges' emotional ratings. The unadjusted emotional rating measure utilises all ratings of the 5 judges. The adjusted emotional rating measure utilises only those ratings found in the confirmatory factor analysis to be tapping the common latent emotion. *$p \leq .05$ & **$p \leq .01$

Testing Hypothesis H3aii. To establish whether the judges' ratings could also detect the experimental treatment that participants were assigned to, a repeated measures MANOVA was performed with the 3 emotion clusters (anxiety, positive emotions, and other negative emotions) and time of recording (onset versus after class) as within-subject factors, while the anxiety eliciting treatment, stereotype salience (salient versus non-salient) was specified as the between-subject factor. Anxiety and depressed affect were clustered to represent *anxiety*, while inspired, confidence and excitement were clustered as *positive emotions*. Restless was isolated as the *other negative emotions*.

Results revealed a main effect of emotion cluster (see Table 7), showing that *anxiety* was the top-rated emotion ($M \pm SE$; 3.40 ± 0.05) relative to both the *positive emotions* (3.08 ± 0.05, $p < .001$) and *other negative emotions* (3.26 ± 0.05, $p = .068$). Importantly, the emotion cluster x salient stereotype interaction was significant (see Table 7) and, simple effect analysis revealed that it occurred because the judges' reported significantly higher *anxiety* ratings for students who were assigned to the salient stereotype treatment (3.55 ± 0.07) compared to those who were assigned to the non-salient stereotype (control) condition (3.26 ± 0.08; $p = .005$).

Table 7

The Effect of Salient Stereotypes on Judges' Emotional Ratings as a Function of Time of Recording and Emotion Type

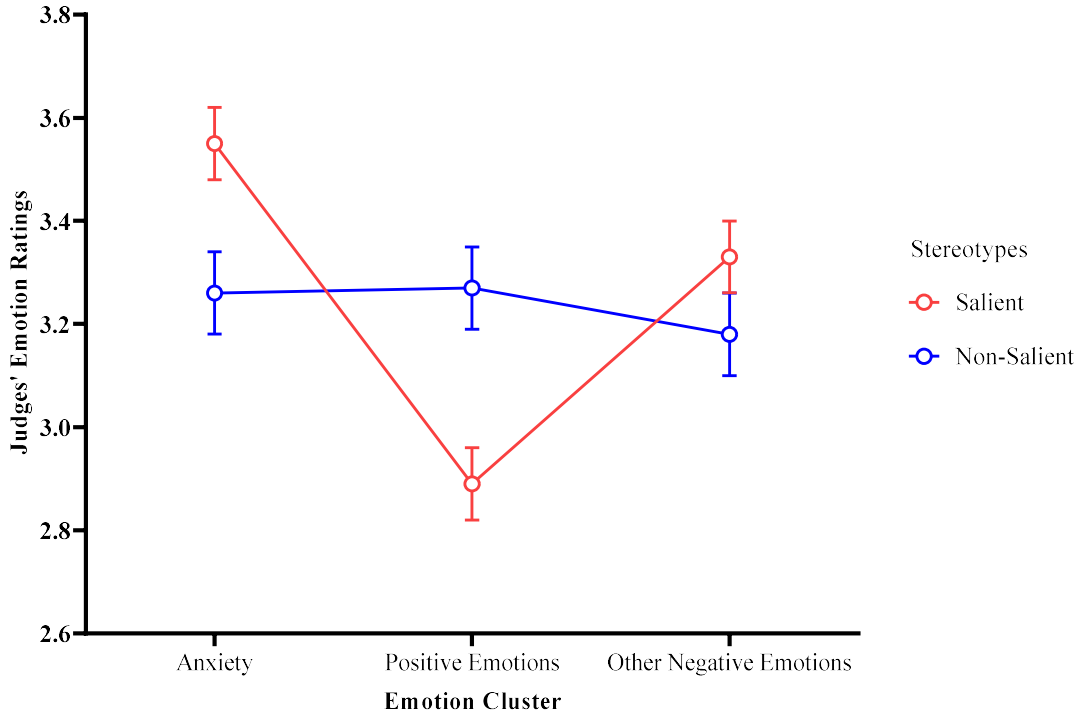
	Sum of Squares	df	Mean Square	F	p	η^2
Emotion Type	12.34	2	6.17	8.16	<.001	0.041
Emotion Type * Salient_stereotypes	14.70	2	7.35	9.72	<.001	0.049
Residual	179.96	238	0.76			
Time	0.30	1	0.30	2.62	0.108	0.001
Time * Salient_stereotypes	0.28	1	0.28	2.43	0.122	0.001
Residual	13.72	119	0.12			
Emotion Type * Time	0.53	2	0.27	1.43	0.240	0.002
Emotion Type * Time * Salient_stereotypes	0.47	2	0.23	1.26	0.286	0.002
Residual	43.96	238	0.19			

Note. Type 3 Sums of Squares

Additionally, and to resolve any lingering doubt that the judges accurately estimated the relevant emotions across the salient stereotype treatment, a difference in their ratings of *other negative emotions* (restlessness) in the salient (3.33 ± 0.07) versus non-salient (3.18 ± 0.08, $p = .155$) stereotype conditions was not observed, while ratings of *anxiety* continued to surpass *other negative emotions* in the salient stereotype condition ($p = .034$). Finally, the judges were even clearer in their report that students in the salient stereotype condition experienced fewer *positive emotions* (2.89 ± 0.07) than their counterparts in the non-salient stereotype condition (3.27 ± 0.08, $p < .001$, see Figure 4). Therefore, to summarise the data presented in Figure 4, judges perceived significantly higher anxiety in the vocal recordings of students in the salient stereotype condition compared to those in the control condition although they were blind to the experimental conditions. These judges also perceived lower positive emotions in the same condition while reporting no differences in other negative emotions.

Figure 4

Judges' Emotional Ratings Depended on Whether the Students Received the Salient versus Non-salient Stereotype Treatment



Note: Error bars = standard errors.

Stereotype salience manipulation check. Concerning *stereotype salience*, results from an independent-samples *t* test confirmed that stereotype-relevant word-fragment completions were significantly higher in the salient than in the non-salient condition (see Table 8) indicating that participants in the stereotype salient condition would have been primed with the relevant threat inducing stereotypes (Gilbert & Hixon, 1991; Sinclair & Kunda, 1999). In relation to the *professor-related anxiety*, results from an independent-samples *t* test revealed that self-reported teacher-related anxiety was significantly greater for those students exposed to a formal-acting than a non-formal-acting professor (see Table 8). This is in accordance with the expectation that perceived formality of a teacher could be a

source of anxiety that is potentially unrelated to stereotypes as it could introduce the element of social-evaluative threat (Allen, et al., 2017; Frisch et al., 2015; Vors et al., 2018). Lastly, onset vocal jitter was positively correlated with after class vocal jitter ($r = .33$, $p < .001$), suggesting that both measurements tap a common underlying process, but further investigation would be needed to know exactly what the underlying process is especially how the experimental conditions affect the onset and after class vocal jitter (or anxiety).

Table 8

Stereotype Salience Manipulation Check Results

Number of stereotype-related word fragment completions			**Self-reported teacher-related classroom anxiety**		
Salient stereotype condition	Non-salient stereotype condition	d_{Cohen} [95% CI]	Formal condition	Non-formal condition	d_{Cohen} [95% CI]
4.82	1.79	3.41	3.06	2.34	1.33
(0.51)	(1.18)	[3.26, 3.56]	(0.63)	(0.41)	[1.24, 1.42]

Note. Numbers in parenthesis are standard deviations and numbers outside parenthesis are means.

Testing SIM's Version of ST

Students' Vocal Jitter at the Beginning of Class. Full ANOVA model results are presented in Table 9 and show a main effect of stereotype salience on anxiety (% vocal jitter). Consistent with the ST perspective (Steele, 1997; 2010) and agreeing with the ST literature in other contexts (Delgado & Prieto, 2008; Hess et al., 2003; Stone et al., 1999), vocal jitter was significantly higher when negative stereotypes were salient ($M \pm SE$: 5.50 ± 0.19) compared

to when they were non-salient (4.47 ± 0.18). Also consistent with the salience x identity interaction prediction, salience was qualified by professor identity (see Table 9). Stereotype-induced accentuation of vocal jitter was apparent only when Malaysian female students were exposed to an outgroup White male professor, $F(1, 131) = 31.53$, $p < .001$, $\eta_p^2 = .16$, but not an ingroup Malaysian male professor, $F(1, 131) = 0.02$, $p = .886$, $\eta_p^2 < .001$ (see Figure 5). The presence of an outgroup member has been commonly used and has been reported to activate ST because "it facilitates the activation of the negative stereotype" that targets the ingroup (Laurin, 2020, p. 1915). In the outgroup professor condition, the students could be anxious about confirming the negative stereotypes that they think are in the minds of their professor and so they may feel that they could be judged badly by these professors because of that. In other words, the presence of the outgroup professor could have triggered stereotype threat (Desombre et al., 2018) or as Shapiro and Neuberg (2007) put it, this threat relates to a threat from outgroup members.

The main effect of professor formality (4.63 ± .20 versus 5.53 ± .22, formal versus non-formal respectively, $p < .001$) was also contingent upon the professor's social identity (see Table 9). Interestingly, against expectation, this interaction occurred because students' vocal jitter was reliably more elevated when the professor was *less* rather than more formal acting, but this occurred only when the professor was an outgroup member, $F(1, 131) = 20.16$, $p < .001$, $\eta_p^2 = .133$. This effect was absent when the professor was an ingroup member, $F(1, 131) = 0.01$, $p = .909$, $\eta_p^2 < .001$. Although this is contrary to the formality prediction which anticipated that teachers who adopt a strictly formal demeanour in the classroom could heighten anxiety in students (Allen, et al., 2017; Frisch et al., 2015; Vors et al., 2018), it is consistent with the inoculation hypothesis that ingroup experts could reduce anxiety (Dasgupta et al., 2015).

None of the results changed meaningfully: (a) in a subsequent bootstrap simulation with 20,000 resamples of the raw data that somewhat helps to address the issue of sample adequacy and statistical power (Table 9, Model 2; Preacher & Hayes, 2004); (b) when the potential effect of emotional contagion from the TA's vocal modulations was controlled for (Table 9, Model 3); or (c) when students' self-reported experience of teacher related anxiety was accounted for (Table 9, Model 4).

Table 9

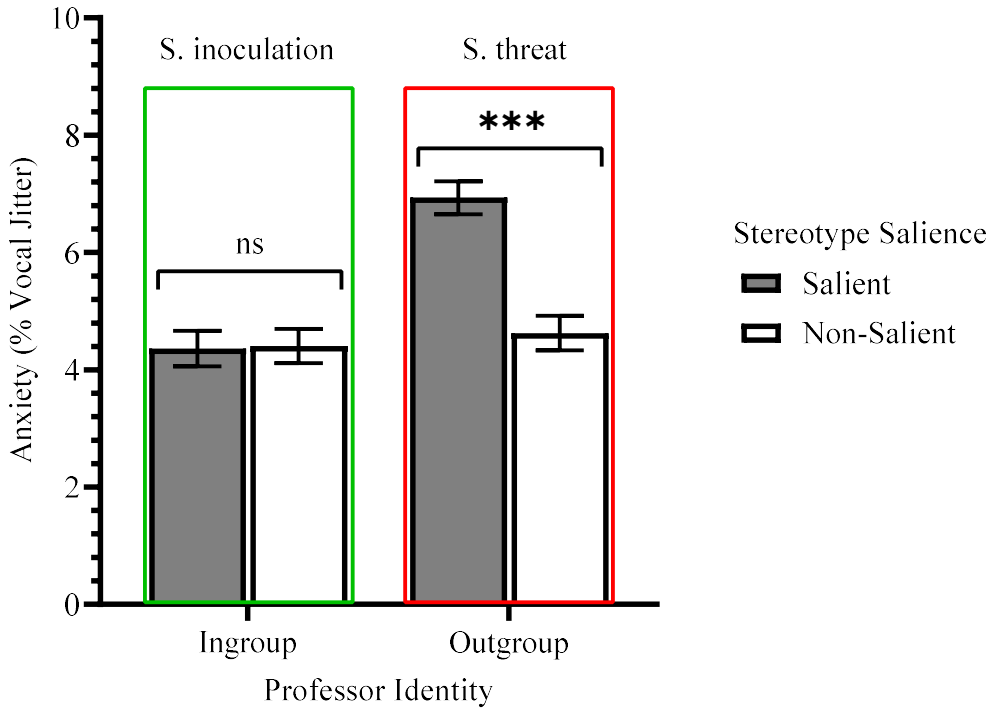
The Effects of Professor Identity, Formality and Stereotype Salience on Students' Anxiety (Vocal Jitter)

	Model 1 No controls (raw data)			Model 2 Bootstrap simulation (N = 20,000)			Model 3 Controlling for TA's vocal jitter			Model 4 Controlling for teacher-related anxiety		
	F	*p*	ηp^2	*F*	*p*	ηp^2	*F*	*p*	ηp^2	*F*	*p*	ηp^2
Professor identity (PI)	22.80	<.001	.148	22.80	<.001	.148	20.67	<.001	.137	24.68	<.001	.160
Stereotype salience (SS)	14.88	<.001	.102	14.88	<.001	.102	13.59	<.001	.095	14.68	<.001	.101
PI * SS	16.04	<.001	.109	16.04	<.001	.109	14.31	<.001	.099	15.15	<.001	.104
Professor formality (PF)	9.31	.003	.066	9.31	<.001	.066	7.36	.008	0.54	10.65	.001	.076
PI * PF	10.34	.002	.073	10.34	.001	.073	10.08	.002	.072	9.71	.002	.069
Control variable	--	--	--	--	--	--	0.12	.736	.001	1.97	.163	.015

Note. $N = 138$. TA = Teaching assistant. Professor identity (-1 = ingroup, 1 = outgroup); Negative ingroup stereotypes (-1 = non-salient, 1 = salient); Professor formality (-1 = non-formal, 1 = formal). Bootstrap simulation was conducted in SPSS.

Figure 5

The Effects of Stereotype Salience on Students' Vocal Jitter (Anxiety) When Their Professor Was Either Ingroup or Outgroup



Note: S. inoculation = stereotype inoculation. S. threat = stereotype threat. Error bars = standard errors. *ns* = non-significant; ***$p < .001$.

Students' Vocal Jitter at the End of Class. Results revealed neither a main nor an interactive effect of stereotype salience, perceived professor's formality, or professor identity on vocal jitter after class ($ps > .100$). Hence, the treatments did not exert an influence on students' anxiety beyond the onset of class. As was expected, participants' anxiety at the onset of class were affected by the experimental conditions. This was the point at which impression management concerns would heighten the pressure to disconfirm stereotypes (i.e., stereotype disconfirmation anxiety) especially when meeting a new professor (Gilrane et al., 2019).

Summary and Conclusions

The analyses to validate vocal jitter as a measure for anxiety showed: (a) consistent with hypothesis H3ai, that the independent judges used vocal jitter as an acoustic heuristic to estimate the anxiety in the students' voices (as evidenced in the positive correlation between the judges' anxiety ratings and vocal jitter) and; (b) the same judges could detect the anxiety evoking treatment that the students were assigned to, with elevated detection of this emotion in the stereotype salient versus non-salient conditions, consistent with hypothesis H3aii. Therefore, in this initial part of the current study, clear and convincing evidence has been presented to validate vocal jitter as an index of anxiety.

Apart from validating vocal jitter as a measure for anxiety, the current study (Study 3) aimed to examine if there are positive or negative impacts of an ingroup (versus outgroup) expert/role model on anxiety as indexed by vocal jitter among stereotyped STEM students as a function of stereotype threat (RQ3). The study looked to the seemingly conflicting theories of *ingroup spotlighting thesis* and the *stereotype inoculation model* to see which one provides better explanations for the poorer outcomes of underrepresented groups in the STEM context of higher education (RQ4). While looking at these objectives, Study 3 also answered the questions of whether vocal jitter provides a suitable index for anxiety (RQ5).

Confirming the initial hypotheses, the study revealed that: (a) stereotype salience does increase anxiety amongst Malaysian female students (H3b), but only (b) when they were exposed to an outgroup (not an ingroup) professor, i.e., salience x identity interaction (H3c; Delgado & Prieto, 2008; Desombre, et al., 2018; Hess et al., 2003; Laurin, 2020; Stone et al., 1999). Apart from that, the student also showed that these effects on anxiety appear to be momentary and did not extend beyond the immediate encounter between the students and their new professor in line with impression management concerns (H3d; Gilrane et al., 2019). Finally, although a professor's formality could also be a source of anxiety in the classroom, a

stereotype-related anxiety was still visible even after controlling for the anxiety caused by this rather obvious source. Alongside these conclusions, this study also establishes that stereotypes produce unique effects on a novel and validated measure of anxiety (vocal jitter; H3ai and H3aii), especially in the context of outgroup but not ingroup professors. This pattern supports the inoculation thesis.

Study 3 provided supportive evidence that stereotype-related anxieties operate within the classroom context too and corroborates Study 1 in Chapter 4, by showing (with a novel vocal jitter measure that directly captures anxiety) that such anxieties are more visible in the context of outgroup (than of ingroup) teachers. A central question of the current thesis, however, is what the impact of these stereotype-relevant anxieties might be with regards to the performance of vulnerable students from underrepresented groups who may believe that STEM is not for them (i.e., those with an imposter syndrome). According to the stereotype inoculation model, such anxieties are unlikely to harm class outcomes of vulnerable students when they are shielded by the buffering effect of expert ingroup teachers. However, according to the newer ingroup spotlighting thesis that is being advanced here, such anxieties are likely more detrimental to the classroom outcomes of vulnerable students from underrepresented groups, precisely when they are exposed to ingroup STEM experts who quash the stereotype that STEM isn't for one's ingroup. In short, the next chapter will further explore the predictions from SIM and those generated from the ingroup spotlighting thesis.

Chapter Seven

Examining Predictions from Stereotype Inoculation versus Ingroup Spotlighting

Based on the findings from Study 3 (Chapter 6) which found that stereotypes exacerbate Malaysian female students' anxiety (via elevated vocal jitter) within the context of stereotype salience in the presence of outgroup professors, the goal in this chapter is to test the hypotheses derived from the SIM and the ingroup spotlighting thesis using the novel measure of anxiety that was validated in Chapter 6. For example, based on the SIM, it was hypothesised that the specific identity of the outgroup professor should be irrelevant to a test of the inoculative benefits of ingroup experts. What matters, however, is the extent to which students from stereotyped backgrounds are vulnerable to imposter feelings. SIM's inoculation thesis would predict that students with imposter syndrome, and elevated vocal jitter (i.e., highly anxious) would perform better when exposed to an ingroup professor. However, according to the ingroup spotlighting thesis posited in this thesis, they might perform worse.

Two classroom experiments, Studies 4 and 5 presented in this chapter, were conducted to unpack the predictions of the SIM and the ingroup spotlighting theses and its predictions. Apart from the effect of imposter syndrome, the longitudinal effects on performance (Study 4), the post-lecture confidence in mathematics (Study 5) and the endorsement of post-quiz attributions (Study 5) were also examined in this chapter.

Lastly, these studies were conducted to also address the need for replications, especially in the light of low sample sizes, so that the stability versus fragility of the relevant effects can be determined as the evidence accumulates (Onghena, 2020).

Study 4: Exploring Professor Group Membership Effects on Vocal Jitter and Actual Performance

According to SIM, the extent to which students from stereotyped backgrounds are vulnerable to imposter feelings would affect the inoculative benefits of ingroup experts/role

models. Previous studies have operationalised imposter syndrome using various measurement scales such as the 14-item scale developed by Harvey (1981), the more popular 20-item Clance imposter phenomenon scale (CIPS; Clance, 1985), the 51-item Perceived Fraudulence Scale (Kolligian & Sternberg, 1991), the dichotomous 8-item Young Imposter Syndrome Quiz (Villwock et al., 2016), or even the 7-item Leary Imposter Scale (Leary et al., 2000). Although these scales have demonstrated moderate to high internal consistency, questions have been raised concerning their suitability especially since a definitive conceptualisation and dimensionality of imposterism has not been properly identified (Mak et al., 2019).

According to Dasgupta (2011), those who suffer from an imposter syndrome tend to "be *unsure of their ability, be anxious*, and have low expectations of repeated future success compared to others who do not feel like imposters" (p. 232). Therefore, this study relied on this operational definition of the imposter syndrome to obtain participants' self-reported efficacy in biology prior to the class. SIM's inoculation thesis would predict that students with low scores on this scale (i.e., imposter syndrome), and elevated vocal jitter (i.e., highly anxious) would perform better when exposed to an ingroup professor (Stout et al., 2011). However, according to the ingroup spotlighting thesis proposed here, they might perform worse. As such, in these studies, a performance measure was added by way of a multiple-choice question (MCQ) test.

At the same time, Study 4 also unpacks the issue of whether the expected effects would extend beyond the immediate classroom situation answering the question of how long these effects would persist. Several papers in the literature have suggested the need for longitudinal studies to better the understanding of the long-lasting effects of the exposure to ingroup role models (Gonzalez-Perez et al., 2020; Leroy et al., 2022; O'Brien, et al., 2017). This is important because research on ST effects have focused primarily on *situational or immediate* performance and have largely overlooked the crucial question as to whether such

effects linger (Block et al., 2011). The long-lasting effect was explored by way of theory development within this study with the expectation that both the stereotype inoculation and ingroup spotlighting theses would have lasting longitudinal effects on performance. This expectation is grounded in the fact that both theses have an impact on deeper more long-term characteristics such as attitudes, belonging, self-concept, self-efficacy, and aspirations (Asgari et al., 2010; Dasgupta, 2011), which if altered, could have a longer lasting impact.

A summary of the research questions and hypotheses for Study 4 in this chapter are as follows:

RQ3 - Are there positive or negative impacts of an ingroup (versus outgroup) expert/role model on STEM performance in underrepresented groups as a function of stereotype threat and anxiety?

RQ4 - Do the *ingroup spotlighting thesis* or the *stereotype inoculation model* provide better explanations for the poorer outcomes of underrepresented groups in the STEM context of higher education?

H4ai - Anxious students will perform worse on a STEM test in the presence of an outgroup teacher compared to an ingroup teacher (SIM).

H4aii - Anxious students with imposter syndrome will perform worse on a STEM test in the presence of an ingroup teacher compared to an outgroup teacher (ingroup spotlighting).

H4b - Both stereotype inoculation and ingroup spotlighting would have longitudinal effects on performance.

Method

Participants and Design. The experimental protocol was identical to the one described in Study 3 where a *2 x 2 x 2 between-subjects design* was used with participants being randomly assigned to different conditions of a stereotype salience manipulation (salient

versus non-salient); professor formality conditioning (formal-acting versus non-formal-acting) and professor *ethnic* identity treatment (*Malaysian ingroup versus White/Black outgroup*). The primary dependent measure was vocal jitter, which was used as a direct behavioural/physiological index of anxiety. However, for this study, participants also (N = 163, M_{age} = 18.83, SD_{age} = 0.65) completed a multiple-choice question (MCQ) test on biology after the class, while being exposed to an ingroup versus an outgroup professor who delivered a biology lecture. *Participants were students from the Foundation in Science program in the Faculty of Science and Engineering at a private higher education institution in Selangor, Malaysia*. *These were different participants from the previous studies.*

*An a priori power analysis was conducted using G*Power* (Faul et al., 2007) *to determine the minimum sample size required to test the study hypotheses. The output indicated that the minimum required sample size to achieve 95% power for detecting a medium effect at a significance of α = .05, was N = 210. Therefore, the obtained sample size of N = 163 was lower than required.* As the current cohort contained only a modest number of women (n = 119), the sample size was augmented with the inclusion of men (n = 44) while focusing on race.

A notable adjustment that was made was that this study attempted to mirror the heterogeneity of outgroup professors in the institution of study (and indeed in the real world), by exposing Malaysian students to two British educated professors who were visibly different in terms of race: one was White, and the other was Black. Meanwhile, the Malaysian professor was educated locally. Once again, it was ensure that the participants had no exposure to the professors before the study.

Procedure. This study relied on the operational definition of the imposter syndrome provided in the preamble to obtain participants' self-reported efficacy in biology prior to the class, using a 15-item scale that was adapted from Miller et al. (1996), and Pintrich and De

Groot (1990): e.g., “I don’t think I will be successful in biology” [reverse scored]; “I am confident I can do well in biology”; 1 = *strongly disagree*, 5 = *strongly agree*, Cronbach’s α = .85).

Participants were also given a 10-item multiple choice test that was based on the materials presented to them. Participants were told that the person with the best score will win RM50 to make the task more meaningful for participants. Ten multiple choice items were chosen from a 30-item list that was piloted to gauge their level of difficulty and to ensure as much variation as possible across students (a description of the standardisation protocol is under the next heading). ST is theorised as impairing cognitive function via the experience of anxiety because it diverts cognitive resources away from cognitive tasks. Questions that were not attempted say little about “cognitive disruption” in the manner that answered (but incorrect) questions do. Unfinished questions could be due to the time constraint imposed on the quiz (11 seconds per question) and not a cognitive impairment *per se*. *Calculating the number of questions answered correctly was also avoided as it could be an effect of chance.* Hence, the focus was only on the number of incorrect answers as the index of performance impairment.

Participants completed the same manipulation checks for stereotype salience (i.e., word fragments) as in Study 3. They also completed the same teacher-related anxiety measure as in Study 3, to verify the assumptions that were made about the anxiety-eliciting potency of the professor formality manipulation (Cronbach’s α = .80). As before, vocal jitter was extracted from participants’ voice prior to lecture starting (onset jitter) and again after the class (after class jitter).

A final consideration with regards to the divergent predictions of the SIM inoculation and ingroup spotlighting theses was whether the anticipated effects would extend beyond the immediate classroom situation (i.e., how long the effects would persist). Therefore, the same

MCQ test was repeated four weeks later (with no further exposure to the experimental treatments) to unpack these issues on an exploratory (theory development) basis. The shortest time gap possible was used between the first and second tests for two reasons: a) to reduce the effect of something else occurring at about the same time that could influence the change in outcome (effects of history), b) to avoid the participants being older and more mature within the time gap between the tests (effects of maturation; Gonzalez-Perez et al., 2020). Participants were not given any prior notification of the test to ensure that the retention effects were purely from the exposure to the conditions during the workshop and not from participants' own preparations for the test. Participants were asked to complete the test within the same amount of time with the same offer of RM50 for the person with the best score to replicate the same test conditions. The same amount of time was given so the levels of difficulty of the test is standardised and what is tested is purely retention.

Preparation of the Performance Measure. A total of 25 participants (23 females and 2 males; M_{age} = 19.84, SD_{age} = 0.47) were recruited from among Malaysian students who have completed the Foundation in Science programme at a higher education institution in Malaysia to answer a set of 30 multiple choice questions after going through the PowerPoint slides used for the lecture in these studies.

To achieve a test that is diagnostic of mastery of the lecture materials, the 30 questions were piloted with varying levels of difficulty (see Appendix B). This was done to closely mirror the empirical distribution of scores in a regular classroom test in the same subject area as that of the lecture. Of the 30 questions, 10 were selected based on the following selection criteria: (a) seven questions were selected because 20% of the participants got it wrong, and this indicated a level of difficulty in those questions that allows for differentiation in the test, (b) two were selected from a further three items that proved difficult for 20% of the participants based on the time it took the average participant to

complete the task – using only those two with the longest response latencies, and (c) the tenth question was selected based on an even distribution of participants who got it right and wrong. All these were done to obtain questions with the highest level of differentiation for the test.

Results and Discussion

Assumption/Manipulation Checks. With regards to the *stereotype salience* manipulation, results confirmed that stereotype relevant word-fragment completions were significantly higher in the salient than in the non-salient condition (see Table 10), once more indicating that participants in the stereotype salient condition would have been primed with the relevant threat inducing stereotypes (Gilbert & Hixon, 1991; Sinclair & Kunda, 1999). Regarding teacher-related anxiety, as expected, results showed that self-reported anxiety was significantly greater for those students exposed to a formal acting professor compared to a less formal-acting professor (see Table 10). Finally, as in Study 3, onset vocal jitter was also positively correlated with after class vocal jitter ($r = .41$, $p < .001$).

Table 10

Stereotype Salience Manipulation Check Results

Number of stereotype-related word fragment completions			**Self-reported teacher-related classroom anxiety**		
Salient stereotype condition	Non-salient stereotype condition	d_{Cohen} [95% CI]	Formal condition	Non-formal condition	d_{Cohen} [95% CI]
4.38 (0.93)	1.33 (1.16)	2.92 [2.53, 3.31]	3.17 (0.77)	2.34 (0.62)	1.17 [0.90, 1.49]

Note. Numbers in parenthesis are standard deviations and numbers outside parenthesis are means.

Anxiety: Testing SIM's Version of ST. In both the White and Black professor conditions, the effect of salient stereotypes was similar in direction for onset vocal jitter (White professor: 3.53 ± .14 versus 3.21 ± .17; Black professor: 3.88 ± .22 versus 2.91 ± .26 salient versus non-salient conditions respectively) and after class jitter (White professor: 3.44 ± .16 versus 3.17 ± .20; Black professor: 3.50 ± .20 versus 3.07 ± .25, salient versus non-salient conditions). Meanwhile the effect was in the reverse direction in the ingroup professor condition for both onset (3.07 ± .26 versus 3.67 ± .19, salient versus non-salient) and after class (3.19 ± .26 versus 3.61 ± .20, salient versus non-salient conditions) vocal jitter. In short, students' reactions to both outgroup professors were identical, so this informed the decision to combine the two outgroup conditions into one.

However, an unanticipated equipment failure occurred during the after-class recording of students' vocal responses in the Black professor condition, and although the TA improvised with a different mobile device (a Blackberry Bold 9780), it was not possible to guarantee that the acoustic measurements were the same. Consequently, the after-class recordings in the Black professor condition were excluded from analysis. This meant that for the analysis, the after-class vocal jitter measurements from only the White versus Malaysian professor conditions could be used with regards to test performance.

The same analytical approach (ANOVA) used in Study 3 was also used here in Study 4 in order to address the question of how the experimental treatments affected students' actual experience of anxiety via their vocal modulations. Once again, the impact of the experimental treatments on vocal jitter at the onset of class and then after class (prior to the quiz) was examined.

Anxiety at Onset of Class (Stereotype-Relevant). Model results are presented in Table 11, showing a non-significant main effect of stereotype salience. However, in line with SIM's salience x identity thesis (Dasgupta, 2011), the null effect of stereotype salience was qualified

by professor identity (see Table 11, and Figure 6). Simple effects analysis revealed that, in the outgroup professor condition (i.e., when exposed to British professors), vocal jitter was significantly higher when stereotypes were salient (3.71 ± .14) compared to when they were non-salient (3.13 ± .16), $F(1, 152) = 7.33$, $p = .008$, $\eta_p^2 = .05$. However, stereotype salience did not have any reliable effect on vocal jitters in the ingroup professor condition, $F(1, 152) = 3.59$, $p = .061$, $\eta_p^2 = .04$ (see Figure 6). Hence, the patterns in Study 3 were replicated, by showing that the presence of ingroup experts can inoculate students from stereotyped backgrounds against the effects of stereotype salience (Stout et al., 2011). The main effect of professor formality was not significant (3.38 ± .13 versus 3.40 ± .14, formal versus non-formal conditions respectively), nor was its interaction with professor identity (see Table 11).

Once again, none of these results changed meaningfully: (a) in a subsequent bootstrap simulation with 20,000 resamples of the raw data (see Table 11, Model 2); (b) when the potential effect of emotional contagion from the TA's vocal jitter was controlled for (see Table 11, Model 3); or (c) when participants' self-reported experience of teacher-related anxiety was accounted for (see Table 11, Model 4).

Table 11

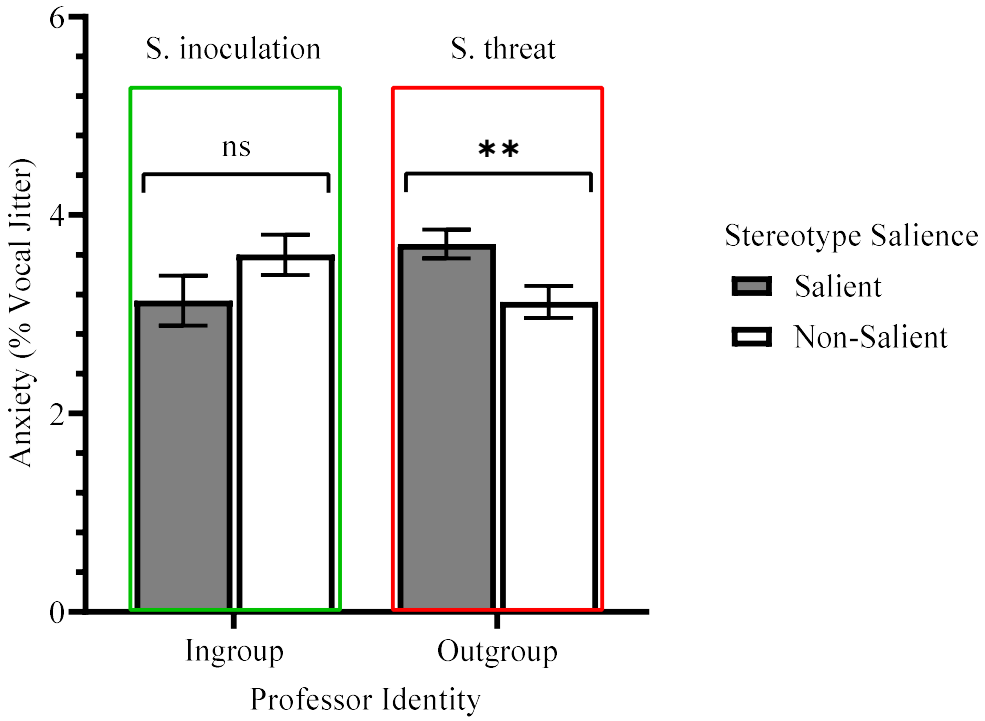
The Effects of Professor Identity, Formality and Stereotype Salience on Students' Anxiety (Vocal Jitter)

	Model 1 No controls (raw data)			Model 2 Bootstrap simulation (N = 20,000)			Model 3 Controlling for TA's vocal jitter			Model 4 Controlling for teacher-related anxiety		
	F	*p*	ηp^2	*F*	*p*	ηp^2	*F*	*p*	ηp^2	*F*	*p*	ηp^2
Professor identity (PI)	0.06	.805	.000	0.06	.805	.000	0.46	.498	.003	0.02	.890	.000
Stereotype salience (SS)	0.94	.759	.001	0.94	.759	.001	.003	.959	.000	0.08	.782	.001
PI * SS	7.21	.008	.045	7.21	.008	.045	5.75	.018	.037	7.05	.009	.045
Professor formality (PF)	0.01	.906	.000	0.01	.906	.000	0.04	.846	.000	0.25	.620	.002
PI * PF	3.15	.078	.020	3.15	.078	.020	3.95	.049	.025	2.68	.104	.018
Control variable	--	--	--	--	--	--	1.36	.246	.009	0.44	.508	.003

Note. N = 163. TA = Teaching assistant. Professor identity (-1 = ingroup, 1 = outgroup); Negative ingroup stereotypes (-1 = non-salient, 1 = salient); Professor formality (-1 = non-formal, 1 = formal). Bootstrap simulation was conducted in SPSS.

Figure 6

The Effects of Stereotype Salience on Students' Vocal Jitter (Anxiety) When Their Professor Was Either Ingroup or Outgroup



Note: S. inoculation = stereotype inoculation. S. threat = stereotype threat. Error bars = standard errors. *ns* = non-significant; **$p < .010$.

Anxiety After Class but Prior to Quiz (Exam Fever). Consistent with the inoculation thesis, results revealed a stereotype salience x professor identity interaction, $F(1, 152) = 5.11$, $p = .026$, $\eta_p^2 = .05$, which occurred because exposure to ingroup experts significantly *reduced* anxiety (vocal jitter) when stereotypes were salient (3.05 ± .19) compared to when they were non-salient (3.55 ± .16, $p = .049$). This effect was not statistically significant in the outgroup professor conditions ($p = .238$). No other effect reached statistical significance ($ps > .30$). Hence, like Study 3, the outgroup condition did not influence students' anxiety beyond the

onset of class when impression management concerns should be highest (Gilrane et al., 2019). However, for those exposed to an ingroup professor, students had the lowest anxiety (vocal jitter) in the stereotype salient versus non-salient condition, suggesting some continued immunity that an ingroup professor provided with regards to anxiety going into the exam (Asgari et al., 2010).

Quiz Performance Errors: Testing SIM's Inoculation Thesis versus the Ingroup Spotlighting Thesis. To resolve the theoretical disagreement between SIM's inoculation thesis and the ingroup spotlighting thesis, a moderated regression was performed in which either of the two behavioural anxiety measures (mean centered) were specified as focal predictors of test performance, while professor identity (effect coded: with outgroup as reference) and efficacy in biology (mean centered) were moderators. Number of incorrect answers was the outcome. That is, whether students' low and high subject efficacy (i.e., high, and low imposterism) impacted their quiz performance as a function of anxiety (vocal jitter) when they were taught by an ingroup versus outgroup professor, was investigated using a moderated regression approach with 20,000 bootstrap samples of the raw data.

With regards to after class anxiety, consistent with SIM's inoculation hypothesis, a main effect of professor identity was found, b = -1.35, se = .40, p = .001. The number of incorrect answers were lower when students were taught by an ingroup (2.25 ± .28) than by an outgroup (3.60 ± .28) professor (Lawner et al., 2019). Also, consistent with previous findings (Flore et al., 2018), increases in vocal jitter (anxiety) were associated with increases in number of incorrect answers, b = .54, se = .26, p = .040, indicating similar patterns suggested by the Yerkes-Dodson law, whereby very high anxiety can lead to the impediment of performance (Chapell, et al., 2005; Yerkes & Dodson, 1908)

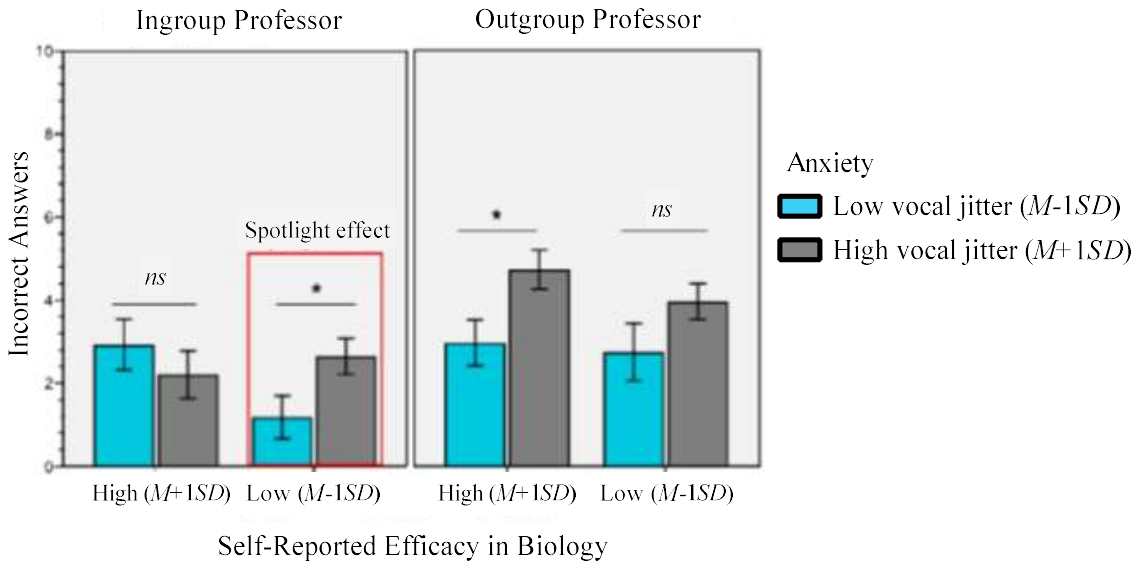
Importantly, the vocal jitter x biology efficacy x professor identity interaction approached statistical significance, b = -1.79, se = 1.02, p = .083 (see Figure 7). A breakdown

of this interaction showed that a vocal jitter x self-efficacy in biology interaction occurred only when the professor was ingroup, b = -1.43, se = .65, p = .031, but not outgroup, b = .36, se = .79, p = .649. Consistent with the ingroup spotlighting thesis proposed here, when self-efficacy in biology was low (M-1SD; i.e., high imposter feelings), the number of incorrect answers was higher for those students who were more (M+1SD) rather than less (M-1SD) anxious, b = .84, se = .35, p = .019, while this trend was absent when self-efficacy in biology was high (M+1SD), b = -.44, se = .49, p = .896. The theory behind these findings was that stereotyped students who have a lower sense of self-efficacy or, in this case, who feel like an imposter within the STEM field, may be particularly concerned about the evaluation of the ingroup teacher which is at the same time anxiety inducing (Hornsey et al., 2002; Hoyt, 2013). Further to that, the ingroup expert/role model puts more pressure on the stereotyped student to excel within the same domain as it leaves no room for excuses and creates an attributional ambiguity that could lead to the self-esteem being harmed (Crocker et al., 1991).

Apart from that, there was neither a main effect of onset vocal jitter on test performance nor did professor identity and self-efficacy in biology qualify this null effect (ps > .05).

Figure 7

Number of Inaccurate Answers to a Biology Quiz as a Function of Anxiety (% Vocal Jitter) and Biology Efficacy (Imposter Feelings) When Malaysian Students Were Exposed to Ingroup versus Outgroup Professor



Note: Reported are estimated marginal means. Error bars are standard errors. *ns* = non-significant, *$p < .050$.

Longitudinal Follow-Up. Students' test scores, four weeks after the initial one, were subjected to the same moderated regression as before, with after class vocal jitter, professor identity and self-efficacy in biology (and their interaction terms) as predictors. Result from this analysis revealed only a main effect of professor identity, $b = -.94$, $se = .35$, $p = .009$. Consistent with SIM's inoculation thesis, the number of incorrect answers were significantly lower amongst students who were taught by an ingroup professor (4.66 ± .26) compared to those who were taught by an outgroup professor (5.60 ± .24). All other main and interactive effects did not approach statistical significance (*p*s > .25). Results were the same when 'after class vocal jitter' was substituted in the model with 'onset vocal jitter' (*p*s > .15), except for

professor identity being the only reliable predictor (p = .031). Hence, contrary to the initial expectation, the ingroup spotlighting effect seems to be time-sensitive, occurring only following the ingroup spotlighting event, while the beneficial or inoculating effects of ingroup experts endured over a 4-week period irrespective of the students' initial imposter feelings. As mentioned earlier, there is ample space and need to develop this further in a future investigation of a longitudinal nature to better understand the long-lasting effects of the exposure to ingroup role models (Gonzalez-Perez et al., 2020; Leroy et al., 2022; O'Brien, et al., 2017) and also the time-sensitive effects of the ingroup spotlighting.

Summary of Key Findings

Study 4 replicated the main findings from Study 3 in that stereotype salience was related to higher levels of anxiety when exposed to outgroup professors compared to ingroup professors. Furthermore, Study 4 revealed: (a) that ingroup experts can cause a reduction in the number of incorrect answers provided by Malaysian students on a biology quiz and, for the first time, demonstrated the longitudinal potency of this social identity vaccine (partial H4b). That is, retention rate was significantly improved when students were taught by a professor who shared a social identity with them (H4ai). Results also showed (b) that situations that accentuate the potential for ingroup spotlighting can thwart the positive effect of this identity-based vaccine, especially for those vulnerable students likely experiencing imposter syndrome (H4aii). This latter finding is particularly important because, until now, such students are the ones theorised to be especially likely to benefit from the inoculative effects of ingroup experts (Stout et al., 2011). So, although the current evidence is largely consistent with the idea that ingroup experts do provide immunity against stereotype threat, it also corroborates the ingroup spotlighting thesis when it comes to stereotyped students likely experiencing imposter syndrome. It is important to note, however, that it was not the anxiety at the onset of class (when concerns about ingroup stereotypes were more relevant) that

significantly impaired students' test scores, but the corresponding feeling prior to the exam. Hence, like Flore et al. (2018), students' test performance errors here were not easily attributable to the salience of negative ingroup stereotypes (and the anxiety that follow it).

Study 5: Replication of Experimental Conditions in the Gender Context

Although, students' test performance errors in Study 4 could not be clearly attributed to the salience of negative ingroup stereotypes and the anxiety that is associated with it, it may be too early to discard the idea that situational salience of stereotypes undermines students' STEM performance based on the outcome of a single study. Therefore, a replication in the gender context was conducted in Study 5, where STEM stereotypes are more visible and for the sake of a closer comparison with Flore et al.'s (2018) data, the study was conducted in the Netherlands.

Study 4 provided a narrow test of the inoculation thesis given its focus on test performance alone, especially when previous studies have shown that an identity-based inoculation is most potent in its effect on students' subsequent confidence in a relevant STEM subject (Asgari et al., 2010; Dasgupta et al., 2015; Stout et al., 2011). To mitigate this potential limitation, a follow up study (Study 5) was conducted which included a measure of post-lecture confidence in mathematics, to examine anxious female students' efficacy in mathematics following exposure to a female (ingroup) versus male (outgroup) professor.

A second consideration, this time with regards to the ingroup spotlighting thesis, is that it is yet to be demonstrated that students are, to a greater extent, motivated to search for excuses for (potential) failure when they have been exposed to ingroup experts (Taylor et al., 2011). Here, it was anticipated that students will endorse attributions that preserve self-esteem and control (internal, controllable, unstable attributions) versus those that imply forces beyond their control (external, uncontrollable, stable attributions) (Reyna, 2008).

Based on the ingroup spotlighting thesis proposed here, it is reasoned that self-preserving attributions will increase when anxious students are exposed to an ingroup expert.

A summary of the research questions and hypotheses for Study 5 in this chapter are as follows:

RQ3 - Are there positive or negative impacts of an ingroup (versus outgroup) expert/role model on STEM performance in underrepresented groups as a function of stereotype threat and anxiety?

RQ4 - Do the *ingroup spotlighting thesis* or the *stereotype inoculation model* provide better explanations for the poorer outcomes of underrepresented groups in the STEM context of higher education?

H5ai - Anxious students would have lower post-lecture confidence in mathematics when exposed to an ingroup teacher compared to an outgroup teacher (ingroup spotlighting).

H5aii - Anxious students would have higher post-lecture confidence in mathematics when exposed to an ingroup teacher compared to an outgroup teacher (SIM).

H5b - Anxious students would endorse more self-esteem preserving attributions when exposed to an ingroup teacher compared to an outgroup teacher.

Method

Participants and Design. The experimental protocol here again largely mirrored the approach that were described in Study 4 *with the 2 x 2 x 2 between-subjects design* including: the salient stereotype, professor *gender* identity, and formality manipulations, as well as the corresponding manipulation checks (for salient stereotypes, *four word fragments*: "unintelligent"; "illogical"; "bad at math"; "inconsistent"; and for formality induction using the same teacher-related anxiety that were described in Studies 3 and 4, Cronbach's α = .87).

However, there were several alterations done to improve the design of this subsequent study. Firstly, in Study 5, participants, *although ethnically diverse*, were all female (N = 113, M_{age} = 19.76, SD_{age} = 2.05) to maintain a more direct focus on the gender gap in STEM. Secondly, rather than an in-person lecture (as in Studies 3 and 4), first-year psychology students at a university in *Amsterdam*, Netherlands were exposed to a 10-minute pre-recorded lecture on introductory statistics (ANOVA), that was taught by a male versus female Dutch professor with identical lecture notes. *Like the previous studies, the participants did not have any exposure to the professors prior to this study.* As mentioned in the preamble, Study 5 was conducted in the gender context where STEM stereotypes are more visible, and in the Netherlands (a European context instead of Asian) to have a closer comparison with Flore et al.'s (2018) data which was collected in the Netherlands.

*As with the previous study, the a priori power analysis conducted using G*Power* (Faul et al., 2007) *to determine the minimum sample size required to test the study hypotheses indicated the required N = 210. Therefore, the obtained sample size of N = 113 was smaller than required. The primary dependent measure was still vocal jitter, which was used as a direct behavioural/physiological index of anxiety.*

Procedure. Prior to this lecture, participants completed the same 15-item subject-specific efficacy measure described in Study 4, modified to reflect the current focus on mathematics (e.g., "I am confident I can do well in mathematics."; 1 = *strongly disagree*, 5 = *strongly agree*, Cronbach's α = .94). This time however, the same 15-item measure of math efficacy was again administered after the lecture to examine the effect of the identity-based vaccine on students' post lecture *confidence in math* (α = .94).

Students were tested in groups of 12-17 students. They were welcomed by a TA who seated them in a lecture room so that they could view a projection screen. The TA in each session (one male and one female Masters student) was matched with the gender identity of

the professor that students were exposed to in order to eliminate the potential for cross contamination of gender effects. TAs also matched the professor with regards to the perceived formality manipulation, so that in the formal-acting professor condition, they dressed formally and conducted themselves with utmost professionalism, while in the non-formal-acting condition, they wore casual attire and were less formal with the students.

As before, recordings of students' voices were collected at the onset of lecture, having exposed them to a short biography and image of the professor displayed on-screen[1] (see Appendix C) and after the class, but prior to the quiz that followed. As in Study 3, vocal jitter from the recordings were extracted as the index of anxiety, and students completed a 10-item MCQ test after class that underwent a similar calibration for test difficulty that was described

---

[1] The literature on stereotype threat and the SIM suggests that even virtual or imagined "exposure" to outgroup teachers is sufficient to trigger evaluative concerns that can harm the outcomes of students with stereotyped identities (Latu et al., 2013; Stout et al., 2011). Based on these findings, an introductory slide was used at the start of the workshop to introduce the professor of the day. The professor in the formal condition enjoyed his or her work and reading while the professor in the informal condition enjoyed socialising and talking to people. The professor in the informal condition also welcomed being called an informal name with a smiley. The formal condition had a photo of the professor not smiling and in formal attire while the informal condition had a photo of the professor smiling and in a t-shirt. All professors gave the impression that they studied Mathematics at the university by mentioning it and also worked at the university by providing a university email address to be contacted. All professors also gave the impression that they would see the students after the workshop.

in Study 4. MCQ performance impairment scoring was identical to the approach described in Study 4, with the focus being on incorrect answers (see next heading).

Lastly, excuses for potential failure were measured after the quiz using an attribution scale from the classic stereotype threat literature (Steele & Aronson, 1995). One item asked about factors within their control ("how much stress have you been under lately") and two items asked about factors beyond their control (e.g., at the test: "how tricky did you find the test"; and "how unfair did you find the test" $r = .42$, $p < .001$). Responses were obtained on a 7-point scale (1 = *not at all stressed*/*tricky*/*unfair*, 7 = *very stressed*/*tricky*/*unfair*).

Preparation of the Performance Measure. The lecture assessment tool (a series of multiple-choice questions on statistical problems) was piloted amongst 56 female students (M_{age} = 19.06, SD_{age} = 0.68) who were enrolled in the Foundation in Science programme at a higher education institution in Malaysia. Thirty multiple choice questions (MCQs) *on introductory statistics (ANOVA), which was the topic covered in the lecture*, were created with varying levels of difficulty (see Appendix D). This was done to closely mirror the empirical distribution of scores in a regular classroom test about the lecture. Of the 30 questions, 10 were selected based on the following selection criteria: (a) eight questions were selected because more than 60% of the participants got it wrong, and this indicated a level of difficulty in those questions that allows for differentiation in the test, and (b) two were selected from a further three questions that proved difficult for 55% of the participants based on the time it took the average participant to complete the task, using only those two with the longest response latencies.

Results and Discussion

Assumption/Manipulation Checks. The pattern of results with regards to *stereotype salience* were replicated, confirming that stereotype-relevant word-fragment completions were significantly higher in the salient than in the non-salient condition (see Table 12).

Results from Studies 3 and 4 with regards to *teacher-related anxiety* were also replicated: showing significantly greater anxiety for those students exposed to a formal-acting versus a non-formal-acting professor (see Table 12). Once again, onset vocal jitter was positively but this time marginally correlated with after class vocal jitter ($r = .17$, $p = .066$).

Table 12

Stereotype Salience Manipulation Check Results

Number of stereotype-related word fragment completions			**Self-reported teacher-related classroom anxiety**		
Salient stereotype condition	Non-salient stereotype condition	d_{Cohen} [95% CI]	Formal condition	Non-formal condition	d_{Cohen} [95% CI]
3.37	0.93	2.70	3.65	2.33	2.94
(1.05)	(0.76)	[2.19, 3.21]	(0.52)	(0.37)	[2.86, 3.02]

Note. Numbers in parenthesis are standard deviations and numbers outside parenthesis are means.

Anxiety: Testing SIM's Version of ST. As in Study 4, the test was conducted in two stages. First, the impact of the experimental treatments on vocal jitter at onset of class and then after class was examined.

Anxiety at Onset of Class (Stereotype-Relevant). The results show a non-significant main effect of stereotype salience (contra ST). Importantly, and in line with SIM's inoculation hypothesis, an effect of stereotype salience was contingent upon the professor's identity (see Table 13; cf. Flore et al., 2018), so that a stereotype-induced increase of vocal jitter was apparent only when the participating female students were exposed to an outgroup

(male) professor, $F(1, 107) = 3.94$, $p = .050$, $\eta_p^2 = .04$, but not to an ingroup (female) professor, $F(1, 107) = 1.35$, $p = .248$, $\eta_p^2 = .01$ (see Figure 8).

Table 13

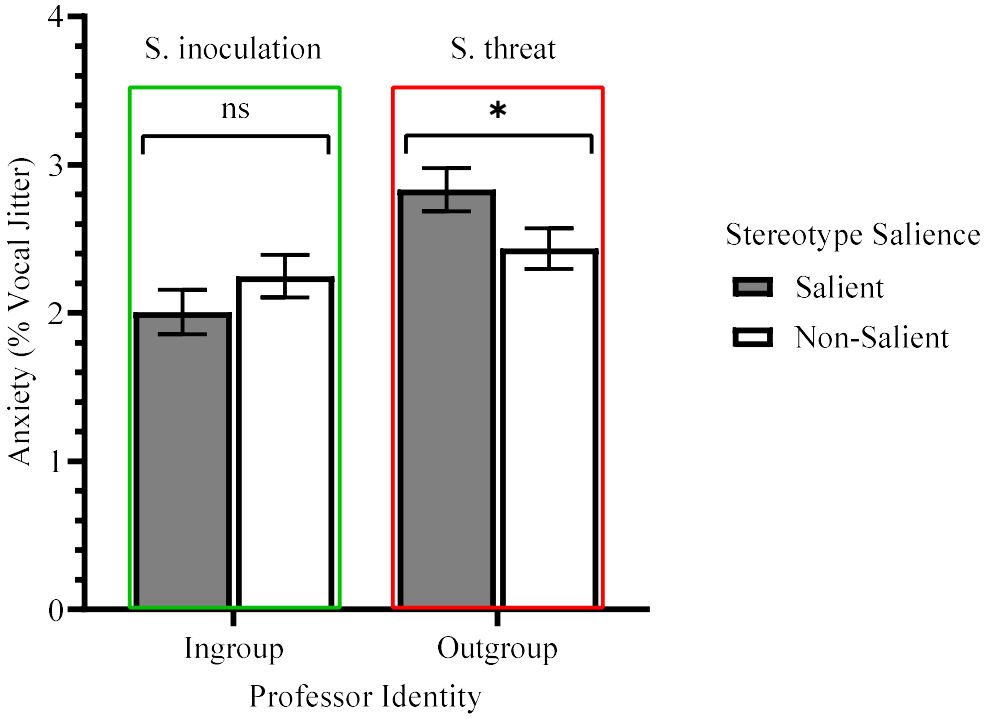
The Effects of Professor Identity, Formality and Stereotype Salience on Students' Anxiety (Vocal Jitter)

	Model 1 No controls (raw data)			Model 2 Bootstrap simulation (N = 20,000)			Model 3 Controlling for TA's vocal jitter			Model 4 Controlling for teacher-related anxiety		
	F	*p*	ηp^2	*F*	*p*	ηp^2	*F*	*p*	ηp^2	*F*	*p*	ηp^2
Professor identity (PI)	12.32	.001	.103	12.32	.001	.103	12.61	.001	.106	9.49	.003	.087
Stereotype salience (SS)	0.29	.590	.003	0.29	.590	.003	0.34	.560	.003	0.25	.621	.002
PI * SS	4.90	.029	.044	4.90	.029	.044	4.96	.028	.045	6.67	.011	.063
Professor formality (PF)	8.38	.005	.073	8.38	.005	.073	8.63	.004	.075	0.33	.566	.003
PI * PF	4.64	.033	.042	4.64	.033	.042	3.91	.051	.036	4.70	.033	.045
Control variable	--	--	--	--	--	--	0.36	.548	.003	0.84	.362	.008

Note. $N = 113$. TA = Teaching assistant. Professor identity (-1 = ingroup, 1 = outgroup); Negative ingroup stereotypes (-1 = non-salient, 1 = salient); Professor formality (-1 = non-formal, 1 = formal). Bootstrap simulation was conducted in SPSS.

Figure 8

The Effects of Stereotype Salience on Students' Vocal Jitter (Anxiety) When Their Professor Was Either Ingroup or Outgroup



Note: S. inoculation = stereotype inoculation. S. threat = stereotype threat. Error bars = standard errors. *ns* = non-significant; *$p \leq .050$.

The main effect of professor's formality on vocal jitter was consistent with expectation this time (2.59 ± .10 versus 2.17 ± .10, respectively for formal-acting versus non-formal-acting professor conditions, see Table 13), but was qualified by professor identity. A breakdown of this interaction revealed that an accentuation of vocal jitter in the formal-acting (relative to the non-formal-acting) professor condition was limited to the outgroup professor condition (3.00 ± .14 versus 2.27 ± .14, $p < .001$, $\eta_p^2 = .11$), and did not manifest in the ingroup professor condition (2.18 ± .14 versus 2.07 ± .15, $p = .609$, $\eta_p^2 = .002$). Hence, once

again a social identity vaccine (ingroup expert) provides immunity against anxiety even in the context of a professor's formality.

Again, replicating the trends from the previous studies, none of these results changed meaningfully: (a) in a subsequent bootstrap simulation with 20,000 resamples of the raw data (see Table 13, Model 2); (b) when the potential effect of emotional contagion from the TAs' vocal jitter was controlled for (see Table 13, Model 3); or (c) when students' self-reported experience of teacher-related anxiety was accounted for (Table 13, Model 4).

Anxiety after Class, but Prior to Quiz (Exam Fever). Results revealed neither a main nor interactive effect of stereotype salience, perceived professor formality and professor identity (*p*s > .10). Hence, the treatments did not exert an influence on students' anxiety beyond the onset of class when impression management concerns should be operational. *Some of these variations in the results for Study 5 as compared to Study 4 could be attributed to the difference in the way the session was conducted whereby one had a lecturer present in the classroom while the other was a recording. Further investigation would be needed to confirm this assumption.*

Quiz Performance Errors: Testing SIM's Inoculation Thesis versus the Ingroup Spotlighting Thesis. With regards to *onset anxiety (stereotype-relevant)*, results from a moderated regression involving vocal jitter, professor identity and self-efficacy in mathematics using 20,000 bootstrap samples revealed no significant main effect of vocal jitter on number of incorrect answers, and professor identity and self-efficacy in math did not qualify this null effect (*p*s > .05).

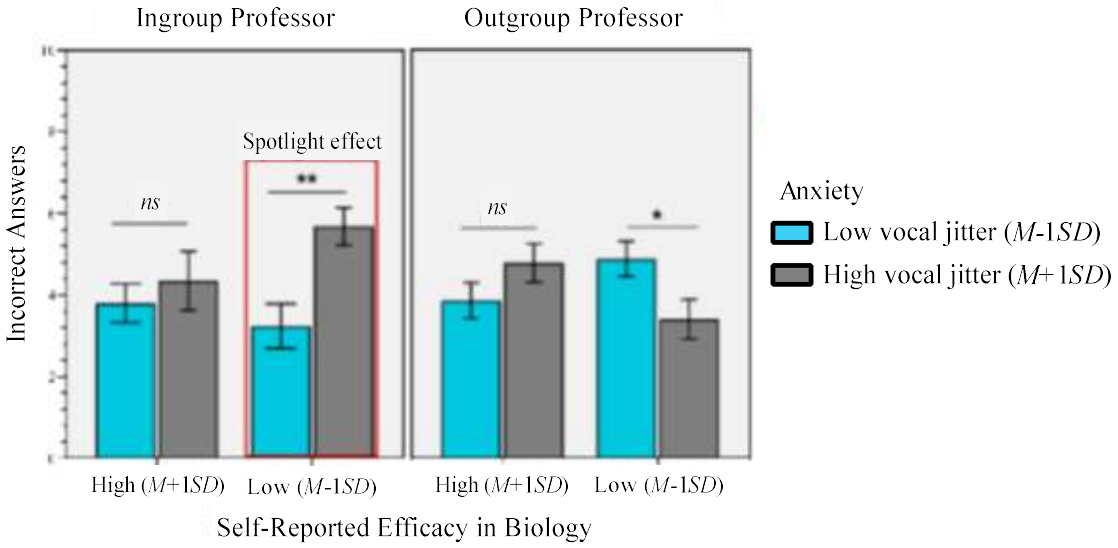
With regards to *after class anxiety* (exam fever), the main effects of professor identity and vocal jitter did not emerge (*p*s > .15). These null effects were, however, qualified by a significant professor identity x vocal jitter interaction, b = -.42, se = .20, p = .040, showing that highly anxious women (M+1SD) gave more incorrect answers than their less anxious

counterparts (*M*-1*SD*), but only when the professor was ingroup, $b = 1.06$, $se = .34$, $p = .002$, but not outgroup, $b = -.17$, $se = .18$, $p = .334$. Importantly, the expected vocal jitter x math efficacy x professor identity interaction emerged significant, $b = -.54$, $se = .19$, $p = .006$. A breakdown of this interaction showed that a vocal jitter x math efficacy interaction largely occurred in both the ingroup professor, $b = -.48$, $se = .24$, $p = .053$, and the outgroup professor, $b = 6.01$, $se = 2.97$, $p = .046$, conditions.

Ingroup (Female) Professor Condition. Consistent with the ingroup spotlighting thesis, results from a simple slope analysis revealed that highly anxious women (*M*+1*SD*) with low math efficacy (*M*-1*SD*), gave more incorrect answers than their less anxious counterparts (*M*-1*SD*) when they were exposed to an ingroup expert, $b = 1.20$, $se = .38$, $p = .002$ (see Figure 9). Meanwhile, highly anxious women whose efficacy in math was strong (*M*+1*SD*), did not perform any worse than their less anxious counterparts, $b = .25$, $se = .50$, $p = .611$, in terms of the number of incorrect answers (see Figure 9).

Figure 9

Number of Inaccurate Answers to a Math Quiz as a Function of Anxiety (Vocal Jitter) and Math Efficacy, When Female Students Were Exposed to an Ingroup versus an Outgroup Professor



Note. Reported are estimated marginal means. Error bars are standard errors. *ns* = nonsignificant, *p = .050, **p < .010.

Outgroup (Male) Professor Condition. Interestingly, the number of incorrect answers were substantially reduced amongst highly (relative to less) anxious women with low math efficacy when they were exposed to an outgroup (male) professor, b = -.72, se = .32, p = .027, but was absent for their counterparts with high math efficacy, b = .48, se = .37, p = .191 (see Figure 9).

After Class Confidence in Mathematics. There were two predictions here: An increase in anxiety should cause (a) a *drop* in math-specific confidence amongst anxious students who are exposed to an ingroup relative to an outgroup professor (following the ingroup spotlighting thesis; H5ai), but (b) an *increase* in confidence amongst anxious students who are exposed to an ingroup versus an outgroup professor (following SIM's

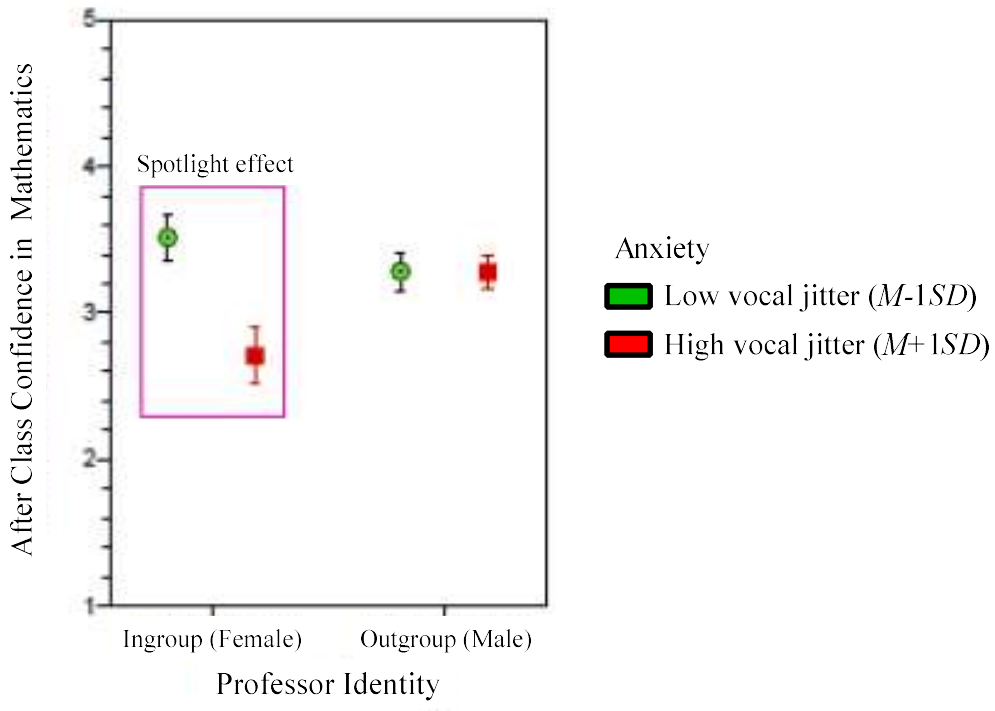
inoculation thesis; H5aii). The analysis primarily focused on the professor identity x vocal jitter interaction because students with an initially low level of confidence in math should have little room to further downgrade their self-efficacy.

Onset Vocal Jitter. Results from a moderated regression involving professor identity and onset vocal jitter using 20,000 bootstrap samples revealed no significant main effect of onset vocal jitter on post-lecture confidence in mathematics and, professor identity did not qualify this null effect (*p*s > .05).

After-Class Vocal Jitter. Results revealed that vocal jitter (anxiety) undermined students' confidence in mathematics, $b = -.19$, $se = .08$, $p = .018$, although this effect was qualified by professor identity in a 2-way interaction, $b = .40$, $se = .16$, $p = .015$. The pre- and post-quiz confidence in mathematics for highly anxious students in the ingroup professor condition were respectively, $2.96 \pm .19$ versus $2.71 \pm .19$, $p = .017$; while the equivalent estimates in the outgroup professor condition were, $3.48 \pm .11$ versus 3.28 ± 1.25, $p = .001$. It is interesting that confidence in mathematics was greater for students who were exposed to an outgroup (male) professor than to an ingroup (female) professor both at the pre-test stage ($p = .019$) and post-class measurement ($p = .012$). This suggests, consistent with the ingroup spotlighting thesis, that the mere sight of an ingroup (female) professor dampened female students' confidence in mathematics and this effect lingered after the class. Also consistent with the ingroup spotlighting thesis, anxiety undermined students' post lecture confidence in mathematics only when they were exposed to an ingroup professor, $b = -.40$, $se = .14$, $p = .006$, but not when they were exposed to an outgroup professor, $b = -.002$, $se = .08$, $p = .979$ (see Figure 10).

Figure 10

The Interactive Effect of Professor Identity and After Class Anxiety (Vocal Jitter) in Predicting Post-Lecture Confidence in Mathematics



Note. Error bars are standard errors.

Post-Quiz Search for Excuses. Consistent with the ingroup spotlighting thesis, results from a paired *t* test showed that students were, overall, more likely to endorse self-esteem preserving attributions that implied their outcomes were unstable and under their control (4.89 ± 1.54) than excuses that implied lack of control (3.76 ± 1.32) for potentially poor performance on the quiz, $t(105) = 5.58$, $p < .001$, $d_{Cohen} = .78$. Nonetheless, convincing evidence of the ingroup spotlighting thesis should demonstrate that a motivated endorsement of these excuses was higher when students were exposed to an ingroup versus an outgroup professor. Hence, we combined the controllable and uncontrollable excuses to generate an index of post-quiz attributions. A moderated regression revealed that professor identity and

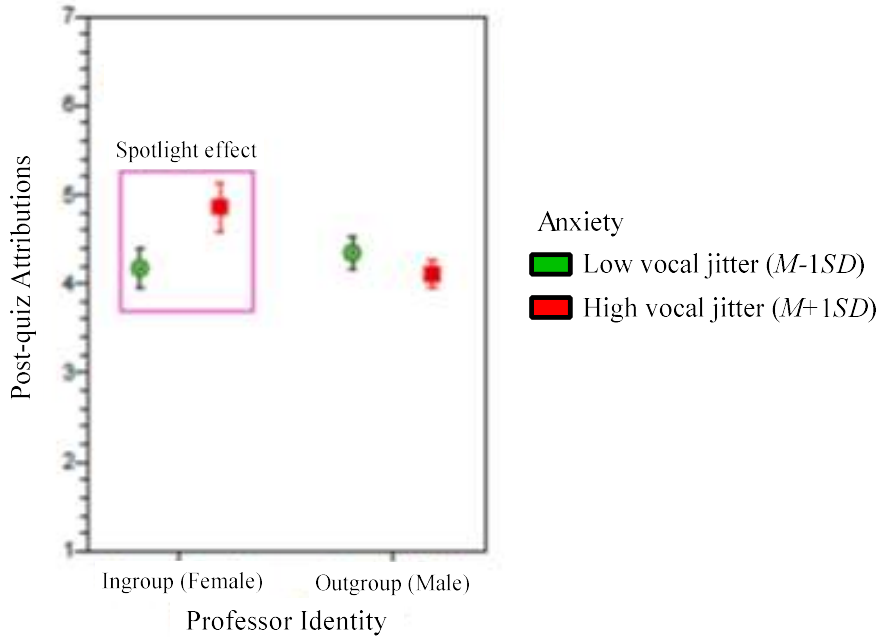
after-class anxiety (vocal jitter) interacted to predict post-quiz search for excuses, b = -.45, se = .22, p = .046. Consistent with the ingroup spotlighting thesis, simple slope results showed that post-quiz attributions were higher in the ingroup than in the outgroup professor condition, when students were highly anxious, b = -.78, se = .33, p = .018, but not when students experienced lower levels of anxiety, b = .17, se = .29, p = .551 (see Figure 11). The pattern was similar when the roles of anxiety and professor identity were reversed: in the ingroup professor condition, anxious students' self-serving attributions were marginally higher than the corresponding attributions amongst their counterparts with low levels of anxiety, b = .33, se = .20, p = .093. Meanwhile this anxiety effect was absent for those students who were exposed to an outgroup (male) professor, b = -.12, se = .10, p = .266. In fact, a follow-up mediational analysis in which the main and interactive effects of professor identity and after-class anxiety were specified as predictor, with attributions as mediator and inaccuracies in quiz performance as outcome, revealed that (a) the self-serving attributions explained greater number of inaccuracies in quiz performance of those high (versus low) anxiety students who were exposed to an ingroup (female) professor, b_{IE} = 1.08, se = .72, 95% CI = [0.01, 2.93], but not those exposed to an outgroup professor, b_{IE} = -.38, se = .44, 95% CI = [-1.69, 0.22]. When the roles of quiz performance (now as mediator) and post-quiz attributions (now as outcome) were reversed, the indirect effect of exposing anxious students to ingroup professors was non-significant, bIE = .12, se = .08, 95% CI = [-0.004, 0.32], as was exposure to outgroup professors, bIE = -.02, se = .02, 95% CI = [-0.09, 0.02]: confirming that self-serving attributions are the mechanism that drives the ingroup spotlighting performance impairment, and not the other way round.

The corresponding moderated regression analysis with onset vocal jitter (as the index of anxiety), revealed neither a main effect of professor identity (p = .591) nor a professor identity x anxiety interaction (p = .923). Interestingly, however, onset vocal jitter predicted

reduced levels of post-quiz attributions regardless of professor identity, b = -.34, se = .12, p = .007.

Figure 11

The Interactive Effect of Professor Identity and After Class Anxiety (Vocal Jitter) in Predicting Post-Quiz Attributions



Note. Error bars are standard errors.

Summary of Key Findings

Study 5 once more replicated the main findings from Studies 3 and 4 in that stereotype salience was related to higher levels of anxiety when exposed to outgroup professors compared to ingroup professors. Further to that, Study 5 revealed that: (a) ingroup (female) experts/role models can cause a decrease in confidence in mathematics for highly anxious students (H5ai). The initial sight of an ingroup professor caused a decrease in

female students' confidence in mathematics and this effect continued after the class. Results also showed (b) that post-quiz attributions were higher in the ingroup than in the outgroup expert condition for anxious students (H5b). These self-preserving attributions also explained decrease in performance for anxious students when exposed to ingroup experts. Put together, the key findings continue to provide more evidence in support for the ingroup spotlighting effect while still supporting the predictions of the SIM which is expected to be true even in a Malaysian context which was the original context of the earlier studies. Based on the findings, an ingroup Malaysian professor could cause anxious Malaysian students who are suffering from imposter syndrome and are highly anxious within a STEM related class to decrease in confidence instead of increasing as predicted by SIM and this ingroup spotlighting performance impairment is driven by self-serving attributions.

Summary and Conclusions

Chapter 7 presented two studies (Studies 4 and 5) that aimed to test the predictions derived from the SIM and the ingroup spotlighting thesis (RQ4) especially regarding performance (RQ3), having established in Chapter 6 via Study 3 that stereotypes worsen the anxiety of Malaysian female students within the context of outgroup professors. Apart from replicating the main findings of Study 3 in that students demonstrated higher levels of anxiety when stereotypes were salient and they were exposed to outgroup professors, Study 4 revealed that ingroup experts could improve performance (H4ai) and the effect could have longitudinal potency with an increased retention rate (partial H4b). However, the study also showed that this identity-based vaccine was not effective among students experiencing the imposter syndrome, the very students who are theorised to be beneficiaries of the inoculative effects that ingroup experts provide (H4aii; Dasgupta, 2011; Stout et al., 2011). So, the study provides evidence that is consistent with the SIM while at the same time corroborating the spotlighting thesis when it comes to stereotyped underrepresented students experiencing

imposter syndrome. Study 5 not only provided the necessary opportunity to replicate the findings for both SIM and the ingroup spotlighting thesis within the gender context, it also further showed that ingroup experts can cause a decrease in confidence in mathematics for highly anxious students and self-preserving attributions increased when anxious students were exposed to an ingroup expert as predicted by the ingroup spotlighting thesis.

Chapter Eight

General Discussions and Conclusion

Overall, this thesis set out to establish if teachers, specifically those that are of a different social identity from the students (i.e., the outgroup), always threaten stereotyped underrepresented STEM students' academic related outcomes *when the social identity is salient*. To do that, it was important to validate the assumptions regarding the sources of anxiety within the classroom and examine whether such anxiety had an impact on students' confidence in the classroom (RQ1). Study 1 presented in Chapter 4 showed that anxiety within a classroom setting is commonly reported among the students surveyed, who also reported concerns over their social image/identity as the largest source of that anxiety was traced to shame and embarrassment (H1a). Considering that shame and embarrassment are emotions connected to self-image and self-presentation in a social setting and are reactions to how individuals are perceived by others (Chekroun & Nugier, 2011), this finding was pertinent and begged further investigation regarding who the students were concerned about within their classroom setting.

Study 1 also showed that the anxiety the students reported was highest in the context of teachers as compared to friends and parents/guardians (H1b). Particularly, teachers' involvement in assessing students' performance could potentially be a concern because other-as-source stereotype threats tend to emerge due to perceptions of how others might evaluate the students' performance (Shapiro & Williams, 2012). Not surprising, the reported anxiety was shown to potentially have a detrimental effect on students' confidence, and stemming from that, possibly even classroom outcomes such as classroom participation, engagement, and performance (H1c; Chapell, et al., 2005). From the literature, this would seem to be the tip of the iceberg as other consequences such as individuals wanting to hide away or to avoid

the social situation altogether (Tangney, 1991; McGregor & Elliot, 2005), and feelings of incompetence or self-deficiency (Sagar & Stoeber, 2009) have also been reported.

Having established that there were social image related anxieties among students in the context of their teachers, the next step involved identifying the specific content of the meta-stereotypes (beliefs about how the students were viewed by their teachers) in the student versus teacher context and whether the stereotype concerns differed depending on the social identity of the teacher (outgroup versus ingroup; RQ2). Study 2 which was reported in Chapter 5 showed that meta-stereotypic beliefs among the students studied were mostly negative (H2a) which agrees with the findings reported by Vorauer et al. (1998, 2000) but within a different context and examining varying characteristics used for meta-stereotypes. The generally negative meta-stereotypes are salient especially when thinking of teachers who do not share the same social identity in common to them (i.e., the outgroup teacher; H2b) like Gordijn (2010) whereby Malaysian students reported more negative meta-stereotypes when thinking of the higher status outgroup. The findings in Study 2 aligned with the idea that outgroup teachers increase the salience of negative meta-stereotypes and could induce anxiety among students (Steele & Aronson, 1995). In other words, if students think that outgroup teachers hold negative views of them, they are more likely to feel anxious about contacts with these outgroup teachers as compared to ingroup teachers (Kim & Oe, 2009; Stephan & Stephan, 1985; Vorauer et al., 2000).

Moreover, the competence-related meta-stereotypes were of particular concern when students were faced with the outgroup teachers rather than the ingroup teachers (H2c). As status is linked to competence (Fasel et al., 2021), it was not surprising that the students were more aware of competence-related meta-stereotypes in the context of the higher status outgroup teachers. The danger, however, is that status also reflects access to resources, influence, and respect, and it has implications on performance and learning outcomes (Fasel

et al., 2021). Therefore, as stereotyped students report more negative competence-related meta-stereotypes, they could experience drop in performance due to ST effects (Stephan, 2014) and the anxiety caused by it (Ma et al., 2022), as well as lowered self-esteem (Gordijn, 2010; Vorauer et al., 1998, 2000).

The results in Studies 1 and 2 informed the focus of Study 3 and beyond on the effects of activating the negative meta-stereotypes that are held by outgroup (compared to ingroup) teachers on student's classroom outcomes. Decades ago, Steele and Aronson (Steele & Aronson, 1995) advanced the thesis that situational salience of negative ingroup stereotypes can cause an increase in anxiety that undermines performance in domains that are tied to the stereotype. Of particular interest were stereotypes that portray racial/ethnic minorities and women as deficient in math and science related fields. Although there were several explanations for this achievement gap, the impacts of stereotypes on motivation and performance have received marked attention. However, some recent negative tests of this idea (Flore et al., 2018), have cast doubt on ST as a robust explanation for the achievement gap often seen amongst students from stereotyped backgrounds (e.g., women). Studies that provided conflicting evidence for a stereotype-threat explanation for identity-based achievement gaps in STEM have often tested the relevant assumption under conditions that could be explained by other situational factors such as: (a) other sources of anxiety were frequently conflated in the testing situation (teacher's formality, exam fever, teacher social identity, emotional contagion and so on) and; (b) the primary mechanism of anxiety, theorised to cause ST effects, was often measured using self-reports that are vulnerable to normative pressures (e.g., feminist/egalitarian norms), rendering the actual causes of identity-based differences hard to discern.

The last three studies addressed these issues in the context of testing the effectiveness of an identity based "vaccine" (i.e., ingroup experts) on STEM outcomes of students from

stereotyped backgrounds. The tests were grounded in SIM's inoculation hypothesis and introduced a new perspective: the ingroup spotlighting thesis. The last three empirical studies of this thesis, Studies 3, 4 and 5, had the overarching goal of closely examining the predictions that could be deduced from the more established stereotype inoculation model (SIM) by Dasgupta (2011) and many others (Asgari et al., 2012; Marx & Ko, 2012; Stout et al., 2011) as compared to the ingroup spotlighting thesis being posited here (RQ4). These studies specifically and systematically, investigated the impact of an ingroup versus outgroup expert on STEM performance of underrepresented students as a function of stereotype threat (ST) and anxiety (RQ3). Based on these studies which are found in Chapters 6 and 7, these stereotype-related anxieties can harm classroom outcomes of vulnerable students who also come from stigmatised (or underrepresented) social groups (e.g., women), especially when they are exposed to ingroup (rather than outgroup) teachers, which contradicts SIM but supports the newer ingroup spotlighting thesis that is being advanced here. In the process of answering these empirical and theoretical research questions, Study 3 in Chapter 6 also helped address an important methodological question: regarding whether vocal jitter could provide an accurate, less biased, and more conclusive test of stereotype threat effects on anxiety and performance (RQ5; H3ai and H3aii).

In these three studies, the preponderance of the evidence indicated some support for ST as a psychological reality, in that: a) the salience of negative ingroup stereotypes largely increased anxiety, especially in an intergroup setting, but only during initial encounters with the outgroup, when impression management concerns were more relevant (see Studies 3, 4 & 5 results for onset vocal jitter); b) removing or minimizing this threat (e.g. via exposure to ingroup experts) largely reduced the number of incorrect answers to a quiz immediately following the treatment (Studies 4-5): A feat that lingered four weeks after this identity-based

vaccine was administered (Study 3). A summary of the research hypotheses and outcomes are presented in Table 14.

To recap, specifically, and systematically, this thesis sought to answer the following research questions in a series of five studies:

RQ1 - Is there empirical evidence to validate the assumptions regarding the sources of anxiety within the classroom? Does such anxiety have an impact on students' confidence in classroom outcomes? (As presented in Chapter 4)

RQ2 - What is the specific content of the meta-stereotypes in the student versus teacher context? Do students' stereotype concerns differ depending on the social identity of the teacher? (As presented in Chapter 5)

RQ3 - Are there positive or negative impacts of an ingroup (versus outgroup) expert/role model on STEM performance in underrepresented groups as a function of stereotype threat and anxiety? (As presented in Chapters 6 and 7)

RQ4 - Do the *ingroup spotlighting thesis* or the *stereotype inoculation model* provide better explanations for the poorer outcomes of underrepresented groups in the STEM context of higher education? (As presented in Chapters 6 and 7)

RQ5 - Does vocal jitter provide an accurate, less biased, and more conclusive test of stereotype threat effects on anxiety and performance? (See Chapter 6)

Table 14

Summary of Research Hypotheses and Outcomes

Number	Hypotheses	Outcomes
	Study 1	
H1a	Concerns related to social image/identity would be the most common source of anxiety in the classroom.	Supported
H1b	Concerns related to social image/identity would be the most common source of anxiety in the classroom within the teacher context compared to friends or family.	Supported
H1c	Concerns related to social image/identity within the teacher context would negatively affect confidence in the classroom.	Supported
	Study 2	
H2a	Meta-stereotype content among Malaysian students regarding their teachers would be generally negative.	Supported
H2b	Meta-stereotype content among Malaysian students would be more negative in the Caucasian teacher (outgroup) condition than Malaysian teacher (ingroup)	Supported
H2c	More competence-related meta-stereotypes would be reported in the outgroup condition.	Supported
	Study 3	
H3ai	The actual vocal jitter extracted from the recordings would be significantly correlated to judges' ratings of anxiety from students' recorded responses.	Supported
H3aii	*The experimental conditions would affect the judges' ratings of anxiety, corresponding to its effect on vocal jitter.*	Supported
H3b	Students' anxiety (vocal jitter) would increase when negative ingroup stereotypes are salient compared to non-salient.	Supported
H3c	Students' anxiety (vocal jitter) would increase when negative ingroup stereotypes are salient only when taught by outgroup (not ingroup) teachers.	Supported
H3d	Students' anxiety would be affected at the onset of the class but not at the end of the class.	Supported
	Study 4	
H4ai	Anxious students will perform worse on a STEM test in the presence of an outgroup teacher compared to an ingroup teacher (SIM).	Supported
H4aii	Anxious students with imposter syndrome will perform worse on a STEM test in the presence of an ingroup teacher compared to an outgroup teacher (ingroup spotlighting).	Supported
H4b	Both stereotype inoculation and ingroup spotlighting would have longitudinal effects on performance.	Partial Support (SIM only)
	Study 5	
H5ai	Anxious students would have lower post-lecture confidence in mathematics when exposed to an ingroup teacher compared to an outgroup teacher (ingroup spotlighting).	Supported
H5aii	Anxious students would have higher post-lecture confidence in mathematics when exposed to an ingroup teacher compared to an outgroup teacher (SIM).	Not Supported
H5b	Anxious students would endorse more self-esteem preserving attributions when exposed to an ingroup teacher compared to an outgroup teacher.	Supported

Stereotype Inoculation Versus Ingroup Spotlighting Theses

One of the primary contributions of the research presented in this thesis is that it raises important questions about the boundary conditions of stereotype inoculation for students from stereotyped backgrounds (see Akcinar et al., 2011; Pelham & Hardin, 2011; Tse et al., 2011; for discussions on some of those boundary conditions). Indeed, a cornerstone of one of the most influential models in stereotype threat reduction within classrooms (i.e., SIM and role model interventions), is that ingroup experts can immunise students from stigmatised backgrounds from stereotype-related anxieties, enabling them to achieve better academic outcomes in those STEM fields affected by negative ingroup stereotypes (Marx & Ko, 2012; Stout et al., 2011). Much of the evidence from the current investigation corroborate these assumptions. Ingroup teachers reduced students' anxiety when reminded of negative ingroup stereotypes, whereas outgroup teachers did not serve this palliative function (H3c). This pattern was rather consistent between the studies in this thesis and the results from past literature such as by Delgado and Prieto (2008), Hess et al. (2003), Laurin (2020) and Stone et al. (1999). In Study 4, ingroup experts also helped to reduce the number of incorrect answers that Malaysian students gave on a quiz posed to them (H4ai; Asgari et al., 2010; Marx & Ko, 2012; Stout et al., 2011), and this effect lingered four weeks from the initial treatment (partial H4b). In Study 5, women made fewer errors in a math-related quiz when taught by a female teacher, provided their level of anxiety was low. That this latter effect emerged when situational experience of anxiety was low aligns with the ST perspective that eliminating situations that trigger threat/anxiety (e.g., through inoculation) helps to improve academic outcome of students from stereotyped backgrounds (Beasley & Fischer, 2012; Steele et al., 2002).

However, the data also somewhat departs from SIM particularly with regards to students from stigmatised backgrounds who are likely suffering from an imposter syndrome

(i.e., high anxiety and low subject efficacy). SIM's inoculation thesis predicts that students with imposter syndrome and highly anxious would perform better when exposed to an ingroup professor (Stout et al., 2011). Stereotyped students who feel like an imposter within the domain they are in, may be concerned about the ingroup teacher evaluating them and, therefore, feel anxious (Hornsey et al., 2002; Hoyt, 2013). The ingroup expert/role model would also increase the pressure to perform on the stereotyped individual because these individuals have no space for excuses and because it creates an attributional ambiguity that could have detrimental effects on the self-esteem (Crocker et al., 1991). The search for excuses for a lack in performance could reduce cognitive resources and stereotyped individuals would feel the burden to prove themselves, especially if there are feeling like an imposter and are anxious about their place in STEM to begin with.

As such, the evidence shown in Studies 4 and 5 highlights a harmful effect of ingroup experts for those students likely suffering from an imposter syndrome. Specifically, these studies showed that an ingroup experts could place a spotlight on such students (even when students were exposed to lecture recordings online), causing them to achieve worse (not better) outcomes in STEM because an ingroup expert provides a glaring example that underrepresented groups can overcome barriers within the domain that they are a minority in (Marx et al., 2005, 2009). As far as reasonably possible, the author is not aware of any other study that has systematically documented the potential side-effect of this type of identity-based vaccine in both a classroom and online lecture scenarios, and this thesis provides a novel acoustic tool for measuring anxiety that can guide future exploration in this area.

Policy Implications

There has been an increasing push in the education sector around the globe towards greater representation of women and minority faculty (e.g., affirmative action in the US, the Athena Swan awards in the UK) especially in the STEM-related fields. What implication

might the current findings have for policy and intervention programs? Given that the current data concur with the burgeoning literature revealing that shared social identities can have palliative benefits (see Haslam et al., 2009), the temptation might be to power ahead with interventions that promote greater representation, especially in education where success or failure in STEM classes could impact future career choices and opportunities.

However, the present findings regarding the ingroup spotlighting effects suggest that the benefits might be more nuanced than the literature currently portrays, and any intervention should take into account both benefits and challenges. For example, while ingroup experts reduce anxiety overall, the impact on the most vulnerable students—students with high anxiety and low self-efficacy (imposter syndrome)—should be considered and addressed via additional supports. *A common method to provide support is mentorship programmes. However, any mentorship programmes should include a diversity of mentors, and this could be strengthened by establishing mentor networks and connections between students and professionals that could help enhance academic outcomes* (Brown, Magaña, Crespo, & White, 2021). It should be noted that the ingroup spotlighting effects emerged during a brief encounter with a teacher and the longitudinal data suggest that it could fade over time, so it is possible that prolonged exposure to positive experts could override the ingroup spotlighting effect (Nhundu, 2007). However, this will need to be investigated in future research.

Any affirmative action policies would need to be considered and evaluated more closely. The assumption that increasing the representation and visibility of the ingroup role model for underrepresented students would automatically help, may not necessarily be true for all underrepresented students. Policies that actively promote diversity and inclusion, ensuring fair representation of both ingroup and outgroup role models for underrepresented students would probably work more effectively as it caters to a wider range of students. Even

the sharing of success stories of individuals needs to have a representation from both ingroup and outgroup to help counteract both stereotypes and inspire confidence. Incorporating role models from diverse backgrounds would better inspire a sense of belonging for underrepresented students. In fact, according to the social role theory (Eagly & Wood, 2012)*, exposure to individuals in non-traditional roles has the potential to disrupt and reshape traditional gender and racial stereotypes, promoting more inclusive attitudes and reshape traditional gender and racial stereotypes. This will in turn promote more inclusive attitudes and behaviours among students who look up to those role models.*

Furthermore, some of the present data indicated that on some occasions, exposure to outgroup teachers could offer tangible benefits for female students who may have an imposter syndrome, in so far as such situations may not evoke as much ingroup spotlighting as perhaps an ingroup teacher context might. Interventionists should exercise caution, though, in applying a broad-brush approach here because the ingroup spotlighting findings are novel and require more research to fully understand. Nevertheless, this thesis suggests that adopting a more nuanced approach that recognizes the specific needs of subgroups of students from stereotyped communities could be fruitful. The shift in education to more online modes of delivery in the post COVID-19 era might provide an opportunity for this sort of nuance in the form of more tailored instruction.

Limitations and Recommendations for Future Research

Convenience Sampling

Caution needs to be exercised in generalising the results presented in this thesis outside of the study population due to the use of a relatively small convenience sample. Apart from being one of the most popular sampling methods used in psychology (Hultsch et al., 2002; Zhao, 2021), convenience sampling was used here because it was the most approachable and accessible, easiest, and cheapest method considering the limited resource

and time available (Forster, 2001). It is after all the practice whereby participants are recruited because of ease of access and recruitment (Zhao, 2021). However, one main limitation of convenience sampling to this research is that it may not necessarily be representative of the general population, therefore, limiting the generalisability (Staetsky, 2019) *especially since participants were chosen based on ease of access rather than a random or systematic method. Considering that all studies except one have participants from one programme in one university (although different cohorts), the findings may not generalise to the broader student population and there is a need to employ a more rigorous sampling method for robust and reliable results.*

Although the limitations of convenience sampling are acknowledged, the technique used to collect data for this thesis yielded a very high response rate among the students sampled. This helped to reduce both non-response and voluntary response biases (Smith & Dundes, 2008). Convenience sampling has also been helpful in surveying the general attitudes and opinions of the students surveyed to provide some credence to the ingroup spotlighting hypothesis to be tested and developed further with more rigorous studies (Galloway, 2005). Therefore, recommendations for future research are to obtain a larger, random sample, in order to increase the confidence in the external validity of the results reported here.

Sample Size

Reduction in the power of the study and increase in the margin of error could be the inevitable results of a small sample size and admittedly, especially for Studies 3 to 5, small sample size was a limitation in this research. With a smaller sample size, the findings become decreasingly representative of the entire population (van Lissa, 2020). Three methods have been applied to augment the studies and maximise the findings despite this limitation. Firstly, dealing with missing data as in Study 2. According to (Hopkin et al., 2015), “Any strategy

that can reduce the number of missing cases or make use of the incomplete information provided by some cases without the introduction of bias into the parameter estimates and standard errors will yield an increase in statistical power without additional sampling." (p. 951).

Secondly, the use of bootstrapping as in Studies 3 to 5. Preacher and Hayes (2004), and Bollen and Stine (1990) agree that bootstrapping could be used for small sample sizes while (Yung & Chan, 1999) concluded that simulation studies indicate that bootstrapping generally does perform well. However, it is acknowledged that while Preacher and Hayes (2004) suggest that bootstrapping could help with the low sample sizes, it is not by a lot.

And finally, the replication of the study as in Study 5. The replication in Study 5 was used to accumulate evidence of the relevant effects in the previous studies in order to strengthen the conclusions. Replication for this purpose meant a new study targeting the same scientific questions to see if it will produce consistent results (Anderson & Kelley, 2022). And replications can be ideal to further develop "useful, testable, and generalizable predictions" for the phenomenon being tested in this thesis (Simons, 2014).

Meta-stereotype Content Measurement

Although the meta-stereotype content measure used in this study is an established measure used in a seminal paper by Vorauer et al. (1998), careful consideration needs to be given to the most appropriate meta-stereotype content measurement especially since in more recent times, questions have been raised regarding this by researchers. Finkelstein et al. (2015) considered whether it was more appropriate to use a trait list rating method as compared to an open-ended question method especially since there were different characteristics that emerged between the two methods (Finkelstein et al., 2013). According to Finkelstein et al. (2015), it may be more appropriate to use an open-ended approach for initial testing such as was done in Study 2, while reserving the close-ended approach used in Study

2 for when more information regarding specific meta-stereotypes is already available. However, the benefit of the approach taken is that the students surveyed would be less concerned about their responses especially since they are being asked to predict how outgroup members evaluate their ingroup instead of being asked about their own attitudes about outgroup members (Kim & Oe, 2009).

Furthering the Research

Apart from what has been mentioned in the previous sections, further studies could be conducted to expand the research. Study 2 on the meta-stereotypes revealed that students generally felt their teachers perceived them more negatively especially in the context of an outgroup teacher versus the ingroup teacher and in the context of competence related traits. Considering that meta-stereotypes are beliefs that the outgroup has a certain opinion of one's own ingroup (Owuamalam & Zagefka, 2014)*, it may be that it does not correspond with what the teachers actually think of the students. Further studies could be done to explore whether students' meta-stereotypes have any basis in reality. The question would be whether the teachers think such thoughts about them to begin with. Considering that students who believe that their outgroup teachers have negative views of them are more likely to feel anxious about contacts with those outgroup teachers, it could be really helpful if statistically, it could be shown that the teachers don't have those views of them.*

In Study 3, the validation of vocal jitter as an index for anxiety was done via independent judges who were blind to the experimental treatments, who then rated the anxiety in the voice recordings that were collected from participants. Statistical analysis was employed to ensure the robustness of the validation via correlations as well as convergent and discriminant validations. The measure could be further validated by comparing it with other established methods or measures to confirm its accuracy such as the State-Trait Anxiety Inventory (Spielberger, 1983)*. Other physiological measures such as heart rate*

variability (Chalmers et al., 2014), electrodermal activity (Rahma et al., 2022), and cortisol levels (Hek et al., 2013) could also be used for a comparison as they serve as other indicators of anxiety. Ultimately, neuroimaging techniques such as functional magnetic resonance imaging (Holzschneider & Mulert, 2011) *can also provide insights into the neural correlates of anxiety to confirm the validity of vocal jitter as a measure for anxiety.*

Conclusions

This thesis started with empirically validating anxiety within the classroom setting is a reality and such anxiety can have important consequences for students' confidence in their classroom outcomes (Study 1 in Chapter 4). This anxiety is most relevant when students are faced with teachers rather than friends or parents/guardians, especially those teachers who do not share the same social identity with the students (outgroup). Further study on the meta-stereotypes that could contribute to these tensions between teachers and students revealed that students generally felt their teachers perceived them more negatively especially in the context of an outgroup teacher versus the ingroup teacher (Study 2 in Chapter 5). This concern seemed more pronounced in the context of competence related traits. Subsequently, this thesis examined the positive and potentially negative impacts of an ingroup (versus outgroup) expert on STEM performance in racial groups and women as a function of stereotype threat and anxiety (Studies 3, 4 and 5 in Chapters 6 and 7). The answers to three key questions summarise the primary contributions of the investigation. First, does the situational salience of negative ingroup stereotypes cause anxiety to increase amongst students from stereotypes backgrounds? The answer to this question is *somewhat yes*. But this trend is mostly reliable when students are exposed to teachers with a different social identity to themselves. Second, does an identity-based intervention, such as exposure to ingroup experts, help to reduce anxiety and performance errors? The answer is, *overall, yes*. Ingroup identity effects of teachers were fairly reliable across the studies. Third, does this inoculation effect benefit all

underrepresented students, especially students from stereotyped backgrounds likely experiencing an imposter syndrome? The answer to this question is *no*, based on the current data. Although such an identity-based "vaccine" seems useful for some students, it may not be as effective for those anxiety-prone students who doubt their place in the relevant STEM field. This latter finding calls for greater caution in both advocacy and intervention around stereotype-based underachievement in STEM for students who come from stigmatised backgrounds.

Glossary

Terms	Definition
anxiety	an emotion characterised by feelings of tension, worried thoughts, and physical changes like increased blood pressure (or in this thesis, voice jitter)
classroom outcomes	classroom participation, engagement and/or performance
ingroup spotlighting	being given heightened attention by the increased exposure to exemplary counter-stereotypical ingroup role models
meta-stereotypes	the stereotypes that individuals think that those outside of their group hold of a social group to which they belong, or a belief or thinking or even awareness that an outgroup that is relevant has a certain opinion of the ingroup
role model	individuals who influence achievements, motivation, and goals by acting as behavioral models, representations of the possible, and/or inspirations
stereotype inoculation	social “vaccines” that immunise stigmatised students against the experience of anxiety that causes stereotype threat effects

References

Abd Rahim, N., Mohamed, Z. B., Amrin, A., & Mohammad, R. (2019). Women's dual roles and career growth: A preliminary study of Malaysian female talents in science, engineering and technology (SET). *Journal of Physics: Conference Series, 1174*, 012013. doi:10.1088/1742-6596/1174/1/012013

Abd Wahab, R. I. (2022, March 7). Being Female in Male-Dominated Industries: Levelling the Playing Field. *Bernama: Thougths*. Retrieved from https://www.bernama.com/en/thoughts/news.php?id=2059337

Abrams, D., Crisp, R. J., Marques, S., Fagg, E., Bedford, L., & Provias, D. (2008). Threat inoculation: experienced and imagined intergenerational contact prevents stereotype threat effects on older people's math performance. *Psychology and Aging, 23*(4), 934-939. doi:10.1037/a0014293

Abrams, D., Eller, A., & Bryant, J. (2006). An age apart: The effects of intergenerational contact and stereotype threat on performance and intergroup bias. *Psychology and Aging, 21*(4), 691-702. doi:https://doi.org/10.1037/0882-7974.21.4.691

Akcinar, E. N., Carr, P. B., & Walton, G. M. (2011). Interactions with men and whites matter too. *Psychological Inquiry, 22*(4), 247-251. doi:http://dx.doi.org/10.1080/1047840X.2011.625606

Allen, A. P., Kennedy, P. J., Dockray, S., Cryan, J. F., Dinan, T. G., & Clarke, G. (2017). The Trier Social Stress Test: Principles and practice. *Neurobiology of Stress, 6*, 113-126. doi:10.1016/j.ynstr.2016.11.001

Allen, E. C., & Collisson, B. (2020). Do aspirational role models inspire or backfire? Perceived similarity mediates the effect of role models on minority students' college choices. *Journal of Marketing for Higher Education, 30*(2), 221-238. doi:https://doi.org/10.1080/08841241.2020.1723780

Alrabai, F. (2015). The influence of teachers' anxiety-reducing strategies on learners' foreign language anxiety. *Innovation in Language Learning and Teaching, 9*(2), 163-190. doi:https://doi.org/10.1080/17501229.2014.890203

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12*(5), 385-390. doi:https://doi.org/10.1111/1467-9280.00371

American Psychological Association. (2022). *Anxiety*. Retrieved November 17, 2022, from American Psychological Association: https://www.apa.org/topics/anxiety

American Psychological Association. (n.d.). *APA Dictionary of Psychology*. Retrieved December 3, 2022, from American Psychological Association: https://dictionary.apa.org/group-identification

American Psychological Association. (n.d.). *APA Dictionary of Psychology*. Retrieved December 14, 2022, from American Psychological Association: https://dictionary.apa.org/unobtrusive-measure

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology, 9*(3), 272-279. doi:10.1037/h0025907

Anderson, S. F., & Kelley, K. (2022). Sample Size Planning for Replication Studies: The Devil Is in the Design. *Psychological Methods*, Advance online publication. doi:https://doi.org/10.1037/met0000520

Anne, D. (2022, September 16). *Of Education, Race and Change in Malaysia*. Retrieved from IDEAS: https://www.ideas.org.my/of-education-race-and-change-in-malaysia/

Armenta, B. E. (2010). Streotype 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, E., Turner, J. A., & Carlsmith, J. M. (1963). Communicator credibility and communication discrepancy as determinants of opinion change. *Journal of Abnormal and Social Psychology, 67*(1), 31-36. doi:10.1037/h0045513

Aronson, J., & Dee, T. (2011). Stereotype threat in the real world. In M. Inzlicht, & T. Schmader (Eds.), *Stereotype Threat: Theory, Process, and Application* (pp. 264-279). New York: Oxford University Press. doi:https://doi.org/10.1093/acprof:oso/9780199732449.003.0017

Aronson, J., & Steele, C. M. (2005). Stereotypes and the fragility of academic competence, motivation, and self-concept. In A. J. Elliot, & C. S. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 436-456). New York: The Guilford Press.

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. doi:10.1006/jesp.1998.1371

Aronson, J., Quinn, D. M., & Spencer, S. J. (1998). Stereotype threat and the academic underperformance of minorities and women. In J. K. Swim, & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 83-103). New York: Academic Press.

Asgari, S., Dasgupta, N., & Cote, N. G. (2010). When Does Contact with Successful Ingroup Members Change Self-Stereotypes? *Social Psychology, 41*(3), 203-211. doi:https://doi.org/10.1027/1864-9335/a000028

Asgari, S., Dasgupta, N., & Stout, J. G. (2012). When do counterstereotypic ingroup members inspire versus deflate? The effect of successful professional women on young women's leadership self-concept. *Personality and Social Psychology Bulletin, 38*(3), 370-383. doi:10.1177/0146167211431968

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 48*, 5-37. doi:10.1016/j.jsp.2009.10.001

Beasley, M. A., & Fischer, M. J. (2012). Why they leave: the impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Social Psychology of Education, 15*, 427-448. doi:https://doi.org/10.1007/s11218-012-9185-3

Beede, D., Julian, T., Langdon, D., McKittrick, G., Khan, B., & Doms, M. (2011). *Women in STEM: A gender gap to innovation (ESA Issue Brief #04-11.* Washington, DC: U.S. Department of Commerce.

Bell, A. E., Spencer, S. J., Iserman, E., & Logel, C. E. (2003). Stereotype threat and women's performance in engineering. *Journal of Engineering Education, 92*(4), 307-312. doi:https://doi.org/10.1002/j.2168-9830.2003.tb00774.x

Birtel, M. D., & Crisp, R. J. (2012). Imagining intergroup contact is more cognitively difficult for people higher in intergroup anxiety but this does not detract from its effectiveness. *Group Processes & Intergroup Relations, 15*(6), 744-761. doi:https://doi.org/10.1177/1368430212443867

Blascovich, J. (2000). Using physiological indexes of psychological processes in social psychological research. In H. T. Reis, & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (pp. 117-137). Cambridge, UK: Cambridge University Press.

Bledsoe, T. S., & Baskin, J. J. (2014). Recognizing student fear: the elephant in the classroom. *College Teaching, 62*(1), 32-41. doi:10.1080/87567555.2013.831022

Block, C. J., Koch, S. M., Liberman, B. E., Merriweather, T. J., & Roberson, L. (2011). Contending with stereotype threat at work: A model of long-term responses. *The Counseling Psychologist, 39*(4), 570-600. doi:10.1177/0011000010382459

Board of Engineers Malaysia. (2019). *Anual Report Board of Engineers Malaysia.* Kuala Lumpur: Board of Engineers Malaysia. Retrieved from http://bem.org.my/documents/20181/163234/Laporan+Tahunan+LJM+2019+.pdf/c04821d4-5794-45c1-8d8a-7ff1634c2216

Bochner, S., & Insko, C. A. (1966). Communicator discrepancy, source credibility, and opinion change. *Journal of Personality and Social Psychology, 4*(6), 614-621. doi:10.1037/h0021192

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.70) [Computer software]. Retrieved from http://www.praat.org/

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology, 20*, 115-140. doi:https://doi.org/10.2307/271084

Boniecki, K. A., & Moore, S. (2003). Breaking the Silence: Using a token economy to reinforce classroom participation. *Teaching of Psychology, 30*(3), 224-227. doi:10.1207/S15328023TOP3003_05

Bosson, J. K., Haymovitz, E. L., & Pinel, E. C. (2004). When saying and doing diverge: The effects of stereotype threat on self-reported versus non-verbal anxiety. *Journal of Experimental Social Psychology, 40*(2), 247-255. doi:https://doi.org/10.1016/S0022-1031(03)00099-4

Botes, E., Dewaele, J.-M., & Greiff, S. (2020). The foreign language classroom anxiety scale and academic achievement: An overview of the prevailing literature and a meta-analysis. *Journal for the Psychology of Language Learning, 2*(1), 26-56. Retrieved from https://www.jpll.org/index.php/journal/article/view/botesetal

Branscombe, N. R. (1998). Thinking about one's gender group's privileges or disadvantages: Consequences for well-being in women and men. *British Journal of Social Psychology, 37*(2), 167-184. doi:10.1111/j.2044-8309.1998.tb01163.x

Bravata, D. M., Watts, S. A., Keefer, A. L., Madhusudhan, D. K., Taylor, K. T., Clark, D. M., . . . Hagg, H. K. (2020). Prevalence, predictors, and treatment of imposter syndrome: a systematic review. *Journal of General Internal Medicine, 35*(4), 1252-1275. doi:10.1007/s11606-019-05364-1

Brescoll, V. L. (2016). Leading with their hearts? How gender stereotypes of emotion lead to biased evaluations of female leaders. *The Leadership Quarterly, 27*(3), 415-428. doi:https://doi.org/10.1016/j.leaqua.2016.02.005

Brown , L. M. (1998). Ethnic stigma as a contextual experience: A possible selves perspective. *Personality and Social Psychology Bulletin, 24*(2), 163-172. doi:https://doi.org/10.1177/0146167298242005

Brown, L. M., & Dobbins, H. (2004). Students' of color and European American students' stigma-relevant perceptions of university instructors. *Journal of Social Issues, 60*(1), 157-174. doi:https://doi.org/10.1111/j.0022-4537.2004.00104.x

Brown, R. (2000). Social identity theory: Past achievements, current problems and future challenges. *European Journal of Social Psychology, 30*(6), 745-778. doi:10.1002/1099-0992(200011/12)30:6<745::AID-EJSP24>3.0.CO;2-O

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psyhcology, 39*(6), 626-633. doi:https://doi.org/10.1016/S0022-1031(03)00039-8

Brown, W. C., Magaña, L., Crespo, C. J., & White, W. B. (2021). Mentoring Underrepresented Minoritized Students for Success. *Pedagogy in Health Promotion, 7*(1_suppl), 20S-22S. doi:https://doi.org/10.1177/23733799211054086

Buss, D. M. (1995). Psychological sex differences: origins through sexual selection. *American Psychologist, 50*(3), 164-168. doi:https://doi.org/10.1037/0003-066X.50.3.164

Cadaret, M. C., Hartung, P. J., Subich, L. M., & Weigold, I. K. (2017). Stereotype threat as a barrier to women entering engineering careers. *Journal of Vocational Behavior, 99*, 40-51. doi:https://doi.org/10.1016/j.jvb.2016.12.002

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105.

Casad, B. J., & Bryant, W. J. (2016). Addressing stereotype threat is critical to diversity and inclusion in organizational psychology. *Frontiers in Psychology, 7*, 8. doi:doi: 10.3389/fpsyg.2016.00008

Casad, B. J., Oyler, D. L., Sullivan, E. T., McClellan, E. M., Tierney, D. N., Anderson, D. A., . . . Flammang, B. J. (2018). Wise psychological interventions to improve gender and racial equality in STEM. *Group Processes & Intergroup Relations, 21*(5), 767-787. doi:https://doi.org/10.1177/1368430218767034

Catalyst. (2020, August 4). Quick take: Women in science, technology, engineering, and mathematics (STEM). Retrieved from https://www.catalyst.org/research/women-in-science-technology-engineering-and-mathematics-stem/

Chalmers, J. A., Quintana, D. S., Abbott, M. -A., & Kemp, A. H. (2014). Anxiety Disorders are Associated with Reduced Heart Rate Variability: A Meta-Analysis. *Frontiers in Psychiatry, 5*, 80. doi:https://doi.org/10.3389/fpsyt.2014.00080

Chandler, J., & Shapiro, D. (2016). Conducting Clinical Research Using Crowdsourced Convenience Samples. *Annual Reviews of Clinical Psychology, 12*, 53-81. doi:10.1146/annurev-clinpsy-021815-093623

Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*(2), 268-274. doi: 10.1037/0022-0663.97.2.268

Chekroun, P., & Nugier, A. (2011). "I'm ashamed because of you, so please, don't do that!": Reactions to deviance as a protection against a threat to social image. *European Journal of Social Psychology, 41*(4), 479-488. doi:https://doi.org/10.1002/ejsp.809

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin, 143*(1), 1-35. doi:https://doi.org/10.1037/bul0000052

Chesebro, J. L. (2003). Effects of teacher clarity and nonverbal immediacy on student learning, receiver apprehension, and affect. *Communication Education, 52*(2), 135-147. doi:10.1080/03634520302471

Clance, P. R. (1985). *The impostor phenomenon: Overcoming the fear that haunts your success.* Atlanta: Peachtree Publishers.

Clance, P. R., & Imes, S. A. (1978). The imposter phenomenon in high achieveing women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, Research & Practice, 15*(3), 241-247. doi:https://doi.org/10.1037/h0086006

Clark, A. H. (2018, December 18). *7 Ways to Use Anxiety to Improve Performance*. Retrieved November 17, 2022, from Psychology Today: https://www.psychologytoday.com/us/blog/hack-your-anxiety/201812/7-ways-use-anxiety-improve-performance

Connelly, L. M. (2017). Unobstructive Measures. *Medsurg Nursing, 26*(1), 59-61. Retrieved from https://www.proquest.com/scholarly-journals/unobtrusive-measures/docview/1870846541/se-2?accountid=135175

Conroy, D. E. (2001). Progress in the development of a multidimensional measure of fear of failure: The Performance Failure Appraisal Inventory (PFAI) . *Anxiety, Stress, and Coping, 14*, 431-452. doi:10.1080/10615800108248365

Conroy, D. E., & Elliot, A. J. (2004). Fear of failure and achievement goals in sport: Addressing the issue of chicken and the egg. *Anxiety, Stress, and Coping, 17*, 271-285. doi:10.1080/1061580042000191642

Conroy, D. E., Kaye, M. P., & Fifer, A. M. (2007). Cognitive Links Between Fear Of Failure And Perfectionism. *Journal of Rational-Emotive & Cognitive-Behavior Therapy, 25*, 237-253. doi:https://doi.org/10.1007/s10942-007-0052-7

Conroy, D. E., Metzler, J. N., & Hofer, S. M. (2003). Factorial invariance and latent mean stability of performance failure appraisals. *Structural Equation Modeling, 10*, 401-422. doi:10.1207/S15328007SEM1003_4

Conroy, D. E., Willow, J. P., & Metzler, J. N. (2002). Multidimensional fear of failure measurement: The performance failure appraisal inventory. *Journal of Applied Sport Psychology, 14*(2), 76-90. doi:10.1080/10413200252907752

Crisp, R. J., Bache, L. M., & Maitner, A. T. (2009). Dynamics of social comparison in counter-stereotypic domains: Stereotype boost, not stereotype threat, for women engineering majors. *Social Influence, 4*(3), 171-184. doi:10.1080/15534510802607953

Crocker, J., Major, B., & Steele, C. M. (1998). Social Stigma. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 2, pp. 504-553). New York: Academic Press.

Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology, 60*(2), 218-228. doi:https://doi.org/10.1037/0022-3514.60.2.218

Croizet, J. C., & Claire, T. (1998). Extending the concept of stereotype threat to social class: The intellectual underperformance of students from low socioeconomic backgrounds. *Personality and Social Psychology Bulletin, 24*(6), 588-594. doi:10.1177/0146167298246003

Croizet, J.-C., Despres, G., Gauzins, M.-E., Huguet, P., Leyens, J.-P., & Meot, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin, 30*(6), 659-799. doi:https://doi.org/10.1177/0146167204263961

Crosby, F. J. (1982). *Relative deprivation and working women.* New York: Oxford University Press.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the bias map. *Advances in Experimental Social Psychology, 40*, 61-149. doi:10.1016/S0065-2601(07)00002-0

Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model. *Psychological Inquiry, 22*, 231-246. doi:10.1080/1047840X.2011.607313

Dasgupta, N., Scircle, M. M., & Hunsinger, M. (2015). Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proceedings of the National Academy of Sciences, 112*(16), 4988-4993. doi:https://doi.org/10.1073/pnas.1422822112

Davidson, B. I., Shaw, H., & Ellis, D. A. (2022). Fuzzy constructs in technology usage scales. *Computers in Human Behavior, 133*, 107206. doi:https://doi.org/10.1016/j.chb.2022.107206

Delgado, A. R., & Prieto, G. (2008). Stereotype threat as validity threat: The anxiety-sex-threat interaction. *Intelligence, 36*(6), 635-640. doi:https://doi.org/10.1016/j.intell.2008.01.008

Dennehy, T. C., & Dasgupta, N. (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences, 114*(23), 5964-5969. doi:https://doi.org/10.1073/pnas.1613117114

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* (pp. 309-344). New York: Cambridge University Press. doi:https://doi.org/10.1017/9781316811764

Department of Education. (2022, November 10). *Department of Education, Australian Government*. Retrieved from https://www.education.gov.au/australian-curriculum/national-stem-education-resources-toolkit/introductory-material-what-stem/why-stem-important

Department of Statistics Malaysia. (2023). *Labour Force Statistics Report.* Putrajaya: Department of Statistics Malaysia. Retrieved from https://storage.googleapis.com/dosm-public-publications/monthly_2023-06_labour_en.pdf

Desombre, C., Anegmar, S., & Delelis, G. (2018). Stereotype threat among students with disabilities: the importance of the evaluative context on their cognitive performance. *European Journal of Psychology of Education, 33*, 201-214. doi:https://doi.org/10.1007/s10212-016-0327-4

Dos Santos, E. D., Albahari, A., Diaz, S., & De Freitas, E. C. (2022). 'Science and Technology as Feminine': Raising awareness about and reducing the gender gap in STEM careers. *Journal of Gender Studies, 31*(4), 505-518. doi:https://doi.org/10.1080/09589236.2021.1922272

Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences, 47*, 103-116. doi:https://doi.org/10.1016/j.lindif.2015.12.018

Dumas, J. E., Johnson, M., & Lynch, A. M. (2002). Likableness, familiarity, and frequency of 844 person-descriptive words. *Personality and Individual Differences, 32*, 523-531. doi:10.1016/S0191-8869(01)00054-X

El Zein, M., Wyart, V., & Grezes, J. (2015). Anxiety dissociates the adaptive functions of sensory and motor response enhancements to social threats. *eLife, 4*, e10274. doi:https://doi.org/10.7554/eLife.10274

Enders, C. K. (2003). Using the Expectation Maximization Algorithm to Estimate Coefficient Alpha for Scales With Item-Level Missing Data. *Psychological Methods, 8*(3), 322-337. doi:10.1037/1082-989X.8.3.322

Enders, C. K. (2010). *Applied Missing Data Analysis.* New York: The Guilford Press.

England, B. J., Brigati, J. R., & Schussler, E. E. (2017). Student anxiety in introductory biology classrooms: Perceptions about active learning and persistence in the major. *PLoS ONE, 12*(8), e0182506. doi:https://doi.org/10.1371/journal.pone.0182506

Estrada, M., Burnett, M., Campbell, A. G., Campbell, P. B., Denetclaw, W. F., Gutierrez, C. G., . . . Zavala, M. (2016). Improving underrepresented minority student persistence in STEM. *CBE-Life Sciences Education, 15*(3), 1-10. doi:https://doi.org/10.1187/cbe.16-01-0038

Evans, L., Haeberlein, K., Chang, A., & Handal, P. (2021). Convergent Validity and Preliminary Cut-Of Scores for the Anxiety and Depression Subscales of the DASS-21 in US Adolescents. *Child Psychiatry & Human Development, 52*, 579-585. doi:https://doi.org/10.1007/s10578-020-01050-0

Ewing, K. M., Richardson, T. Q., James-Myers, L., & Russell, R. K. (1996). The relationship between racial identity attitudes, worldview, and African American graduate students' experience of the imposter phenomenon. *Journal of Black Psychology, 22*(1), 53-66. doi:https://doi.org/10.1177/00957984960221005

Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science, 326*(5952), 535-538. doi:https://doi.org/10.1126/science.1168244

Fasel, N., Vauclair, C.-M., Lima, M. L., & Abrams, D. (2021). The relative importance of personal beliefs, meta-stereotypes and societal stereotypes of age for the wellbeing of older people. *Ageing & Society, 41*(12), 2768-2791. doi:10.1017/S0144686X20000537

Faul, F., Erdfelder, E., Lang, A. -G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:https://doi.org/10.3758/BF03193146

Feldman, R. S., & Saletsky, R. D. (1986). Nonverbal interracial teacher-student interaction. In R. S. Feldman (Ed.), *The Social Psychology of Education: Current Research and Theory* (pp. 115-131). Cambridge, UK: Cambridge University Press.

Finchilescu, G. (2010). Intergroup anxiety in interracial interaction: The role of prejudice and metastereotypes. *Journal of Social Issues, 66*(2), 334-351. doi: https://doi.org/10.1111/j.1540-4560.2010.01648.x

Finkelstein, L. M., King, E. B., & Voyles, E. C. (2015). Age metastereotyping and cross-age workplace interactions: A meta view of age stereotypes at work. *Work, Aging, and Retirement, 1*(1), 26-40. doi:https://doi.org/10.1093/workar/wau002

Finkelstein, L. M., Ryan, K. M., & King, E. B. (2013). What do the young (old) people think of me? Content and accuracy of age-based meta-stereotypes. *European Journal of Work and Organizational Psychology, 22*(6), 633-657. doi:10.1080/1359432X.2012.673279

Finkelstein, L. M., Voyles, E. C., Thomas, C. L., & Zacher, H. (2020). A Daily Diary Study of Responses to Age Meta-stereotypes. *Work, Aging and Retirement, 6*(1), 28-45. doi:https://doi.org/10.1093/workar/waz005

Finnigan, K. M., & Corker, K. S. (2016). Do performance avoidance goals moderate the effect of different types of stereotype threat on women's math performance? *Journal of Research in Personality, 63*, 36-43. doi:https://doi.org/10.1016/j.jrp.2016.05.009

Fisher, D., Fraser, B., & Cresswell, J. (1995). Using the "questionnaire on teacher interaction" in the professional development of teachers. *Australian Journal of Teacher Education, 20*(1), 8-18. doi:10.14221/ajte.1995v20n1.2

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-902. doi:10.1037/0022-3514.82.6.878

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of Social Psychology, 53*(1), 25-44. doi:https://doi.org/10.1016/j.jsp.2014.10.002

Flore, P. C., Mulder, J., & Wicherts, J. M. (2018). The influence of gender stereotypes threat on mathematics test scores of Dutch high school students: a registered report. *Comprehensive Results in Social Psychology, 3*(2), 140-174. doi:10.1080/23743603.2018.1559647

Forbes, C. E., Schmader, T., & Allen, J. J. (2008). The role of devaluing and discounting in performance monitoring: A neurophysiological study of minorities under threat. *Social Cognitive and Affective Neuroscience, 3*(3), 253-261. doi:https://doi.org/10.1093/scan/nsn012

Forster, J. J. (2001). Sample surveys: nonprobability sampling. In N. J. Smelser, & P. B. Baltes (Eds.), *nternational Encyclopedia of the Social & Behavioral Sciences* (pp. 13467-13470). Oxford: Elsevier. doi:http://dx.doi.org/10.1016/B0-08-043076-7/00499-X

Fowler, C., & Gasiorek, J. (2020). Implications of metastereotypes for attitudes toward intergenerational contact. *Group Processes & Intergroup Relations, 23*(1), 48-70. doi:10.1177/1368430217744032

Frey, F. E., & Tropp, L. R. (2006). Being seen as individuals versus as group members: Extending research on metaperception to intergroup contexts. *Personality and Social Psychology Review, 10*(3), 265-280. doi:10.1207/s15327957pspr1003_5

Frisch, J. U., Häusser, J. A., & Mojzisch, A. (2015). The Trier Social Stress Test as a paradigm to study how people respond to threat in social interactions. *Frontiers in Psychology, 6*, 14. doi:10.3389/fpsyg.2015.00014

Frymier, A. B., & Weser, B. (2001). The role of student predispositions on student expectations for instructor communication behaviour. *Communication Education, 50*(4), 314-326. doi:10.1080/03634520109379258

Fuller, B. F., Horii, Y., & Conner, D. A. (1992). Validity and reliability of nonverbal voice measures as indicators of stressor-provoked anxiety. *Research in Nursing and Health, 15*(5), 379-389.

Funk, C., & Parker, K. (2018, January 9). Women and Men in STEM Often at Odds Over Workplace Equity. *Pew Research Center*. Retrieved December 16, 2021, from https://www.pewresearch.org/social-trends/2018/01/09/women-and-men-in-stem-often-at-odds-over-workplace-equity/

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology, 4*(1), 1-26. doi:10.1080/14792779343000004

Galloway, A. (2005). Non-probability sampling. In K. Kempf-Leonard (Ed.), *Encyclopedia of Social Measurement* (pp. 859-864). Elsevier. doi:https://doi.org/10.1016/B0-12-369398-5/00382-0

Gerwing, T. G., Rash, J. A., Allen Gerwing, A. M., Bramble, B., & Landine, J. (2015). Perceptions and Incidence of Test Anxiety. *The Canadian Journal for the Scholarship of Teaching and Learning, 6*(3), Article 3. Retrieved from http://ir.lib.uwo.ca/cjsotl_rcacea/vol6/iss3/3

Gibbons, E. X. (1981). The social psychology of mental retardation: What's in a label? In S. S. Brehm, S. M. Kassin, & E. X. Gibbons (Eds.), *Developmental Social Psychology* (pp. 249-270). New York: Oxford University Press.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60*(4), 509-517. doi:https://doi.org/10.1037/0022-3514.60.4.509

Gilovich, T., Medvec, V., & Savitsky, K. (2000). The spotlight effectin social judgment: An egocentric bias in estimates of the salienceof one's own actions and appearance. *Journal of Personality and Social Psychology, 78*(2), 211-222. doi:10.1037//0022-3514.78.2.211

Gilrane, V. L., Wessel, J. L., Cheung, H. K., & King, E. B. (2019). The consequences of making the right impressions for STEM women: Metastereotypes, impression management, and supervisor ratings. *Archives of Scientific Psychology, 7*(1), 22-31. doi:https://doi.org/10.1037/arc0000065

Gladstone, J. R., & Cimpian, A. (2021). Which role models are effective for which students? A systematic review and four recommendations for maximizing the effectiveness of role models in STEM. *International Journal of STEM Education, 8*(59). doi:https://doi.org/10.1186/s40594-021-00315-x

Goff, P. A., Steele, C. M., & Davies, P. G. (2008). The space between us: Stereotype threat and distance in interracial contexts. *Journal of Personality and Social Psychology, 94*(1), 91-107. doi:https://doi.org/10.1037/0022-3514.94.1.91

Gomez, A. (2003). *Meta-stereotypes and intergroup relations.* Madrid: National University of Distance Education.

Gong, R. (2023, March 8). *Malaysia's gender gap in STEM education and employment.* Kuala Lumpur: Khazanah Research Institute. Retrieved from https://www.krinstitute.org/assets/contentMS/img/template/editor/2303%20Gender%20Gap%20in%20STEM.pdf

Gonzalez-Perez, S., Mateos de Cabo, R., & Sainz, M. (2020). Girls in STEM: Is it a female role-model thing? *Frontiers in Psychology, 11*(2204). doi:https://doi.org/10.3389/fpsyg.2020.02204

Gordijn, E. H. (2010). When thinking that you are fat makes you feel worthless: Activation and application of meta-stereotypes when appearance matter. *Social Cognition, 28*(1), 20-39. doi:https://doi.org/10.1521/soco.2010.28.1.20

Gordijn, E. H., & Boven, G. (2009). Loneliness among people with HIV in relation to locus of control and negative meta-stereotyping. *Basic and Applied Social Psychology, 31*(2), 109-116. doi:10.1080/01973530902880266

Gordijn, E. H., Finchilescu, G., Brix, L., Wijnants, N., & Koomen, W. (2008). The influence of prejudice, cultural stereotypes and meta-stereotypes on anticipated affect about interaction with a member of another ethnic group. *South African Journal of Psychology, 38*(4), 589-601. Retrieved from https://hdl.handle.net/10520/EJC98517

Gordijn, E. H., Vacher, L., & Kuppens, T. (2017). "To serve and protect" when expecting to be seen negatively: The relation between police officers' contact with citizens, meta-stereotyping, and work-related well-being. *Journal of Community & Applied Social Psychology, 27*(3), 253-268. doi:10.1002/casp.2310

Gotlib, I. H. (1984). Depression and general psychopathology in university students. *Journal of Abnormal Psychology, 93*(1), 19-30. doi:https://doi.org/10.1037/0021-843X.93.1.19

Goy, S. C., Wong, Y. L., Low, W. Y., Noor, S. N., Fazli-Khalaf, Z., Onyeneho, N., . . . GinikaUzoigwe, A. (2018). Swimming against the tide in STEM education and gender equality: A problem of recruitment or retention in Malaysia. *Studies in Higher Education, 43*(11), 1793-1809. doi:https://doi.org/10.1080/03075079.2016.1277383

Gregerson, T., & Horwitz, E. K. (2002). Language learning and perfectionism: Anxious and non-anxious language learners' reactions to their own oral performance. *The Modern Language Journal, 86*(4), 562-570. doi:10.1111/1540-4781.00161

Grotenhuis, M., Pelzer, B., Eisinga, R., Nieuwenhuis, R., Schmidt-Catran, A., & Konig, R. (2017). A novel method for modelling interaction between categorical variables. *International Journal of Public Health, 62*(3), 427-431. doi:https://doi.org/10.1007/s00038-016-0902-0

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience, 14*, 488-501. doi:https://doi.org/10.1038/nrn3524

Guyll, M., Madon, S., Prieto, L., & Scherr, K. C. (2010). The potential roles of self-fulfilling prophecies, stigma consciousness, and stereotype threat in linking Latino/a ethnicity and educational outcomes. *Journal of Social Issues, 66*(1), 113-130. doi:https://doi.org/10.1111/j.1540-4560.2009.01636.x

Halim, L., Lay, A. N., & Mohd Shahali, E. H. (2021). *STEM Education in Malaysia.* Routledge.

Halperin, K., Snyder, C. R., Shenkel, R. J., & Houston, B. K. (1976). Effects of source status and message favorability on acceptance of personality feedback. *Journal of Applied Psychology, 61*(1), 85-88. doi:10.1037/0021-9010.61.1.85

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*(1), 1-51. doi:https://doi.org/10.1111/j.1529-1006.2007.00032.x

Hand, S., Rice, L., & Greenlee, E. (2017). Exploring teachers' and students' gender role bias and students' confidence in STEM fields. *Social Psychology of Education, 20*, 929-945. doi:https://doi.org/10.1007/s11218-017-9408-8

Hanson, S. (2009). *Swimming against the tide: African American girls and science education.* Philadelphia: Temple University Press.

Harris, R. B., Grunspan, D. Z., Pelch, M. A., Fernandes, G., Ramirez, G., & Freeman, S. (2019). Can test anxiety interventions alleviate a gender gap in an undergraduate STEM course? *CBE - Life Sciences Education, 18*(3), 1-9. doi:https://doi.org/10.1187/cbe.18-05-0083

Hayes, A. F. (2013). *Introduction to Mediation, Moderation and Conditional Process Analysis.* New York: The Guildford Press.

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60*(6), 895-910. doi:10.1037/0022-3514.60.6.895

Heilman, M. E., & Chen, J. J. (2005). Same Behavior, Different Consequences: Reactions to Men's and Women's Altruistic Citizenship Behavior. *Journal of Applied Psychology, 90*(3), 431-441.

Hek, K., Direk, N., Newson, R. S., Hofman, A., Hoogendijk, W. J., Mulder, C. L., & Tiemeier, H. (2013). Anxiety disorders and salivary cortisol levels in older adults: a population-based study. *Psychoneuroendocrinology, 38*(2), 300-305. doi:https://doi.org/10.1016/j.psyneuen.2012.06.006

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58*(1), 47-77. doi:https://doi.org/10.3102/00346543058001047

Heng, Y. T., Wagner, D. T., Barnes, C. M., & Guarana, C. L. (2018). Archival research: Expanding the methodological toolkit in social psychology. *Journal of Experimental Social Psychology, 78*, 14-22. doi:https://doi.org/10.1016/j.jesp.2018.04.012

Henry, M. A., Shorter, S., Charkoudian, L. K., Heemstra, J. M., Le, B., & Corwin, L. A. (2021). Quantifying fear of failure in STEM: modifying and evaluating the Performance Failure Appraisal Inventory (PFAI) for use with STEM undergraduates. *International Journal of STEM Education, 8*, 43. doi:https://doi.org/10.1186/s40594-021-00300-4

Herrmann, S. D., Adelman, R. M., Bodford, J. E., Graudejus, O., Okun, M. A., & Kwan, V. S. (2016). The effects of a female role model on academic performance and persistence of women in STEM courses. *Basic and Applied Social Psychology, 38*(5), 258-268. doi:10.1080/01973533.2016.1209757

Hess, T. M., Auman, C., Colcombe, S., & Rahhal, T. (2003). The impact of stereotype threat on age differences in memory performance. *Journals of Gerontology Series B - Psychological Sciences and Social Sciences, 58*(1), P3-P11. doi:10.1093/geronb/58.1.P3

Hill, C., Cheong, K. C., Leong, Y. C., & Fernandez-Chung, R. (2014). TNE - Trans-national education or tensions between national and external? A case study of Malaysia. *Studies in Higher Education, 39*(6), 952-966. doi:10.1080/03075079.2012.754862

Hill, C., Corbett, C., & Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics.* Washington: American Association of University Women.

Hill, K. T., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can be done about it. *The Elementary School Journal, 85*(1), 105-126. doi:http://www.jstor.org/stable/1001622

Hirschman, C. (2016). Gender, the Status of Women, and Family Structure in Malaysia. *Malaysian Journal of Economic Studies, 53*(1), 33-50.

Hoferichter, F., Kulakow, S., & Hufenbach, M. C. (2021). Support from parents, peers, and teachers is differently associated with middle school students' well-being. *Frontiers in Psychology, 12*, 758226. doi:https://doi.org/10.3389/fpsyg.2021.758226

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88*(1), 28-54. doi:10.1002/sce.10106

Hogg, M. A., & Vaughan, G. M. (2011). *Social Psychology* (6th Edition ed.). Essex: Pearson.

Holzschneider, K., & Mulert, C. (2011). Neuroimaging in anxiety disorders. *Dialogues in Clinical Neuroscience, 13*(4), 453-461. doi:10.31887/DCNS.2011.13.4/kholzschneider

Hopkin, C. R., Hoyle, R. H., & Gottfredson, N. C. (2015). Maximizing the Yield of Small Samples in Prevention Research: A Review of General Strategies and Best Practices. *Prevention Science, 16*(7), 950-955. doi:10.1007/s11121-014-0542-7

Hopkins, N., Reicher, S., Harrison, K., Cassidy, C., Bull, R., & Levine, M. (2007). Helping to improve the group stereotype: On the strategic dimension of prosocial behaviour. *Personality and Social Psychology Bulletin, 33*(6), 776-788. doi:https://doi.org/10.1177/0146167207301023

Hornsey, M. J. (2008). Social Identity Theory and Self-categorization Theory: A Historical Review. *Social and Personality Psychology Compass, 2*(1), 204-222. doi:https://doi.org/10.1111/j.1751-9004.2007.00066.x

Hornsey, M. J., Oppes, T., & Svensson, A. (2002). "It's OK if we say it, but you can't": Responses to intergroup and intragroup criticism. *European Journal of Social Psychology, 32*(3), 293-307. doi:https://doi.org/10.1002/ejsp.90

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125-132. doi:10.1111/j.1540-4781.1986.tb05256.x

Hox, J., & McNeish, D. (2020). Small samples in multilevel modeling. In R. van de Schoot, & M. Miočević (Eds.), *Small sample size solutions: A guide for applied researchers and practitioners* (pp. 215-225). London: Routledge.

Hoyt, C. L. (2013). Inspirational or self-deflating: The role of self-efficacy in elite role model effectiveness. *Social Psychological and Personality Science, 4*(3), 290-298. doi:10.1177/1948550612455066

Huguet, P., & Regner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology, 99*(3), 545-560. doi:10.1037/0022-0663.99.3.545

Hultsch, D. F., MacDonald, S. W., Hunter, M. A., Maitland, S. B., & Dixon, R. A. (2002). Sampling and generalisability in developmental research: Comparison of random and convenience samples of older adults. *International Journal of Behavioral Development, 26*(4), 345-359. doi:10.1080/0165025014300024

Hurtado, S., Cabrera, N. L., Lin, M. H., Arellano, L., & Espinosa, L. L. (2009). Diversifying Science: Underrepresented Student Experiences in Structured Research Programs. *Research in Higher Education, 50*(2), 189-214. doi:https://doi.org/10.1007/s11162-008-9114-7

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139-155.

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparison of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly, 14*(3), 299-324. doi:https://doi.org/10.1111/j.1471-6402.1990.tb00022.x

Iervolino, A. C., Hines, M., Golombok, S. E., Rust, J., & Plomin, R. (2005). Genetic and environmental influences on sex-typed behavior during the preschool years. *Child Development, 76*(4), 826-840. doi:https://doi.org/10.1111/j.1467-8624.2005.00880.x

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*(5), 365-371. doi:10.1111/1467-9280.00272

Inzlicht, M., & Schmader, T. (Eds.). (2012). *Stereotype Threat: Theory, Process, and Application.* New York: Oxford University Press.

Ireland, D. T., Freeman, K. E., Winston-Proctor, C. E., DeLaine, K. D., Lowe, S. M., & Woodson, K. M. (2018). (Un)Hidden figures: A synthesis of research examining the intersectional experiences of black women and girls in STEM education. *Review of Research in Education, 42*(1), 226-254. doi:https://doi.org/10.3102/0091732X18759072

Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology, 17*(162). doi:10.1186/s12874-017-0442-1

Jefferson, J. W. (2001). Anxiety Disorder: More Than Just a Little Shyness. *Primary Care Companion to the Journal of Clinical Psychiatry, 3*(1), 4-9. doi:10.4088/pcc.v03n0102

Jensen, L. E., & Deemer, E. D. (2020). Attachment style and self-handicapping: The mediating role of the imposter phenomenon. *Social Psychology of Education, 23*, 1259-1276. doi:https://doi.org/10.1007/s11218-020-09580-0

Johnson, S. K., Murphy, S. E., Zewdie, S., & Reichard, R. J. (2008). The strong, sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organizational Behavior and Human Decision Processes, 106*(1), 39-60.

Johnson, W. F., Emde, R. N., Scherer, K. R., & Klinnert, M. D. (1986). Recognition of emotion from vocal cues. *Archives of General Psychiatry, 43*(3). doi:10.1001/archpsyc.1986.01800030098011

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis, & J. Haviland (Eds.), *The handbook of emotion* (pp. 220-235). New York: Guilford.

Jussim, L., & Honeycutt, N. (2023). Psychology as Science and as Propaganda. *Psychology Learning & Teaching, 22*(3), 237-244. doi:https://doi.org/10.1177/14757257231195347

Kamans, E., Gordijn, E. H., Oldenhuis, H., & Otten, S. (2009). What I think you see is what you get: Influence of prejudice on assimilation to negative meta-stereotypes among Dutch Moroccan teenagers. *European Journal of Social Psychology, 39*, 842-851. doi:10.1002/ejsp.593

Kamiloğlu, R. G., Fischer, A. H., & Sauter, D. A. (2020). Good vibrations: A review of vocal expressions of positive emotions. *Psychonomic Bulletin & Review, 27*, 237-265. doi:https://doi.org/10.3758/s13423-019-01701-x

Kanamori, Y., Xu, Y. J., Harrell-Williams, L. M., & Lightsey Jr, O. R. (2022). Intergroup contact, intergroup anxiety, and anti-transgender prejudice: An examination using structural equation modeling. *Archives of Sexual Behavior, 51*, 1943-1958. doi:https://doi.org/10.1007/s10508-021-02192-5

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology, 64*(5), 402-406. doi:10.4097/kjae.2013.64.5.402

Kang, S. K., & Chasteen, A. L. (2009). The moderating role of age-group identification and perceived threat on stereotype threat among older adults. *The International Journal of Aging and Human Development, 69*(3), 201-220. doi:10.2190/AG.69.3.c

Keller, J. (2002). Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles, 47*, 193-198. doi:https://doi.org/10.1023/A:1021003307511

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617-627. doi:10.1001/archpsyc.62.6.617

Khanna, M. M. (2015). Ungraded Pop Quizzes: Test-Enhanced Learning Without All the Anxiety. *Teaching of Psychology, 42*(2), 174-178. doi:https://doi.org/10.1177/0098628315573144

Kim, J., & Oe, T. (2009). Meta-stereotype as an indicator of intergroup attitude: How Japanese perceive they are viewed by Koreans. *Japanese Psychological Research, 51*(4), 279-285. doi:https://doi.org/10.1111/j.1468-5884.2009.00411.x

King, J. E., & Cooley, E. L. (1995). Achievement orientation and the impostor phenomenon among college students. *Contemporary Educational Psychology, 20*(3), 304-312. doi:https://doi.org/10.1006/ceps.1995.1019

King, N. S., & Pringle, R. M. (2019). Black girls speak STEM: Counterstories of informal and formal learning experiences. *Journal of Research in Science Teaching, 56*, 539-569. doi:https://doi.org/10.1002/tea.21513

Kirschbaum , C., Prike, K.-M., & Hellhammer, D. H. (1993). The “Trier Social Stress Test”–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1-2), 76-81. doi:10.1159/000119004

Kitano, K. (2001). Anxiety in the college Japanese language classroom. *The Modern Language Journal, 85*(4), 549-566. doi:10.1111/0026-7902.00125

Klein, O., & Azzi, A. E. (2001). The strategic confirmation of meta-stereotypes: How group members attempt to tailor an out-group's representation of themselves. *British Journal of Social Psychology, 40*(2), 279-293. doi:https://doi.org/10.1348/014466601164759

Koch, S. C., Müller, S. M., & Sieverding, M. (2008). Women and computers. Effects of stereotype threat on attribution of failure. *Computers & Education, 51*(4), 1795-1803. doi:https://doi.org/10.1016/j.compedu.2008.05.007

Kolligian, J., & Sternberg, R. J. (1991). Perceived fraudulence in young adults: Is there an "impostor syndrome"? *Journal of Personality Assessment, 56*(2), 308-326. doi:10.1207/s15327752jpa5602_10

Kray, L. J., Thompson, L., & Galinsky, A. (2001). Battle of the sexes: Gender stereotype confirmation and reactance in negotiations. *Journal of Personality and Social Psychology, 80*(6), 942-958. doi:10.1037//0022-3514.80.6.942

Krendl, A. C., Richeson, J. A., Kelley, W. M., & Heatherton, T. F. (2008). The negative consequences of threat: a functional magnetic resonance imaging investigation of the neural mechanisms underlying women's underperformance in math. *Psychological Science, 19*(2), 168-175. doi:10.1111/j.1467-9280.2008.02063.x

Kricorian, K., Seu, M., Lopez, D., Ureta, E., & Equils, O. (2020). Factors influencing participation of underrepresented students in STEM fields: Matched mentors and mindsets. *International Journal of STEM Education, 7*, 16. doi:https://doi.org/10.1186/s40594-020-00219-2

Krueger, J. (1996). Personal beliefs and cultural stereotypes about racial characteristics. *Journal of Personality and Social Pscyhology, 71*(3), 536-548.

Kumar, S., & Jagacinski, C. M. (2006). Imposters have goals too: The imposter phenomenon and its relationship to achievement goal theory. *Personality and Individual Differences, 40*(1), 147-157. doi:https://doi.org/10.1016/j.paid.2005.05.014

Laher, H., & Finchilescu, G. (2010). Meta-stereotypes: Intergroup anxiety in interracial contact among students. *New Voices in Psychology, 6*(1), 4-16. doi:https://doi.org/10.1016/j.neuroimage.2016.05.075

Lammers, J., Gordijn, E. H., & Otten, S. (2008). Looking through the eyes of the powerful. *Journal of Experimental Social Psychology, 44*(5), 1229-1238. doi:10.1016/j.jesp.2008.03.015

Lane, J. E. (2011). Importing private higher education: international branch campuses. *13*(4), 367-381. doi:10.1080/13876988.2011.583106

Latu, I. M., Mast, M. S., Lammers, J., & Bombari, D. (2013). Successful female leaders empower women's behavior in leadership tasks. *Journal of Experimental Social Psychology, 49*(3), 444-448.

Laurin, R. (2020). An investigation of the roles of group identification, perceived ability, and evaluative conditions in stereotype threat experiences. *Social and Personality Psychology, 123*(5), 1904-1918. doi:http://dx.doi.org/10.1177/0033294119884013

Lawner, E. K., Quinn, D. M., Camacho, G., Johnson, B. T., & Pan-Weisz, B. (2019). Ingroup role models and underrepresented students' performance and interest in STEM: A meta-analysis of lab and field studies. *Social Psychology of Education, 22*, 1169-1195. doi:https://doi.org/10.1007/s11218-019-09518-1

Lawton, W., & Katsomitros, A. (2012, January 12). *International branch campuses: data and developments.* Retrieved from Observatory of Borderless Higher Education: http://www.obhe.ac.uk/documents/view_details?id=894

Leary, M. R., Patton, K. M., Orlando, A. E., & Funk, W. W. (2000). The impostor phenomenon: Self-perceptions, reflected appraisals, and interpersonal strategies. *Journal of Personality, 68*(4), 725-756. doi:https://doi.org/10.1111/1467-6494.00114

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a Unifying Social Cognitive Theory of Career and Academic Interest, Choice, and Performance. *Journal of Vocational Behavior, 45*(1), 79-122. doi:https://doi.org/10.1006/jvbe.1994.1027

Leroy, N., Max, S., & Pansu, P. (2022). Is Emma or Liam the Top Scorer in Math? The Effects of a Counter-Stereotypical Role Model on Math Achievement. *Sex Roles, 86*, 587-603. doi:https://doi.org/10.1007/s11199-022-01289-y

Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *Journal of Abnormal Psychology, 102*(1), 133-144. doi:10.1037//0021-843x.102.1.133

Lewis, K. M. (2000). When leaders display emotion: How followers respond to negative emotional expression of male and female leaders. *Journal of Organizational Behavior, 21*(2), 221-234. doi:https://doi.org/10.1002/(SICI)1099-1379(200003)21:2%3C221::AID-JOB36%3E3.0.CO;2-0

Lindberg, S. M., Hyde, J. S., Peterson, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136*(6), 1123-1135. doi:10.1037/a0021276

Lippa, R. A. (2010). Gender differences in personality and interests: When, where, and why? *Social and Personality Psychology Compass, 4*(11), 1098-1110. doi:https://doi.org/10.1111/j.1751-9004.2010.00320.x

Little, R. J., & Rubin, D. B. (2020). *Statistical Analysis with Missing Data* (3rd ed.). Hoboken: Wiley.

Littleford, L. N., Wright, M. O., & Sayoc-Parial, M. (2005). White students' intergroup anxiety during same-race and interracial interactions: A multimethod approach. *Basic and Applied Social Psychology, 27*(1), 85-94. doi:https://doi.org/10.1207/s15324834basp2701_9

Liu, M. (2006). Anxiety in Chinese EFL students at different proficiency levels. *System, 34*(3), 301-316. doi:10.1016/j.system.2006.04.004

Liu, S., Liu, P., Wang, M., & Zhang, B. (2021). Effectiveness of stereotype threat interventions: A meta-analytic review. *Journal of Applied Pscyhology, 106*(6), 921-949. doi:10.1037/APL0000770

Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology, 73*(1), 91-103. doi:https://doi.org/10.1037/0022-3514.73.1.91

Ma, Y., Dang, Q., & Zhang, B. (2022). Effect of metastereotypes on female cooperative behaviour and the mediating role of intergroup anxiety. *Asian Journal of Social Psychology, 25*(2), 155-169. doi:https://doi.org/10.1111/ajsp.12474

Madon, S., Willard, J., Guyll, M., & Scherr, K. C. (2011). Self-fulfilling prophecies: Mechanisms, power, and links to social problems. *Social and Personality Psychology Compass, 5*(8), 578-590. doi:https://doi.org/10.1111/j.1751-9004.2011.00375.x

Mak, K. K., Kleitman, S., & Abbott, M. J. (2019). Impostor phenomenon measurement scales: A systematic review. *Frontiers in Psychology, 10*, 671. doi:https://doi.org/10.3389/fpsyg.2019.00671

Makarova, E., Aeschlimann, B., & Herzog, W. (2019). The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations. *Frontiers in Education, 4*(60). doi:10.3389/feduc.2019.00060

Malcom, S. M., Hall, P. Q., & Brown, J. W. (1976). *The Double Bind: The Price of Being a Minority Woman in Science.* Washington: AAAS Publication.

Marrelli, A. F. (2007). Unobtrusive Measures. *Performance Improvement, 46*(9), 43-47. doi:https://doi.org/10.1002/pfi.164

Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology, 44*(4), 645-657. doi:10.1348/014466604X17948

Marx, D. M., & Ko, S. J. (2011). Refocusing or Recycling? The Stereotype Inoculation Model and Its Relationship With Research on Ingroup Role Models. *Psychological Inquiry, 22*(4), 280-284. doi:10.1080/1047840X.2011.621050

Marx, D. M., & Ko, S. J. (2012). Superstars “like” me: The effect of role model similarity on performance under threat. *European Journal of Social Psychology, 42*, 807-812. doi:10.1002/ejsp.1907

Marx, D. M., & Roman, J. S. (2002). Female role models: Protecting women’s math test performance. *Personality and Social Psychology Bulletin, 28*(9), 1183-1193. doi:10.1177/01461672022812004

Marx, D. M., Ko, S. J., & Friedman, R. A. (2009). The “Obama effect”: How a salient role model reduces race-based performance differences. *Journal of Experimental Social Psychology, 45*(4), 953-956. doi:10.1016/j.jesp.2009.03.012

Marx, D. M., Stapel, D. A., & Muller, D. (2005). We Can Do It: The Interplay of Construal Orientation and Social Comparisons Under Threat. *Journal of Personality and Social Psychology, 88*(3), 432-446. doi:https://doi.org/10.1037/0022-3514.88.3.432

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon’s Mechanical Turk. *Behavior Research Methods, 44*, 1-23. doi:https://doi.org/10.3758/s13428-011-0124-6

Matera, M., Verde, S. D., & Meringolo, P. (2015). I like you more if I think you like me: The effect of metastereotypes on attitudes toward people with deafness. *Journal of Social Psychology, 155*(4), 381-394. doi:10.1080/00224545.2015.1018857

Maylor, U. (2009). 'They do not related to Black people like us': Black teachers as role models for Black pupils. *Journal of Education Policy, 24*(1), 1-21. doi:https://doi.org/10.1080/02680930802382946

McCambridge, J., de Bruin, M., & Witton, J. (2012). The effects of demand characteristics on research participant behaviours in non-laboratory settings: A systematic review. *PLoS One, 7*(6), e39116. doi:10.1371/journal.pone.0039116

McGlone, M. S., & Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *Journal of Applied Developmental Psychology, 27*(5), 486-493. doi:https://doi.org/10.1016/j.appdev.2006.06.003

McGregor, H. A., & Elliot, A. J. (2005). The shame of failure: Examining the link between fear of failure and shame. *Personality and Social Psychology Bulletin, 31*(2), 218-231. doi:10.1177/0146167204271420

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women’s mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology, 39*(1), 83-90. doi:10.1016/S0022-1031(02)00513-9

McIntyre, R. B., Paulson, R. M., & Lord, C. G. (2003). Alleviating women's mathematics stereotype threat through salience of group achievements. *Journal of Experimental Social Psychology, 39*, 83-90. doi:10.1016/S0022-1031(02)00513-9

McKown, C., & Weinstein, R. S. (2003). The development and consequences of stereotype consciousness in middle childhood. *Child Development, 74*(2), 498-515. doi:https://doi.org/10.1111/1467-8624.7402012

McLaughlin, L. A. (2005). Spotlighting: Emergent gender bias in undrgraduate engineering education. *Journal of Engineering Education, 94*(4), 373-381. doi:https://doi.org/10.1002/j.2168-9830.2005.tb00865.x

McLeod, S. A. (2012, January 14). *Experimental Psychology*. Retrieved from Simply Psychology: www.simplypsychology.org/experimental-method.html

McNeish, D. (2017). Missing data methods for arbitrary missingness with small samples. *Journal of Applied Statistics, 44*(1), 24-39. doi:10.1080/02664763.2016.1158246

Mendez, E., Gomez, A., & Tropp, L. R. (2007). When metaperceptions are affected by intergroup processes. *International Journal of Psychology and Psychological Therapy, 7*(2), 237-250.

Ministry of Education Malaysia. (2013). *Malaysia Education Blueprint 2013-2025.* Putrajaya: Kementerian Pendidikan Malaysia.

Ministry of Education Malaysia. (2015, February 24). *National Higher Education Strategic Plan*. Retrieved from Ministry of Education Malaysia Official Website: http://www.moe.gov.my/en/pelan-strategik-pengajian-tinggi-negara

Ministry of Education Malaysia. (2018). *Quick Facts 2018: Malaysia Educational Statistics.* Putrajaya: Ministry of Education Malaysia.

Ministry of Higher Education. (2013). *National Education Statistic: Higher Education Sector 2012.* Putrajaya: Government Printers. Retrieved from http://www.mohe.gov.my/web_statistik/Perangkaan_SPT_2012.pdf

Mok, K. H. (2011). The quest for regional hub of education: Growing heterarchies, organizational hybridization, and new governance in Singapore and Malaysia. *Journal of Education Policy, 26*(1), 61-81. doi:10.1080/02680939.2010.498900

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38*(4), 379-387. doi:https://doi.org/10.1037/0003-066X.38.4.379

More varsities coming. (2012, February 23). New Straits Times. Retrieved from http://www2.nst.com.my/top-news/more-varsities-coming-1.50776

Morgenroth, T., Ryan, M. K., & Peters, K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology, 19*(4), 465-483. doi:http://dx.doi.org/10.1037/gpr0000059465

Morin, A. (2017, August 12). *A new study says anxiety can help you perform better - but only if you know how to respond to it*. Retrieved November 17, 2022, from Forbes: https://www.forbes.com/sites/amymorin/2017/08/12/a-new-study-says-anxiety-can-help-you-perform-better-but-only-if-you-know-how-to-respond-to-it/?sh=24bc724920f6

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences, 109*(41), 16474-16479. doi:https://doi.org/10.1073/pnas.1211286109

Mrazek, M. D., Chin, J. M., Schmader, T., Hartson, K. A., Smallwood, J., & Schooler, J. W. (2011). Threatened to distraction: Mind-wandering as a consequence of stereotype threat. *Journal of Experimental Social Psychology, 47*(6), 1243-1248. doi:10.1016/j.jesp.2011.05.011

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52*, 431-462. doi:10.1007/BF02294365

Muzzatti, B., & Agnoli, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. *Developmental Psychology, 43*(3), 747-759. doi:10.1037/0012-1649.43.3.747

Nadler, J. T., & Clark, M. H. (2011). Stereotype threat: A meta-analysis comparing African Americans to Hispanic Americans. *Journal of Applied Social Psychology, 41*(4), 872-890. doi:https://doi.org/10.1111/j.1559-1816.2011.00739.x

Nair, S. (2021, May 24). Talent: Digital job vacancies almost tripled in Malaysia. *The Edge Malaysia*. Retrieved from https://www.theedgemarkets.com/article/talent-digital-job-vacancies-almost-tripled-malaysia

National Center for Science and Engineering Statistics. (2021). *Women, Minorities, and Persons with Disabilities in Science and Engineering.* Alexandria, VA: National Science Foundation.

National Institute of Mental Health. (2022, April). *Anxiety Disorders*. Retrieved November 17, 2022, from National Institute of Mental Health: https://www.nimh.nih.gov/health/topics/anxiety-disorders#part_2223

National Science Board. (2015). *Revisiting the STEM Workforce: A Companion to Science and Engineering Indicators 2014.* National Science Foundation.

National Science Foundation. (2022, January). *The State of U.S. Science and Engineering 2022*. Retrieved November 29, 2022, from Science & Engineering Indicators: https://ncses.nsf.gov/pubs/nsb20221/u-s-and-global-stem-education-and-labor-force

Neuville, E., & Croizet, J.-C. (2007). Can salience of gender identity impair math performance among 7-8 years old girls? The moderating role of task difficulty. *European Journal of Psychology of Education, 22*(3), 307-316. doi:https://doi.org/10.1007/BF03173428

Ng, S. B. (2019). *Exploring STEM competences for the 21st century.* UNESCO International Bureau of Education. Retrieved from https://unesdoc.unesco.org/in/documentViewer.xhtml?v=2.1.196&id=p::usmarcdef_0000368485&file=/in/rest/annotationSVC/DownloadWatermarkedAttachment/attach_import_c61a24e2-f736-421d-a16c-f4e85ea1b3bf%3F_%3D368485eng.pdf&locale=en&multi=true&ark=/ark:/48223/p

Ngo, B. (2006). Learning from the margins: The education of Southeast and South Asian Americans in context. *Race Ethnicity and Education, 9*(1), 51-65. doi:https://doi.org/10.1080/13613320500490721

Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314-1334. doi:10.1037/a0012702

Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314-1334. doi:10.1037/a0012702

Nhundu, T. J. (2007). Mitigating gender-typed occupational preferences of Zimbabwean primary school children: The use of biographical sketches and portrayals of female role models. *Sex Roles, 56*, 639-649. doi:https://doi.org/10.1007/s11199-007-9204-6

O'Brien, L. T., Hitti, A., Shaffer, E., Van Camp, A. R., Henry, D., & Gilbert, P. N. (2017). Improving Girls' Sense of Fit in Science: Increasing the Impact of Role Models. *Social Psychology and Personality Science, 8*(3), 301-309. doi:https://doi.org/10.1177/1948550616671997

Office of the Chief Scientist. (2014). *Science, Technology, Engineering and Mathematics: Australia's Future.* Canberra: Australian Government.

Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion Drives Attention: Detecting the Snake in the Grass. *Journal of Experimental Psychology: General, 130*(3), 466-478. doi:10.1037//0096-3445.130.3.466

Olsson, M., & Martiny, S. E. (2018). Does exposure to counterstereotypical role models influence girls' and women gender stereotypes. *9*, 2264. doi:10.3389/fpsyg.2018.02264

Olsson, M., & Martiny, S. E. (2018). Does exposure to counterstereotypical role models influence girls' and women's gender stereotypes and career choices? A review of social psychological research. *Frontiers in Psychology, 9*(2264). doi:https://doi.org/10.3389/fpsyg.2018.02264

Onghena, P. (2020). One by one: The design and analysis of replicated randomized single-case experiments. In R. van de Schoot, & M. Miočević (Eds.), *Small sample size solutions: A guide for applied researchers and practitioners* (pp. 87-101). London: Routledge. doi:https://doi.org/10.4324/9780429273872

O'Rourke, B. (2021, November 18). Growing gap in STEM supply and demand. *The Harvard Gazette*. Retrieved December 16, 2021, from https://news.harvard.edu/gazette/story/2021/11/increasing-access-and-opportunity-in-stem-crucial-say-experts/

Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology, 26*(3), 291-310. doi:https://doi.org/10.1006/ceps.2000.1052

Osborne, J. W. (2001). Testing stereotype threat: Does anxiety explain race and sex differences in achievement? *Contemporary Educational Psychology, 26*(3), 291-310. doi:https://doi.org/10.1006/ceps.2000.1052

Osborne, J. W. (2007). Linking stereotype threat and anxiety. *Educational Psychology, 27*(1), 135-154. doi:https://doi.org/10.1080/01443410601069929

Owuamalam, C. K., & Matos, A. S. (2019). Do egalitarians always help the disadvantaged more than the advantaged? Testing a value-norm conflict hypothesis in Malaysia. *Asian Journal of Social Psychology, 22*(2), 151-162. doi:https://doi.org/10.1111/ajsp.12351

Owuamalam, C. K., & Rubin, M. (2014). When do low status groups help high status groups? The moderating effects of ingroup indentification, audience group membership, and perceived reputational benefit. *Journal of Social and Political Psychology, 2*(1), 289-312. doi:https://doi.org/10.5964/jspp.v2i1.33

Owuamalam, C. K., & Rubin, M. (2017). Fuming with rage! Do members of low status groups signal anger more than members of high status groups? *Scandinavian Journal of Psychology, 58*(5), 458-467. doi:https://doi.org/10.1111/sjop.12388

Owuamalam, C. K., & Zagefka, H. (2011). Downplaying a compromised social image: The effect of metastereotype valence on social identification. *European Journal of Social Psychology, 41*(4), 528-537. doi:10.1002/ejsp.805

Owuamalam, C. K., & Zagefka, H. (2014). On the psychological barriers to the workplace: when and why metastereotyping undermines employability beliefs of women and ethnic minorities. *Cultural Diversity and Ethnic Minority Psychology, 20*(4), 521-528. doi:http://dx.doi.org/10.1037/a0037645

Owuamalam, C. K., Tarrant, M., Farrow, C. V., & Zagefka, H. (2013). The effect of metastereotyping on judgements of higher-status outgroups when reciprocity and social image improvement motives collide. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 45*(1), 12-23. doi:10.1037/a0030012

Owuamalam, C., & Zagefka, H. (2012). We'll never get past the glass ceiling! Meta-stereotyping, world-views and perceived relative group-worth. *British Journal of Psychology, 104*(4), 543-562. doi:https://doi.org/10.1111/bjop.12006

Oxford Reference. (n.d.). *non-reactive measure*. Retrieved December 14, 2022, from Oxford Reference: https://www.oxfordreference.com/display/10.1093/oi/authority.20110803100237876;jsessionid=F003F05AD32E14861979BF9BA0F4AE5A#:~:text=Any%20measure%20of%20behaviour%20that,is%20being%20observed%20or%20measured.

Ozdas, A., Shiavi, R. G., Silverman, S. E., Silverman, M. K., & Wilkes, D. M. (2004). Investigation of vocal jitter and glottal flow spectrum as possible cues for depression and near-term suicidal risk. *IEEE Transactions on Biomedical Engineering, 51*(9), 1530-1540.

Ozer, D. J., & Reise, S. P. (1994). Personality assessment. *Annual Review of Psychology, 45*, 357-388. doi:https://doi.org/10.1146/annurev.ps.45.020194.002041

Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. In A. M. Gallagher, & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 294-315). Cambridge: Cambridge University Press.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*(5), 411-419. doi:10.1017/S1930297500002205

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of Research Methods in Personality Psychology* (pp. 224-239). New York: Guilford.

Pelham, B. W., & Hardin, C. (2011). Opportunities for Sharpening the Stereotype Inoculation Model. *Psychological Inquiry, 22*(4), 285-290. doi:10.1080/1047840X.2011.621517

Peng, C. Y., Harwell, M., Liou, S. M., & Ehman, L. H. (2006). Advances in missing data methods and implications for educational research. In S. S. Sawilowsky (Ed.), *Real data analysis* (pp. 31-78). Charlotte, North Carolina: Information Age Pub.

Picho, K., Roriguez, A., & Finnie, L. (2013). Exploring the moderating role of context on the mathematics performance of females under stereotype threat: A meta-analysis. *The Journal of Social Psychology, 153*(3), 299-333. doi:10.1080/00224545.2012.737380

Pietri, E. S., Drawbaugh, M. L., Lewis, A. N., & Johnson, I. R. (2019). Who encourages Latina women to feel a sense of identity-safety in STEM environments? *Journal of Experimental Social Psychology, 84*, 103827. doi:https://doi.org/10.1016/j.jesp.2019.103827

Pigott, T. D. (2001). A Review of Methods for Missing Data. *Educational Research and Evaluation, 7*(4), 353-383. doi:10.1076/edre.7.4.353.8937

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology, 76*(1), 114-128.

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology, 76*(1), 114-128. doi:https://doi.org/10.1037/0022-3514.76.1.114

Pinel, E. C. (2004). You're just saying that because I'm a woman: Stigma consciousness and attributions to discrimination. *Self and Identity, 3*(1), 39-51. doi:https://doi.org/10.1080/13576500342000031

Planalp, S., DeFrancisco, V. L., & Rutherford, D. (1996). Varieties of cues to emotion in naturally occurring situations. *Cognition and Emotion, 10*(2), 137-153. doi:10.1080/026999396380303

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*(2), 243-281. doi:10.1111/j.1559-1816.2004.tb02547.x

Preacher, K. J. (2001, April). *Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]*. Retrieved from http://quantpsy.org

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*(4), 717-731.

Pringle, R. M., Brkich, K. M., Adams, T. L., West-Olatunii, C., & Archer-Banks, D. A. (2012). Factors Influencing Elementary Teachers' Positioning of African American Girls as Science and Mathematics Learners. *School Science and Mathematics, 112*(4), 217-229. doi:https://doi.org/10.1111/j.1949-8594.2012.00137.x

Purdie-Vaughns, V., Steele, C. A., Davies, P. G., Ditlmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology, 94*(4), 615-630. doi:10.1037/0022-3514.94.4.615

Rafidi, R. (2020, June). *Teaching in the new normal.* Retrieved November 24, 2020, from New Straits Times: https://www.nst.com.my/education/2020/06/599580/teaching-new-normal

Raghunathan, T. E. (2004). What do we do with missing data? Some options for analysis of incomplete data. *Annual Review of Public Health, 25*, 99-117. doi:10.1146/annurev.publhealth.25.102802.124410

Rahma, O. N., Putra, A. P., Rahmatillah, A., Putri, Y. K., Fajriaty, N. D., Ain, K., & Chai, R. (2022). Electrodermal Activity for Measuring Cognitive and Emotional Stress Level. *Journal of Medical Signals and Sensors, 12*(2), 155-162. doi:https://doi.org/10.4103/jmss.JMSS_78_20

Randsley de Moura, G., Leicht, C., Leite, A. C., Crisp, R. J., & Gocłowska, M. A. (2018). Leadership diversity: Effects of counterstereotypical thinking on the support for women leaders under uncertainty. *Journal of Social Issues, 74*(1), 165-183.

Raykov, T. (2011). Evaluation of convergent and discriminant validity with multitrait-multimethod correlations. *British Journal of Mathematical and Statistical Psychology, 64*(1), 38-52. doi:https://doi.org/10.1348/000711009X478616

Reidy, J., & Keogh, E. (1997). Testing the discriminant and convergent validity of the mood and anxiety symptoms questionnaire using a British sample. *Personality and Individual Differences, 23*(2), 337-344. doi:https://doi.org/10.1016/S0191-8869(97)00048-2

Reis, H. T., & Gosling, S. D. (2010). Social psychological methods outside the laboratory. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., Vol. 1, pp. 82-114). Wiley.

Reyes, A. (2017). *Language, identity, and stereotype among southeast asian american youth.* New York: Routledge. doi:https://doi.org/10.4324/9781315092027

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science, 14*(3), 287-290. doi:https://doi.org/10.1111/1467-9280.03437

Richmond, V. P., & McCroskey, J. C. (1998). *Communication apprehension, avoidance and effectiveness.* Boston: Allyn & Bacon.

Robins, R. W., Norem, J. K., & Cheek, J. M. (1999). Naturalizing the self. In L. A. Pervin, & O. P. John (Eds.), *Handbook of Personality: Theory and Research* (2nd ed., pp. 443-477). New York: Guilford Press.

Robnett, R. D. (2015). Gender Bias in STEM Fields: Variation in Prevalence and Links to STEM Self-Concept. *Psychology of Women Quarterly, 40*(1), 65-79. doi:10.1177/0361684315596162

Rocca, K. A. (2010). Student participation in the college classroom: An extended multidisciplinary literature review. *Communication Education, 59*(2), 185-213. doi:10.1080/03634520903505936

Ross, C. E., & Mirowsky, J. (2008). Age and the balance of emotions. *Social Science and Medicine, 66*(12), 2391-2400. doi:10.1016/j.socscimed.2008.01.048

Roth, J., Steffens, M. C., & Vignoles, V. L. (2018). Group membership, group change, and intergroup attitudes: A recategorization model based on cognitive consistency principles. *Frontiers in Pscyhology, 9*, 479. doi:10.3389/fpsyg.2018.00479

Rothgerber, H., & Wolsiefer, K. (2014). A naturalistic study of stereotype threat in young female chess players. *Group Processes & Intergroup Relations, 17*(1), 79-90. doi:https://doi.org/10.1177/1368430213490212

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*, 581-592. doi:10.1093/biomet/63.3.581

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* New York: Wiley.

Rueff-Lopes, R., Navarro, J., Caetano, A., & Silva, A. J. (2015). A Markov chain analysis of emotional exchange in voice-to-voice communication: Testing for the mimicry hypothesis of emotional contagion. *Human Communication Research, 41*(3), 412-434. doi:https://doi.org/10.1111/hcre.12051

Ryan, K. M., King, E. B., & Finkelstein, L. M. (2015). Younger workers' metastereotypes, workplace mood, attitudes, and behaviors. *Journal of Managerial Psychology, 30*(1), 54-70. doi:10.1108/JMP-07-2014-0215

Sackett, P. R., & Ryan, A. M. (2012). Concerns about generalizing stereotype threat research findings to operational high-stakes testing. In M. Inzlicht, & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application* (pp. 249-263). New York: Oxford University Press.

Sackett, P. R., Hardison, C. M., & Cullen, M. J. (2004). On Interpreting Stereotype Threat as Accounting for African American-White Differences on Cognitive Tests. *American Psychologist, 59*(1), 7-13. doi:https://doi.org/10.1037/0003-066X.59.1.7

Sagar, S. S., & Stoeber, J. (2009). Perfectionism, fear of failure, and affective responses to success and failure: The central role of fear of experiencing shame and embarrassment. *Journal of Sport and Exercise Psychology, 31*(5), 602-627. doi:10.1123/jsep.31.5.602

Salamone, J. D., & Correa, M. (2012). The Mysterious Motivational Functions of Mesolimbic Dopamine. *Neuron, 76*(3), 470-485. doi::https://doi.org/10.1016/j.neuron.2012.10.021

Samuel, T. S., Buttet, S., & Warner, J. (2022). “I Can Math, Too!”: Reducing Math Anxiety in STEM-Related Courses Using a Combined Mindfulness and Growth Mindset Approach (MAGMA) in the Classroom. *Community College Journal of Research and Practice*. doi:10.1080/10668926.2022.2050843

Schafer, J. L., & Olsen , M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research, 33*(4), 545-571. doi:10.1207/s15327906mbr3304_5

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99*(2), 143-165. doi:https://doi.org/10.1037/0033-2909.99.2.143

Scherer, K. R. (1995). Expressions of emotions in voice and music. *Journal of Voice, 9*(3), 235-248. doi:10.1016/S0892-1997(05)80231-0

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*, 227-256.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76-92. doi:10.1177/0022022101032001009

Schmader, T., & Beilock, S. (2012). An integration of processes that underlie stereotype threat. In M. Inzlicht, & T. Schmader (Eds.), *Stereotype threat: Theory, process, and application* (pp. 34-50). New York: Oxford University Press.

Schwarzer, R. (1990). Current trends in anxiety research. In P. D. Drenth, J. A. Sergeant, & R. J. Takens (Eds.), *European perspectives in psychology* (Vol. 2, pp. 225-244). Chichester: Wiley.

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4*(1), 27-41. doi:https://doi.org/10.1080/08917779108248762

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology, 39*(1), 68-74. doi:https://doi.org/10.1016/S0022-1031(02)00508-5

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: why undergraduates leave the sciences.* Boulder: Westview Press.

Shao, J., MacCallum, J. K., Zhang, Y., Sprecher, A., & Jiang, J. J. (2010). Acoustic analysis of the tremulous voice: Assessing the utility of the correlation dimension and perturbation parameters. *Journal of Communication Disorders, 43*(1), 35-44. doi:https://doi.org/10.1016/j.jcomdis.2009.09.001

Shapiro, J. R., & Neuberg, S. L. (2007). From stereotype threat to stereotype threats: implications of a multi-threat framework for causes, moderators, mediators, consequences, and interventions. *Personality and Social Psychology Review, 11*(2), 107-130. doi:https://doi.org/10.1177/108886830629479

Shapiro, J. R., & Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles, 66*, 175-183. doi:https://doi.org/10.1007/s11199-011-0051-0

Shelton, J. N., Richeson, J. A., & Salvatore, J. (2005). Expecting to be the target of prejudice: Implications for interethnic interactions. *Personality and Social Psychology Bulletin, 31*(9), 1189-1202. doi:https://doi.org/10.1177/0146167205274894

Shelton, J. N., Richeson, J. A., & Vorauer, J. D. (2006). Threatened identities and interethnic interactions. *European Review of Social Psychology, 17*(1), 321-358. doi:https://doi.org/10.1080/10463280601095240

Shelton, J. N., Richeson, J. A., & Vorauer, J. D. (2006). Threatened identities and interethnic interactions. *European Review of Social Psychology, 17*(1), 321-358. doi:10.1080/10463280601095240

Shewach, O. R., Sackett, P. R., & Quint, S. (2019). Stereotype threat effects in settings with features likely verses unlikely in operational test settings: A meta-analysis. *Journal of Applied Psychology, 104*(12), 1514-1534. doi:http://dx.doi.org/10.1037/apl0000420

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80-83. doi:https://doi.org/10.1111/1467-9280.00111

Shin, J. E., Levy, S. R., & London, B. (2016). Effects of role model exposure on STEM and non-STEM student engagement. *Journal of Applied Social Psychology, 46*(7), 410-427. doi:https://doi.org/10.1111/jasp.12371

Shin, J. L., Levy, S. R., & London, B. (2016). Effects of role model exposure on STEM and non-STEM student engagement. *Journal of Applied Social Psychology, 46*(7), 410-427. doi:https://doi.org/10.1111/jasp.12371

Sigelman, L., & Tuch, S. A. (1997). Metastereotypes: Blacks' perceptions of Whites' stereotypes of Blacks. *The Public Opinion Quarterly, 61*(1), 87-101. Retrieved from http://www.jstor.org/stable/2749513

Sigelman, L., & Tuch, S. A. (1997). Metastereotypes: Blacks' perceptions of Whites' stereotypes of Blacks. *The Public Opinion Quarterly, 61*(1), 87-101. Retrieved from https://www.jstor.org/stable/2749513

Silber-Varod, V., Kreiner, H., Lovett, R., Levi-Belz, Y., & Amir, N. (2016). Do social anxiety individuals hesitate more? The prosodic profile of hesitation disfluencies in Social Anxiety Disorder individuals. *Proceedings of Speech Prosody 2016*, (pp. 1211-1215). Boston. doi:10.21437/SpeechProsody.2016-249

Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science, 9*(1), 76-80. doi:https://doi.org/10.1177/1745691613514755

Šimunović, M., & Babarović , T. (2021). The role of parental socializing behaviors in two domains of student STEM career interest. *Research in Science Education, 51*, 1055-1071. doi:https://doi.org/10.1007/s11165-020-09938-6

Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology, 77*(5), 885-904. doi:https://doi.org/10.1037/0022-3514.77.5.885

Sinharay, S., Stern, H. S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods, 6*(4), 317-329. doi:10.1037/1082-989X.6.4.317

Sirat, M. (2005). *Transnational higher education in Malaysia: balancing benefits and concerns through regulations (Working Paper).* Retrieved from Department of Higher Education: jpt.mohe.gov.my

Smerdon, D., Hu, H., McLennan, A., Von Hippel, W., & Albrecht, S. (2020). Female chess players show typical stereotype-threat effects: Commentary on Stafford (2018). *Psychological Science, 31*(6), 756-759. doi:https://doi.org/10.1177/0956797620924051

Smith, M. K., & Dundes, L. (2008). The Implications of Gender Stereotypes for the Dentist-Patient Relationship. *Journal of Dental Education, 72*(5), 562-570. doi:https://doi.org/10.1002/j.0022-0337.2008.72.5.tb04520.x

Snyder, C. R., Lassegard, M. A., & Ford, C. E. (1986). Distancing after group success and failure: Basking in reflected glory and cutting off reflected failure. *Journal of Personality and Social Psychology, 51*(2), 382-388. doi:10.1037/0022-3514.51.2.382

Solbes-Canales, I., Valverde-Montesino, S., & Herranz-Hernandez, P. (2020). Socialization of gender stereotypes related to attributes and professions among young Spanish school-aged children. *Frontiers in Psychology, 11*, 609. doi:https://doi.org/10.3389/fpsyg.2020.00609

Sonnert, G., Fox, M. F., & Adkins, K. (2007). Undergraduate Women in Science and Engineering: Effects of Faculty, Fields, and Institutions Over Time. *Social Science Quarterly, 88*(5), 1333-1356. doi:10.1111/j.1540-6237.2007.00505.x

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. doi:10.1006/jesp.1998.1373

Spielberger, C. D. (1983). *Manual for the State-Trait Inventory STAI (Form Y).* Palo Alto: Mind Garden.

Staetsky, D. L. (2019). Can Convenience Samples be Trusted? Lessons From the Survey of Jews in Europe, 2012. *Contemporary Jewry, 39*, 115-153. doi:https://doi.org/10.1007/s12397-019-09280-8

Stafford, T. (2018). Female chess players outperform expectations when playing men. *Psychological Science, 29*(3), 429-436. doi:https://doi.org/10.1177/0956797617736887

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behaviour Research Methods, Instruments & Computers, 31*(1), 137-149.

Statistics, B. o. (n.d.). *Occupational outlook handbook: Fastest growing occupations.* Retrieved September 4, 2019, from https://www.bls.gov/ooh/fastest-growing.htm

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629. doi:10.1037/0003-066X.52.6.613

Steele, C. M. (2010). *Whistling Vivaldi: And other clues to how stereotypes affect us.* New York: Norton.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797-811. doi:10.1037/0022-3514.69.5.797

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology, 34*, 379-440. doi:https://doi.org/10.1016/S0065-2601(02)80009-0

Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology, 34*, 379-440. doi:doi:10.1016/S0065-2601(02)80009-0

Stephan, W. G. (2014). Intergroup anxiety: Theory, research, and practice. *18*(3), 239-255. doi:https://doi.org/10.1177/1088868314530518

Stephan, W. G. (2014). Intergroup anxiety: Theory, research, and practice. *Personality and Social Psychology Review, 18*(3), 239-255. doi:https://doi.org/10.1177/1088868314530518

Stephan, W. G., & Stephan, C. W. (1985). Intergroup Anxiety. *Journal of Social Issues, 41*(3), 157-175. doi:https://doi.org/10.1111/j.1540-4560.1985.tb01134.x

Stewart-Williams, S., & Halsey, L. G. (2021). Men, women and STEM: Why the differences and what should be done? *European Journal of Personality , 35*(1), 2021. doi:https://doi.org/10.1177/0890207020962326

Stockard, J., Rohlfing, C. M., & Richmond, G. L. (2021). Equity for women and underrepresented minorities in STEM: Graduate experiences and career plans in chemistry. *Proceedings of the National Academy of Sciences, 118*(4), e2020508118. doi:https://doi.org/10.1073/pnas.2020508118

Stoet, G., & Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Review of General Psychology, 16*(1), 93-102. doi:10.1037/a0026617

Stoet, G., & Geary, D. C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence, 48*, 137-151. doi:https://doi.org/10.1016/j.intell.2014.11.006

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology, 77*(6), 1213-1227. doi:10.1037/0022-3514.77.6.1213

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology, 100*(2), 255-270. doi:10.1037/a0021385

Strack, J., Lopes, P., Esteves, F., & Fernandez-Berrocal, P. (2017). Must we suffer to succeed? *Individual Differences, 38*(2), 113-124. doi:https://doi.org/10.1027/1614-0001/a000228

Stroud, C., & Wee, L. (2006). Anxiety and identity in the language classroom. *RELC Journal, 37*(3), 299-307. doi:https://doi.org/10.1177/0033688206071311

Sultana Alam, M., Sajid, S., Kok, J., Rahman, M., & Amin, A. (2021). Factors that influence high school female students' intentions to pursue science, technology, engineering and mathematics (STEM) education in Malaysia. *Pertanika Journals Social Sciences & Humanities, 29*(2), 839-867. doi:https://doi.org/10.47836/pjssh.29.2.06

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.). Needham Heights: Allyn & Bacon.

Taherdoost, H. (2016). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research in Management, 5*(3), 28-36. doi:https://dx.doi.org/10.2139/ssrn.3205040

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin, & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey: Brooks/Cole.

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel, & W. Austin (Eds.), *Psychology of intergroup relations* (pp. 7-24). Chicago: Nelson Hall.

Tangney, J. P. (1991). Moral affect: The good, the bad, and the ugly. *Journal of Personality and Social Psychology, 61*(4), 598-607. doi:https://doi.org/10.1002/ejsp.809

Tangney, J. P. (2002). Perfectionism and the self-conscious emotions: Shame, guilt, embarrassment and pride. In G. L. Flett, & P. L. Hewitt (Eds.), *Perfectionism: Theory, research, and treatment* (pp. 199-216). Washington DC: American Psychological Association.

Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1999). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology, 70*(6), 1256-1269. doi:10.1037/0022-3514.70.6.1256

Taylor, C. A., Lord, C. G., McIntyre, R. B., & Paulson, R. M. (2011). The Hillary Clinton effect: When the same role model inspires or fails to inspire improved performance under stereotype threat. *Group Processes & Intergroup Relations, 14*(4), 447-459. doi:https://doi.org/10.1177/1368430210382680

Tempel, T., & Neumann, R. (2014). Stereotype threat, test anxiety, and mathematics performance. *Social Psychology of Education, 17*, 491-501. doi:https://doi.org/10.1007/s11218-014-9263-9

Tham, S. Y. (2013). Internationalizing higher education in Malaysia: government policies and university's response. *Journal of studies in international education, 17*(5), 648-662. doi:10.1177/1028315313476954

The British Psychological Society. (2014). *Code of Human Research Ethics.* Leicester: The British Psychological Society. Retrieved from http://www.bps.org.uk/system/files/Public%20files/inf180_web.pdf

The jamovi project. (2019). jamovi (Version 1.0.5.0). Retrieved from https://www.jamovi.org

Titze, I. R. (1991). A model of neurologic sources of aperiodicity in vocal fold vibration. *Journal of Speech, Language and Hearing Research, 34*, 460-472.

Titze, I. R. (1994). Toward standards in acoustic analysis of voice. *Journal of Voice, 8*(1), 1-7. doi:10.1016/S0892-1997(05)80313-3

Tomeh, D. H., & Sackett, P. R. (2022). On the Continued Misinterpretation of Stereotype Threat as Accounting for Black-White Differences on Cognitive Tests. *Personnel Assessment and Decisions, 8*(1), 1-14. doi:https://doi.org/10.25035/pad.2022.01.001

Trawalter, S., Adam, E. K., Chase-Lansdale, P. L., & Richeson, J. A. (2012). Concerns about appearing prejudiced get under the skin: Stress responses to interracial contact in the moment and across time. *Journal of Experimental Social Psychology, 48*(3), 682-693. doi:10.1016/j.jesp.2011.12.003

Tse, C., Logel, C., & Spencer, S. J. (2011). Stereotype Inoculation: An Elegant and Practical Model. *Psychological Inquiry, 22*(4), 296-298. doi:10.1080/1047840X.2011.623266

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the Social Group: A self-categorization Theory.* New York: Blackwell.

U.S. Bureau of Labor Statistics. (2022, September 8). *Fastest Growing Occupations*. Retrieved from Occupational Outlook Handbook: https://www.bls.gov/ooh/fastest-growing.htm

UNESCO. (2012). *World Atlas of Gender Equality in Education.* Retrieved from http://www.uis.unesco.org/Education/Documents/unesco-world-atlas-gender-education-2012.pdf

UNESCO. (2017). *Cracking the code: Girls' and women's education in STEM.* UNESCO.

Unraveling the double bind: Women of color in STEM. (2011). *Harvard Educational Review, 81*(2), 157-161. Retrieved from https://www.proquest.com/scholarly-journals/unraveling-double-bind-women-color-stem/docview/874155305/se-2

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin, 134*(3), 383-403. doi:10.1037/0033-2909.134.3.383

van Lissa, C. J. (2020). Small sample meta-analyses: Exploring heterogeneity using MetaForest. In R. van de Schoot, & M. Miočevic (Eds.), *Small sample size solutions: A guide for applied researchers and practitioners* (pp. 186-202). London: Routledge.

Van Mier, H. I., Schleepen, T. M., & Van den Berg, F. C. (2019). Gender differences regarding the impact of math anxiety on arithmetic performance in second and fourth graders. *Frontiers in Psychology, 9*, 2690. doi:https://doi.org/10.3389/fpsyg.2018.02690

Van Zomeren, M., Fischer, A. H., & Spears, R. (2007). Testing the Limits of Tolerance:How Intergroup Anxiety Amplifies Negative and Offensive Responses to Out-Group-Initiated Contact. *Personality and Social Psychology Bulletin, 33*(12), 1686-1699. doi:https://doi.org/10.1177/0146167207307485

Vezzali, L. (2017). Valence matters: Positive metastereotypes and interethnic interactions. *Journal of Social Psychology, 157*(2), 247-261. doi:10.1080/00224545.2016.1208140

Villwock, J. A., Sobin, L. B., Koester, L. A., & Harris, T. M. (2016). Impostor syndrome and burnout among American medical students: A pilot study. *International Journal of Medical Education, 7*, 364-369. doi:10.5116/ijme.5801.eac4

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, 483-493. doi:https://doi.org/10.1016/j.jad.2017.11.048

von Hippel, P. T. (2016). New Confidence Intervals and Bias Comparisons Show That Maximum Likelihood Can Beat Multiple Imputation in Small Samples. *Structural Equation Modeling: A Multidisciplinary Journal , 23*(3), 422-437. doi:10.1080/10705511.2015.1047931

Vorauer, J. D. (2006). An information search model of evaluative concerns in intergroup interaction. *Psychological Review, 113*(4), 862-886. doi:10.1037/0033-295X.113.4.862

Vorauer, J. D., Hunter , A. J., Main, K. J., & Roy, S. A. (2000). Meta-stereotype activation: Evidence from indirect measures for specific evaluative concerns experienced by members of dominant groups in intergroup interaction. *Journal of Personality and Social Psychology, 78*(4), 690-707. doi:10.1037/0022-3514.78.4.690

Vorauer, J. D., Hunter, A. J., Main, K. J., & Roy, S. A. (2000). Meta-stereotype activation: Evidence from indirect measures for specific evaluative concerns experienced by members of dominant groups in intergroup interaction. *Journal of Personality and Social Psychology, 78*(4), 690-707. doi:10.1037/0022-3514.78.4.690

Vorauer, J. D., Main, K. J., & O'Connell, G. B. (1998). How do individuals expect to be viewed by members of lower status groups? Content and implications of meta-stereotypes. *Journal of Personality and Social Psychology, 75*(4), 917-937. doi:10.1037/0022-3514.75.4.917

Vors, O., Marqueste, T., & Mascret, N. (2018). The Trier Social Stress Test and the Trier Social Stress Test for groups: Qualitative investigations. *PLoS ONE, 13*(4), e0195722. doi:10.1371/journal.pone.0195722

Voyles, E., Finkelstein, L., & King, E. (2014). A tale of two theories: Stereotype threat and metastereotypes. *Industrial and Organizational Psychology, 7*(3), 419-422. doi:10.1111/iops.12171

Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39*(5), 456-467. doi:10.1016/S0022-1031(03)00019-2

Walton, G. M., & Spencer, S. J. (2009). Latent ability: Grades and test scores systematically underestimate the intellectual ability of negatively stereotyped students. *Psychological Science, 20*(9), 1132-1139. doi:10.1111/j.1467-9280.2009.02417.x

Walton, G. M., Murphy, M. C., & Ryan, A. M. (2015). Stereotype threat in organizations: Implications for equity and performance. *Annual Review of Organizational Psychology and Organizational Behavior, 2*, 523-550. doi:https://doi.org/10.1146/annurev-orgpsych-032414-111322

Weaver, R. R., & Qi, J. (2005). Classroom organization and participation: College students' perceptions. *The Journal of Higher Education, 76*(5), 570-601. doi:10.1353/jhe.2005.0038

Weem, G. H., & Onwuegbuzie, A. J. (2001). The impact of midpoint responses and reverse coding on survey data. *Measurement and Evaluation in Counseling and Development, 34*(3), 166-176.

Wei, T. E. (2012). Sticks, stones, words, and broken bones: New field and lab evidence on stereotype threat. *Educational Evaluation and Policy Analysis, 34*(4), 465-488. doi:10.3102/0162373712452629

Wilson, T. D., Aronson, E., & Carlsmith, K. (2010). The Art of Laboratory Experimentation. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., Vol. 1, pp. 51-81). Hoboken: Wiley.

Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2019). Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency. *Lancet, 393*(10171), 531-540. doi:10.1016/S0140-6736(18)32611-4

Witt-Rose, D. L. (2003). *Student self-efficacy in college science: an investigation of gender, age, and academic achievement.* University of Wisconsin-Stout.

Woodcock, A., Hernandez, P. R., Estrada, M., & Schultz, W. P. (2012). The consequences of chronic stereotype threat: Domain disidentification and abandonment. *Journal of Personality and Social Psychology, 103*(4), 635-646. doi:https://doi.org/10.1037/a0029120

Wu, L., & Jing, W. (2013). Leadership hurdles. *Nature, 493*, 125-126. doi:https://doi.org/10.1038/nj7430-125a

Xu, L., Sun, L., Li, J., Zhao, H., & He, W. (2021). Metastereotypes impairing doctor-patient relations: The roles of intergroup anxiety and patient trust. *PsyCh Journal, 10*(2), 275-282. doi:https://doi.org/10.1002/pchj.408

Yerkes, R. M., & Dodson, J. D. (1908). The relationship of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459-482. doi:https://doi.org/10.1002/cne.920180503

Young , J. L., Young, J. R., & Ford, D. Y. (2017). Standing in the gaps: Examining the effects of early gifted education on Black girl achievement in STEM. *Journal of Advanced Academics, 28*(4), 290-312. doi:https://doi.org/10.1177/1932202X17730549

Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal, 75*(4), 426-439. doi:10.1111/j.1540-4781.1991.tb05378.x

Young, D. M., Rudman, L. A., Buettner, H. M., & McLean, M. C. (2013). The Influence of Female Role Models on Women's Implicit Science Cognitions. *Psychology of Women Quarterly, 37*(3), 283-292.

Young, J. L., Ero-Tolliver, I., Young, J. R., & Ford, D. Y. (2017). Maximizing Opportunities to Enroll in Advanced High School Science Courses: Examining the Scientific Dispositions of Black Girls. *Journal of Urban Learning, Teaching, and Research, 13*, 174-183.

Young, J. L., Young, J. R., & Ford, D. Y. (2017). Standing in the gaps: Examining the effects of early gifted education on Black girl achievement in STEM. *Journal of Advanced Academics, 28*(4), 290-312. doi:https://doi.org/10.1177/1932202X17730549

Yung, Y.-F., & Chan, W. (1999). Statistical analyses using bootstrapping: Concepts and implementation. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 82-108). Thousand Oaks: SAGE.

Yzerbyt, V. Y., Judd, C. M., & Muller, D. (2009). How do they see us? The vicissitudes of meta-perception in intergroup relations. In S. Demoulin, J. Leyens, & J. F. Dovidio (Eds.), *Intergroup misunderstandings: Impact of divergent social realities* (pp. 63-84). New York: Psychology Press.

Zhao, K. (2021). Sample representation in the social sciences. *Synthese, 198*, 9097-9115. doi:https://doi.org/10.1007/s11229-020-02621-3

Appendix A: Questionnaire on Relative Salience of Different Sources of Fear among Malaysian Students

Section 1

Overall, how confident are you when working in groups during a lecture/session?

Please circle only one of the following:

Not confident at all Very confident

-3	-2	-1	0	1	2	3

On the whole, how confident are you when answering questions in class?

Please circle only one of the following:

Not confident at all Very confident

-3	-2	-1	0	1	2	3

Section 2

Read each statement below and indicate the extent to which you agree or disagree with them in relation to academics.

1	2	3	4	5	6
Strongly Disagree					**Strongly Agree**

Fear of Experiencing Shame and Embarrassment

When I am making mistakes, I get down on myself easily.

When I am making mistakes, it is embarrassing if my TEACHERS/FRIENDS/PARENTS or GUARDIANS are there to see it.

When I am making mistakes, I believe that my TEACHERS/FRIENDS/PARENTS or GUARDIANS know I am making mistakes.

When I am making mistakes, I believe that my doubters feel that they were right about me.

When I am making mistakes, I worry about what my TEACHERS/FRIENDS/PARENTS or GUARDIANS think about me.

When I am making mistakes, I worry that my TEACHERS/FRIENDS/PARENTS or GUARDIANS may think I am not trying.

Fear of Devaluing One's Self-Estimate

When I am making mistakes, it is often because I am not smart enough to get things right.

I begin to lose confidence when I don't get things right.

When I am making mistakes, I blame my lack of talent.

When I am making mistakes, I am more afraid to try new things.

When I am making mistakes, I am afraid that I might not have enough talent.

When I am making mistakes, I feel far from perfect.

When I am making mistakes, I hate the fact that I am not in control of the outcome.

Fear of Having an Uncertain Future

When I am making mistakes, my future seems uncertain.

When I am making mistakes, I believe that my future plans will change.

When I am making mistakes, it upsets my "plan" for the future.

When I am making mistakes, I am not worried about it affecting my future plans. (Reverse)

Fear of Important Others Losing Interest

When I am not getting things right, my TEACHERS/FRIENDS/PARENTS or GUARDIANS are less interested in me.

When I am not getting things right, my TEACHERS/FRIENDS/PARENTS or GUARDIANS seem to want to help me less.

When I am not getting things right, my TEACHERS/FRIENDS/PARENTS or GUARDIANS tend to leave me alone.

When I am not getting things right, my TEACHERS/FRIENDS/PARENTS or GUARDIANS are not interested in me anymore.

When I am not getting things right, my value decreases for my TEACHERS/FRIENDS/PARENTS or GUARDIANS.

Fear of Upsetting Important Others

When I am making mistakes, it upsets my TEACHERS/FRIENDS/PARENTS or GUARDIANS.

When I am making mistakes, I expect to be criticized by my TEACHERS/FRIENDS/PARENTS or GUARDIANS.

When I am making mistakes, I lose the trust of my TEACHERS/FRIENDS/PARENTS or GUARDIANS.

When I am making mistakes, my TEACHERS/FRIENDS/PARENTS or GUARDIANS are not happy.

When I am making mistakes, my TEACHERS/FRIENDS/PARENTS or GUARDIANS are disappointed.

Appendix B: Multiple Choice Question Bank

Note: Answers are highlighted in bold and questions 6, 8, 11, 13, 16, 17, 23, 26, 27 and 29 were selected for the tests conducted after the workshop.

1. According to the National Health Service (NHS) website, what is dyspepsia?
    a. Lower abdominal discomfort or pain.
    b. Upper abdominal discomfort or pain.
    c. Lower abdominal swelling and pressure.
    d. Upper abdominal swelling and pressure.
2. What does the prefix *dys-* mean?
    a. Pleasurable
    b. Good
    c. Bad
    d. Useless
3. What does the suffix *-pepsia* mean?
    a. A state of the digestion
    b. A condition of the intestines
    c. A problem with the stomach
    d. An issue with the abdomen
4. What does the word *dyspepsia* mean?
    a. Something's good about the digestive system.
    b. Something's not so right with the digestive system.
    c. Something's interesting about the digestive system.
    d. All is well with the digestive system.

5. According to the pharmacological point of view, what is the major problem of dyspepsia?
    a. Dizziness
    b. Heart aches
    c. Difficulty in breathing
    d. Acidity
6. What are the two chemical basis of treatment for dyspepsia?
        i. Reduced acidity in the stomach
        ii. Prevention of acid reflux
        iii. Reduced stomach discomfort
        iv. Maintained digestive functions
    a. i
    b. i and ii
    c. i, ii and iii
    d. all of the above
7. To reduce acidity in the stomach, what are the two modes of action?
    a. Physicochemical and pharmacological
    b. Physiochemical and photocological
    c. Proteomical and photochemical
    d. Thermochemical and pharmacological
8. What do antacids do?
    a. Decrease the gastric pH
    b. Neutralise the gastric acid
    c. Prevent the production of acid
    d. Maintain the gastric pH

9. What do anti-foaming agents do?
    a. Prevents the formation of foaming agents
    b. Prevents the formation of foam
    c. Increase the gastric pH
    d. Anti-reflux formation
10. What do raft-forming agents do?
    a. Anti-reflux formation
    b. Prevents the formation of rafts
    c. Decrease the gastric pH
    d. Neutralise the gastric acid
11. Which of the following is not the content of antacid products?
    a. Sodium carbonate
    b. Calcium carbonate
    c. Magnesium carbonate
    d. Potassium carbonate
12. What do you get when acid reacts with a base?
    a. Base and water
    b. Salt and water
    c. Acid and water
    d. Gas and water
13. Based on the word dysfunctional, what do you think the prefix *dys-* means?
    a. Sufficient
    b. Bad
    c. Good
    d. Poor

14. What does neutralization mean?

    a. Producing pH 7

    b. Neutralizing the excess acid

    c. Obtaining a higher pH level

    d. Producing more base

15. The following are the active ingredients of raft-forming agents, except

    a. Sodium alginate

    b. Sodium bicarbonate

    c. Calcium carbonate

    d. Magnesium arginate

16. What are the two main ingredients of most raft-forming agents?

        i. An alginate

        ii. A hydroxide

        iii. A gas forming agent

        iv. A gas reducing agent

    a. i and ii

    b. i and iii

    c. i and iv

    d. i, ii and iii

17. According to British Pharmacopoeia polyuronic acids are composed of residues of

    a. D-mannuronic and L-guluronic acids

    b. L-mannuronic and D-guluronic acids

    c. D-gannuronic and L-muluronic acids

    d. L-gannuronic and D-muluronic acids

18. Alginic acid is mainly used for

    a. Treatment of intestinal reflux disease

    b. Treatment of gaseous reflux disease

    c. Treatment of salivary gland disease

    d. Treatment of gastro-oesophageal reflux disease

19. What is the pK_a value for alginic acid?

    a. 2.5

    b. 3.0

    c. 3.5

    d. 4.0

20. The raft that floats on the stomach content gains its buoyancy due to

    a. 0_2 gas

    b. N_2 gas

    c. CO gas

    d. CO_2 gas

21. Based on the name, what do you think anti-foaming agents do?

    a. Break foam

    b. Prevent agents from forming

    c. Stop usage of soap

    d. Avoid foaming agents

22. What is an active ingredient of anti-foaming agents?

    a. Silicone

    b. Dimeticone

    c. Simeticone

    d. Ice-cream cone

23. The active ingredient of anti-foaming agents is believed to function by

    a. Lowering the surface tension of small gas bubbles so they form larger bubbles.

    b. Lowering the surface tension of large gas bubbles so they form smaller bubbles.

    c. Increasing the surface tension of small gas bubbles so they form larger bubbles.

    d. Increasing the surface tension of large gas bubbles so they form smaller bubbles.

24. The H2-receptor antagonists function by

    a. Blocking the acid pumps on stomach wall.

    b. Blocking the action of histamine H2 receptors in the stomach.

    c. Blocking the acid pump's H2 receptors in the stomach.

    d. Blocking the action of stomach pumps.

25. The _____ reduces production of acid by blocking the acid pumps on the stomach wall.

    a. Photon pump inhibitors

    b. Stomach pump inhibitors

    c. Proton pump inhibitors

    d. Python pump inhibitors

26. What are parietal cells?

    a. Base secreting cells of the stomach wall

    b. Mucus secreting cells of the stomach wall

    c. Water secreting cells of the stomach wall

    d. Acid secreting cells of the stomach wall

27. How many ways can we suppress acid production via the pharmacological mode of action?

    a. 1

    b. 2

    c. 3

    d. 4

28. _____ inhibits the proton pump in the parietal cells.

    a. Omeprazole

    b. Nomeprazole

    c. Gomeprizole

    d. Nomeprizole

29. Acid secretion can be suppressed by inhibitors until

    a. New proton pumps have been produced.

    b. A base can be transported to the cell membrane.

    c. Acid production has increased.

    d. Cell membranes have been regenerated.

30. Simeticon consists of

        i. Decepticon

        ii. Dimeticone

        iii. Silicon dioxide

        iv. Ometicon

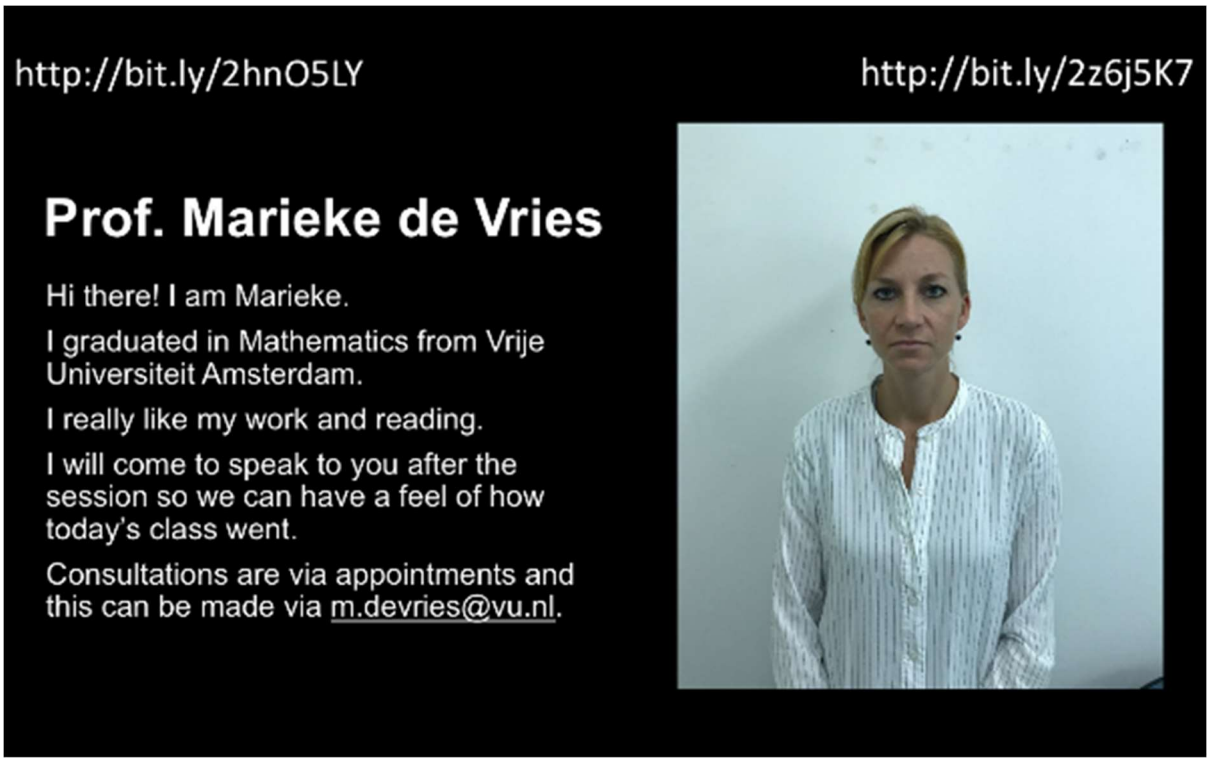
    a. i and ii

    b. ii and iii

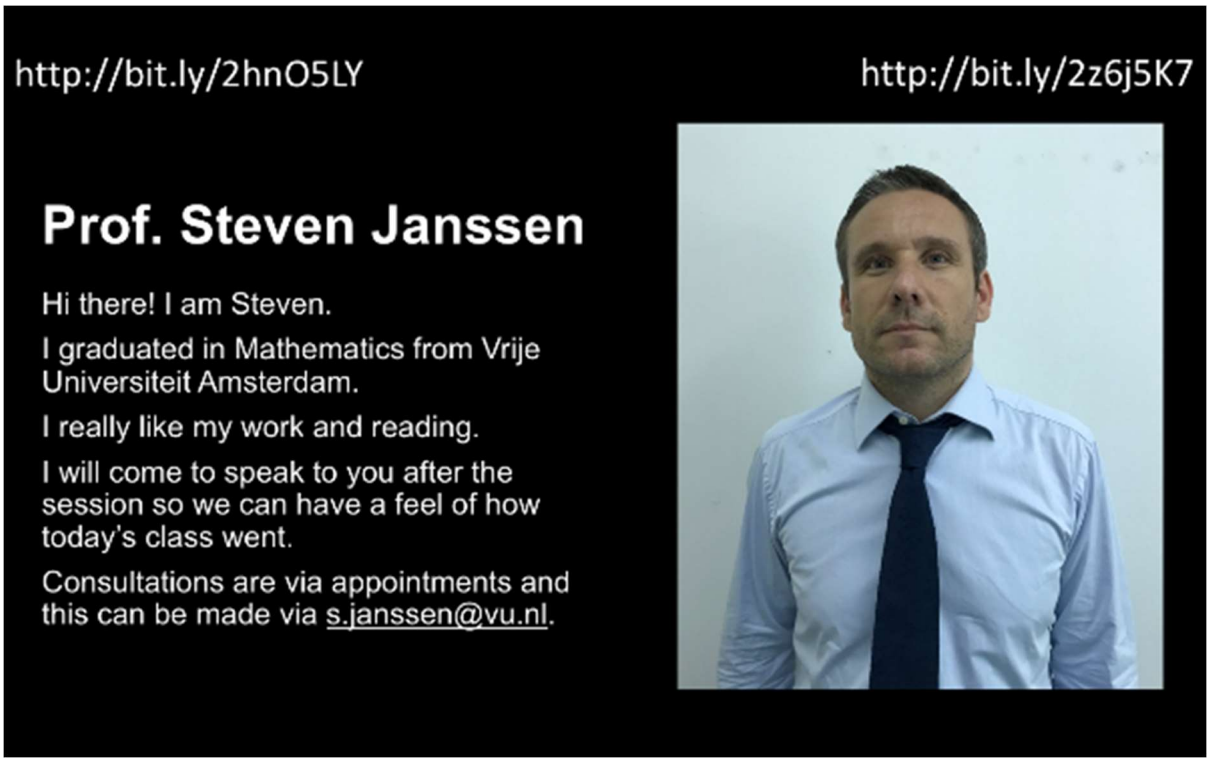
    c. ii and iv

    d. iii only

Appendix C: Introductory Slide to Introduce the Professor for the Session



Slide 1: Formal Ingroup Professor



Slide 2: Formal Outgroup Professor

Slide 3: Informal Ingroup Professor

Slide 4: Informal Outgroup Professor

Appendix D: Multiple Choice Question Bank

Note: Answers are highlighted in bold and questions 2, 3, 4, 5, 11, 14, 21, 27, 28 and 30 were selected for the tests conducted after the workshop.

1. What does ANOVA represent?
    a. **Analysis of variance**
    b. Multivariance of analysis
    c. Multivariate analysis of variance
    d. Covariance analysis
2. Which of the following is **TRUE** regarding analysis of variance?
    a. It compares the total amount of variance with the variance explained.
    b. **It is a way to compare multiple conditions in a single test.**
    c. It was invented by Roland Fisher.
    d. It is also called a t-test.
3. For between-subjects design
    a. **the explained variance is the variance between groups.**
    b. the unexplained variance is the variance between groups.
    c. the explained variance is the variance within groups.
    d. the unexplained variance is the variance across groups.
4. An extension of ANOVA that includes continuous variables is called
    a. ANOCVA
    b. MANOVA
    c. MACNOVA
    d. **ANCOVA**

5. Which of the following is an **INCORRECT** match?
    a. One-way ANOVA - One independent variable and one dependent variable
    b. **Multi-way ANOVA - One independent variable and multiple dependent variable**
    c. MANOVA - Multiple dependent variable
    d. ANCOVA - continuous variables
6. Which of these is **NOT** an assumption of ANOVA?
    a. **The scores are not sampled randomly.**
    b. Roughly a normal distribution is required.
    c. Roughly equal variance for each group or condition is a requirement.
    d. Roughly equal number of participants in the groups or observations in the conditions is needed.
7. The variance between and within groups is often referred to as
    a. **The mean squared error (MSE)**
    b. The mean squared (MS)
    c. The mean squared earning (MSE)
    d. The multiple squaring error (MSE)
8. ANOVA requires
    a. five degree-of-freedom values.
    b. four degree-of-freedom values.
    c. three degree-of-freedom values.
    d. **two degree-of-freedom values.**

9. Which of the following is the **CORRECT** way of reporting:
    a. We found a significant difference between the scores of the four groups, $f(3,32) = 15.30$, $p < .001$.
    b. **We found a significant difference between the scores of the four groups, $F(3,32) = 15.30$, $p < .001$.**
    c. We found a significance between the scores of the four groups, $F(3.32) = 15.30$, $p < .001$.
    d. We found a significant difference between the scores of the four groups, $F(3,32) = 15.30$ - $P < .001$.
10. ANOVA is a
    a. metric test.
    b. non-parametric test.
    c. **parametric test.**
    d. significant test.
11. MANOVA stands for
    a. Mean Analysis of Variance
    b. Multivariate Analysis of Covariate
    c. **Multivariate Analysis of Variance**
    d. Mean Analysis of Multivariate
12. How many methods to conduct ANOVA on SPSS were shown?
    a. 1
    b. **2**
    c. 3
    d. 4

13. Which of the following are the two ways of running ANOVA on SPSS?

    i. Analyze > Compare Means > ANOVA

    ii. Analyze > Compare Means > One-way ANOVA

    iii. Analyze > General Linear Model > Univariate

    iv. Analyze > General Linear Model > Multivariate

    a. i and iii

    b. i and iv

    c. ii and iii

    d. ii and iv

14. Which of the following is the more intuitive method of conducting an ANOVA in SPSS?

    a. Analyze > Compare Means > ANOVA

    b. **Analyze > Compare Means > One-way ANOVA**

    c. Analyze > General Linear Model > Univariate

    d. Analyze > General Linear Model > Multivariate

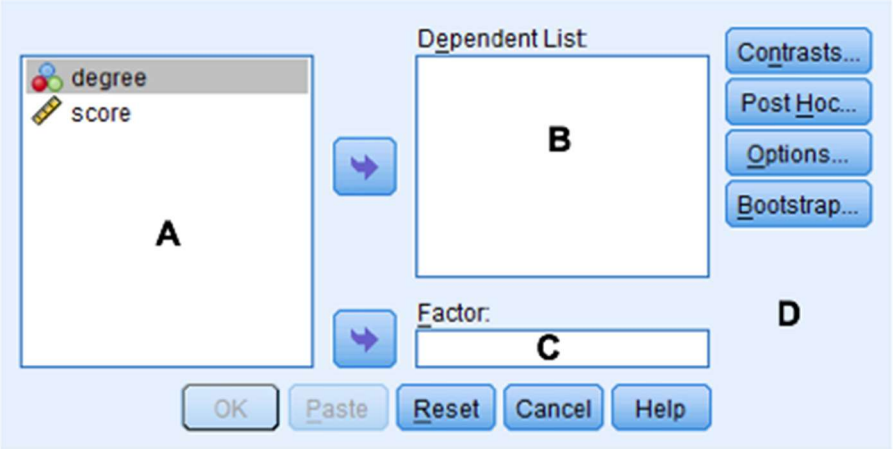
15. Which of the following allows for multiple independent variables?

    a. Analyze > Compare Means > ANOVA

    b. Analyze > Compare Means > One-way ANOVA

    c. **Analyze > General Linear Model > Univariate**

    d. Analyze > General Linear Model > Unilateralvariate



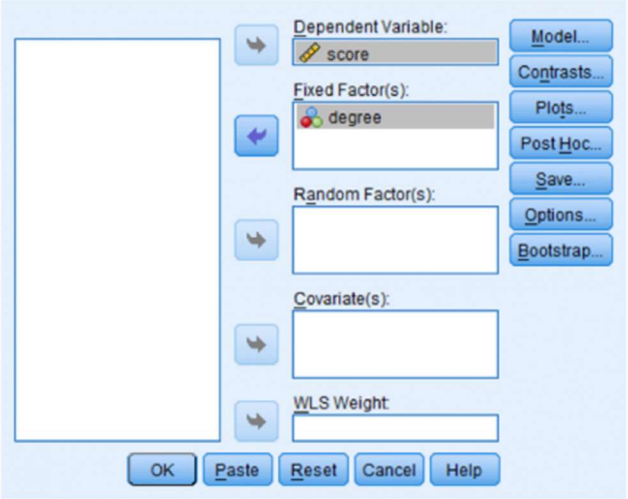
16. Where in the diagram above should the independent variable be placed?

    a. A

    b. B

    c. **C**

    d. D



17. Under which of the following will you be able to select descriptive statistics?

    a. Contrasts…

    b. Post Hoc…

    c. **Options…**

    d. Bootstrap…

18. Which post hoc test was used in the analysis example in the slides?

    a. LSD

    b. **Bonferroni**

    c. Tukey

    d. Scheffe

ANOVA

score

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	504.083	2	252.042	3.269	.058
Within Groups	1618.875	21	77.089		
Total	2122.958	23			

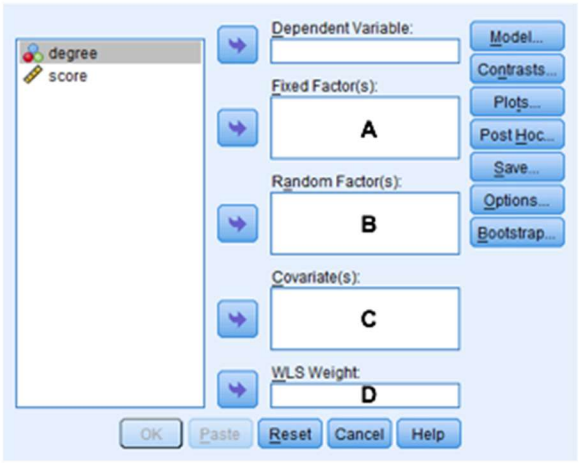
19. According to the analysis output above, how should the F value be reported?

    a. $F(21,2) = 3.27$

    b. $\boldsymbol{F(2,21) = 3.27}$

    c. $F(2,23) = 3.27$

    d. $F(21,23) = 3.27$



20. Where in the diagram above should the independent variable be placed?

    a. **A**

    b. B

    c. C

d. D

Tests of Between-Subjects Effects

Dependent Variable: score

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	504.083[a]	2	252.042	3.269	.058
Intercept	78784.858	1	78784.858	1021.995	.000
degree	504.083	2	252.042	3.269	.058
Error	1618.875	21	77.089		
Total	82743.000	24			
Corrected Total	2122.958	23			

a. R Squared = .237 (Adjusted R Squared = .165)

21. In the output above, the number 21 indicates:

    a. standard error

    b. degrees of freedom between conditions

    c. **residual degrees of freedom**

    d. total number of observations

22. If *k* = number of conditions and *N* = total number of observations, degrees of freedom between conditions is which of the following?

    a. **k – 1**

    b. N – k

    c. 1 – k

    d. k – N

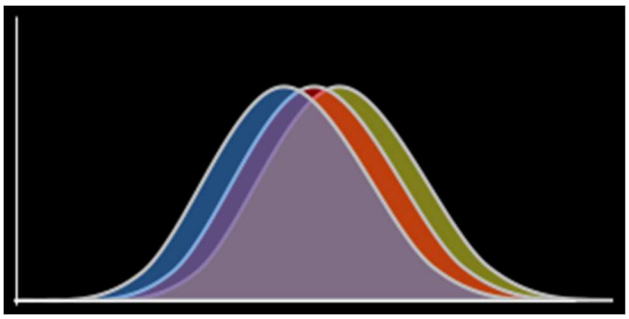
23. If *k* = number of conditions and *N* = total number of observations, residual degrees of freedom is which of the following?

    a. k – 1

    b. 1 – k

    c. **N – k**

    d. k – N



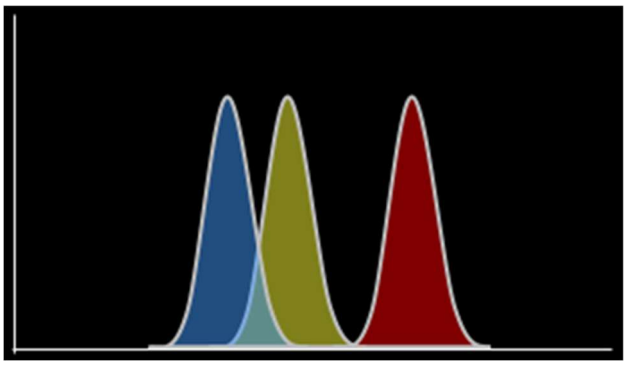
24. Which of the following is **TRUE** of the diagram above?

    a. **variance between groups is smaller than variance within each group**

    b. variance between groups is larger than variance within each group

    c. variance between groups is the same as variance within each group

    d. none of the above



25. Which of the following is **TRUE** of the diagram above?

    a. variance between groups is smaller than variance within each group

    b. **variance between groups is larger than variance within each group**

    c. variance between groups is the same as variance within each group

    d. none of the above

26. The MSE between groups is also known as

    a. **the effect**

    b. the noise

    c. the peak

    d. the error

27. The MSE within a group is also known as

    a. the effect

    b. **the residual**

    c. the degree

    d. the variance

28. Which of the following ANOVA assumptions does not apply to the MANOVA?

    a. The scores are sampled randomly

    b. Roughly a normal distribution

    c. Roughly equal variance for each group or condition

    d. **None of the above**

29. Which of the following is the **CORRECT** way of reporting mean and standard deviation values?

    a. (M = 32.2, SD = 1.2)

    b. (SD = 1.2, M = 32.2)

    c. (*M* = 32.2, *SD* = 1.2)

    d. (*SD* = 1.2, *M* = 32.2)

30. A one-way ANOVA is used for the following, **EXCEPT**

    a. between-subjects design

    b. within-subjects design

    c. **mixed design**

    d. none of the above