

THE CREATION OF STEM ROLE MODELS:
AN EXPLORATORY STUDY ON THE DESIGN OF MENTORING CHARACTERS
FOR A STEM GAMING WEBSITE

By

Leticia Lana Cherchiglia

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Information and Media–Doctor of Philosophy

2019

ABSTRACT

THE CREATION OF STEM ROLE MODELS: AN EXPLORATORY STUDY ON THE DESIGN OF MENTORING CHARACTERS FOR A STEM GAMING WEBSITE

By

Leticia Lana Cherchiglia

The use of interactive activities in classrooms (e.g., digital games) has been linked to a boost in students' motivation, interest, and learning. Such mediated environments usually include visual representations of the user (e.g., avatars) and/or mentoring characters (e.g., virtual mentors). It has been suggested that the psychological connection between users and their avatar (or virtual mentors) can potentially increase the effects of positive educational outcomes.

When considering the context of STEM (Science, Technology, Engineering, and Mathematics) education, the lack of effective STEM role models has been connected with the reinforcement of stereotypes in STEM, which in turn have negative psychological and academic effects in students - such as lower performance and lower interest in STEM subjects, as well as feelings of unbelonging to STEM fields. Negative outcomes are stronger among minority groups in STEM (i.e., women and non-white men) and can affect students even at a young age, undermining their interest in pursuing STEM careers in the future.

The current research project aims to explore if virtual mentors can be used as STEM role models for middle school students in a STEM gaming website. This project aims to contribute to the broad field of Human Computer Interaction (HCI) by investigating how different designs for virtual mentors (i.e., STEM-looking or non-STEM looking) in a STEM gaming website can affect 1) middle schoolers' perceptions of virtual mentors as STEM role models and 2) middle schoolers' Growth Mindset behavior, interest in STEM skills, and self-efficacy related to learning STEM topics (STEM learning self-efficacy) and being successful in STEM subjects

(STEM academic self-efficacy). User Experience (UX) principles guided the design of the virtual mentors and the STEM gaming website; the psychological connection between students and their virtual mentors was drawn from previous literature focused mainly in learning theories (e.g., Social Cognitive Theory, Growth Mindset), stereotypes (e.g., Stereotype Threat), and avatars (e.g., Proteus effect).

Results suggest that girls and boys perceive and interact with the STEM gaming website in similar ways, but differences exist when considering pre-to-post change in STEM metrics. While all participants showed an overall increase in Growth Mindset and STEM learning self-efficacy after website use, when compared to boys, girls showed a greater increase in STEM learning self-efficacy. Regarding STEM skill interest, girls and boys demonstrated an opposite behavior: girls showed an increase in STEM skill interest, while boys showed a decrease.

Regarding the design of the virtual mentors, during interviews all participants were more inclined to choose a STEM virtual mentor and to perceive STEM virtual mentors as better role models (i.e., more successful and better in facilitating learning) than non-STEM virtual mentors. However, when considering STEM metrics, it seems that boys would benefit more positively from having STEM virtual mentors while for girls it would be better to have non-STEM virtual mentors. Finally, there are reasons to believe that identification with the virtual mentor can indeed impact middle schoolers' STEM metrics and such impact is different for girls and boys; thus, future research should consider the effects of choosing and/or customizing a virtual mentor - both features were suggested by participants as improvements for the website.

This exploratory study is a first step towards the understanding of the psychological connection between users and their virtual mentors in a STEM gaming website through the lenses of both learning and avatar theories.

This dissertation is dedicated to so many people, scattered around the globe but all united in my heart. I'm so blessed to have you in my life, and I can only wish that all your dreams come true.
You know who you are.

ACKNOWLEDGMENTS

Without the support from the following people, I would have never finished this dissertation. To Tom, who so patiently stood by my side always bringing a smile to my face. To my parents, brother, sister-in-law, my whole Brazilian family, for their unconditional love and understanding. To my mother-in-law and sister-in-law, for being my home away from home. To all my friends, wherever you are: thank you for all the messages, warm hugs, and laughter.

I'm also extremely grateful to Robby, who was always there to help, motivate, and guide me. To my committee, who has always been very supportive and helpful, specially Carrie who is such an emotional anchor - all the hugs and sparkles of joy to you. To Dr. C. and Dr. V., the best work supervisors I could ask for. To Dr. Bauer, always so concerned with students' wellbeing. I must not forget to thank the faculty who contributed to my academic growth and the department staff who so proactively helped me in everything I needed - besides the so needed chit-chat in the hallway about crafting or video games releases.

Special thanks to Mrs. Jorae for her kindness and hard work; to Mrs. DiOrio, students' parents, and the students involved in data collection; to Jack and Aaron for helping me with the development of the experimental websites; to all advisors at Laspau and OISS who helped me navigate the not-so-easy Ph.D. road; to those who made the Science Without Borders program possible, to all the hard working Brazilians, and CNPq for funding this work (207633/2014-2).

Finally, to all the hidden figures who somehow contributed to my Ph.D. journey making it easier and/or possible – even though I might not know you, and even though I may never meet you, with all my heart: thank you.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	x
CHAPTER 1	1
INTRODUCTION	1
A Semantic Clarification.....	3
CHAPTER 2	5
LITERATURE REVIEW AND HYPOTHESES	5
Stereotypes in STEM	5
Implicit Bias, Stereotype Threat, and Growth Mindset	6
The Gender Gap.....	8
Interest in STEM topics and STEM careers	10
Self-efficacy.....	12
The Importance of Role Models in STEM	15
The Psychological Connection Between Self and Avatars.....	16
Avatar Effects	17
Pedagogical Agents.....	19
Pedagogical Agents as STEM Role Models	22
CHAPTER 3	27
DESIGNING THE STIMULI: A USER EXPERIENCE APPROACH.....	27
The STEM Game Crew Website	27
Virtual Mentors’ Design	32
The Experimental Website.....	37
CHAPTER 4	44
METHODS	44
Quantitative Study Design	44
Qualitative Study Design	47
Participants.....	49
Measures	51
CHAPTER 5	55
RESULTS	55
Quantitative Data	55
Website Data.....	70
Qualitative Data	75

CHAPTER 6	80
DISCUSSION	80
Interpretation of Results.....	81
Theoretical and Design Implications	87
Limitations and Future Research	89
CHAPTER 7	91
FINAL REMARKS	91
APPENDICES	92
APPENDIX A: VIRTUAL MENTORS' DESIGN	93
APPENDIX B: INTERVIEW IMAGES	100
APPENDIX C: WEBSITE WIREFRAME.....	104
APPENDIX D: SCALES FOR QUANTITATIVE STUDY	105
APPENDIX E: INTERVIEW QUESTIONS FOR QUALITATIVE STUDY.....	108
APPENDIX F: SUMMARY OF FINDINGS FOR QUANTITATIVE DATA	109
BIBLIOGRAPHY.....	111

LIST OF TABLES

Table 1:	Feedback given to participants in the experimental website, based on self-reported success and effort after playing the game; adapted from Mindset Works' (2017) Growth Mindset Feedback Tool for Learners	42
Table 2:	Listing of Participants' IDs, Experimental Groups, Website Versions, and Pre-Survey URLs	45
Table 3:	Quantitative data: participants' distribution among experimental groups according to virtual mentor (VM) gender and participants' reported gender	50
Table 4:	Qualitative data: participants' distribution among experimental groups according to virtual mentor (VM) gender and participants' gender	51
Table 5:	Means and Standard Error on the Measures of Pre/Post Growth Mindset, STEM self-efficacies and STEM Skill Interest within gender groups	57
Table 6:	Means and Standard Deviation on the a 5-point Likert Measures of Virtual Mentors (VM)' Perceptions within gender groups and virtual mentor type	63
Table 7:	Distribution of Participants by Game Played, Dimensions, and Pre-Post Gameplay Perceptions	72
Table 8:	Distribution of Participants by Gender, Dimensions, and Pre-Post gameplay perceptions	73
Table 9:	Distribution of Participants by Gender, Game Played, and Growth Mindset Perceptions	75
Table 10:	Coding Scheme Used for the Qualitative Data	76
Table 11:	Growth Mindset scale	105
Table 12:	STEM learning self-efficacy (L) and STEM academic self-efficacy (A) scale	105
Table 13:	STEM Skills Interest scale	106
Table 14:	Virtual Mentor's ability to facilitate learning (FL) and Virtual Mentor's credibility (C) scale	106
Table 15:	Virtual Mentor Similarity Identification (SI) and Virtual Mentor Wishful Identification (WI) scale	107
Table 16:	Questions for the Semi-Structured Interview	108

Table 17: Summary of Findings for Quantitative Data organized by STEM Variables of Interest	109
Table 18: Summary of Findings for Quantitative Data organized by Virtual Mentor’s Variables of Interest	110

LIST OF FIGURES

Figure 1: PALD model (Heidig & Clarebout, 2011)	22
Figure 2: Theoretical Model	26
Figure 3: Main page of the <i>STEM Game Crew</i> website	29
Figure 4: Example of game page showing the scientific method applied to gameplay in the <i>STEM Game Crew</i> website	30
Figure 5: Avatar location on the <i>STEM Game Crew</i> website	31
Figure 6: Possible construction for avatars in the <i>STEM Game Crew</i> website	31
Figure 7: Isbister's (2006) interpersonal circumplex for character design	32
Figure 8: Baby face (left) versus non-baby face (right) - Isbister (2006)	33
Figure 9: Virtual mentor's head design	34
Figure 10: Virtual mentor's body design	34
Figure 11: Examples of STEM virtual mentor's images related to feedback	35
Figure 12: Examples of STEM (left) and non-STEM (right) virtual mentor scenes	36
Figure 13: Examples of a Mathematician (left) and a web developer (right)	36
Figure 14: Different skin colors for the virtual mentor's head design	37
Figure 15: Example of pop-up window showing information about a STEM career	40
Figure 16: Welcome page for participants with a STEM virtual mentor (woman)	41
Figure 17: Feedback/fun fact page for participants with a STEM virtual mentor (woman)	43
Figure 18: Representations of virtual mentors used in the experimental website	48
Figure 19: Examples of more options for the virtual mentors (i.e., different skin colors and different STEM professions)	48
Figure 20: STEM academic self-efficacy behavior for the group who had STEM virtual mentors related to participants' gender	59

Figure 21: STEM academic self-efficacy behavior for the group who had non-STEM virtual mentors related to participants' gender	60
Figure 22: STEM skill interest behavior for the group who had STEM virtual mentors related to participants' gender	61
Figure 23: STEM skill interest behavior for participants who had non-STEM virtual mentors related to participants' gender	61
Figure 24: STEM learning self-efficacy behavior for boys related to median-split values of wishful identification	64
Figure 25: STEM learning self-efficacy behavior for girls related to median-split values of wishful identification	65
Figure 26: STEM learning self-efficacy behavior for boys related to median-split values of similarity identification	66
Figure 27: STEM learning self-efficacy behavior for girls related to median-split values of similarity identification	67
Figure 28: STEM academic self-efficacy behavior for boys related to median-split values of wishful identification	68
Figure 29: STEM academic self-efficacy behavior for girls related to median-split values of wishful identification	69
Figure 30: STEM academic self-efficacy behavior for boys related to median-split values of similarity identification	69
Figure 31: STEM academic self-efficacy behavior for girls related to median-split values of similarity identification	70
Figure 32: Virtual mentor's color scheme	93
Figure 33: STEM virtual mentor's images: welcome and last pages	94
Figure 34: Non-STEM virtual mentor's images: welcome and last pages	94
Figure 35: STEM virtual mentor's images: explaining the website and STEM careers	95
Figure 36: Non-STEM virtual mentor's images: explaining the website and STEM careers	95
Figure 37: STEM virtual mentor's images: game claim (hypotheses)	96
Figure 38: Non-STEM virtual mentor's images: game claim (hypotheses)	96

Figure 39: STEM virtual mentor's images: game evidence (play the game)	97
Figure 40: Non-STEM virtual mentor's images: game evidence (play the game)	97
Figure 41: STEM virtual mentor's images: game reasoning (conclusions)	98
Figure 42: Non-STEM virtual mentor's images: game reasoning (conclusions)	98
Figure 43: STEM virtual mentor's images: feedback	98
Figure 44: Non-STEM virtual mentor's images: feedback	99
Figure 45: STEM virtual mentor's images: presenting fun fact	99
Figure 46: Non-STEM virtual mentor's images: presenting fun fact	99
Figure 47: Non-STEM virtual mentor: women	100
Figure 48: Non-STEM virtual mentor: men	100
Figure 49: STEM worker virtual mentor: women	100
Figure 50: STEM worker virtual mentor: men	101
Figure 51: Chemist (Science) virtual mentor: women	101
Figure 52: Chemist (Science) virtual mentor: men	101
Figure 53: Web developer (Technology) virtual mentor: women	102
Figure 54: Web developer (Technology) virtual mentor: men	102
Figure 55: Civil engineer (Engineering) virtual mentor: women	102
Figure 56: Civil engineer (Engineering) virtual mentor: men	103
Figure 57: Mathematician (Math) virtual mentor: women	103
Figure 58: Mathematician (Math) virtual mentor: men	103
Figure 59: Complete experimental website wireframe	104

CHAPTER 1

INTRODUCTION

Students' professional choices are steering away from Science, Technology, Engineering, and Mathematics (STEM) fields, a phenomenon stronger among women. As evidence, there is the foreseen shortage of professionals in STEM (U.S. Bureau of Labor Statistics, 2014) and the fact that women represent only 24% of current STEM workforce (U.S. Department of Commerce, 2017). Explanations to this problem include students' lack of interest in STEM topics, the existence of different types of negative stereotypes related to STEM fields, and the lack of effective role models in STEM. All these factors are stronger among women (Hughes, 2016; Johnson, Pietri, Fullilove, & Mowrer, 2019, Kafai, 2016; Richard, 2016; Steinke, 2005).

Traditional education fails to teach STEM topics in ways that can potentially increase students' interest and self-efficacy (i.e., the belief that a specific behavior can be accomplished; Bandura, 1977) related to STEM careers and learning STEM concepts. As a matter of fact, textbooks and lab classes are ranked by middle schoolers as their least preferred method to learn scientific topics, and most students declare themselves unmotivated to pursue scientific careers (Marino, Israel, Beecher, & Basham, 2013). STEM interest is lower for girls when compared to boys (Steinke, Applegate, Lapinski, Ryan, & Long, 2012); this might be an outcome of cultural norms socially rewarding girls who construct their identities based on social interaction and attractiveness instead of technology-based expertise (Hayes, 2016). Another explanation is that girls are affected by negative stereotypes regarding their perceived ability in STEM fields (e.g., "girls are worse at math than boys") in addition to stereotypes that affect all students when learning STEM topics (e.g., the belief that math and science are hard subjects).

Stereotypes can be created and reinforced by implicit bias (i.e., an unconscious bias towards a specific social group; Flanagan & Kaufman, 2016). One way of reducing implicit bias is using effective role models (Blanton, Crocker, & Miller, 2000; Marx & Roman, 2002). Especially for girls, the lack of effective STEM role models might help explain their low interest in STEM careers (Kafai, 2016; Soldner, Rowan-Kenyon, Inkelas, Garvey, & Robbins, 2012). Effective role models can help change girls' mental representation of the belongingness of women in STEM (Flanagan & Kaufman, 2016).

The use of mediated interactive activities in the classroom (such as games or Internet-based activities) can make STEM-related disciplines more interesting (Clark, Nelson, Sengupta, & D'Angelo, 2009). There is evidence that students do prefer mediated interactive activities than traditional learning (Marino et al., 2013). At the same time, the use of avatars (i.e., visual representations of users) in mediated environments has been related to positive educational outcomes such as increased motivation and learning (Falloon, 2010). In gaming contexts, players share a psychological connection with their avatars which can impact behaviors during and post gaming (Peña, 2011; Ratan & Dawson, 2016; Yee, Bailenson, & Ducheneaut, 2009).

The study of pedagogical agents (i.e., characters providing support, instructions and/or motivations to users) could benefit from the avatar perspective. Previous studies in educational settings have yielded mixed results in regards with pedagogical agents' ability to bolster users' learning and motivation (Heidig & Clarebout, 2011; Schroeder, Adesope, & Gilbert, 2013). These studies did not consider the potential identification that users might feel after interacting with their pedagogical agent. Designing the pedagogical agent with a gaming perspective in mind might yield positive and stronger results, also allow the outcomes to extrapolate an educational setting in order to relate to users' own identity. Moreover, as pointed out by Heidig

and Clarebout (2011), pivotal factors to be considered when evaluating pedagogical agent's effects on users are their design, the instructional context where they are immersed, and learners' characteristics. We propose that, after a scrupulous design process, virtual mentors (a type of pedagogical agent) have the potential to become effective role models for students when engaging in STEM gaming platforms.

Thus, the main goal of this research is to investigate whether different types of virtual mentors in a STEM gaming website can serve as effective STEM roles models for middle schoolers, and if the use of virtual mentors can lead to an increase in students' self-efficacy related with STEM, and interest in STEM skills. Another goal is to examine differences between groups of interest (e.g., girls or boys). This research is composed by quantitative and qualitative components; while the former aims to investigate the connections between the different constructs mentioned above, the latter aims to gather a better understanding of students' preferences when it comes to the design of the virtual mentors in the STEM gaming website, as well as students' perceptions of the website itself.

A Semantic Clarification

As mentioned before, STEM is an acronym for Science, Technology, Engineering, and Mathematics, defined by the Encyclopedia Britannica as “a field and curriculum centered on education in the disciplines of science, technology, engineering, and mathematics”. The origin of the term (first labeled as SMET) is attributed to Judith Ramaley, who in 2001 was the director of the National Science Foundation's Education and Human Resources Division (Christenson, 2011; Hallinen, 2015). Still, the term only became frequently mentioned in political, educational, and research settings in the last decade (Loewus, 2015).

Because of the recency of the term STEM, this research project also cites studies based on the broad understanding of the words “science” and “scientists”. We believe such literature is still valid. According to the Merriam Webster dictionary, the word “science” is defined as “knowledge or a system of knowledge covering general truths or the operation of general laws especially as obtained and tested through scientific method” or “a system or method reconciling practical ends with scientific laws”; the word “scientist” is defined as “a scientific investigator”. Following this logic, all professionals in STEM can be seen as scientists in the broad field of Science.

Such argument holds strength in educational settings. It seems that the definition of knowledge, skills, and way of thinking related with STEM by Siekmann and Korbel (2016) overlaps with the definition of science literacy by the National Academies of Sciences (2016). Overall, STEM skills are summarized as “data analysis and interpretation, research and experimental design, testing hypotheses, analysis and problem-solving, and technical skills” (Bosworth, Lyonette, Wilson, Bayliss, & Fathers, 2013).

Moreover, when thinking about K-12 curriculum in the U.S., science is traditionally composed by physical sciences (e.g., Physics, Chemistry), life and human sciences (e.g., Biology, Veterinary), and earth and space sciences (e.g., Geography, Astronomy). STEM encompass other “types of sciences” outside the field of science itself such as computer sciences (i.e., related with Technology), applied sciences (i.e., related with Engineering), and mathematical sciences (i.e., related with Math). Therefore, from this point on, we make the differentiation between science (the subject in schools) and Science (a broader meaning of science as in STEM).

CHAPTER 2

LITERATURE REVIEW AND HYPOTHESES

Stereotypes in STEM

Scientists are broadly stereotyped as white older men – often depicted as very intelligent but quite eccentric and not very attractive - wearing white lab coats and glasses and working in research labs (Losh, 2010; Steinke et al., 2007). Entertainment media (i.e., television, movies, and video games) usually conforms to this stereotype (Dudo, et al. 2011; Dudo, Cicchirillo, Atkinson, & Marx, 2014), adding to the scene the depiction of women scientists as attractive and intelligent, but extremely career-focused and usually mistreated in the workplace by men (Steinke, 2005). It is important to consider the role of media and technology as both are socially constructed (Cote, 2015), thus can influence people’s perceptions of Science and scientists.

Such stereotype of a scientist (i.e., a man working in a research lab, wearing white lab coats and glasses) assumes that all scientists are professionals related to very specific careers in Science (e.g., Chemistry, Biology, Veterinary). Not all professionals in those careers will conform to such stereotype because some professionals might not even work in research labs. We acknowledge the fact that specific careers in STEM have their own stereotypes that do not traditionally conform to the one of a scientist (e.g., the idea of the “geek” in technology-related careers). Still, because the stereotyped view of scientists can make it harder to “increase personal respect for scientists or interest in science careers” (Losh, 2010; p. 381) while limiting children’s mindsets when visualizing themselves as future scientists in STEM, stereotypes in Science should be taken into consideration when designing STEM educational experiences.

There is evidence that the stereotyped image of a scientist is consistent with what most children and adolescents believe to be a scientist. A common way of assessing children's and adolescents' perceptions of scientists is asking them to draw a scientist, what is known as the Draw-A-Scientist-Test (DAST), proposed by Finson, Beaver, and Cramond (1995). A recent meta-analysis of the DAST literature based on almost 80 different studies suggest that, although children nowadays are more likely to depict women scientists than in the past, the proportion is still as low as 28% (Miller, Nolla, Eagly, & Uttal, 2018). Another finding from that study is that children are more likely to stereotype their drawings as they grow old, and boys are more likely to stereotype their drawings of scientists than girls. These results suggest that not only stereotypes are learned as children mature but school settings hold a pivotal role in terms of exposing images of scientists to children.

Because of all the arguments presented so far, we believe that middle school is a pivotal time to introduce students to STEM educational activities, but more importantly, such activities should be designed in ways that minimize the creation and/or reinforcement of stereotypes. In order to do so, we must investigate factors that might cause stereotypes in STEM, as well as consequences of such stereotyping. After all, stereotypes in STEM can potentially impact students' interest in STEM, and their perceived ability in STEM related disciplines.

Implicit Bias, Stereotype Threat, and Growth Mindset

At the core of stereotypes lies implicit bias (i.e., an unconscious bias towards a specific social group, created by either learning negative rules about that social group both from others or making observations in the world; Flanagan & Kaufman, 2016). Implicit bias can create negative expectations about learning STEM related disciplines (e.g., “math and science are hard subjects”, “math and science are not for everyone”) and negative gender-stereotypical expectations

targeting girls in STEM related settings (e.g., “girls are not interested in games”, “girls are worse at math than boys”, Jenson & De Castell, 2010). Implicit bias can be endorsed unconsciously by parents, educators, game designers, friends, and even self-endorsed, leading students (especially girls) to unconsciously conform to stereotypes created for them (Flanagan & Kaufman, 2016).

Dweck’s (2006) extensive research on fixed mindsets (i.e., the belief that a certain trait is unchangeable due to genetics) versus growth mindsets (i.e., the belief that a certain trait can be changed through effort) might cast a light on stereotypes. Negative stereotypes regarding ability would be examples of fixed mindset beliefs; also, the more individuals who are affected by a negative stereotype hold a fixed mindset (versus a growth mindset) the more likely these individuals are to be affected by the stereotype itself (Dweck, 2008). STEM learning stereotypes are related with fixed mindsets and often fruit of cultural norms, which might be learned and/or reinforced in educational settings. For example, when students fail an exam in STEM related disciplines such as math or science, parents and/or teachers might say “You are just not a math/science person” or “It’s because math and science are too hard”. Such fixed mindset feedbacks help perpetuate the myth of inborn genetic abilities in STEM, as it leads students to believe that, no matter how hard they try, there is nothing they can do to improve their academic performance, in turn leading to self-justification for future poor performance (i.e., “I’m just not a math/science person”, “math/science is not for me”; Dweck, 2008; Kimball & Smith, 2013).

Research in gender development (Hyde, 2014; Martin & Ruble, 2010) suggests that the scientist stereotype can indeed draw girls away from interest in STEM-related activities and STEM careers (Bian, Leslie, & Cimpian, 2017; Weisgram, 2016). Also, by learning that scientists are more agentic (e.g., independent) than communal (e.g., sociable), girls are susceptible to a cultural mismatch since society associates communal traits with perceptions of

femininity and agentic traits with perceptions of masculinity (Carli, Alawa, Lee, Zhao, & Kim, 2016).

Such argument has been connected with girls' lack of interest in STEM careers (Diekman, Steinberg, Brown, Belanger, & Clark, 2017) and might contribute to the phenomenon known as stereotype threat: when individuals are reminded of a negative stereotype regarding their gender or ethnicity in a performance setting, their performance will be lower, regardless of actual skill level (Steele & Aronson, 1995). Another phenomenon connected with stereotypes and performance is called stereotype lift, or a performance boost caused by the activation of a negative out-group stereotype (Lee, Nass, & Bailenson, 2014). For example, if before a math-related task both boys and girls are reminded that "girls are worse at Math than boys", girls could potentially suffer stereotype threat (thus performing worse than what was expected for their skill level) and boys could potentially suffer stereotype lift (then performing better than what was expected for their skill level).

The Gender Gap

One of the consequences of negative stereotypes in STEM is the gender gap in STEM: according to data from the National Science Foundation (2014), starting from middle school, girls consistently show lower scores on standardized science and math tests compared with boys of the same grade. We believe that the gender gap in STEM can be seen as a vicious cycle driven by negative stereotypes regarding girls' abilities in STEM: because girls are led to believe they will perform worse than boys, they become less interested in pursuing STEM activities, which in turn detracts them from their actual performance in these domains. This supports the stereotype and the gap widens. However, when girls are told that both genders can perform equally in

STEM related disciplines such as math, girls perform as well as boys (American Association of University Women, 2010).

The gender gap in STEM is perpetuated throughout high school and college, culminating in the low percentage of women (24%) in current STEM workforce (U.S. Department of Commerce, 2017; National Science Foundation, 2014). According to a recent report (Elsevier, 2017), other factors that can explain gender inequities in STEM include “persistent bias in hiring, authorship, recognition, and promotion” (p. 11). This argument is strengthened by Holman, Stuart-Fox, and Hauser (2018)’s study which analyzed over than 6,000 journals and found that the gender gap in academia is especially strong in hard sciences fields, prestigious journals, and senior research positions.

The gender gap in STEM is a phenomenon not restricted to the United States. Different countries show varying degrees of gender inequality in STEM fields, most likely due to cultural and social differences. Although proportions of women in STEM fields have been slightly growing worldwide when analyzing almost 20 years of data in over 12 different countries, women scientists are still minority (Elsevier, 2017). Literature related to stereotype threat for women pursuing Engineering majors suggests that stereotype threat is a result of societal factors influencing women’s interest in STEM careers, including social interaction (Dell, Verhoeven, Christman, & Garrick, 2017; Voigt, Hocevar, & Hagedorn, 2007; Wentling & Camacho, 2008). This finding is consistent with another study showing that although STEM majors are usually perceived as harsh, pursuing a STEM major can be made easier for college students through academic conversations with peers, interaction with faculty, and socially supportive environments (Soldner et al., 2012).

Interest in STEM topics and STEM careers

As mentioned before, negative stereotypes related to Science and scientists might decrease students' interest in STEM topics and STEM careers, especially for girls.

Unsurprisingly, at middle-school age, most students declare themselves unmotivated to pursue Science careers (Marino et al., 2013) and girls are even less interested in Science and technology than boys (Hayes, 2016). Another factor contributing to girls' lack of interest in STEM is cultural: historically, girls have received social rewards when showing interest in the culture of beauty and romance and/or when improving their communication skills; technology-related interest and skills do not yield such social rewards (Hayes, 2016).

One example is the case of video games: since a young age, girls have less access to video games and video games consoles at home and are more likely to be regulated by parents than boys (Jenson & De Castell, 2010). Such social norm can make girls feel less comfortable to approach video games and less interested in them, because other leisure activities are going to be more easily accessible, familiar, and potentially less-parentally regulated. Both issues may not only make girls less interested in STEM but can potentially undermine their performance in STEM related courses (Hughes, 2016). As a matter of fact, a recent study has suggested that girls who are considered heavy gamers (i.e., more than 9h per week of gameplay) are three times more likely to pursue a STEM career than non-gamers, a finding connected to girls' ability to build and reinforce their identity through gaming (Hosein, 2019). For boys, previous gaming behavior had no effect or only a weak effect.

Regardless of gender, middle school is a very important time for students to get interested in STEM: students pursuing Science majors believe their interest in Science began before or during middle school, and that school was responsible for sparking such interest (Maltese & Tai,

2010), a finding stronger for girls (52%) than boys (33%). Teachers' attitudes in the classroom (Maltese & Tai, 2010) and the use of mediated interactive activities in the classroom for science learning (such as games or Internet-based activities) can be responsible for making science-related topics more interesting (Clark et al., 2009). Moreover, informal settings can serve as important environments for Science learning, thus helping children to become more enthusiastic about STEM and to be able to perceive scientific topics as less complicated and more ubiquitous in the real world (National Research Council, 2009).

The approval and support of parenting figures are also pivotal to children's and adolescents' interest in pursuing Science careers (Maltese & Tai, 2010). However, this is a complex and delicate topic. For example, between 1983 and 2001 there has been an increase in adults' positive images of Science careers, endorsing Science careers for sons and daughters, and considering pursuing Science careers themselves (Losh, 2010). At the same time, it is not unusual for parents to educate their children in ways that prevent or minimize scientific discovery (Big Think, 2013), or to discredit scientific discoveries that are not aligned with specific political or ideological views (Kahan, Braman, Cohen, Gastil, & Slovic, 2010; National Academies of Sciences, 2016). These factors together with parenting styles which promote fixed mindsets related to STEM might negatively influence children's interest in STEM careers. Because it is extremely difficult to control for such parental influence, we do not include such construct into this research. Instead, we focus on the use of STEM educational activities in the context of middle school education, while investigating factors related with the self (e.g., self-efficacy) that might be connected with the lack of students' interest in STEM careers.

Finally, it is worth mentioning that although many studies have examined students' interest in pursuing STEM careers (Dabney et al., 2012; Kier, Blanchard, Osborne, & Albert,

2014; Oh, Jia, Lorentson, & LaBanca, 2013; Tyler-Wood, Knezek, & Christensen, 2010), there might be an issue when using this construct in relation to children. After all, middle schoolers might be uncertain of their interest in future careers in STEM because each professional field might be too broad or abstract for them at that point in time. Besides, asking students to rate their interest in a plethora of scientific careers (e.g., Biologist, Chemist) might lead to unwanted results as students might carry over stereotypes related to Science and scientists when considering themselves as professionals in STEM. Thus, we believe the best approach is to focus on students' interest in STEM skills (i.e., those related but not limited to “data analysis and interpretation, research and experimental design, testing hypotheses, analysis and problem-solving, and technical skills”; Bosworth et al., 2013).

STEM skills are closely associated with STEM careers (i.e., if a middle schooler has no interest in programming, it is unlikely that this student will pursue a career in computer science in the future. Also, several STEM skills can be learned and/or developed while in school (e.g., using logic to solve problems, programming, working with a microscope, writing reports). Thus, we define *STEM skill interest* as the interest in skills related with STEM careers (e.g., doing experiments in a laboratory, solving puzzles and/or riddles, thinking of new ways to do things).

Self-efficacy

Bandura's (1989) Social Cognitive Theory (SCT) states that human functioning is dynamic, and it will be influenced concomitantly by three main factors: personal determinants, behavioral determinants, and environmental determinants. According to SCT, one can learn by observing models and building self-efficacy, defined as “the conviction that one can successfully execute the behavior required to produce the outcomes” (Bandura, 1977, p. 193). The three determinants mentioned above (personal, behavioral, and environmental) can be connected to the

concept of self-efficacy because 1) personally, one's level of self-efficacy will determine their willingness to perform the behavior, 2) behaviorally, the outcome of the behavior will shape one's level of self-efficacy, and 3) environmentally, aspects in the environment (such as reinforcements) can influence the successfulness of the behavior (Bandura, 1977).

Bandura (1989) identifies four factors responsible for building self-efficacy: 1) mastery experiences, 2) physical and emotional states, 3) social modeling, and 4) social persuasion. It is logical to assume that successful experiences can lead to increased self-efficacy and vice-versa, but it is important to note that achieving success through low effort (when compared to high effort) can lead to the expectation of rapid results, thus making a future failure even more discouraging (Bandura, 2008) – an argument connected to the importance of having a Growth Mindset instead of a fixed one. Also, one's physical and emotional states can influence how success and failure are interpreted, biasing what would be otherwise an accurate judgment (e.g., a smart student fails an exam because of anxiety or sickness; self-efficacy might decrease despite of the high level of skill).

Social factors are extremely important when considering how to overcome negative stereotypes in STEM fields. Social modeling and social persuasion can increase self-efficacy because of their relationship with role models. Specifically, social modeling relates to observing a role model who can demonstrate self-efficacy; this can serve as inspiration and motivation to change one's own self-efficacy. Social persuasion relates to having a role model who is “knowledgeable and practice[s] what they preach” (Bandura, 2008); role models might be able to provide opportunities for mastery experiences and to persuade one to believe in themselves. Staples, Hulland, and Higgins (1998) suggest that mastery experience is the main source of

information when one is forming self-efficacy judgments, followed by social modeling, social persuasion, and physical and emotional states.

Literature based on Social Cognitive Career Theory (SCCT; Lent, Brown, & Hackett, 1994) suggests that self-efficacy can predict one's interest in pursuing a specific career. There is a broad body of literature encompassing traditional learning environments and interest in STEM fields and/or STEM self-efficacy (Diekman et al., 2010; MacPhee, Farro, & Canetto, 2013; Rittmayer & Beier, 2009; Soldner et al., 2012). Overall, women/girls report lower STEM interest and/or STEM self-efficacy than men/boys, although some studies show evidence that this gap might be closing (Britner & Pajares, 2006; Kenney-Benson, Pomerantz, Ryan, & Patrick, 2006). Regardless, there seems to exist a theoretical gap regarding the construct of STEM self-efficacy as it is defined and/or measured differently over several studies. Factors related with learning, performance, or expected outcomes (i.e., "a person's estimate that a given behavior will lead to certain outcomes; Bandura, 1977, p. 193) sometimes are all included in the definition of STEM self-efficacy, while sometimes the construct only focus on one of them.

In this research project, we believe that it is important to differentiate between two types of self-efficacy related to STEM: 1) *STEM learning self-efficacy*, defined as self-efficacy associated with learning STEM-related topics (e.g., understanding the content of a math lesson), and 2) *STEM academic self-efficacy*, defined as self-efficacy associated with performance in STEM-related disciplines (e.g., earning a good grade in math). After all, self-efficacy is a construct related with one's beliefs in their abilities regarding specific tasks or set of tasks which usually have a success/failure type of outcome (Bandura, 1977); moreover, although potentially related, both STEM self-efficacy constructs are theoretically different in a school setting (e.g., a high-anxious test taker might be an excellent learner despite of low scores in a discipline).

The Importance of Role Models in STEM

According to Kao and Harrell (2016), effective role models are perceived as competent, successful, and similar to the self (i.e., sharing common attributes). Competence and successfulness are connected to self-efficacy, which as discussed before can increase based on role models (Bandura, 2008). Sharing common traits with scientists has been connected with the likelihood that students will actually pursue Science careers (Brush, 1979; Tai, Liu, Maltese, & Fan, 2006; Zahry & Besley, 2017).

While meeting or knowing a scientist can positively impact children's views of Science and scientists thus decreasing stereotyping (Steinke et al, 2012; Woods-Townsend et al., 2015), that is not enough to create effective STEM role models. Scientists sometimes lack proper training when engaging with the public, thus being perceived as boring or too technical (Woods-Townsend et al., 2015). As a matter of fact, scientists rank building trust and sparking excitement as less important than prioritizing informational communication or defending Science from misinformation (Dudo & Besley, 2016). Such behavior only serves to reinforce public's stereotyped view of a scientist as an eccentric "know-all" who lacks social skills. Moreover, finding common traits with scientists might be especially hard for minorities in STEM because of differences in gender, ethnicity, and in perceived proficiency in STEM related disciplines. Unsurprisingly, students more likely to pursue STEM majors are usually male, very proficient in math and science, and with parents showing higher levels of education (National Academies of Sciences, 2016).

In K12 educational settings, teachers can become role models for students thus sparking their interest in science and Science careers, but sometimes teachers can have quite the opposite impact due to their lack of passion, perceived incompetence, or rigid teaching styles (Maltese &

Tai, 2010; Big Think, 2011). As mentioned before, STEM learning activities in mediated environments (e.g., games) can make learning more interesting (Clark et al., 2009). Students apparently prefer mediated interactive activities to traditional science learning (Marino et al., 2013). Thus, if the goal of a STEM learning activity is to encourage students to pursue STEM careers, effective role models need to be enacted while minimizing the potential of stereotype threat - otherwise minorities might feel even less motivated to pursue STEM careers (Pearce, 2016). Bandura (2008) himself raises the point that nowadays people do not necessarily need to draw role models based on their own social surroundings: the pervasiveness of the internet allows role models to arise from mediated environments.

The Psychological Connection Between Self and Avatars

The Proteus effect (Yee & Bailenson, 2007) is a phenomenon seen in mediated environments where individuals behaviorally conform to the behaviors expected from their avatars (i.e., participants with taller avatars are more confident to negotiate than people with shorter avatars). The Proteus Effect can persist in subsequent face-to-face interactions for a short period of time (Yee et al., 2009). The original argument for the Proteus effect comes from self-perception theory: individuals would perceive their avatars as a third-person and, because individuals were embodying this third person in the mediated environment, individuals would believe their behavior should match the behavior of what was perceived as a natural behavior for the third person. However, other scholars (Peña, 2011; Peña, Hancock, & Merola, 2009) defend that the Proteus effect happens due to priming, because individuals would prime schemas related to their avatars when embodying them (also schemas related to the context in which the avatars are immersed in the mediated environment); consequently, individuals would behave in ways consistent with the primed schemas.

Despite the theoretical explanation given to the Proteus effect, there is a psychological bond created between users of the mediated environment and their avatars (Gee, 2008), and this bond will be stronger when feelings of identification (e.g., “my avatar is like me”), embodiment (e.g., “the avatar’s body is an extension of my body”), and idealization (e.g., “I want to be like my avatar”) are high. While Van Looy, Courtois, De Vocht, and De Marez (2012) use these three dimensions in their construct of game character identification, Ratan and Dawson (2016) have created the construct of avatar self-relevance (i.e., how relevant the user perceives their avatar to be); avatar self-relevance includes only users’ identification and embodiment with their avatar. It is suggested that avatar self-relevance is a potential moderator of the Proteus effect: having a customizable avatar of the same gender and ethnicity of the individual can increase avatar self-relevance, thus potentially increasing the strength of the Proteus effect (Ratan & Dawson, 2016). However, in certain scenarios where stereotype threat is likely to occur, using a customizable avatar of a different gender/ethnicity than the self might yield positive results. For example, when solving a math test (where there is the stereotype that women are worse than men in math), women who had a customizable masculine avatar actually performed better than women using feminine avatars (Ratan & Sah, 2016). This argument reinforces the idea defended by Klimmt, Hefner, Vorderer, Roth, and Blake (2010) that players’ identity is affected when players enact roles in video games.

Avatar Effects

Within the Proteus Effect framework, specific uses of avatars have been connected with different positive outcomes in learning environments. For example, engineers showed higher creative performance when using avatars based on famous inventors (Guegan, Buisine, Mantelet, Maranzana, & Segonds, 2016), and male students showed higher performance in educational

avatar-related tasks when using avatars based on their ideal-selves (Ratan et al., 2016). In a science learning game context, Kao and Harrell (2015; 2016) investigated how specific avatar designs could impact players' engagement and interest. These examples suggest that, as mentioned by Falloon (2010), avatars can be responsible for increased engagement in educational activities, but specific categories related with avatars' design should be considered, such as the ones mentioned below.

1. **Avatar's identity.** Several studies have investigated how using avatars based on different avatar's identity could impact participants' performance; examples of these possible identities are: no identity (i.e., abstract shaped), actual-self, ideal-self, and role-models (Guegan et al., 2016; Kao & Harrell, 2015; Kao & Harrell, 2016; Ratan et al., 2016). It is important to note that "role-models" is a broad category, including fictional and non-fictional others such as superheroes, famous scientists, others perceived as inventors, or athletes.
2. **Avatar customization.** According to Ratan and Sah (2016), avatar customization (i.e., the ability to personalize avatar's physical traits) can reinforce avatar self-relevance, because the psychological connection between players-avatars through avatar identification and embodiment has the potential to be higher when players are able to choose (to a certain extent) how they want to represent themselves in the mediated environment through their avatars. Although both girls and boys seem generally interested in the idea of customizing their avatars in games (Marino et al., 2013), there is evidence that girls enjoy and value avatar customization more than boys (Heeter, Egidio, Mishra, Winn, & Winn, 2009), an act that might be connected with the act of playing "dress-up" - a feminine social construct (Jenson & De Castell, 2010).

3. **Avatar idealization.** Avatars are usually designed in ways that elicit wishful feelings outside the mediated environment (i.e., avatars as ideal-selves; Klimmt et al., 2010). Although there is evidence that a certain degree of avatar-idealization is natural to occur when letting players freely customize their avatars in games (Klimmt et al., 2010), too much idealization can lead to lower well-being (Bessière, Seay, & Kiesler, 2007), lower engagement, and performance (Kao & Harrell, 2015). Asking participants to customize their avatars as ideal-selves can lead to better performance for men, but lower performance for women (Ratan et al., 2016).

Depending on the design of the mediated environment, avatars can show different degrees of agency and avatar-likeness. According to Ratan (2017), both constructs are located in a triadic model based on autonomy vs. control, functionality congruent goals vs. self-congruent goals, and own characters traits vs. not-own characters traits. For example, digital vehicles would be an example of avatar-as-object, as they are highly controllable by users and serve a more functionality congruent goal; digital companions would be an example of avatar-as-social other, as they have more autonomy in the mediated environment and usually do not share many characteristics with the user; last but not least, digital selves would be an example of avatar-as-me, as they share many (if not all) characteristics with the user (usually through avatar customization) while serving a more self-congruent goal. It is unclear if Ratan's (2017) conceptualization of avatars would be able to fully understand pedagogical agents (i.e., visual representations of others whose goal is to mentor users in a learning mediated environment) as a type of avatar (i.e., digital companions).

Pedagogical Agents

Overall, pedagogical agents have been defined as characters in a mediated learning environment that guide, mentor, and/or facilitate instruction to users. Since the 90's, and more

recently with the growing introduction of technological devices in classrooms, pedagogical agents have been the focus of many studies in the realm of education, design, and HCI. Early studies on pedagogical agents were derived from research on interface agents (i.e., computer programs designed to aid users in computer tasks) and focused on technological factors related to pedagogical agents' design and development (e.g., perceived intelligence, entertainment value).

Only in the late 90's research started investigating the connection between pedagogical agents and educational factors such as learning or motivation (Clarebout, Elen, Johnson, & Shaw, 2002; Dehn & van Mulken, 2000). Findings indicate that pedagogical agents have been used as a tool to augment engagement, motivation, and learning support in online learning (Augusto, McNair, McCullagh, & McRoberts, 2010; Cantrell, Fischer, Bouzaher, & Bers, 2010). However, Heidig and Clarebout (2011) meta-analysis of 39 previous studies on pedagogical agents and their connection with learners' outcomes (e.g., learning, motivation) pointed out mixed results and the overall lack of a control group.

A more recent meta-analysis investigated 43 studies focused on the relationship between pedagogical agent and participants' learning, with participants ranging from K12 to high school students (Schroeder et al., 2013). Results suggest that indeed the use of pedagogical agents can yield positive outcomes in learning settings, an effect stronger among K12 students than high school students. Another finding was that learning was better facilitated when the pedagogical agent used on-screen text (rather than voice) to communicate with users.

Regardless of evidence in previous literature of the effect that pedagogical agents may (or not) have on learners, several studies have focused on pedagogical agents' visual design and behavior in the mediated experience. After all, drawing from Bandura (1989), Baylor (2011) claims that "the agent's appearance is the most important design feature as it dictates the

learner's perception of the agent as a virtual social model", followed by motivational messages and communication with users. Kim and Baylor (2006) also drew from social-cognitive theories when proposing several design perspectives to be taken into consideration when designing pedagogical agents.

There are communalities between these seven factors and the Pedagogical Agents-Conditions of Use Model (PACU) framework proposed by Heidig and Clarebout (2011). The PACU model takes into consideration four conditions: 1) the learning environment design and context, 2) learners' characteristics, 3) instructional behavior of the pedagogical agent (e.g., feedback, instructions), and 4) the visual design of the pedagogical agent. Because of the complexity associated with the visual design of pedagogical agents (condition 4 mentioned above), another model was proposed, the Pedagogical Agents Levels of Design (PALD). This framework posits three different design levels: global (i.e., human or non-human; static or animated), medium (i.e., choice of character's role and technical decisions), and detail (e.g., age, gender, clothing). The model can be seen in Figure 1 (next page). According to the authors, previous studies related to the design of pedagogical agents fail to address all different design levels, besides providing mixed results in terms of which aspects would yield better results.

Another facet to this issue is to consider users' preference related to pedagogical agents' design features. After all, there might be a mismatch in regards with what designers believe to be the best for learners, and what learners believe to be best for them (consciously and unconsciously). There is evidence that, as learners grow old, similarity with the pedagogical agent (i.e., same gender and ethnicity) becomes less important to participants. K12 students are more likely to choose a pedagogical agent similar to themselves (Johnson, DiDonato, & Reisslein, 2013) than high school students (Kim & Wei, 2011) and college students (Baylor,

Shen, & Huang, 2003; Moreno & Flowerday, 2006). K12 students also have indicated their preference for pedagogical agents perceived as knowledgeable and better suited to facilitate learning (Johnson et al., 2013).

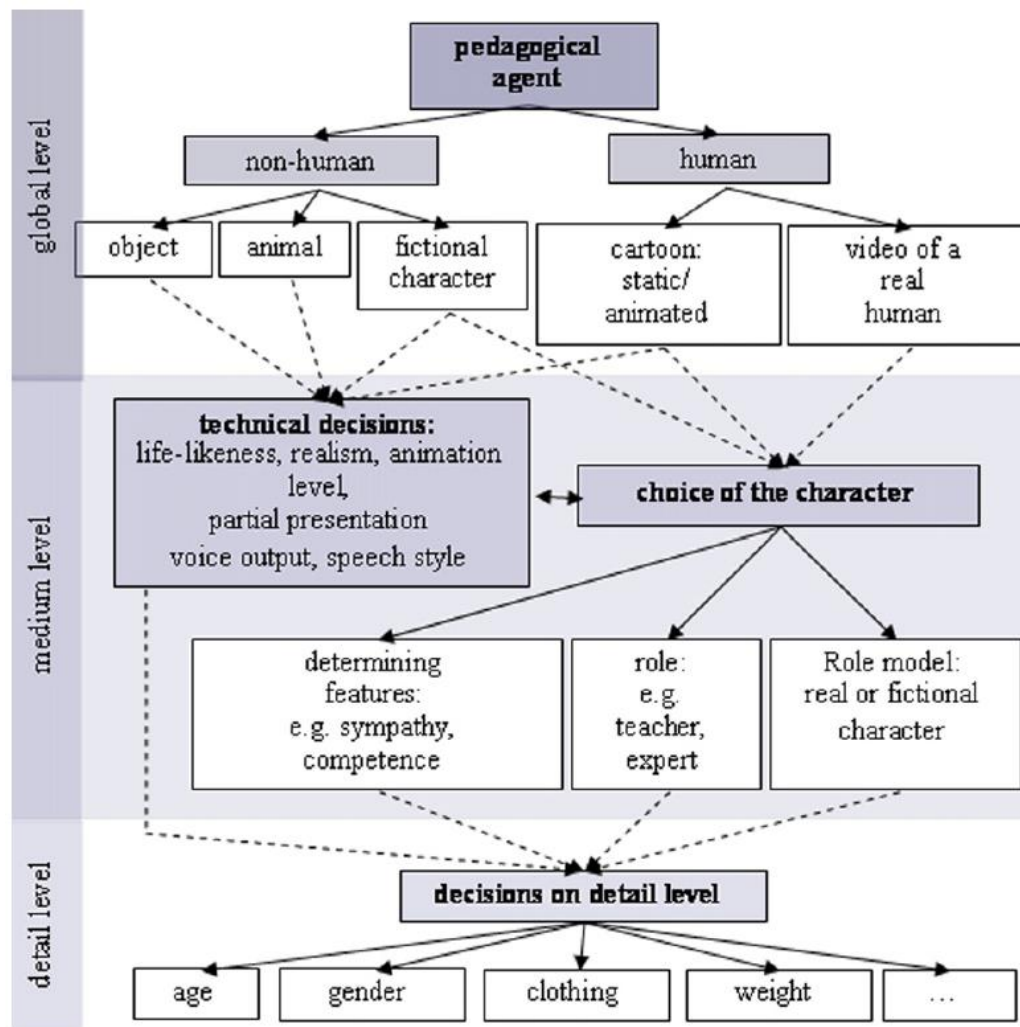


Figure 1. PALD model (Heidig & Clarebout, 2011).

Pedagogical Agents as STEM Role Models

Overall, pedagogical agents have been used in previous literature as a tool to augment engagement, motivation, and learning in online learning environments (Augusto et al., 2010;

Beckem, 2012; Cantrell et al., 2010; Curran, & Chatel, 2013). The theoretical background used in these studies has been focused on learning theories, which leads us to wonder if the use of avatar theories such as the Proteus effect could yield better results or even a more in-depth understanding of 1) how specific design choices for these virtual agents can impact learners, and 2) the potential psychological connection between learners-virtual agents.

In this research project, we want to explore if pedagogical agents in a STEM gaming website can be perceived by middle schoolers as STEM role models. More specifically, we consider if positive STEM outcomes can be achieved after middle schoolers use a STEM gaming website. We refer to such pedagogical agents as virtual mentors, because these virtual human characters will be designed with the primary goals of mentoring middle school students through a Growth Mindset approach, and to increase excitement related to STEM topics and STEM careers.

Given the exploratory nature of this research, we want to investigate if and how different designs for the virtual mentors can yield different results, which might differ also when considering participants' gender. It is very important that the design of such mentoring characters steer away from the stereotypical portrayal of scientists in the media (i.e., eccentric white male), as there is plenty evidence in previous literature that by doing so negative stereotypes in STEM are created and reinforced, especially for girls. We want to avoid the occurrence of stereotype threat (Steele & Aronson, 1995), while promoting a Growth Mindset due to its positive impact in helping to deconstruct negative stereotypes (Dweck, 2008; Kimball & Smith, 2013).

However, we cannot steer completely from the STEM realm because if we want to construct STEM role models, we want middle schoolers to perceive the virtual mentor as someone competent and successful in a STEM learning context, while ideally someone relatable

and with whom students can identify with. Such approach could potentially increase students' STEM self-efficacy and ability to see themselves as belonging to STEM fields, because of the potential positive psychological connection between students and their virtual mentors - drawing from SCT (Bandura, 2008), SCCT (Lent et al., 1994), stereotype threat avatar theories related to role models (Kao & Harrell, 2016), avatar effects (Yee & Bailenson, 2007), and findings related to pedagogical agents (Schroeder et al., 2013).

Because of the gender gap in STEM and the extensive literature suggesting that boys and girls have a different relationship with STEM fields, we have reasons to believe that boys and girls might behave and be affected differently when considering our STEM gaming website and our STEM variables of interest. Thus, we propose the following research question:

RQ1: Are there differences in the behavior of STEM variables of interest (i.e., Growth Mindset, STEM self-efficacies, and STEM skill interest) when considering participants' gender (boys or girls)?

In regards with the design of the virtual mentor, and given the lack of control groups in previous pedagogical agents' literature as pointed out by Heidig and Clarebout (2011) meta-analysis, we decided to include the following hypotheses:

H1: Participants assigned to the version of the STEM gaming website with a virtual mentor (when compared to those without a virtual mentor) will demonstrate higher **a)** Growth Mindset, **b)** STEM self-efficacies, and **c)** STEM skill interest.

Delving further into the aspect of how virtual mentors should be designed, and following a more exploratory nature, we want to investigate if a more “STEM-looking” virtual mentor would

have any kind of impact into the STEM outcomes related to our middle-school participants. Thus, the following research question is proposed:

RQ2: Are there differences in the behavior of STEM variables of interest (i.e., Growth Mindset, STEM self-efficacies, and STEM skill interest) when considering virtual mentor's type (STEM or non-STEM) and participants' gender (boy or girl)?

Given the argument that visual aesthetics can change the way users perceive the virtual mentor as a social model (Baylor, 2011), and the careful design process that will be used when designing the virtual mentors, we expect that participants will perceive STEM-looking virtual mentors as more competent and successful in a STEM teaching environment when compared to non-STEM-looking virtual mentors. However, as identification is a very complex construct that depends not only on participants' own identity but also on their implicit bias towards scientists and how virtual mentors' aesthetics and behavior will be perceived by participants, we do not know how much participants will be able to identify with their virtual mentor. It is also unclear if results will be affected by participants' gender. The following hypotheses and research questions are proposed:

H2: STEM virtual mentors will be perceived as having higher a) ability to facilitate learning, and b) credibility when compared to non-STEM virtual mentors.

RQ3: Are there any effects of virtual mentor type (STEM or non-STEM) in participants' a) similarity identification and b) wishful identification with their virtual mentor?

RQ4: Are there any effects of participants' gender in participants' perceptions of their virtual mentor (i.e., virtual mentor's credibility, virtual mentor's ability to facilitate learning, similarity identification and wishful identification with the virtual mentor)?

Finally, in order to provide more richness to our quantitative study (which has been summarized in a theoretical model – see Figure 2), we decided to include an analysis of data related to gameplay in the STEM gaming website, and qualitative data from interviews as well. Coming from a User Experience and design perspective, it is important to try to understand the user and how they perceive the stimuli. Our main goals with the qualitative study are to explore: 1) if virtual mentors would be something of interest for middle schoolers, 2) which types of design for the virtual mentors would be more appealing for middle-schoolers, and 3) how the website and the virtual mentors could be improved.

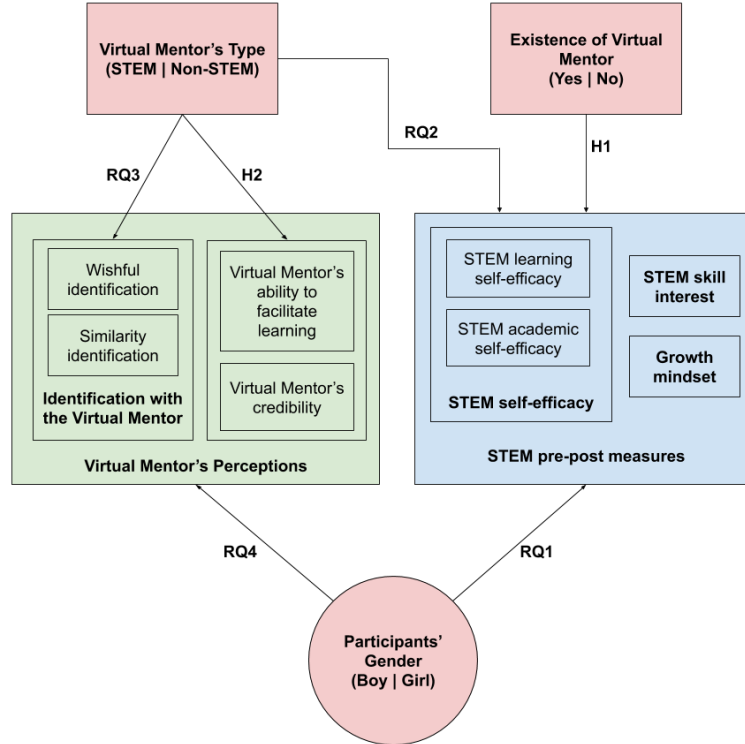


Figure 2. Theoretical Model.

CHAPTER 3

DESIGNING THE STIMULI: A USER EXPERIENCE APPROACH

This exploratory research project has the primary goal of investigating different virtual mentor's designs for a STEM gaming website such as the *STEM Game Crew* website. Thus, the design and development of the experimental websites used as stimuli in this research project focused on 1) the experimental website per se (e.g., look-and-feel, flow, content, structure), and 2) the virtual mentors to be featured in the experimental website.

However, before any design decision was taken, it was necessary to analyze the core elements and the structure of the current *STEM Game Crew* website, which had been used previously in avatar effects research (Cherchiglia, & Ratan, 2018; Cherchiglia et al., 2016a; Cherchiglia et al., 2016b). Then, the virtual mentors were designed, and the new experimental website was designed and built; due to the importance of all visual stimuli, all decisions were based on User Experience (UX) principles and standards, those related not only to website design but also to character design.

In this chapter we cover details regarding: 1) the original *STEM Game Crew Website*, 2) the design of the virtual mentors featured on the experimental website (quantitative study) and those used during interviews (qualitative study), and 3) web design aspects related to the experimental website (e.g., look-and-feel, flow, content, structure).

The STEM Game Crew Website

The *STEM Game Crew* website ¹ was created through a partnership between researchers from Michigan State University (MSU) and WKAR, the public broadcasting station in East

¹ Available at <http://stemgamecrew.org>.

Lansing, Michigan. The website complements WKAR's television program *Curious Crew* ², in which middle-school aged children are invited to explore STEM topics through a hands-on approach (e.g., doing experiments, observing phenomena in the world). The *STEM Game Crew* website showcases 65 STEM digital games that were curated to be topically consistent with *Curious Crew* episodes. It is important to note that the *STEM Game Crew* website does not host nor is involved with the development of the STEM games featured in the website; instead, it provides links to STEM games hosted in external educational websites such as, for example: PBS Kids ³, Nasa Space Place ⁴ and Math Playground ⁵. Figure 3 shows the main page of the *STEM Game Crew* website as of May 21, 2019.

Besides supporting *Curious Crew* and serving as a library for STEM games, the *STEM Game Crew* website uses game tracking as an opportunity to teach players about the scientific method – see Figure 4. Specifically, before playing a game through the website, users are asked to create hypotheses based on the information displayed about the game (e.g., game's name and description). These hypotheses are related to the game being considered fun, simple, capable to teach something, and able to explain science ideas. Then, after playing the game, users return to the website and report reflective observations based on their gameplay experiences. These observations are the same questions asked before in the hypotheses step (i.e., fun, simple, capable to teach something, able to explain science ideas) but now users can support their claims with evidence as they have played the game themselves. Users can also see the average score of that game, a composite measure from all users' ratings. Based on Social Cognitive Theory

² Available at <http://www.wkar.org/programs/curious-crew>

³ Available at <https://pbskids.org>

⁴ Available at <https://spaceplace.nasa.gov>

⁵ Available at <https://www.mathplayground.com>

(Bandura, 1989), applying the scientific method in real-world situations is expected to motivate students to develop an inquisitive mindset and to help augment students' interest in STEM fields.

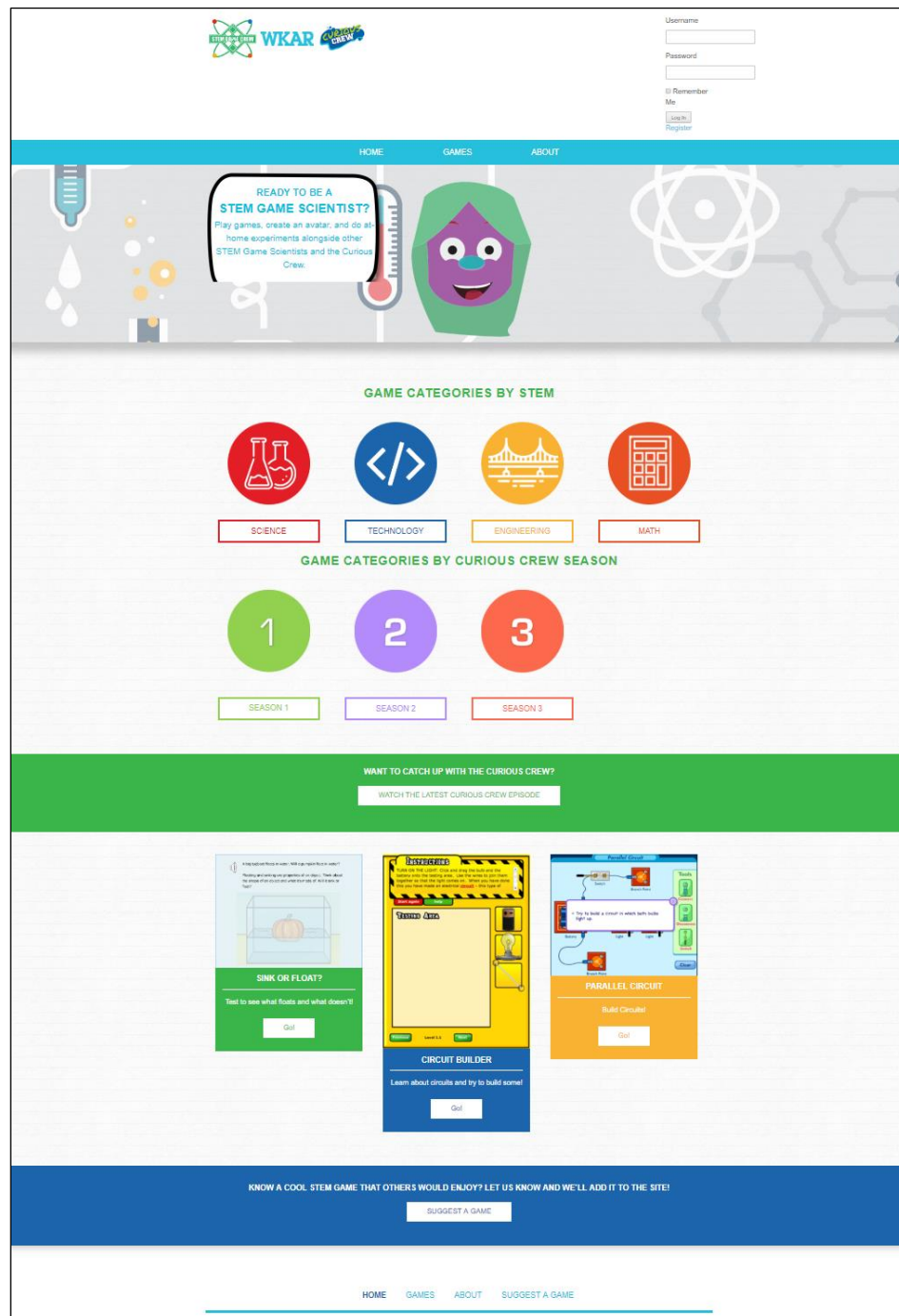


Figure 3. Main page of the *STEM Game Crew* website.





Username

Password

☐ Remember Me

[Register](#)

HOME
GAMES
ABOUT

SINK OR FLOAT?

1. MAKE A HYPOTHESIS



A big lightbulb floats in water. Will a pumpkin float in water?

Floating and sinking are properties of an object. Think about the shape of an object and what it's made of. Will it sink or float?

Test to see what floats and what doesn't!

Based on the picture and description, I think this game will...

...be fun ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

...be simple ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

...teach me something ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

...explain science ideas ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

I made this prediction because the picture looks:

...And the description sounds:

2. TEST YOUR HYPOTHESIS

Test your hypothesis by playing the game.

Don't forget to come back when you're done!

3. MAKE CONCLUSION

After playing, I think that the game...

...was fun ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

...was simple ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

...taught me something ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

...explained science ideas ☐ ☒ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

I also think that this game...

OTHER STEM-GAME SCIENTISTS' OBSERVATIONS

...WAS FUN	☒ ☒ ☒ ☒ ☒ 5/5
...WAS SIMPLE	☒ ☒ ☒ ☒ ☒ 4.5/5
...TAUGHT ME SOMETHING	☒ ☒ ☒ ☒ ☒ 3.5/5
...EXPLAINED SCIENCE IDEAS	☒ ☒ ☒ ☒ ☒ 4/5

HOME
GAMES
ABOUT
SUGGEST A GAME

Figure 4. Example of game page showing the scientific method applied to gameplay in the *STEM Game Crew* website.

To record game tracking, users must complete a free registration step which asks for their username, email, password, nickname, and year of birth. Users must be at least 12 years old and are instructed to not use their real names when registering in the website in order to keep data anonymous. After the registration step, users can customize a virtual representation of themselves (i.e., avatar). A user's avatar will always be displayed in the website's header, on the right top corner (see Figure 5). It is not possible to include the avatars created in the website in the STEM games themselves because, as mentioned before, the *STEM Game Crew* website was not involved in any aspects of the development of the STEM games.

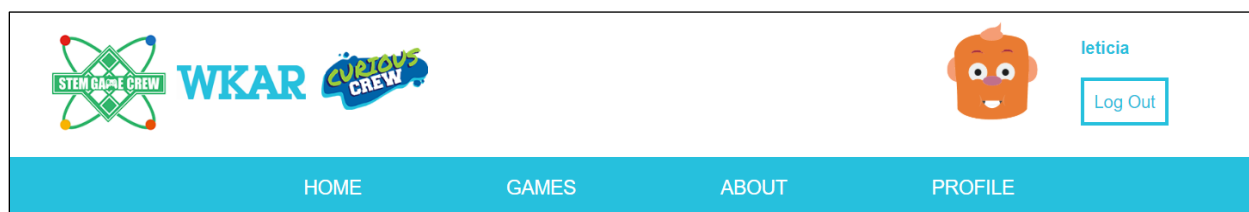


Figure 5. Avatar location on the *STEM Game Crew* website.

Currently, avatars in the *STEM Game Crew* website are colorful, simple-looking, non-gendered, and can be customized in terms of face (4 options), hair (12 options), eyes (10 options), eyebrows (5 options), mouth (2 options), and nose (13 options), thus yielding up to 62,400 different avatars. Examples of possible avatars can be seen in Figure 6 below.



Figure 6. Possible construction for avatars in the *STEM Game Crew* website.

Virtual Mentors' Design

The way people perceive and react to characters has been linked not only to cultural factors but also to characters' aesthetics and social behaviors. According to Isbister (2006), because people unconsciously use visual cues in order to make assumptions about one's role and abilities, a psychological tool such as the interpersonal circumplex is very useful when designing characters. The author suggests an interpersonal circumplex where social behavior is mapped along four axes: dominance, extroversion, friendliness, and connectedness – see Figure 7.

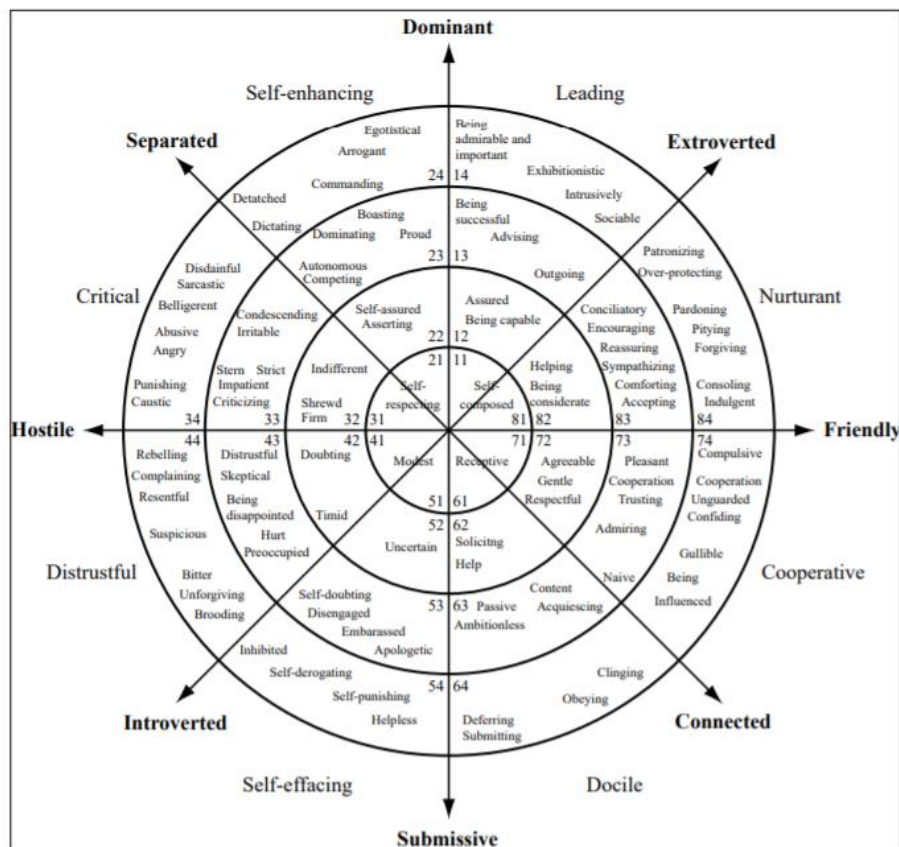


Figure 7. Isbister's (2006) interpersonal circumplex for character design.

Depending on the way characters are drawn (e.g., face, body, clothing), and the type of behaviors these characters demonstrate through the use of verbal and non-verbal language, people will perceive them as more dominant (or submissive), extroverted (or introverted), friendly (or hostile), and connected (or separated). For example, characters who smile often,

show steady but not overly intense eye contact, have an open body stance, and warmer tone of voice are more likely to be perceived as friendly than characters who never smile, show intense eye contact, have a tense body stance, and cold tone of voice. Indifferent characters, on the other hand, show signs of non-engagement or lack of interest, such as avoiding eye contact, having closed body stance, and less energetic tone of voice. Another example is that baby faces (Figure 8, left) have been linked to nurturing feelings such as care and warmth, but not feelings of accountability; if the designer wants the character to be perceived as someone independent and responsible, characters with more mature features (Figure 8, right) would be better suited (Isbister, 2006).

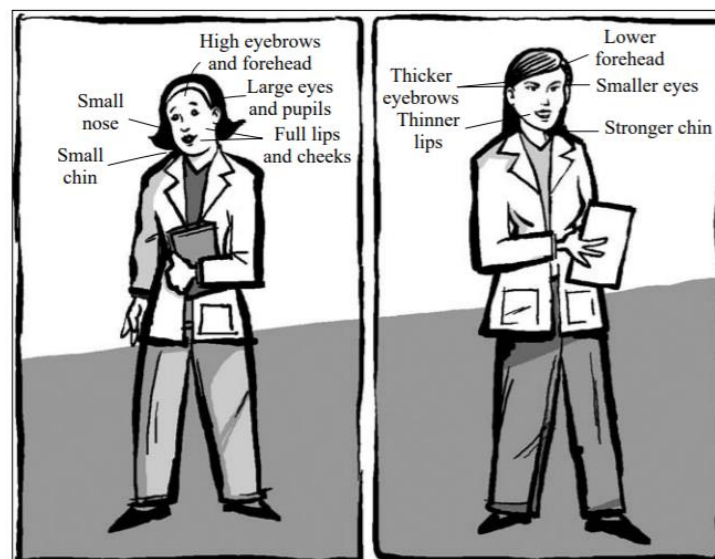


Figure 8. Baby face (left) versus non-baby face (right) - Isbister (2006).

When designing the virtual mentors for this research project, we took such standards into consideration as we wanted participants to be able to perceive their virtual mentors as role models (i.e., someone successful and competent, but also sharing similarities with the participant; Kao & Harrell, 2016). When considering adolescents' wishful identification with scientific characters in the media, there is evidence that boys and girls identify more with female characters who are depicted as respected, caring, and dominant, and with male characters who

are depicted as intelligent, respected, and dominant (Steinke et al., 2012). Another factor to be taken into consideration was the PALD model (Heidig & Clarebout, 2011) and its different levels (global, medium, and detail), which guided our design and the documentation of such design.

Drawing from such factors, and the fact that our target population was composed by middle-schoolers, it seemed that the best approach was to design the virtual mentors (both STEM and non-STEM) in a humanistic and cartoonish way, based on the image of young professional adults (i.e., more mature facial features, not overly attractive, well-groomed, dressed up professionally and conservatively). See Figures 9 and 10 below for examples; all assets were purchased from Good Studio ⁶ under a standard license (modifiable, non-commercial use only) and modified by the researcher through editing software.

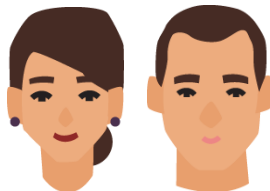
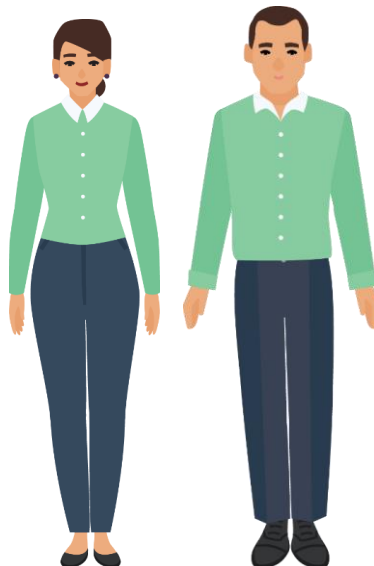


Figure 9. Virtual mentor's head design.



⁶ Available at https://creativemarket.com/Good_Studio

Figure 10. Virtual mentor's body design.

Regarding the social behaviors of the virtual mentors, since only static images could be used - instead of animations - we built scenes with the goal of making the virtual mentors to look like as lively, friendly, and engaging as possible, while still maintaining the role of a knowledgeable mentoring person (i.e., performing actions such as reading, typing in a computer, teaching). Thus, virtual mentors always maintained eye contact with participants or turned their gaze towards the activity they were performing in the scene (e.g., reading), besides being depicted with smiling faces and a relaxed, open body stance. See Figure 12 in the next page for an example of scene, and Appendix A for a listing of all images. The only time virtual mentors were depicted sad was when the feedback given to participants was of a disappointed nature because participants reported not putting enough effort into playing the game - see Figure 11 below for examples of feedback images showed to participants.

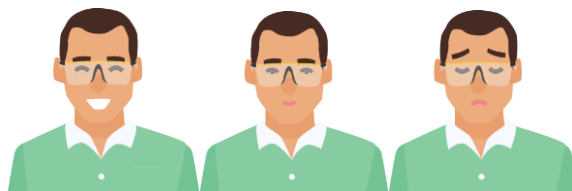


Figure 11. Examples of STEM virtual mentor's images related to feedback.

Another design decision was to try to minimize any kind of stereotypes linked to STEM careers, gender, or race. Therefore, for the STEM virtual mentors in the experimental website (quantitative part of the study), it seemed better to consider the idea of a STEM worker than specific careers in STEM fields (which carry their own stereotypes). Consequently, objects from different STEM professions were included in the same scene (e.g., microscope, civil engineering hat, mathematical tools). Non-STEM virtual mentors were always depicted holding or accompanied by regular common-place office objects (e.g., books, scribbles on the white board). A small backstory was created in order to further engage participants: in the STEM

version of the experimental website, the virtual mentor would present him/herself as a researcher within STEM fields, while in the non-STEM version, the virtual mentor wouldn't mention her/his profession. See Figure 12 for examples of scenes used in the experimental website – all can be found in Appendix A. For the qualitative study we were interested in how the different STEM professions would matter to the participants thus the scenes were designed in a way that objects, clothing and accessories would match the depicted STEM career – see Figure 13 as an example; all figures are listed in Appendix B.



Figure 12. Examples of STEM (left) and non-STEM (right) virtual mentor scenes.

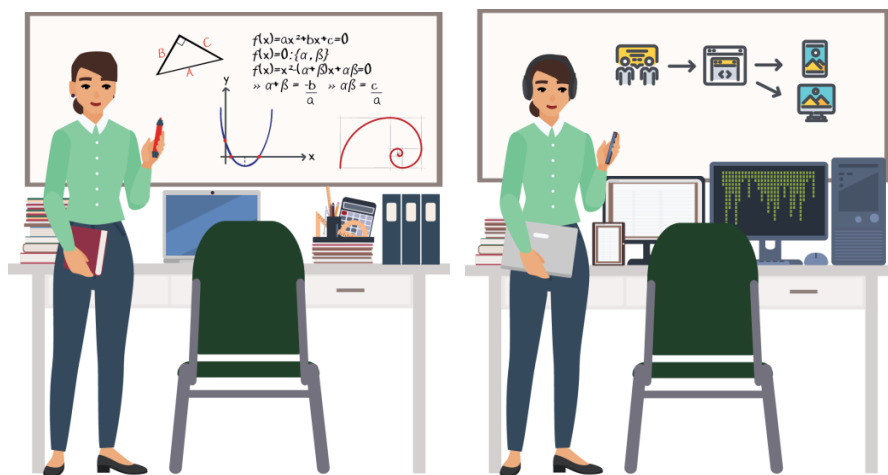


Figure 13. Examples of a Mathematician (left) and a web developer (right).

Regarding gender and race, for the quantitative study we chose for all virtual mentors an olive-toned skin color, together with brown hair and brown eyes; the name Alex was chosen because of its non-gendered nature. For the qualitative study, when creating different skin colors (white and black) we decided to be as consistent and realistic as possible, thus racial features were included (e.g., Black/African Americans usually have natural curly hair and wider noses than White/Caucasians). See Figure 14 below.



Figure 14. Different skin colors for the virtual mentor's head design.

When considering colors for clothing items as well as the website, we chose a neutral-gendered color scheme based on current known standards of web design and character design (i.e., most women and men like the colors blue and green). See Appendix A for a listing of all colors used in the design of the virtual mentors.

The Experimental Website

Look-and-Feel. When designing the experimental website for this research project, we did copy over some elements from the original *STEM Game Crew* website such as the *STEM Game Crew* logo, the background paper-like image, the use of primary colors, and the icons representing Science, Technology, Engineering, and Math. Other elements were slightly altered, for example, the scientific method applied to gameplay (i.e., make a hypothesis, test your hypothesis, make a conclusion) was referred instead as “Claim-Evidence-Reasoning” because such language was used in the middle-school where the experiment would take place. Another

example is that games were rated using a “Yes/No” input instead of a 5-point Likert scale in order to ensure valid and non-neutral responses.

The bulk of the experimental website, however, was built from scratch through Wix in order to fit the research project scope and to be consistent with UX standards for usability and accessibility. Wix was chosen because of its ease of design and use, besides the features of high customization, internal database, and password protected pages. The five versions of the experimental website (one for each condition: 1) no virtual mentor, 2) STEM virtual mentor woman, 3) STEM virtual mentor man, 4) non-STEM virtual mentor woman, 5) non-STEM virtual mentor man) were under premium paid plans which removed ads and provided free domains. The first page of all versions of the website was password protected.

Flow. The first noticeable change was in regards with the flow of the website. While in the *STEM Game Crew* website users can browse through pages with the aid of a menu, the experimental website was built following traditional educational experiences in a survey-like design (i.e., lack of menu, navigation forward via “Next” buttons or linked images). The goal was to minimize any kind of potential distracting factors in order to have a more controlled environment. Such decisions also helped the experiment to happen in a timely manner and to fit the 45-minutes timeframe corresponding to one class period of data collection. The whole website was composed of ten pages, where participants would:

1. **Login:** input the password (given by the researchers);
2. **Welcome:** meet the virtual mentor (if in the virtual mentor conditions) and input their participants’ ID;
3. **Explanation of scientific method:** learn about the scientific method applied to gameplay (Claim-Evidence-Reasoning);

4. **Game choice:** pick one STEM game to be played out of two options;
5. **Claim:** create hypotheses about the game;
6. **Evidence:** be redirected to the game's page and play it for 10 minutes;
7. **Reasoning:** draw conclusions about the game, and report their effort and success related to gameplay;
8. **Feedback and fun fact:** receive feedback from their virtual mentor related to their self-reported effort and success (if in the virtual mentor conditions) and read a scientific fun fact;
9. **STEM Careers:** learn more about STEM careers (i.e., applied math, architecture engineering, biomedical engineering, civil engineering, computer engineering, data processing, digital media, engineering technology, and web design);
10. **Final:** say goodbye to their virtual mentor (if in the virtual mentor conditions) and be redirected to the post-survey.

A wireframe was built via the Realtime Board-Miro UX tool in order to map and prototype the website; the complete wireframe can be seen in Appendix C-**Figure A28** and online⁷. Virtual mentors' images related to all pages are listed in Appendix A.

It is important to note that in the Game Choice page, participants were given the option to choose one out of two STEM games. The games were Feeding Frenzy⁸, an action game about fighting cancer developed by Tinime Games and HopeLab, and Bumper Ducks⁹, a puzzle Physics game developed by Filament Games in partnership with the Smithsonian Science Education Center. These games were chosen because of their solid educational nature, the fact that they were hosted in websites with no ads, and the possibility to be played during a 10-minute

⁷ Available at https://miro.com/app/board/o9J_kzT-SUc=

⁸ Available at http://www.re-mission2.org/games/#/feeding_frenzy

⁹ Available at https://ssec.si.edu/resource_launch/604

period. We still wanted to give participants an option because, according to Bandura (1989), agency is a pivotal factor when considering one's motivation to learn.

Because both games were related to Science, we decided to include fun facts related to technology (i.e., how computers work) and engineering/math (i.e., how rollercoasters are built). Only one fun fact was shown in the feedback and fun fact page, to be automatically chosen randomly by the website once the page loaded.

Due to the survey-like nature of the website and the fact that gameplay was automatically timed (10 minutes), we could have all participants performing the same tasks roughly at the same time; the only variance was regarding participants' reading speed and how much time they decided to spend exploring the STEM careers page. For such page, once participants clicked in a specific STEM career, a pop-up would open showing information about the selected career (e.g., description, average salary, examples of one woman and one man who are famous in the field) – see Figure 15. The information was displayed as cards from the STEM card game Tech Trek (Cherchiglia, Jorae, Zhao, Zhang, & Heeter, 2017), which was designed by the researcher as part of a serious game design course with the goal to make available to middle-schoolers information about technology related careers.

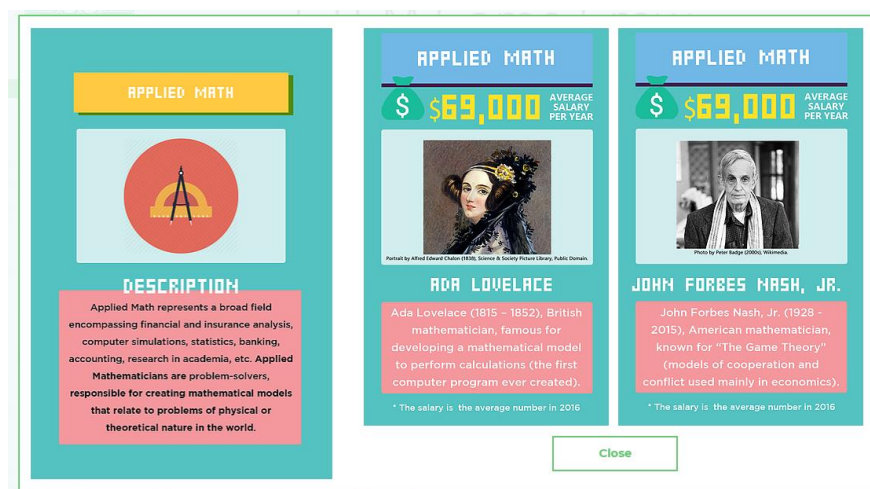


Figure 15. Example of pop-up window showing information about a STEM career.

Virtual mentor's placement in the website. Previous studies regarding avatar design and use for the STEM Game Crew website (Cherchiglia, & Ratan, 2018; Cherchiglia et al., 2016a; Cherchiglia et al., 2016b) suggested that participants could potentially benefit from a stronger visual stimulus (i.e., avatars would not only be displayed in the website's header). Thus, we decided that the virtual mentor would be shown in every page of the experimental website, in the left side of the screen, in an area taking up approximately 40% of the page. The only two exceptions to this rule was the game choice page (virtual mentor displayed on the right side) and the feedback & fun fact page (virtual mentor displayed twice in different positions, i.e., first on the left side of the feedback message, and then on the right side of the fun fact message). See Figure 16 for an example of Welcome page.

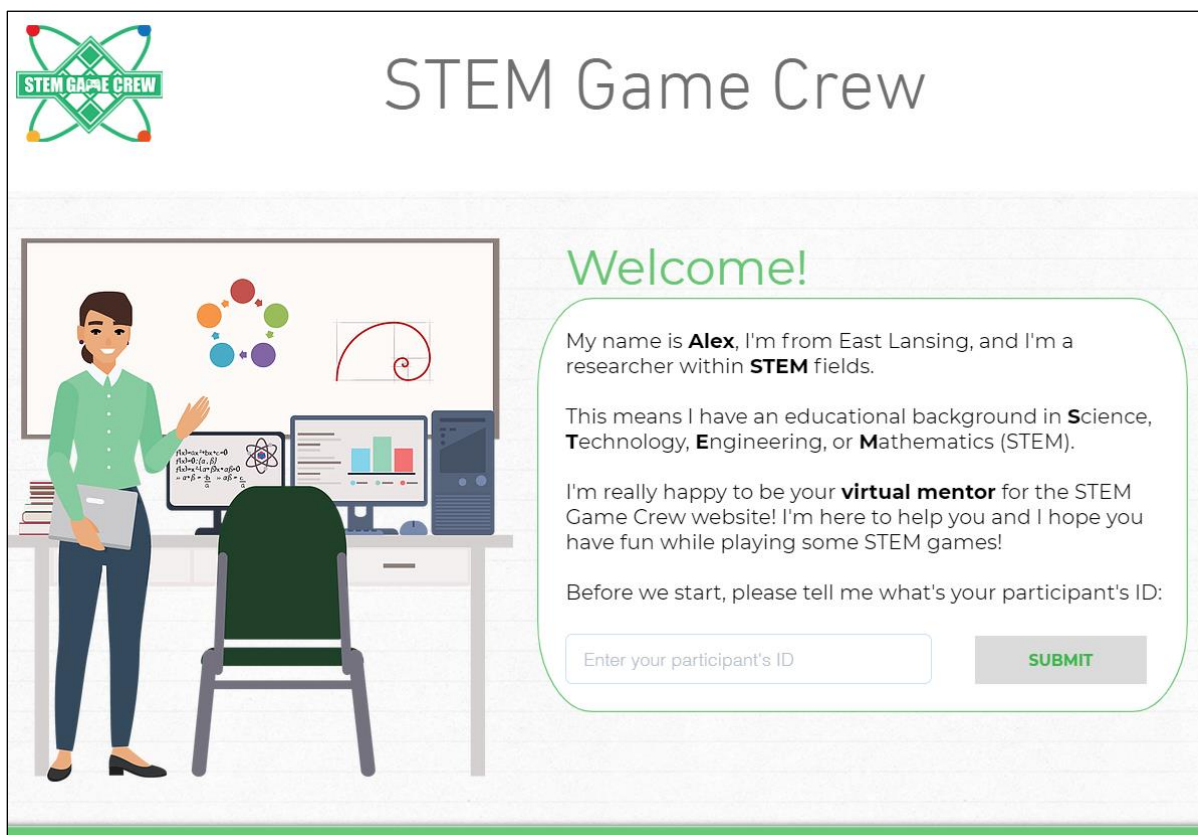


Figure 16. Welcome page for participants with a STEM virtual mentor (woman).

As can be seen in Figure 17, texts were displayed as if they were messages coming from the virtual mentor. For participants without a virtual mentor, all texts were designed as regular instructions in a non-personified way. For example, instead of displaying the message “Before we start, please tell me what's your participant's ID”, the following message would be shown: “Please enter your participant's ID”. Moreover, the feedback messages related to participants’ self-reported perceptions of their gameplay (i.e., how much effort they put into playing the game, and how successful they felt after playing the game) were built based on Growth Mindset standards, as one of the goals of this research project was to potentially increase Growth Mindset. See Table 1 below for all possibilities of feedback given to participants, and Figure 17 for an example of the feedback and fun fact page.

Table 1

Feedback given to participants in the experimental website, based on self-reported success and effort after playing the game; adapted from Mindset Works’ (2017) Growth Mindset Feedback Tool for Learners.

-
1. **When participants succeed with strong effort:** “I'm very proud of you for not giving up, and all the effort you put forth. I hope you are also proud of yourself! I want you to remember how challenging this game might have been when you began, but look how far you've come! Remember, our brains only grow when we challenge ourselves!”
 2. **When participants succeed without effort:** “It's great that you have it down, but it looks like your skills weren't being challenged by this game. I don't want you to be bored for not challenging yourself...I think you're ready for something more difficult. Remember, our brains only grow when we challenge ourselves!”
 3. **When participants did not succeed despite strong effort:** “I admire your persistence and appreciate your mental effort, even though it seems you didn't do as well as you wanted to. When you think you can't do it, remind yourself that you can't do it YET. Don't give up! Let's look at this as an opportunity to learn, a challenge. Remember, our brains only grow when we challenge ourselves.”
 4. **When participants did not succeed and did not put in much effort:** “I'm sorry you're feeling this way. I understand that this game may seem too difficult or too boring at first. But it looks like you're not putting forth much effort, and I would be really happy if you tried harder. After all, if you want to get better in anything, it's going to take effort and practice to get there. Remember, our brains only grow when we challenge ourselves.”
-

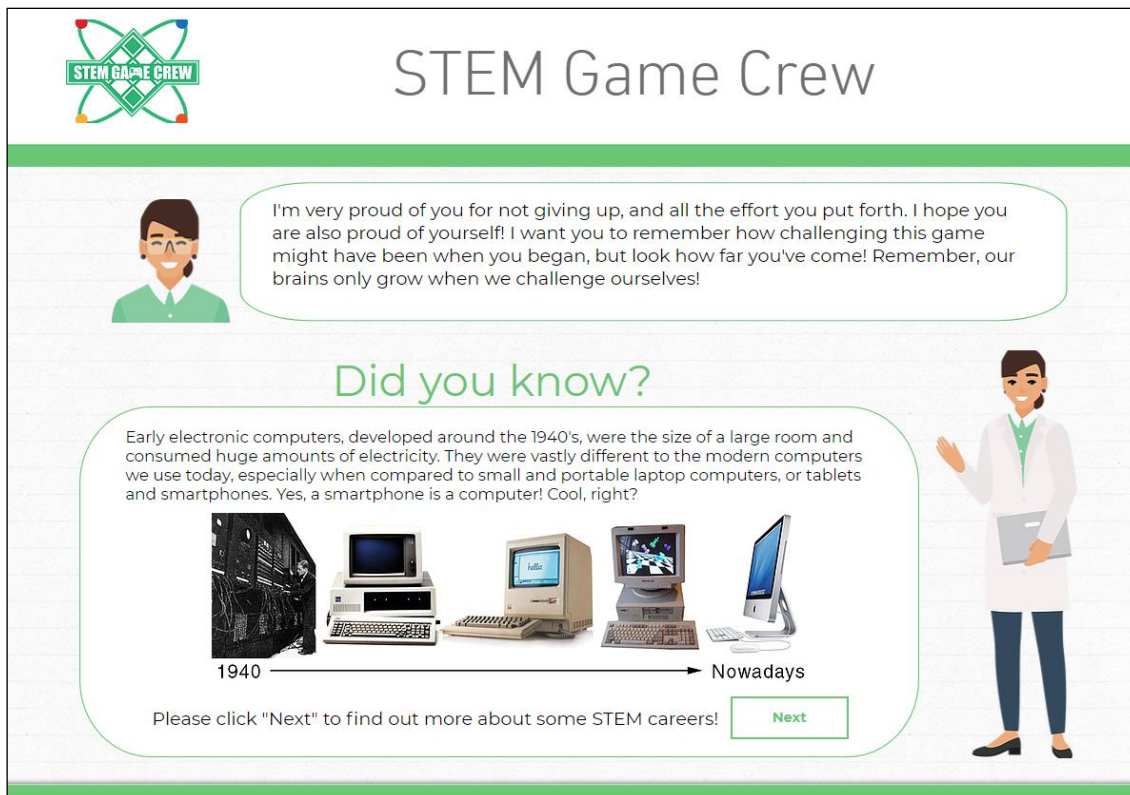


Figure 17. Feedback and fun fact page for participants with a STEM virtual mentor (woman).

Depending on the feedback shown to participants, the image of the virtual mentor would be different. For example, when participants succeed with strong effort, virtual mentors would be very happy, when participants succeed without effort or did not succeed despite strong effort, virtual mentors would be happy, and when participants did not succeed and did not put in much effort, virtual mentors would be disappointed. See Appendix A for images of the virtual mentors in those scenarios.

CHAPTER 4

METHODS

Quantitative Study Design

The quantitative data collection happened on Fall 2018 in a Midwestern Public Middle School. Participants were middle school students from 6th, 7th, and 8th grades attending a mandatory Technology class. Consent forms were sent out two weeks in advance by the Technology teacher, who was also responsible for collecting the signed consent forms prior to data collection. Consent forms were signed by participants' parents or guardians. Participants were asked for their own consent before the experiment started. Participation was voluntary and no incentives were given.

Data collection happened on a regular class day, during five distinct class sessions of 45 minutes each, in a computer lab adjacent to participants' regular Technology classroom. Students not taking part in the study remained in the regular Technology classroom and were assigned other tasks by the Technology teacher. For those taking part in the study, a 2x2 between-subject experiment was conducted by manipulating virtual mentor's (VM) existence (yes or no), and virtual mentor's type (STEM or non-STEM) in the experimental website. Participants were randomly assigned to one out of three conditions (i.e., STEM virtual mentor, Non-STEM virtual mentor, and No virtual mentor); all participants in a class session would have the same virtual mentor's type to avoid possible cross-contamination of the data.

Before starting the experiment, participants were given a piece of paper containing a participant's ID (i.e., "student" followed by a number from 1 to 150) and a shortened web address (URL) for an online pre-survey. To reduce the complexity of this research project, it was

decided that in the two conditions of the website featuring a virtual mentor (i.e., STEM and Non-STEM), the gender of the virtual mentor should match the perceived gender of the participant; the researchers in the room were responsible for making such on-the-spot decision as soon as participants entered the computer lab, thus giving the appropriate piece of paper to participants – see Table 2. Throughout the experiment, participants were asked for their participants' ID, which was used to link all data together.

Table 2.

Listing of Participants' IDs, Experimental Groups, Website Versions, and Pre-Survey URLs.

<u>Participants' ID</u>	<u>Experimental Group</u>	<u>Website Version</u>	<u>Pre-Survey</u>
“student1” to “student30”	STEM VM	STEM VM Woman	URL 1
“student31” to “student60”	STEM VM	STEM VM Man	URL 2
“student61” to “student90”	Non-STEM VM	Non-STEM VM Woman	URL 3
“student91” to “student120”	Non-STEM VM	Non-STEM VM Man	URL 4
“student121” to “student150”	No VM	No VM	URL 5

Regardless of the pre-survey URL given to participants, all pre-surveys contained an introductory page containing information related to participants' consent form, followed by questions about 1) Growth Mindset, 2) STEM learning self-efficacy, 3) STEM academic self-efficacy, and 4) STEM skill interest. The researchers in the room were responsible for reading all questions aloud; participants followed along answering the pre-survey at the same time, a format known to participants because of the middle-school standardized test's approach. This was a suggestion of the Technology teacher in order to ensure proper comprehension of the questions and higher data reliability.

After completing the pre-survey, participants were automatically redirected to one out of five versions of the experimental website (i.e., STEM woman, STEM man, non-STEM woman,

non-STEM man, and no virtual mentor). All versions of the website were password protected; the password was given to participants as soon as they finished the pre-survey. Participants were asked to enter their participants' ID in the first page of the experimental website. The experimental website was covered in detail in Chapter 3, but in summary, the website allowed participants to meet their virtual mentor (if there was one), play one STEM game for 10 minutes (participants' choice), apply the scientific method to the chosen STEM game, and read information about STEM careers.

After using the website for about 25 minutes, participants would press a button on the last page of the website which would automatically redirect them to an online post-survey asking first participants' ID, followed by questions about 1) Growth Mindset, 2) STEM learning self-efficacy, 3) STEM academic self-efficacy, and 4) STEM skill interest. Participants were then asked if they recalled having a virtual mentor in their version of the experimental website; if the answer was positive, a new set of questions would be triggered, namely: 1) recall of virtual mentors' design, 2) virtual mentors' perceived ability to facilitate learning, and 3) virtual mentors' perceived credibility, and 4) feelings of identification (similarity and wishful) with their virtual mentor. Finally, demographic questions were asked to all participants.

The post-survey questions were not read aloud to participants as done with the pre-survey because of the additional virtual mentor questions shown only if participants recalled having a virtual mentor, leading to different post-survey lengths thus different completion times. Although this question added complexity to the study design, it served as an attention check besides potentially increasing the reliability of the answers given to the virtual mentor questions. Once participants were finished with the post-survey, if there was still time, they could choose between either play more of the STEM games or go back to the main Technology classroom.

Qualitative Study Design

The qualitative data collection happened on the same day and classroom of the quantitative data collection, however after class. Consent forms were sent out two weeks in advance by the Technology teacher, who was also responsible for collecting the signed consent forms prior to data collection. Consent forms were signed by participants' parents or guardians.

Due to the exploratory nature of this research project, semi-structured interviews were conducted, and the total time of data collection was of approximately 75 min (15 minutes per interviewed participant). Students not taking part into the experiment were assigned other educational tasks by their Technology teacher and remained in the regular Technology classroom. Light refreshments were provided to all participants.

Upon entering the room, participants were asked for their consent to participate in the study, also to have their interview recorded (audio-only). Then, participants signed a sign-in sheet. The interview was facilitated using 36 printed images numbered in the back showing possible options for the design of the virtual mentors - see Appendix B for a listing of all images. Figure 18 below shows the four images used to represent the virtual mentors seen in the two versions of the experimental website featuring virtual mentors (i.e., STEM or Non-STEM). The remaining 32 images showed more options for the design of virtual mentor in regards with skin color (i.e., light, medium, dark) and STEM professions (i.e., chemist, web developer, civil engineer, mathematician). Figure 19 below shows some examples.



Figure 18. Representations of virtual mentors used in the experimental website.



Figure 19. Examples of more options for the virtual mentors (i.e., different skin colors and different STEM professions).

Interview questions are listed in detail in Appendix E-Table 16; in summary, participants were asked about their virtual mentor in the experimental website (if they had one), their experience using the experimental website, input regarding different designs for the virtual mentors, and suggestions for improvements related to the experimental website and the virtual mentors. All interviews were transcribed with the aid of a paid online audio to text transcription service called Temi¹⁰.

Participants

Quantitative Data. From the original sample of 61 middle-school students, data from five participants was discarded because of concerns over quality and reliability. Regarding the remaining participants (N = 56), most of them were 7th graders (85.7%), followed by 8th graders (12.5%) and only one participant in 6th grade (1.8%). Participants' age ranged from 11 to 14 years old (M = 12.38). Most participants reported their ethnicity as White/Caucasian (73.2%), while 14.2% of the participants preferred not to answer the ethnicity question (other reported groups were: 5.4% Hispanic/ Latino, 3.6% Mixed, 1.8% Native American, and 1.8% Asian). Regarding gender, 48.2% of the participants reported to be a girl, while 46.4% reported to be a boy (other reported groups were: 3.6% for Other, and 1.8% for Prefer not to Answer).

Because gender is a characteristic of interest in this research project, Table 3 (next page) shows participants' distribution among different experimental groups; as a reminder, for the two groups featuring a virtual mentor, the researchers in the room were responsible for matching the gender of the virtual mentor with the perceived gender of the participant.

¹⁰ Available at <https://www.temi.com>.

Table 3.

Quantitative data: participants' distribution among experimental groups according to Virtual Mentor (VM) gender and participants' reported gender.

<u>Experimental Group</u>	<u>VM Gender</u>	<u>Participants' Reported Gender</u>				<u>Total</u>
		<u>Boy</u>	<u>Girl</u>	<u>Other</u>	<u>Prefer not to Answer</u>	
No VM	---	9	6	2	0	17
Non-STEM VM	Woman	0	8	0	0	8
	Man	9	0	0	1	10
	Total	9	8	0	1	18
STEM VM	Woman	0	13	0	0	13
	Man	8	0	0	0	8
	Total	8	13	0	0	21
Total		26	27	2	1	56

Qualitative Data. For the qualitative portion of this research project, originally 14 students were scheduled to be interviewed after class, on the same day of the quantitative data collection. All these students took part into the quantitative data collection, as the interview was going to be about their experience in the STEM gaming website and their evaluation of virtual mentors' images. Only five students attended the interview session, potentially because the interviews happened after class. All participants (N = 5) gave permission to have their interviews recorded (audio-only). All participants were 7th graders. Participants' age ranged from 12 to 13 years old (M = 12.2); all participants were White/Caucasian, and 4 out of 5 participants were boys. Table 4 shows participants' distribution in relation with the previous experimental groups used in the quantitative data collection.

Table 4.

Qualitative data: participants' distribution among experimental groups according to Virtual Mentor (VM) gender and participants' gender.

<u>Experimental Group</u>	<u>VM Gender</u>	<u>Participants' Gender</u>		
		<u>Boy</u>	<u>Girl</u>	<u>Total</u>
No VM	---	1	0	1
Non-STEM VM	Woman	0	0	0
	Man	2	0	2
	Total	2	0	2
STEM VM	Woman	0	1	1
	Man	1	0	1
	Total	1	1	2
Total		4	1	5

Measures

Growth Mindset. We used the Mindset Assessment Profile designed by Dweck and available for free in the Mindset Works website (2017) in order to measure Growth Mindset. Participants were asked to rate how much they agreed with the statements; items were rated on a 6-point Likert-type scale from “Disagree a lot” to “Agree a lot”. The same scale was used in the pre and post surveys. It is important to note that this assessment profile asks questions related to both Growth Mindset (e.g., “No matter how much intelligence you have, you can always change it a good deal.”) and Fixed Mindset (e.g., “You can learn new things, but you cannot really change your basic level of intelligence.”); in our statistical analyses we only considered the four items related to Growth Mindset as this is our construct of interest. The four items were averaged into a single measure for the pre-survey ($\alpha = .61$) and post-survey ($\alpha = .70$). See Appendix D-Table 11 for a listing of all items.

STEM self-efficacies. To measure participants' self-efficacy related to learning STEM topics (STEM learning self-efficacy) and academic achievement in STEM classes (STEM academic self-efficacy), we adapted Pintrich and De Groot (1990)'s self-efficacy scale in order to fit to a STEM context. The scale used asked questions related to Science, Math and Technology classes; Engineering is not a regular middle-school class. Participants were asked "When you think about [Science][Math][Technology] and your [Science][Math][Technology] classes, how much do you agree with the following statements?" with items rated on a 5-point Likert-type scale from "Strongly disagree" to "Strongly agree". An example of item for STEM learning self-efficacy would be "I am confident in my ability to learn new scientific concepts" and an example of item for STEM academic self-efficacy would be "I am confident in my ability to do very well in my Science classes" - see Appendix D-Table 12 for a listing of all items.

The same scale was used in the pre and post surveys. STEM learning self-efficacy was composed by nine items (i.e., three for Science, three for Math, and three for Technology) which were then averaged into a single measure for the pre-survey ($\alpha = .84$) and post-survey ($\alpha = .90$). A similar procedure was done for STEM academic self-efficacy resulting in a single measure for the pre-survey ($\alpha = .89$) and post-survey ($\alpha = .86$).

Although there was a strong Pearson correlation between STEM learning self-efficacy and STEM academic self-efficacy (pre-survey: $r(56) = .854$, $p < .001$; post-survey: $r(56) = .825$, $p < .001$), and the Cronbach's alpha was high when considering the two self-efficacies together (pre-survey: $\alpha = .93$; post-survey: $\alpha = .93$), we still decided to keep both constructs separated as such approach is better aligned with our theoretical framework besides yielding better results in terms of the statistical tests we performed later on.

STEM Skills Interest. As mentioned before, instead of using previous scales for STEM career interest, we decided that the best approach was to adapt the Career Quiz from Washington Career Bridge (2018); more specifically, items related to “Information Technology” and “Science, Technology, Engineering and Mathematics”. Participants were asked to rate how interested they were in fourteen STEM-related activities such as “Doing experiments in a laboratory” or “Thinking of new ways to do things”; items were rated on a 5-point Likert-type scale from “Not interesting at all” to “Extremely interesting”. See Appendix D-Table 13 for a listing of all items. The same scale was used in the pre and post surveys and the 14 items were averaged into a single measure for the pre-survey ($\alpha = .82$) and post-survey ($\alpha = .82$).

Virtual mentor’s ability to facilitate learning and virtual mentor’s credibility. In order to measure participants’ perceptions of their virtual mentor in terms of their ability to facilitate learning and their credibility, we adapted items from two out of four dimensions of the API (Agent Persona Instrument) proposed by Baylor and Ryu (2003). The two dimensions used were “facilitating learning” and “credibility”; “engaging” and “human-like” dimensions were not used in this study. The adapted scale consisted of ten items total and was only used in the post survey, and only answered by participants who recalled having a virtual mentor. These participants were asked to rate the extent to which they agreed to statements such as “The virtual mentor made the instruction interesting” (facilitate learning) and “The virtual mentor was intelligent” (credibility) – see all items in Appendix D-Table 14. Items were rated on a 5-point Likert-type scale from “Strongly disagree” to “Strongly agree”. The five items for facilitate learning were averaged into a single measure ($\alpha = .91$) and the five items for credibility were also averaged into a single measure ($\alpha = .92$).

Identification with the virtual mentor. In order to measure Similarity Identification and Wishful Identification with the virtual mentor, we adapted items from two out of three dimensions of Van Looy et al.'s (2012) scale for game character identification. The two dimensions used were “similarity identification” and “wishful identification”; “embodied presence” was not used. The adapted scale consisted of eleven items total and was only used in the post survey, and only answered by participants who recalled having a virtual mentor. These participants were asked to rate the extent to which they agreed to statements such as “My virtual mentor is like me in many ways” (similarity identification) and “If I could become like my virtual mentor, I would” (wishful identification) – see all items in Appendix D-Table 15. Items were rated on a 5-point Likert-type scale from “Strongly disagree” to “Strongly agree”. The six items for similarity identification were averaged into a single measure ($\alpha = .90$) and the five items for wishful identification were also averaged into a single measure ($\alpha = .95$).

Website measures. In the experimental website, participants were asked to answer before and after gameplay “Yes/No” questions related to the STEM game being considered fun, simple, capable to teach something, and able to explain STEM ideas. Participants also answered, post gameplay only, an open-ended question regarding their reasoning, besides “Yes/No” questions related to how much effort they had put in playing the game, and if they believed they had been successful while playing the game. Moreover, participants were given the choice to pick one out of two STEM games to play for about ten minutes. The website also provided a behavioral measure related to how many and which STEM careers were clicked when participants were in the “Know more about STEM careers” optional page. It’s important to note that in all performed statistical analysis, the website data did not influence the results, thus website measures were not mentioned when reporting results for the quantitative data.

CHAPTER 5

RESULTS

Quantitative Data

A summary of all findings for this subsection can be found in Appendix F-Tables 17 and 18. Before we started our analysis, a series of repeated-measures tests were run across all participants, regardless of their gender. Results suggest that, for all participants, the website was able to promote an increase of 3.2% in perceptions of having a **Growth Mindset** ($F(1,55) = 4.02$, $p = .05$, $\eta^2 = .07$; pre-intervention: $M = 4.64$, $SE = .10$; post-intervention: $M = 4.80$, $SE = .103$). Moreover, all participants demonstrated an increase of 2.75% in their belief of being able to learn STEM topics when comparing pre-post measures of **STEM learning self-efficacy** ($F(1,55) = 4.87$, $p = .03$, $\eta^2 = .08$; pre-intervention: $M = 4.10$, $SE = .07$; post-intervention: $M = 4.21$, $SE = .08$). There was no significant difference in pre-post measures for **STEM academic self-efficacy** and **STEM skill interest** when considering participants altogether.

To answer **RQ1** (“Are there differences in the behavior of STEM variables of interest (i.e., Growth Mindset, STEM self-efficacies, and STEM skill interest) when considering participants’ gender (boys or girls)?”), a series of repeated-measures ANOVA tests were run across all participants looking for a significant gender interaction. No significant gender interaction was found when analyzing pre-to-post change in **Growth Mindset** and **STEM academic self-efficacy**.

Regarding pre-post measures of **STEM learning self-efficacy**, a marginally significant gender interaction was found ($F(1,51) = 3.78$, $p = .06$, $\eta^2 = .07$). Girls showed an increase of 5.75% in their reported belief of being able to learn STEM topics (pre-intervention: $M = 4.00$,

SE = .10; post-intervention: M = 4.23, SE = .11) when compared to boys, who showed a very small increase (0.5%) in STEM learning self-efficacy (pre-intervention: M = 4.28, SE = .10; post-intervention: M = 4.30, SE = .12).

Finally, there was a significant gender interaction in pre-post measures of **STEM skill interest** ($F(1,51) = 4.675$, $p = .04$, $\eta^2 = .08$). Girls reported to be more interested in performing STEM skills due to website use – an increase of 3% (pre-intervention: M = 3.59, SE = .11; post-intervention: M = 3.71, SE = .12). Boys showed a decrease of 2.5% in STEM skill interest (pre-intervention: M = 3.97, SE = .11; post-intervention: M = 3.87, SE = .12).

In order to delve further, we ran these tests **separately for girls and boys**. Results indicated that girls experienced greater **Growth Mindset** ($F(1,26) = 4.58$, $p = .04$, $\eta^2 = .15$), **STEM learning self-efficacy** ($F(1,26) = 7.41$, $p = .01$, $\eta^2 = .22$), and **STEM skill interest** ($F(1,26) = 8.78$, $p = .01$, $\eta^2 = .25$) after using the website than before using it. Specifically, due to website use, girls showed an increase of 3.8% in perceptions of Growth Mindset, an increase of 5.75% in their belief of being able to learn STEM topics, and an increase of 3% in being interested in performing STEM-related skills. There was no significant difference in pre-post measures of **STEM academic self-efficacy** when considering girls only.

When considering boys only, there was no significant differences in pre-post measures of any of the variables of interest. It seems that the results found previously, when considering all participants together, were mostly driven by girls' behavior; moreover, there is evidence that different results might be found when considering boys and girls in different experimental groups (i.e., no virtual mentor, STEM virtual mentor, non-STEM virtual mentor). Table 5 shows means and standard errors for all variables and groups of interest. In summary, results suggest that girls and boys are indeed affected differently by website use, which casts light on **RQ1**.

Table 5.

Means and Standard Error on the Measures of Pre/Post Growth Mindset, STEM self-efficacies and STEM Skill Interest within gender groups.

<u>Measures</u>	<u>Boys (N = 26)</u>		<u>Girls (N = 27)</u>	
	M	SE	M	SE
Pre Growth Mindset (6-point Likert scale)	4.64	.17	4.66*	.13*
Post Growth Mindset (6-point Likert scale)	4.73	.17	4.85*	.12*
Pre STEM Learning self-efficacy (5-point Likert scale)	4.28	.07	4.00*	.12*
Post STEM Learning self-efficacy (5-point Likert scale)	4.30	.09	4.23*	.13*
Pre STEM Academic self-efficacy (5-point Likert scale)	4.30	.08	4.07	.15
Post STEM Academic self-efficacy (5-point Likert scale)	4.24	.10	4.19	.13
Pre STEM Skill Interest (5-point Likert scale)	3.97	.10	3.59*	.12*
Post STEM Skill Interest (5-point Likert scale)	3.87	.10	3.71*	.13*

*p < .05 for pre-post comparison within gender group

To answer **H1** (“Participants assigned to the version of the STEM gaming website with a virtual mentor (when compared to those without a virtual mentor) will demonstrate higher a) Growth Mindset, b) STEM self-efficacies, and c) STEM skill interest”), a series of repeated measures ANOVA tests were performed and results suggest no significant effects of virtual mentor existence on pre-to-post change in **Growth Mindset** ($F(1, 54) = .51, p = .48$), **STEM learning self-efficacy** ($F(1, 54) = .75, p = .39$), **STEM academic self-efficacy** ($F(1, 54) = .58, p = .45$), and **STEM skill interest** ($F(1, 54) = .42, p = .52$). Thus, **H1a**, **H1b**, and **H1c** were not supported.

We decided to add participant’s gender into the model used to examine **H1**. A series of repeated measures ANOVA tests were performed, and results suggest no significant interaction effects of virtual mentor existence and participant’s gender on pre-to-post change in **Growth Mindset**, **STEM learning self-efficacy**, and **STEM academic self-efficacy**. No lower order

interactions were significant for these variables of interest. A significant effect of participant's gender only on pre-post change in **STEM skill interest** was found ($F(1,49) = 4.82, p = .03, \eta^2 = .09$), similar to the finding of **RQ1** thus not considered. Tests for between-subjects differences on the post-test measures were not significant; **H1a**, **H1b**, and **H1c** remain unsupported.

To answer **RQ2** (“Are there differences in the behavior of STEM variables of interest (i.e., Growth Mindset, STEM self-efficacies, and STEM skill interest) when considering virtual mentor's type (STEM or non-STEM) and participants' gender (boy or girl)?”), we must take into account only participants who had a virtual mentor while using the website.

A series of repeated measures ANOVA tests were performed, and results suggest no significant effects of virtual mentor type and participants' gender on pre-to-post change in **Growth Mindset**. No lower order interactions were significant.

There was a significant effect of participants' gender only on pre-to-post change in **STEM learning self-efficacy** ($F(1,34) = 5.76, p = .02, \eta^2 = .15$). When compared to boys, girls showed an increase of 7.25% in their reported belief of being able to learn STEM topics due to website use (pre-intervention: $M = 4.04, SE = .12$; post-intervention: $M = 4.33, SE = .12$). When compared to girls, boys showed a very small decrease (.25%) in STEM learning self-efficacy due to website use (pre-intervention: $M = 4.28, SE = .10$; post-intervention: $M = 4.30, SE = .12$).

There was a marginally significant three-way interaction effect for virtual mentor type and participants' gender on pre-post measures of **STEM academic self-efficacy** ($F(1,34) = 3.37, p = .08, \eta^2 = .09$). No lower order interactions were significant. When considering boys who had a STEM virtual mentor, the website was able to promote an increase of 3.5% in boys' belief of being able to achieve a good performance in STEM-related disciplines (pre-intervention: $M = 4.21, SE = .22$; post-intervention: $M = 4.35, SE = .16$); on the contrary, for boys who had a non-

STEM virtual mentor; there was a decrease of 4% in STEM academic self-efficacy (pre-intervention: $M = 4.46$, $SE = .21$; post-intervention: $M = 4.30$, $SE = .15$;).

Girls showed a different behavior: a higher increase (6%) in STEM academic self-efficacy was seen among girls with a non-STEM virtual mentor (pre-intervention: $M = 4.22$, $SE = .22$; post-intervention: $M = 4.46$, $SE = .16$) while there was also an increase, although smaller (1.5%), in STEM academic self-efficacy when considering girls who had a STEM virtual mentor (pre-intervention: $M = 4.07$, $SE = .17$; post-intervention: $M = 4.13$, $SE = .13$).

The graphs below provide a visualization of pre-post boys' and girls' perceptions of STEM academic self-efficacy, organized by different virtual mentor types: participants with STEM virtual mentors (Figure 20) and those with non-STEM virtual mentors (Figure 21).

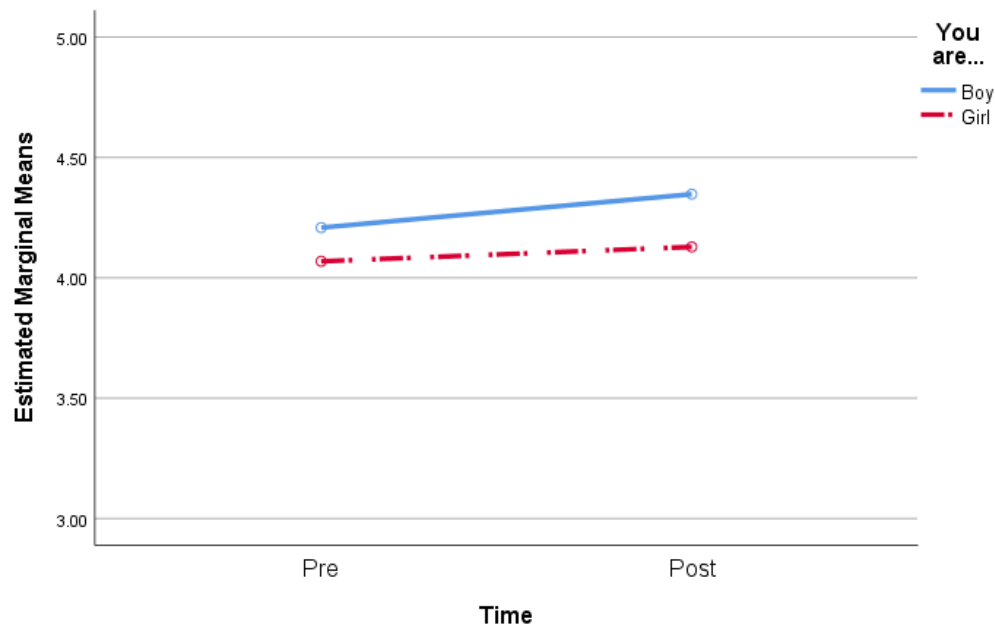


Figure 20. STEM academic self-efficacy behavior for the group who had STEM virtual mentors related to participants' gender.

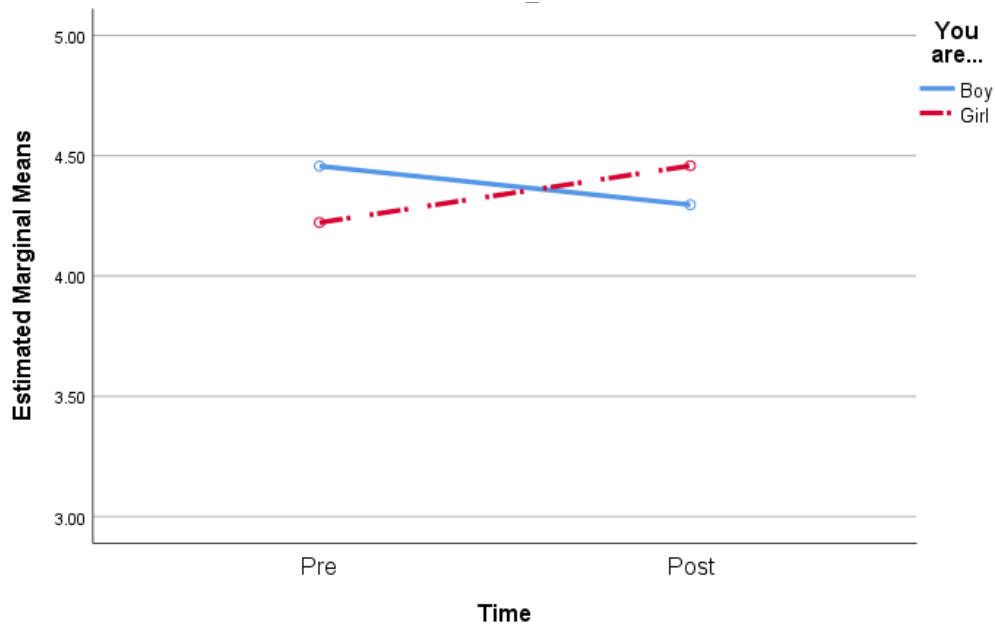


Figure 21. STEM academic self-efficacy behavior for the group who had non-STEM virtual mentors related to participants' gender.

There was a significant three-way interaction effect for virtual mentor's type and participants' gender on pre-to-post change in **STEM skill interest** ($F(1,34) = 5.62, p = .02, \eta^2 = .14$). Boys with STEM virtual mentors reported to be more slightly more interested (1.75% increase) in performing STEM-related skills after website use ($M = 3.71, SE = .21$) when compared to before website use ($M = 3.64, SE = .20$); on the contrary, boys who had non-STEM virtual mentors showed lower values of STEM skill interest after website use ($M = 3.95, SE = .20$) when compared to before website use ($M = 4.12, SE = .19$) – a decrease of 4.25% in STEM skill interest.

For girls, those who had a non-STEM virtual mentor showed an increase of 5% in STEM skill interest due to website use (pre-intervention: $M = 3.68, SE = .20$; post-intervention: $M = 3.88, SE = .21$); there was also an increase, although smaller (1.25%), when considering girls who had a STEM virtual mentor (pre-intervention: $M = 3.53, SE = .15$; post-intervention: $M = 3.58, SE = .16$).

The graphs below provide a visualization of pre-post boys' and girls' perceptions of STEM skill interest, organized by different virtual mentor types: participants with STEM virtual mentors (Figure 22) and those with non-STEM virtual mentors (Figure 23).

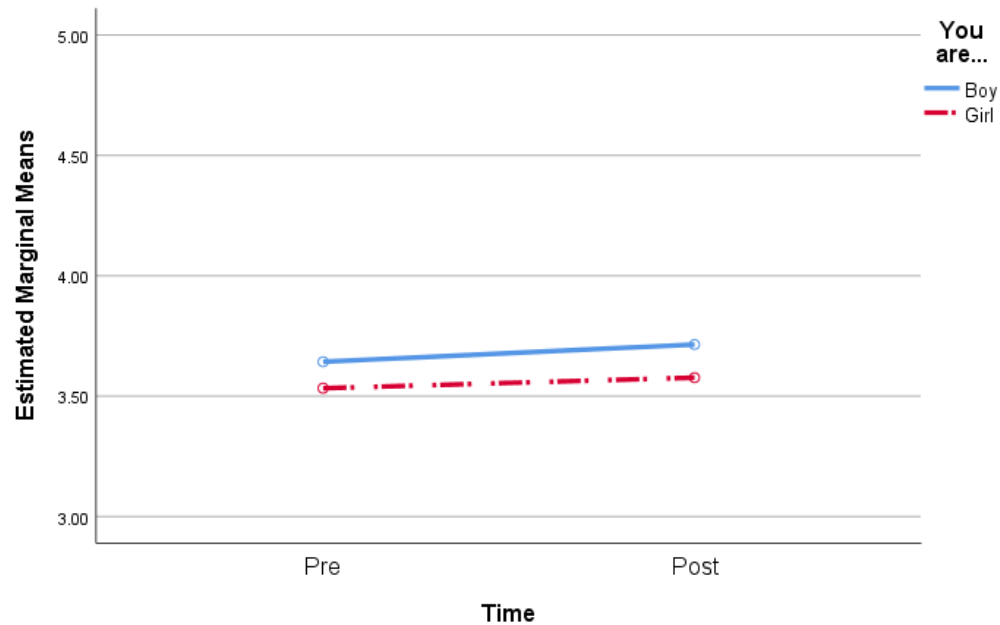


Figure 22. STEM skill interest behavior for the group who had STEM virtual mentors related to participants' gender.

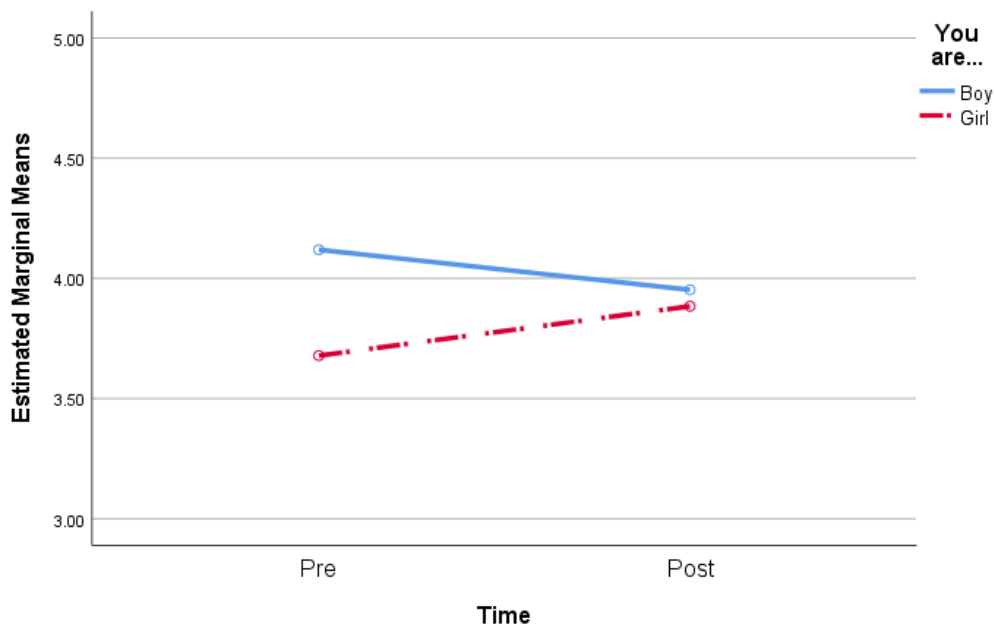


Figure 23. STEM skill interest behavior for the group who had non-STEM virtual mentors related to participants' gender.

There was also a significant interaction effect for participants' gender on pre-to-post change in **STEM skill interest** ($F(1,34) = 4.18, p = .05, \eta^2 = .11$) which follows the same behavior found before for all participants, but it is not interpreted here in light of the significant three-way interaction.

In order to answer **H2** ("*STEM virtual mentors will be perceived as having higher a) ability to facilitate learning, and b) credibility when compared to non-STEM virtual mentors*"), a series of one-way ANOVA tests were performed and results suggest no significant effect of virtual mentor type in participants' perceptions of their virtual mentor in being knowledgeable (i.e., **virtual mentor's credibility**) - ($F(1, 30) = .36, p = .55$). There was also no significant effect of virtual mentor type in participants' perceptions of their **virtual mentor's ability to facilitate learning** ($F(1, 30) = .00, p = .99$). **H2a** and **H2b** are not supported.

In order to answer **RQ3** ("*Are there any effects of virtual mentor type (STEM or non-STEM) in participants' a) similarity identification and b) wishful identification with their virtual mentor?*"), a series of one-way ANOVA tests were performed. Results suggest no significant effect of virtual mentor type in participants' feelings of being similar to the virtual mentor (i.e., **similarity identification**) or participants' feelings of wanting to be like the virtual mentor (i.e., **wishful identification**).

In order to answer **RQ4** ("*Are there any effects of participants' gender in participants' perceptions of their virtual mentor (i.e., virtual mentor's credibility, virtual mentor's ability to facilitate learning, similarity identification and wishful identification with the virtual mentor)?*"), a series of one-way ANOVA tests were performed. Results suggest no significant effect of participants' gender in participants' perceptions of **virtual mentor's credibility, virtual mentor's ability to facilitate learning, participants' similarity identification and wishful**

identification with their virtual mentor. Still, we included Table 6 showing the means and standard error for these variables.

Table 6.

Means and Standard Deviation on the 5-point Likert Measures of Virtual Mentors (VM)' Perceptions within gender groups and virtual mentor type.

<u>Measures</u>	Boys (N = 12)				Girls (N = 18)			
	Non-STEM (N = 6)		STEM (N = 6)		Non-STEM (N = 8)		STEM (N = 10)	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
VM Credibility	3.90	1.11	3.80	1.13	4.28	.71	3.90	.88
VM Ability to Facilitate Learning	3.63	.63	3.80	.90	4.00	.69	3.82	.96
Similarity Identification with VM	2.75	1.06	2.39	.82	3.17	.85	2.72	.89
Wishful Identification with VM	2.77	1.10	2.70	1.32	3.55	.95	3.90	1.00

Having the virtual mentor variables in mind, we decided to try to construct a better model for **RQ2**, by running the series of repeated measures ANOVA tests again but this time considering all participants' perceptions of their virtual mentors as covariates. For **Growth Mindset** and **STEM skill interest**, the new model performed the same or worse than the previous models thus yielding no significant effects.

A better statistical model regarding the effect of participants' gender on pre-to-post change in **STEM learning self-efficacy** was achieved when controlling for similarity identification and wishful identification with the virtual mentor ($F(1,26) = 6.40$, $p = .02$, $\eta^2 = .20$) -- covariates appearing in the model are evaluated at the following values: wishful identification (3.13); similarity identification (2.78). When compared to boys, girls showed an increase of 9% in their reported belief of being able to learn STEM topics due to website use (pre-intervention: $M = 4.0$, $SE = .13$; post-intervention: $M = 4.36$, $SE = .13$). When compared to girls, boys showed

a very small decrease (.75%) in STEM learning self-efficacy due to website use (pre-intervention: $M = 4.28$, $SE = .16$; post-intervention: $M = 4.25$, $SE = .17$).

Both covariates were also significant: wishful identification ($F(1,26) = 4.27$, $p = .05$, $\eta^2 = .14$) and similarity identification ($F(1,26) = 11.91$, $p = .00$, $\eta^2 = .31$). We used a median split for wishful identification ($\tilde{x} = 3.2$) and similarity identification ($\tilde{x} = 3.0$) in order to illustrate directionality.

Considering **wishful identification**, for boys, if wishful identification was high (above median), there was an increase in STEM learning self-efficacy; if wishful identification was low (below median), there was a decrease in STEM learning self-efficacy. In other words, after website use, boys who had higher feelings of wishing to be like their virtual mentor demonstrated an increase in their reported belief of being able to learn STEM topics; on the contrary, boys who had lower feelings of wishing to be like their virtual mentor demonstrated a decrease in their reported belief of being able to learn STEM topics. See Figure 24 below.

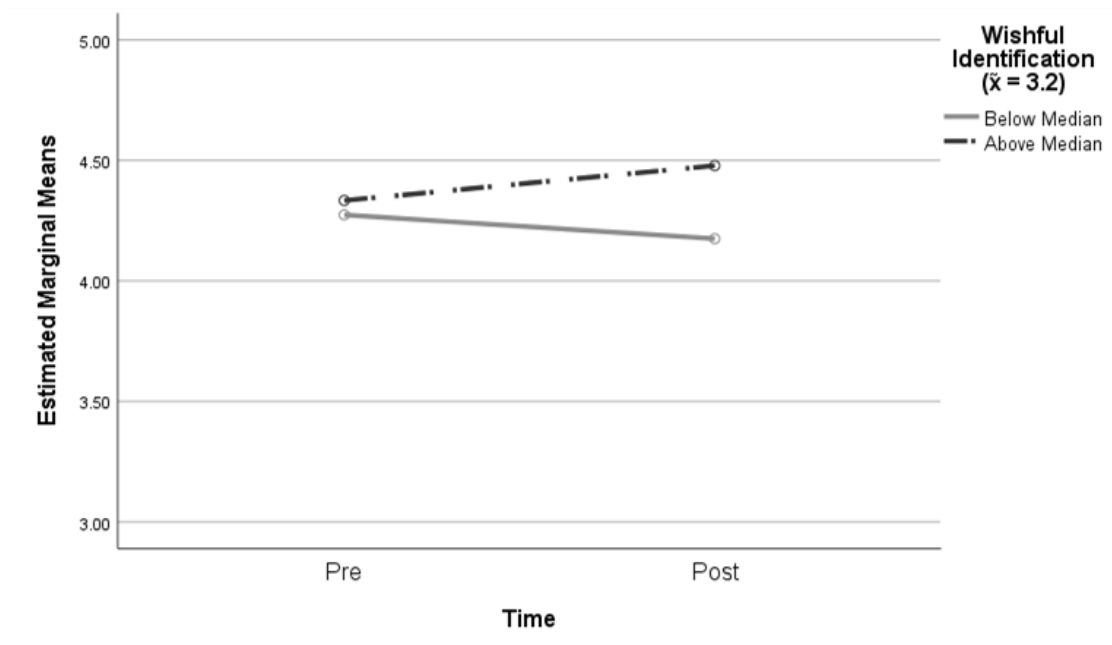


Figure 24. STEM learning self-efficacy behavior for boys related to median-split values of wishful identification.

For girls, there was an increase in STEM learning self-efficacy regardless of wishful identification, although the increase was stronger for those with high wishful identification. In other words, after website use, girls who had higher feelings of wishing to be like their virtual mentor demonstrated a greater increase in their reported belief of being able to learn STEM topics than girls who had lower feelings of wishing to be like their virtual mentor – the increase for those girls was smaller. See Figure 25 below.

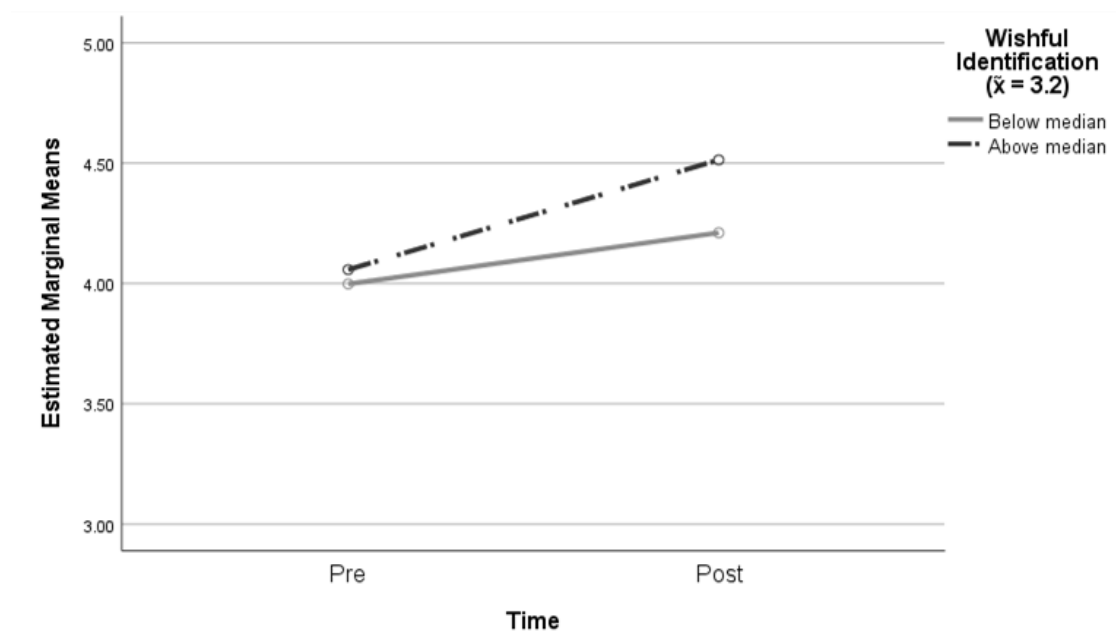


Figure 25. STEM learning self-efficacy behavior for girls related to median-split values of wishful identification.

For **similarity identification**, boys demonstrated an opposite behavior than for wishful identification. Boys with high similarity identification (above median) showed a decrease in STEM learning self-efficacy and boys with low similarity identification (below median) showed an increase in STEM learning self-efficacy. In other words, after website use, boys who had higher feelings of being similar to their virtual mentor demonstrated a decrease in their reported belief of being able to learn STEM topics; on the contrary, boys who had lower feelings of being

similar to their virtual mentor demonstrated an increase in their reported belief of being able to learn STEM topics. See Figure 26 below.

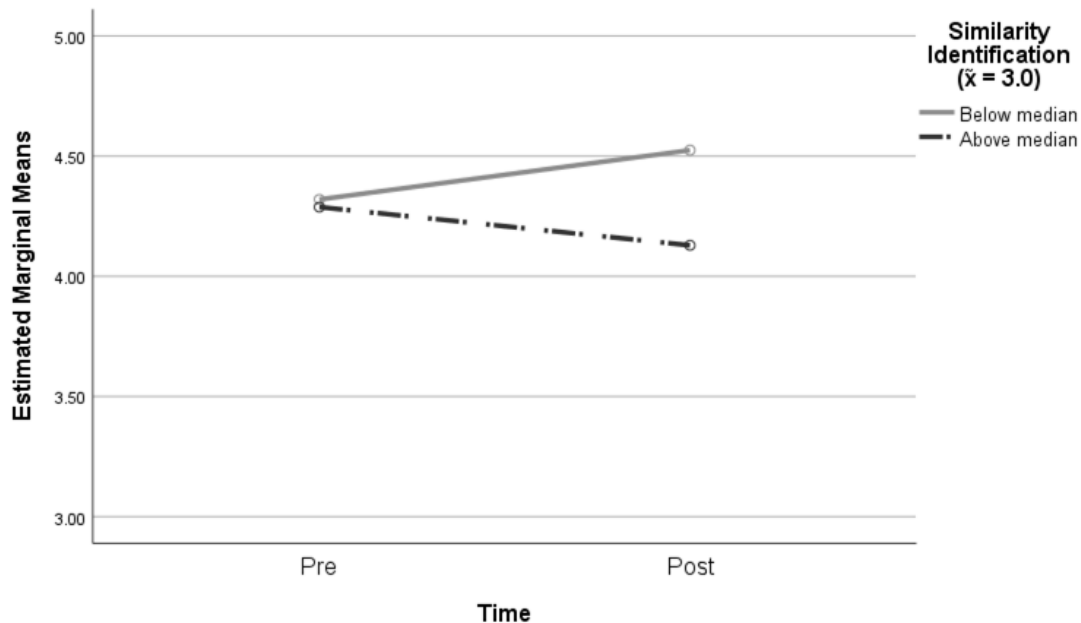


Figure 26. STEM learning self-efficacy behavior for boys related to median-split values of similarity identification.

Similar to their behavior for wishful identification, girls demonstrated an increase in STEM learning self-efficacy regardless of similarity identification. However, girls' behavior when considering low/high similarity identification groups was opposite than for low/high wishful identification groups as the increase this time was stronger for those with low similarity identification. In other words, after website use, girls who had higher feelings of being similar to their virtual mentor demonstrated a smaller increase in their reported belief of being able to learn STEM topics than girls who had higher feelings of being similar to their virtual mentor – the increase for those girls was greater. See Figure 27 below.

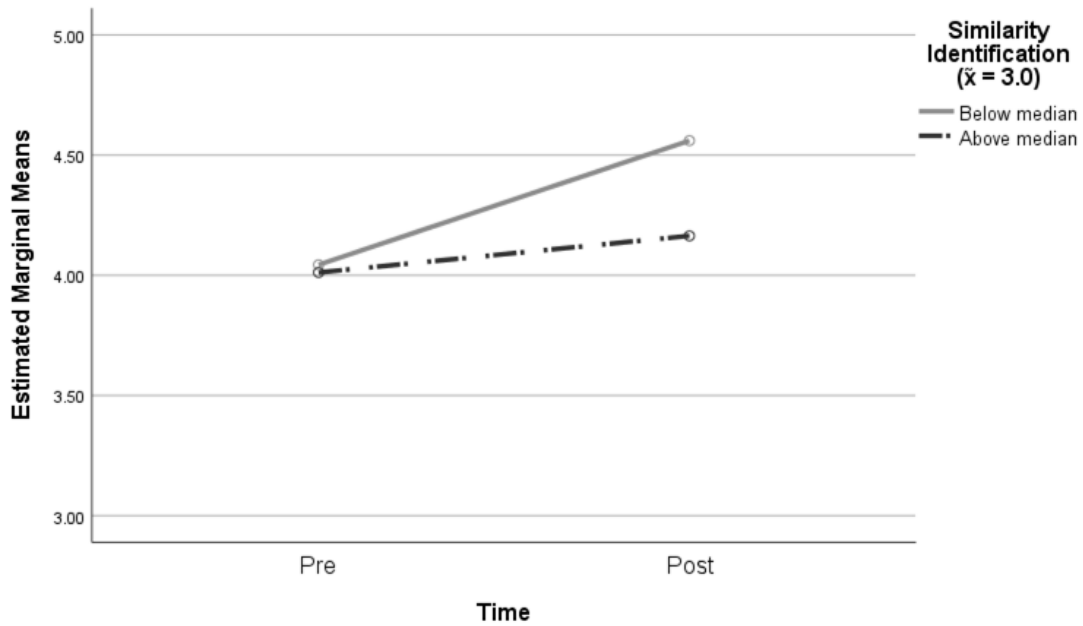


Figure 27. STEM learning self-efficacy behavior for girls related to median-split values of similarity identification.

For **STEM academic self-efficacy**, a better statistical model was achieved when controlling for similarity identification and wishful identification with the virtual mentor; a significant three-way interaction effect for virtual mentor's type and participants' gender was found ($F(1,24) = 4.21, p = .05, \eta^2 = .15$) -- covariates appearing in the model are evaluated at the following values: wishful identification (3.13); similarity identification (2.78). No lower order interactions were significant.

When considering boys who had a STEM virtual mentor, the website was able to promote an increase of 4% in boys' belief of being able to achieve a good performance in STEM-related disciplines (pre-intervention: $M = 4.26, SE = .23$; post-intervention: $M = 4.42, SE = .20$); on the contrary, for boys who had a non-STEM virtual mentor; there was a decrease of 3.75% in STEM academic self-efficacy (pre-intervention: $M = 4.41, SE = .23$; post-intervention: $M = 4.26, SE = .20$). Girls showed a different behavior: a higher increase (6.5%) in STEM academic self-efficacy was seen among girls with a non-STEM virtual mentor (pre-intervention:

M = 4.19, SE = .20; post-intervention: M = 4.45, SE = .18) while there was also an increase, although smaller (2%), in STEM academic self-efficacy when considering girls who had a STEM virtual mentor (pre-intervention: M = 4.12, SE = .18; post-intervention: M = 4.20, SE = .16).

Both covariates were significant: wishful identification ($F(1,24) = 15.08, p = .001, \eta^2 = .386$) and similarity identification ($F(1,24) = 19.36, p < .001, \eta^2 = .45$). We used a median split for wishful identification ($\bar{x} = 3.2$) and similarity identification ($\bar{x} = 3.0$) in order to illustrate directionality; the same value for the median was used for boys and girls. The same patterns found for STEM learning self-efficacy were found for STEM academic self-efficacy when considering boys' and girls' feelings of wanting to be like their virtual mentor and feelings of being similar to their virtual mentor – see Figures 28-31.

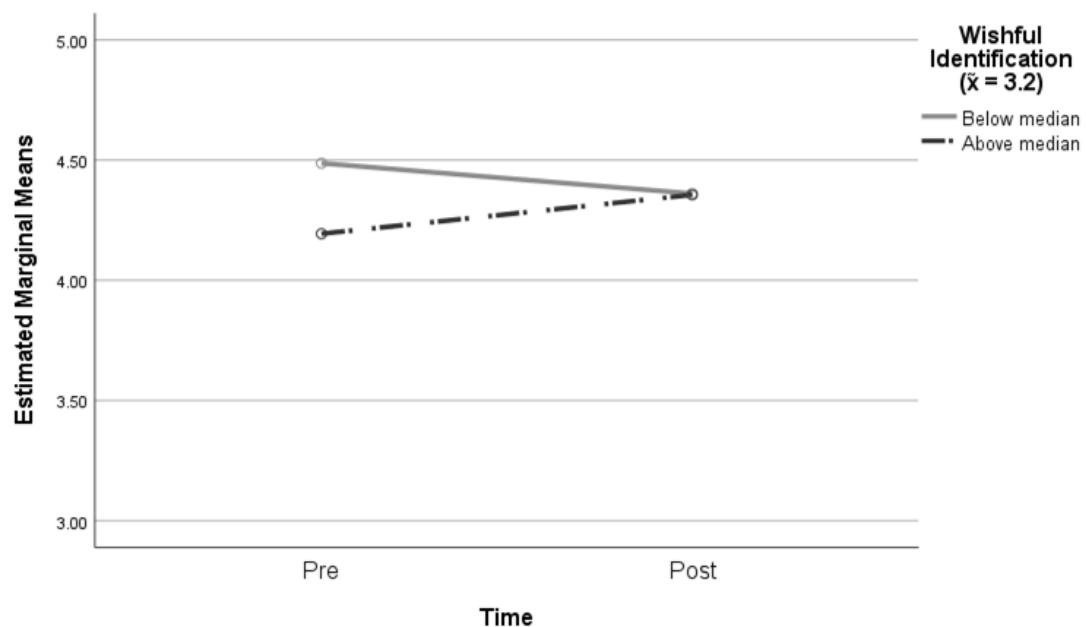


Figure 28. STEM academic self-efficacy behavior for boys related to median-split values of wishful identification.

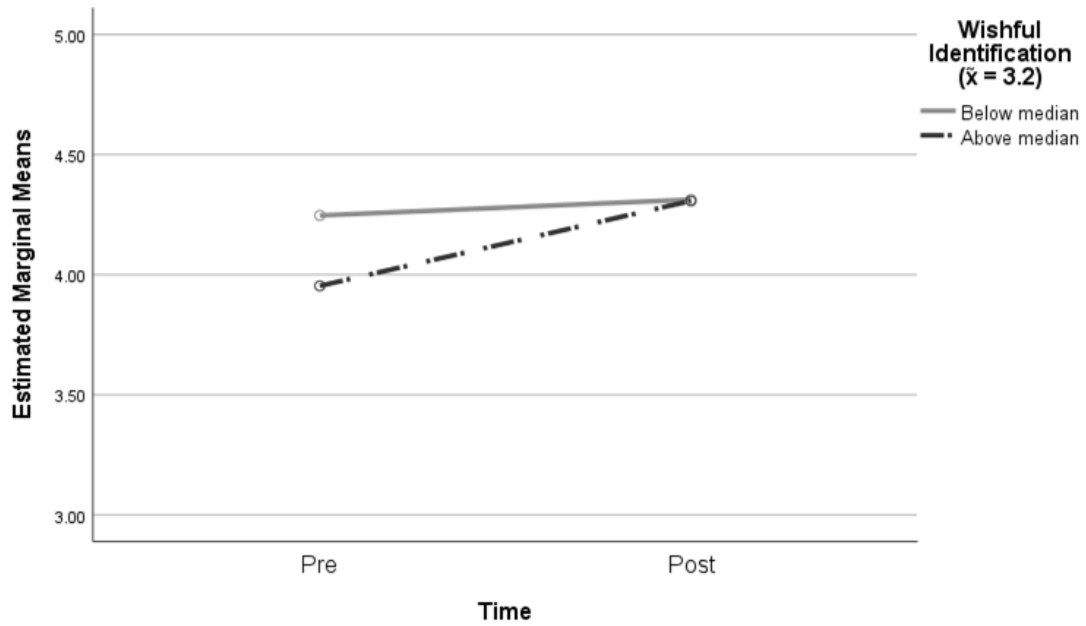


Figure 29. STEM academic self-efficacy behavior for girls related to median-split values of wishful identification.

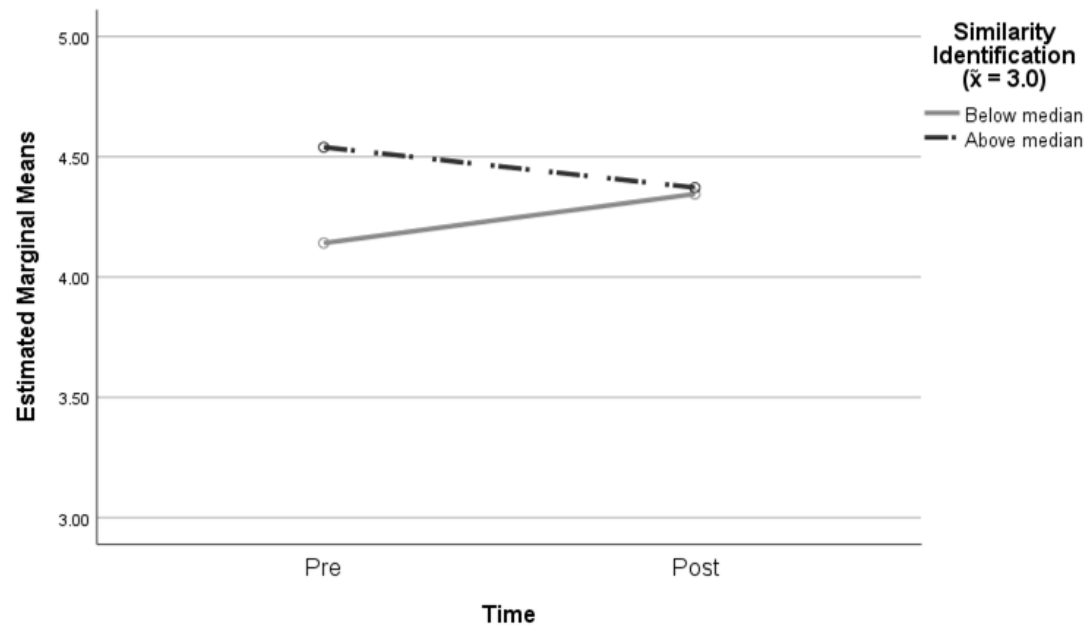


Figure 30. STEM academic self-efficacy behavior for boys related to median-split values of similarity identification

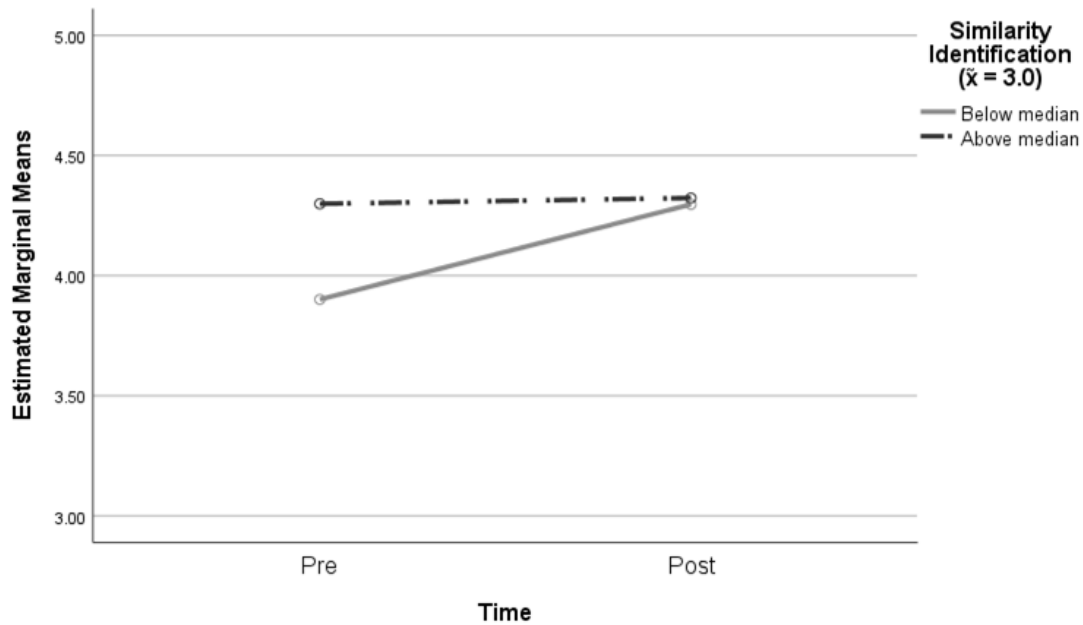


Figure 31. STEM academic self-efficacy behavior for girls related to median-split values of similarity identification

In summary, boys with high wishful identification and low similarity identification had an increase in STEM academic self-efficacy, while boys with low wishful identification and high similarity identification had a decrease. Girls always demonstrated increases in STEM academic self-efficacy, although the increase was greater for those with high wishful identification and low similarity identification.

Website Data

Given the exploratory and design nature of this research project, we decided to include findings regarding the website data in this results section. It is important to note that when considering participants' gender in the analysis, we did not include those who answered "Other" or "Prefer not to Answer" for the gender question. Also, as mentioned before, we conducted exploratory analysis and found no relevant statistical connection between the website data and the pre-post survey data; thus both datasets were analyzed separately.

During the experiment, participants had to choose one out of two STEM game to be played for 10 minutes; 66% of participants chose Bumper Ducks (the puzzle Physics game) while 34% of participants chose Feeding Frenzy (the action cancer-fighting game). A very high proportion of boys chose Bumper Ducks (76.9%) instead of Feeding Frenzy (23.1%), whereas for girls there was an almost even distribution between the two games although Bumper Ducks was still the most preferred (51.9% versus 48.1% of Feeding Frenzy).

Next, we investigated the scientific method applied to gameplay in the STEM gaming website, composed by the “Yes/No” measures asked before and after gameplay. We were interested in participants’ perceptions of the game being fun, simple, capable to teach something, and able to explain STEM ideas, also if those perceptions would change after gameplay. A series of exact McNemar's tests were performed.

When considering Bumper Ducks ($N = 37$), there was a statistically significant difference in participants’ pre-post gameplay perceptions of the game being capable of teaching something ($p = .004$) and able to explain STEM ideas ($p = .001$). Overall, most participants (43.2%) thought Bumper Ducks was capable of teaching something and maintained their opinion after gameplay, followed by 37.8 % of participants who thought the game was not capable of teaching something but ended up changing to a positive opinion after gameplay. Moreover, most participants (43.2%) thought Bumper Ducks was not able to explain STEM ideas but changed to a positive opinion after gameplay, followed by 40.5% of participants who thought the game was able to explain STEM ideas and maintained their opinion after gameplay. For Feeding Frenzy ($N = 19$), no statistically significant difference in participants’ pre-post gameplay perceptions was found. Table 7 shows all distributions for both games.

Table 7.

Distribution of Participants by Game Played, Dimensions, and Pre-Post Gameplay Perceptions.

<u>Game Played</u>	<u>Dimensions</u>	<u>Pre-Post Gameplay Perceptions</u>			
		Yes-Yes	No-Yes	No-No	Yes-No
Bumper Ducks (N = 37)	Fun	86.5	10.8	0	2.7
	Simple	67.6	16.2	8.1	8.1
	Capable of teaching something*	43.2	37.8	13.5	5.4
	Able to explain STEM ideas*	40.5	43.2	10.8	5.4
Feeding Frenzy (N = 19)	Fun	94.7	5.3	0	0
	Simple	42.1	31.6	21.1	5.3
	Capable of teaching something	78.9	5.3	0	15.8
	Able to explain STEM ideas	73.7	15.8	10.5	0

* differences were statistically significant with $p < .05$

When considering boys regardless of game played (N = 26), there was a statistically significant difference in pre-post gameplay perceptions of the game being able to explain STEM ideas ($p = .012$). Most boys (46.2%) thought the STEM game played was able to explain STEM ideas and maintained their opinion after gameplay, followed by 38.5% of boys who thought the game was not able to explain STEM ideas but changed to a positive opinion after gameplay. For girls (N = 27), there was a statistically significant difference in pre-post gameplay perceptions of the game being simple ($p = .021$) and able to explain STEM ideas ($p = .021$).

Most girls (48.1%) thought the STEM game played was simple and maintained their opinion after gameplay, followed by 33.3% of girls who thought the game was not simple but changed to a positive opinion after gameplay. Moreover, most girls (55.5%) thought the STEM game played was able to explain STEM ideas and maintained their opinion after gameplay, followed by 33.3% of girls who thought the game was not able to explain STEM ideas but

changed to a positive opinion after gameplay. Table 8 shows all distributions regardless of game played.

Table 8.

Distribution of Participants by Gender, Dimensions, and Pre-Post gameplay perceptions.

<u>Participants' Gender</u>	<u>Dimensions</u>	<u>Pre-Post Gameplay Perceptions</u>			
		Yes/Yes	No/Yes	No/No	Yes/No
Boys (N = 26)	Fun	88.5	11.5	0	0
	Simple	69.3	11.5	7.7	11.5
	Capable of teaching something	50.0	26.9	7.7	15.4
	Able to explain STEM ideas*	46.2	38.5	11.5	3.8
Girls (N = 27)	Fun	88.8	7.4	0	3.7
	Simple*	48.1	33.3	14.8	3.7
	Capable of teaching something	59.3	25.9	11.1	3.7
	Able to explain STEM ideas*	55.5	33.3	7.4	3.7

* differences were statistically significant with $p < .05$

In order to provide more richness to these findings, we performed a thematic analysis regarding the open-ended question related to post-game gameplay. We looked for four main themes: 1) fun/enjoyment, 2) simple/challenge, 3) teaching something, 4) explanation of STEM ideas. First, we considered data from the 20 boys and 14 girls who played Bumper Ducks. Findings suggest that most boys (70%) and girls (78.5%) thought this game was fun and enjoyed playing it, but one boy and one girl did not consider it fun enough. A considerable proportion of boys and girls considered the game simple (40% of boys and 35.7% of girls), but two boys and one girl considering it too simple. More girls (42.8%) than boys (25%) agreed that the game was able to teach something (e.g., “teach you how to use patterns”, “taught me to try again and use my resources and to keep trying and don’t give up”), although a considerable proportion of boys

(20%) and girls (14.2%) had an opposite opinion. Regarding the games' ability to explain STEM ideas, 30% of boys and 28.5% of girls agreed and even mentioned STEM concepts related to the Physics game (e.g., "how different mass sizes affect other sizes when their force acts upon it", "the physics of momentum and mass and how to use them to get to a certain goal", "you had to think like a scientist would do it by sling shooting the rubber duck across the pond", and "it was showing differences in mass and energy transfer").

Regarding the 6 boys and the 13 girls who played Feeding Frenzy, most boys (83%) and girls (69.2%) thought this game was fun and enjoyed playing it. A considerable proportion of boys (16.7%) and girls (30.7%) thought the game was simple, with one boy and four girls categorizing the game as challenging, while one boy thought the game was confusing. 23% of girls considered the game was able to teach something, while one boy and one girl had an opposite opinion. A considerable proportion of boys (33.3%) and most girls (69.2%) agreed that the game could explain STEM ideas and mentioned STEM concepts related to this Biology game (e.g., "a white blood cell fighting off bacteria as chemo fighting off cancer cells", "scientific bacteria, different blood cells and other characters one could find in an organisms blood stream", "how our body's cells fight off disease and bacteria, which taught me more about science").

Next, we investigated participants' perceptions of being successful in the game and putting effort while playing the game, which were the variables of interest related to Growth Mindset and used by the websites to show different feedback messages to the participants. Chi-square tests were performed, and no significant differences were found between different game types and/or participants' gender. Still, at least 89% participants across all subgroups answered positively to the success and effort questions. Table 9 shows all distributions related to these variables of interest.

Table 9.

Distribution of Participants by Gender, Game Played, and Growth Mindset Perceptions.

<u>Participants' Gender</u>	<u>Game Played</u>	Growth Mindset Perceptions			
		Success		Effort	
		Yes	No	Yes	No
Boys (N = 26)	Bumper Ducks (N = 20)	20	0	16	4
	Feeding Frenzy (N = 6)	5	1	5	1
Girls (N = 27)	Bumper Ducks (N = 14)	14	0	12	2
	Feeding Frenzy (N = 13)	12	1	13	0

Finally, we examined the behavioral measure of clicking in STEM careers to know more about them in the optional website page. Since we had nine different STEM careers, we created four categories related to clicking behavior: no interest (0), low interest (1-3), medium interest (4-6), and high interest (7-9). Most boys (50%) and girls (70.4%) showed a low interest behavior, followed by high interest (boys: 23.1%; girls: 14.8%), no interest (boys: 11.5%; girls: 7.4%), and medium interest (boys: 11.5%; girls: 3.7%). When considering all participants, the STEM careers clicked the most were Architecture Engineering (37.5%) and Digital Media (37.5%), followed by Applied Math (33.9%) and Bio Engineering (33.9%), Web Design (32.1%), Engineering Technology (28.6%) and Civil Engineering (28.6%), Computer Engineering (25%), and Data Processing (19.6%). Chi-square tests were performed, and no significant differences were found between different clicking behavior, different careers, and/or participants' gender.

Qualitative Data

After interview data was transcribed, we performed a thematic analysis looking for the codes shown in Table 10, which are an expansion of the three qualitative questions proposed

before, due to the semi-structure nature of the interviews. In this subsection we discuss our main findings, which do not necessarily relate to all codes. Additionally, it is important to note that we avoided making a distinction between boys and girls since only one participant was a girl, also participants' names were omitted in our analysis (instead we refer to them as P1 through P5).

Table 10.

Coding Scheme Used for the Qualitative Data.

<u>Short Code</u>	<u>Description</u>
Web-Eval	Evaluation of experience when using the website. It can be positive (like) or negative (dislike). It can be related with overall feelings or specific design features.
VM-Recall	Reasons participant recalled virtual mentor used in the website. It can be overall remarks or specific design features.
VM-Pick	Reasons participant picked a specific virtual mentor. It can be overall remarks or specific design features.
VM-Learn	Reasons participant perceives a specific VM as more helpful to facilitate learning than others. It can be overall remarks or specific design features.
VM-Options	Evaluation of why having different options for VM would be good (or bad).
Customization	Evaluation of possible benefits (if any) of VM or background customization. It can be overall remarks or specific design features.
Success	How participant defines success and a successful person.
VM-Success	Reasons participant perceived a specific VM as more successful than others. It can be overall remarks or specific design features.
Web-Improv	Suggestions on how to improve the website. It can be overall remarks or specific design features.
VM-Improv	Suggestions on how to improve the VMs. It can be overall remarks or specific design features.

Regarding participants' evaluation of their experience when using the website, all participants liked the website, mentioning to enjoy playing the games and/or learning about STEM careers. For example, P4 said: "I think it was kind of cool that at the end of your first

game you got to see different careers involving technology and um like, it may make people want to be some of the people that were on the career list”.

Some participants had a hard time recalling their virtual mentor: two out of five participants misremembered their virtual mentor, but one of them was able to recall the right one afterwards, when paying closer attention to details such as background elements. Participants mentioned that what helped them recall their virtual mentor was virtual mentor’s apparel, objects carried by the virtual mentor, and background. When asked about the benefits of having a virtual mentor in the website, two participants mentioned that the virtual mentor served as a social companion because the messages showed in the website seemed more personal (e.g., P1 said: “it looked like he was actually the one that was speaking instead of just having words up on the screen”, P5 said: “it's like talking to you”).

Overall, participants preferred a STEM virtual mentor when compared to a non-STEM Virtual mentor and preferred a virtual mentor who matched their gender. Three out of five participants choose the same virtual mentor they had before in the website. We found evidence of traits of identification with the virtual mentors; for example, P3 said: “Well the spiral is really cool and I also like this right here [points to scientific graph]. Yeah. And I like that she has a book and a laptop because that's like me.”, while P5 said: “The only reason is because he's holding a laptop and nothing else.[...] Because I like technology. He also has some sciency stuff in the background.”. Moreover, some participants mentioned learning reasons when picking their virtual mentor, such as P1, who said: “Well, math is one of my really weak spots in school and so having somebody that does math as a virtual mentor would help a lot and would make me want to do math more.”, and P3, who said: “Well, it features like both math and science, which are like my two favorite subjects...”

Three out of five participants believed the STEM virtual mentor would be better at helping them learn in the website, while one participant (P5) mentioned that “no matter what the virtual mentor is, all the learning is going to be the same”). Interestingly, not all designs for STEM careers were perceived as intended by the researchers. For example, P4 mentioned that the non-STEM virtual mentor looked like a “substitute teacher that will just read, well, and tell you what to do”; moreover, the design chosen for the web developer (i.e., wearing headphones and holding a phone) gave this participant the impression that the virtual mentor was not paying attention, in contrast to the design for the civil engineer and scientist which made such virtual mentors to look prepared to teach STEM to children because of “the diagrams on the background and all their tools on the desk” which would “make it a real life scenario”.

All participants picked STEM virtual mentors instead of non-STEM ones when considering which virtual mentor would be an example of a successful person. Participants sometimes mentioned the objects in the background, specific STEM skills or specific STEM careers as indicators of success. For example: P2 and P5 mentioned coding and “doing science stuff”, while P4 preferred the scientist design because of his ability to “work with all of these chemicals” and “study a lot of stuff”. This is coherent with participants’ vision of what makes someone successful (e.g., mastering a topic, learning from things, being able to teach and explains ideas to someone else).

Regarding participants’ suggestions for improvement in the website, most of them were related to including more educational material along with the games and making the website a “little education game hub” by adding more games to it. Some participants also mentioned that a point system would be interesting in order to unlock virtual mentor’ cosmetics or more games.

Finally, regarding suggestions for improvement related to the virtual mentors, participants demonstrated high interest in being able to pick their own virtual mentor. The connection between user and virtual mentor came through when P1 said that the virtual mentor could be “a person that represents you, instead of having to have like a select person where you cannot change it.”, and P2 mentioned to be interested in choosing a virtual mentor to “fit my person; make me want to hear the person, think like hey this is what they're saying, maybe I should listen to it”; also P4 referred to the virtual mentor as “avatar” many times when brainstorming about customization. Speaking of customization, all participants seemed very interested in being able to customize the virtual mentors’ personality and/or appearance, also the background of the images or items the virtual mentor would be carrying in their hands. Some creative ideas were proposed such as P5’s idea to “coordinate each, um, background and what the people are holding to the thing that you're trying to learn. So if you're trying to learn science then you would put it like on that lab coat and if you trying to learn math, you would put it on that normal one with the math in the background”.

CHAPTER 6

DISCUSSION

Results from the website data analysis suggest that girls and boys interact with the STEM gaming website in similar ways, as both groups: 1) preferred the same STEM game to be played (i.e., Bumper Ducks), 2) reported that both STEM games were fun and able to explain STEM ideas, 3) felt successful and put effort while playing the game, and 4) were not very interested in knowing more about STEM careers on their own. However, quantitative survey data suggests that girls and boys are affected differently by the STEM gaming website when it comes to pre-post change in STEM variables (RQ1). Specifically, while all participants showed an overall increase in Growth Mindset and STEM learning self-efficacy after website use, when compared to boys, girls showed a greater increase in STEM learning self-efficacy than boys, besides an increase in STEM skill interest while boys showed a decrease.

In regards with the design of the virtual mentors, comparison with the control group (i.e., no virtual mentor) yielded no effects (H1). However, increases in both STEM academic self-efficacy and STEM skill interest were seen in boys and girls who had a STEM-looking virtual mentor, while for non-STEM-looking virtual mentors, boys showed a decrease and girls showed an even greater increase (RQ2). Moreover, although no differences were seen in participants' perceptions of the virtual mentor (i.e., credibility, ability to facilitate learning, identification) by virtual mentor's type (H2, RQ3) or participants' gender (RQ4), during interviews all participants preferred STEM virtual mentors which were perceived as more successful and better at facilitating learning. Finally, it seems that high wishful identification and low similarity

identification with the virtual mentor can indeed impact both STEM self-efficacies for both boys and girls.

Interpretation of Results

Findings from **RQ1** suggest that indeed there are differences between girls and boys in relation with pre-to-post changes of some STEM variables of interest. Although when considered together there was an increase in both Growth Mindset and STEM learning self-efficacy, it seems this was a finding driven by girls' behavior.

Given the evidence in the literature connecting girls to negative or lower STEM metrics (e.g., Diekman et al., 2017; Hayes, 2016; MacPhee et al., 2013), it was indeed very exciting to see that girls showed a significant increase in Growth Mindset (3.8%), STEM learning self-efficacy (5.75%), and STEM skill interest (3%) after website use. Data from the website seems to indicate that 33% of girls believed the STEM game was not going to be simple by judging its image and description but changed their opinion after playing it; also, almost all girls reported high feelings of success and putting effort during gameplay. We believe that such findings are consistent with girls' increase in STEM metrics, also a sign that many girls' had low STEM learning self-efficacy before playing the game but were able to change those negative beliefs about themselves after gameplay when realizing their own ability to be successful in and learn from the STEM game.

However, the non-significant change in STEM academic self-efficacy when all other STEM metrics showed an increase might suggest that girls are still unable to connect the positive STEM gaming experience with their ability to perform well in STEM-related school disciplines such as math, science, and technology. Such finding is consistent with literature pointing out that girls might be more modest, harder on themselves, and report lower confidence when it comes to

performance in STEM classes, even when receiving a higher score than boys (Pajares, 2005; Schunk & Pajares, 2002).

Still, when compared to boys, girls always showed an increase in pre-to-post change in STEM variables of interest, a behavior which sometimes was even greater than boys' increase (STEM learning self-efficacy) or in an opposite direction from boys' behavior (STEM skill interest). It is important to note that, however, even with these increases, girls' means are still consistently lower than boys' means before or after website use. This is consistent with literature pointing out that girls demonstrate lower STEM interest than boys (Steinke et al., 2012; Diekman et al., 2017; Hayes, 2016) and overall lower self-efficacy (MacPhee et al., 2013; Rittmayer & Beier, 2009) potentially due to negative stereotypes related to STEM fields and STEM professionals (Bian et al., 2017; Weisgram, 2016). Still, it is exciting to see an increase in STEM metrics for girls after having a single time interaction with a STEM gaming website which displayed motivational Growth Mindset messages. This finding strengthens the argument that schools should invest in STEM gaming activities, especially given the likelihood that girls will not be able to engage in such activities at home (Jenson & De Castell, 2010).

When considering boys' behavior, it is unclear what factors were responsible for no significant change in pre-to-post STEM metrics, and their decrease in some STEM metrics when compared to girls' behavior. Website data might cast a light on these findings. Most boys picked the Physics puzzle game Bumper Ducks (76.9%), which is a relatively more childish-looking and slower pace game than the Biology action game Feeding Frenzy (note that for girls there was an almost even distribution between the two games). When considering the open-ended question related to gameplay in the website, we see that some boys reported to have found this game not challenging enough. Literature related to gameplay styles and gender differences suggest that

boys prefer challenging and competitive settings more than exploratory ones (e.g., Heeter & Winn, 2016), also that their gaming skill is relatively higher than girls' due to a broader and higher exposition to gaming environments (Hughes, 2016; Jenson & De Castell, 2010). Thus, perhaps the game played and the experience in the STEM gaming website were not enough to provoke a significant pre-to-post change regarding STEM metrics for boys.

A puzzling finding of this research study comes from **RQ2** and the perceived differences between boys and girls related to STEM and non-STEM virtual mentors. For participants who had a virtual mentor (N = 31), it seems that having a STEM virtual mentor can be more beneficial to boys than girls when considering pre-to-post changes in STEM academic self-efficacy and STEM skill interest. After all, both boys and girls with STEM virtual mentors showed increased STEM academic self-efficacy and STEM skill interest after website use, but girls' increase was higher when having a non-STEM virtual mentor. On the other hand, boys with a non-STEM virtual mentor showed a decrease in both STEM academic self-efficacy and STEM skill interest. Thus, the main takeaway is that boys seem to benefit more from a STEM virtual mentor, while girls seem to benefit more from having a non-STEM virtual mentor. This is an interesting finding given the argument that people use visual cues from fictional and non-fictional characters/people when making assumptions about their social roles (Baylor, 2011; Isbister, 2006), also literature connecting self-efficacy with role modelling (Bandura, 2008).

One argument could be that having a STEM-looking woman as their virtual mentor potentially reminded girls of the negative stereotypes related to women and STEM, in a stereotype threat-like scenario. However, given the fact that girls with a STEM virtual mentor still showed an increase in these STEM metrics (rather than a decrease), we believe that a better explanation is that the non-STEM woman looked more approachable, relatable and caring than

the STEM woman. Literature suggests that displays of kindness and concern for others help explain higher wishful identification with female scientific characters (Steinke et al., 2012). Additionally, perhaps girls were able to relate more with the non-STEM woman due to an already in place implicit bias towards STEM women (Jenson & De Castell, 2010), the stereotypical negative portrayal of STEM women in the media (Dudo et al., 2011; Steinke, 2005), or the lack of diversity in STEM fields creating a lack of effective STEM role models for girls (Flanagan & Kaufman, 2016).

On the contrary for boys, there are plenty STEM-looking man in the real world who could act as a role model, thus perhaps it was easier to see the STEM-looking man as someone relatable. Moreover, literature suggests that looking intelligent helps explain higher wishful identification with male scientific characters (Steinke et al., 2012), and that boys are more likely to stereotype the image of a scientist as a white intelligent man than girls (Miller et al., 2018). The design for the non-STEM virtual mentor depicted the character reading books and working in a generic office, thus boys who had a non-STEM virtual mentor could have been negatively affected when seeing their mentoring character in a “detrimental” power position according to expected societal gender roles (e.g., “a substitute teacher”, “a secretary”). Boys could also have perceived the non-STEM virtual mentor to be someone less intelligent and less prepared to teach than the STEM-looking one. Although scarce, qualitative data from the interviews seems to strengthen this last argument as STEM virtual mentors were preferred and perceived as more successful and better in facilitating learning than the non-STEM virtual mentors.

However, a complicating factor to this finding is that the survey data was not able to confirm the expectation posited in **H2** that STEM virtual mentors are more likely to be perceived as credible and able to facilitate learning when compared with non-STEM virtual mentors. Given

that qualitative data does points towards such findings, it seems that the small sample size might have hurt our statistical analysis. Another explanation might be, during interviews, participants were allowed plenty time to reflect upon the design of the virtual mentors while during the experiment participants not only had limited time per page but also might have been focusing their attention on other elements of the website such as instructions and forms. Participants might have had a hard time to recall their virtual mentor because of this factor.

RQ3 aimed to explore possible differences in identification with the virtual mentor when considering virtual mentor type, and no significant interaction effects were found; the explanation might be similar to H2 given that qualitative data also suggests that STEM virtual mentors would be the preferred choice if participants had to pick a virtual mentor. **RQ4** considered all virtual mentor perceptions in relation with participants' gender, and no significant interaction effects were found as well. Although scarce, qualitative data indeed shows no difference in regards with virtual mentor perceptions based on different participants' gender, as all participants agreed on preferring STEM virtual mentors, also evaluated them as more credible and better suited to facilitate learning than non-STEM virtual mentors.

In regards with **H1** and the expectation that virtual mentor's existence would positively impact pre-to-post changes in STEM metrics, comparison with the control group (i.e., no virtual mentor) yielded no effects, even though qualitative data suggests an interest in having a virtual mentor (rather than not having one) due its perceived social presence which could increase attention paid to feedback messages and instructions. Such unexpected finding might be a consequence of the small sample size, participants decision to focus more on reading the messages or playing the games than paying attention to the image of their virtual mentor.

Even with these analyses yielding non-significant findings, we believe it is important to try to better understand if identification with the virtual mentor can help explain the finding from RQ2. In order to do so, other models were created using the virtual mentor perceptions as covariates. For Growth Mindset and STEM skill interest, no gender or virtual mentor type interaction was found. However, it seems that identification with the virtual mentor does play a role into pre-to-post changes in STEM self-efficacies, as the most beneficial situation happened when girls and boys demonstrated high wishful identification and low similarity identification with their virtual mentors. However, given previous findings, it seems that it is necessary to have a bigger sample size, a stronger stimulus, and/or a follow-up study in order to better determine the mechanics that help explain the impact of identification with the virtual mentor and STEM self-efficacies.

Drawing from explanations related to social modelling (Bandura, 2008) and avatar theories related to role models (Kao & Harrell, 2006), it is reasonable to expect that an effective STEM role model would be someone aspiring and motivating, capable of eliciting high wishful identification feelings; however, similarity identification (i.e., feelings that the virtual mentor is similar to and resembles the user) would be expected to be high rather than low. Perhaps when using the website, participants experienced difficulties to bond with their virtual mentor in such a way, given the fact that the virtual mentors were modelled after young-adults working in a professional/educational environment. Although participants shared similarities with their virtual mentor (i.e., same gender and skin color), the chosen design might have made the virtual mentors visually too different from the participants (middle school students) in a behavioral way.

Theoretical and Design Implications

One important theoretical implication from this research project is increasing the bulk of literature related to self-efficacy and interest in STEM fields, specifically in relationship with gender differences in a STEM gaming website context. There is evidence that girls and boys perceive and are affected by gaming and STEM learning environments in different ways, given the analysis of participants pre-to-post change in STEM constructs such as Growth Mindset, STEM self-efficacies, and STEM skill interest. By using a STEM gaming website which was designed to elicit Growth Mindset, create excitement towards STEM fields, and feature a mentoring character, girls were able to increase their belief that a trait such as intelligence can be changed through effort (Growth Mindset), become more confident in their ability to learn STEM topics (STEM learning self-efficacy), and intensify their interest in STEM skills which might lead to a higher interest in pursuing STEM careers in the future. Boys, on the other hand, did not achieve such positive results, which might be a consequence of the lack of enough challenge in the STEM gaming website or an already high predisposition to learning from gaming environments (thus the experiment did not provide a strong enough stimulus for these participants). Academia can benefit from such gender-related findings: instead of proposing one theory to fit all middle-schoolers regardless of gender, perhaps different models can be built for girls and boys based on their different ways to perceive and interact with gaming and STEM learning environments.

A theoretical and design implication is that the way pedagogical agents are designed might affect STEM metrics differently for boys and girls. Specifically, boys seem to benefit more from having a STEM virtual mentor, while girls seem to benefit more from having a non-STEM virtual mentor. Although it is important to take such finding with a grain of salt - given

the lack of other supportive findings related to virtual mentor's type - it seems that boys and girls bonded more with virtual mentors who were perceived as more relatable to their gender, a potential conformation to implicit bias and expected gender roles in STEM fields (i.e., most scientists are men, not women).

Additionally, boys and girls seem to benefit more from virtual mentors who elicit high wishful identification and low similarity identification. Perhaps this is a suggestion that virtual mentors can be understood as digital companions according to Ratan's (2017) conceptualization of avatars. After all, a digital companion is an "avatar-as-social other", someone who is not controlled by the user and is perceived to be different than the self (thus sharing less similar characteristics). If participants were able to choose and/or customize their virtual mentor, perhaps similarity identification could be increased, also increasing virtual mentor's relevance to the user - drawing from Ratan and Sah (2016). As a matter of fact, participants mentioned in interviews to be very interested in choosing and/or customizing their virtual mentor because, by doing so, participants could choose someone who better represents them or who can better assist them to learn. Such design decision could potentially allow middle schoolers to choose a virtual mentor better suitable to be a STEM role model (i.e., someone more similar to themselves, and perceived to be credible and able to facilitate learning).

Finally, a last design implication would be that STEM gaming websites should include more STEM games and more diverse STEM games in terms of genre (e.g., puzzle, action) and visual design (e.g., cartoonish, fantastic), so that all users can find a game suitable to their interest and skill. Moreover, the STEM games should accompany some kind of educational material, given that some participants mentioned during interviews or when answering the open-ended question that such pedagogical feature was lacking.

Limitations and Future Research

This research project serves as a preliminary and exploratory study into how mentoring characters can be designed for a STEM gaming website. Some clear limitations to this research are related to the small sample size for both the quantitative data ($N = 56$) and the qualitative data ($N = 5$). Moreover, there is a lack of diversity in the sample, as the majority of the participants were White/Caucasians, only one girl was interviewed, and data collection happened in only one public middle-school located in a university town with strong school systems.

This project found significant results for a single 25 min exposure to a STEM gaming website; longer exposure in a longitudinal context (i.e., multiple data collections with the same students in different points in time) would possibly lead to stronger effects as it would have increased the strength of the stimuli. In order for such situation to be possible, however, data collection would need to happen during a proposed activity in a middle-school summer camp; another possibility would be if the researcher was directly involved with a public middle school and had approval from the principal to adapt the school curriculum to better fit the goals of the research.

Another limitation related to one-single point of data collection is that the design for the virtual mentor and the website could have benefited from iterations, informed by the users, in a more UX design approach. Again, this situation would demand multiple data collections which would increase the complexity, costs, and time associated with this research project.

Future research should consider looking into the use of eye-tracking in order to determine how much of participants' gaze is being directed towards the virtual mentors in the website, and if different designs can detract or further engage participants. If participants' attention is not

being directed towards the virtual mentor during website use, there is a great probability that weak or no effects will be seen.

Additionally, future research should investigate the effects of adding the functionality of choosing and/or customizing the virtual mentor, as findings from this research project together with evidence in avatar literature suggest that higher identification with the virtual mentor might be better achieved when having such feature available. This feature could also allow participants to choose someone perceived as more knowledgeable and able to facilitate learning; specific design details associated with different cosmetics and items should be examined. Perhaps when registering in the website, users can be given a more approachable (i.e., less STEM-looking) virtual mentor, and after playing games and interacting with the website, users can gain points towards “levelling-up” their virtual mentors, a more co-constructive way of learning which has the potential of increasing feelings of similarity with the virtual mentor and STEM self-efficacies, because of the psychological bond created between users and their virtual mentors.

Finally, given the different findings associated with STEM learning self-efficacy and STEM academic self-efficacy, it seems that these constructs should be kept separate in terms of conceptualization - one’s belief that learning is possible might be differently enough than one’s belief that performing well in a discipline is possible. Previous literature lacks a clear definition regarding the overall construct of STEM self-efficacy, thus future research could further investigate the relationship between both constructs in order to make a stronger stand for the need of keeping them separate.

CHAPTER 7

FINAL REMARKS

This research project examined if and how virtual mentors could be seen as STEM role models for middle schoolers in a STEM gaming website. Specifically, we wondered if virtual mentors could positively impact middle schoolers' metrics related to STEM such as Growth Mindset, STEM self-efficacies, and STEM skill interest. Gender differences were investigated, and results suggest that girls and boys use (and are affected by) STEM gaming environments in different ways. Girls showed a behavior of increase in STEM metrics after website use which in some cases was stronger than boys' increase, or in a different direction than boys' behavior.

Although future research considering virtual mentor customization might help better explain the psychological connection between users and their virtual mentors, there is reason to believe that some designs of virtual mentors are able to elicit more beneficial behaviors for participants depending on participants' gender, and that identification with the virtual mentor might be one of the mechanics behind changing one's STEM self-efficacies. Specifically, it seems that boys are more likely to benefit from having a STEM virtual mentor while girls are more likely to benefit from having a non-STEM virtual mentor; both girls and boys seem to benefit from having high wishful and low similarity identification with their virtual mentors.

This research study is a first step towards the understanding of virtual mentors as digital companions, thus bringing learning and avatar theories together.

APPENDICES

APPENDIX A: VIRTUAL MENTORS' DESIGN

	Visuals	Color Hex Code
White skin		#FBC8AA #FBC2A0 #F4AF93
Olive skin		#EAAB7D #E79E6D
Black skin		#995A4D #8C4C3F
Hair		#452A24
Eyes		#000
Lips		Women: #8F2726 Men: #ED8585 (white/olive) #7E3C2E (black)
Shirt		#7FDCA5 #74C193
Pants		#35485D #293C4F
Shoes		#323336 #1E1F21

Figure 32. Virtual mentor's color scheme

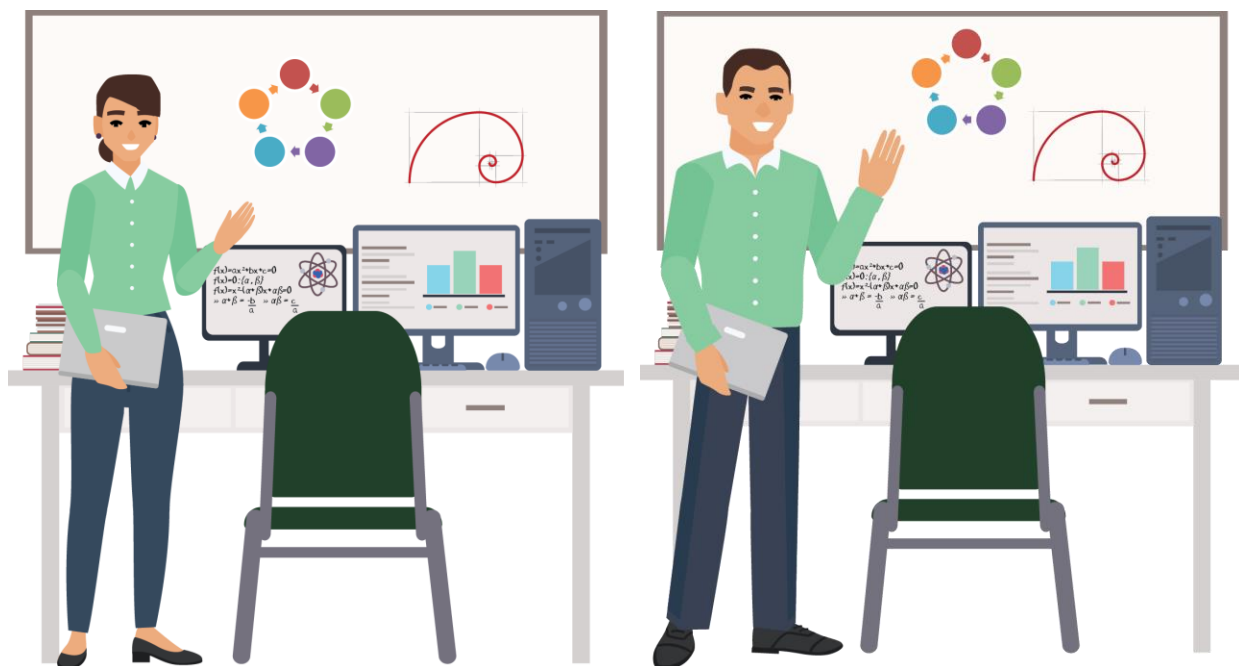


Figure 33. STEM virtual mentor's images: welcome and last pages



Figure 34. Non-STEM virtual mentor's images: welcome and last pages

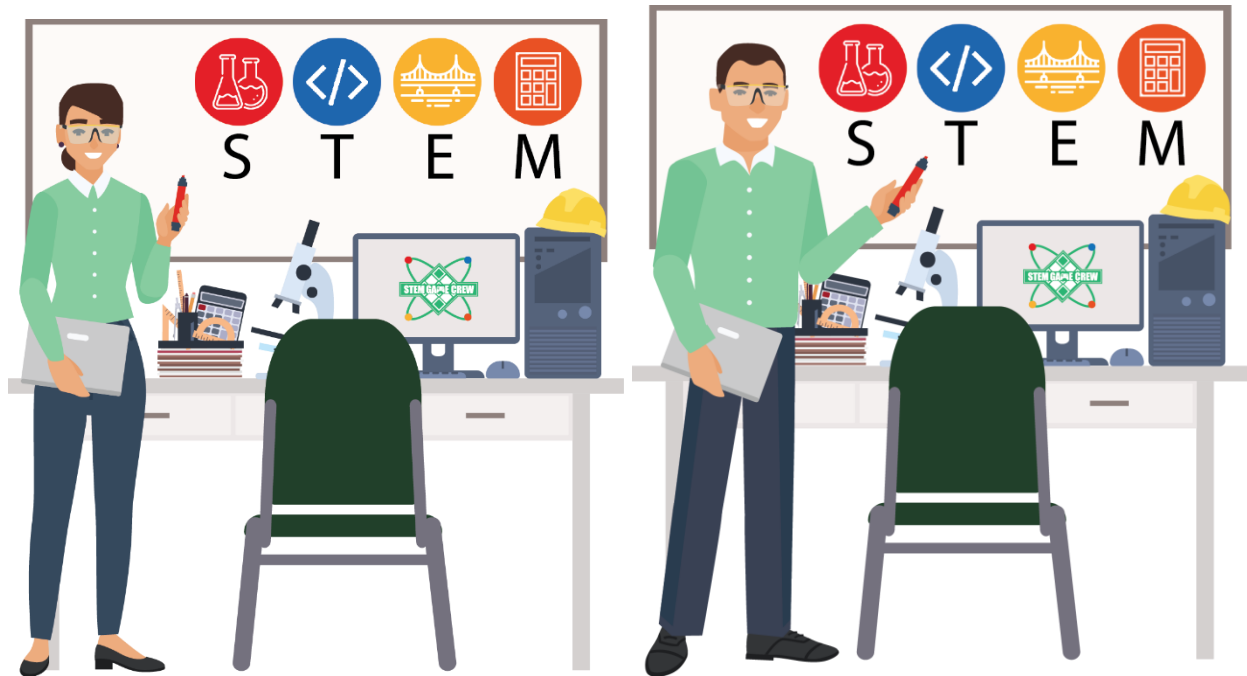


Figure 35. STEM virtual mentor's images: explaining the website and STEM careers

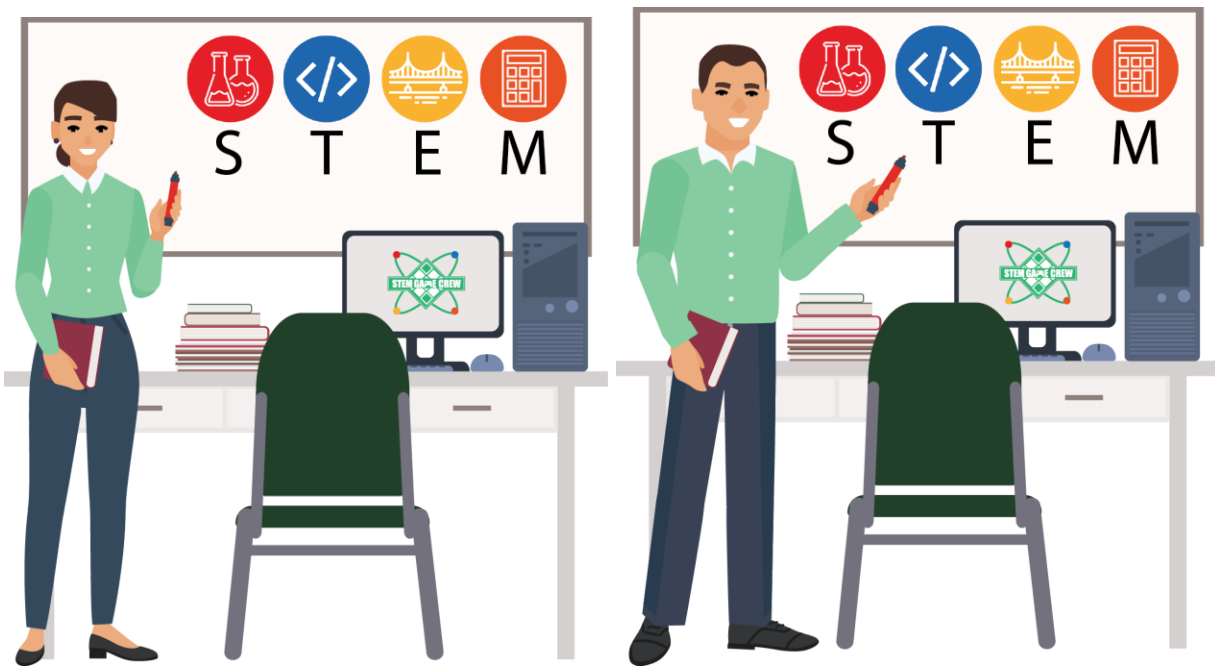


Figure 36. Non-STEM virtual mentor's images: explaining the website and STEM careers



Figure 37. STEM virtual mentor's images: game claim (hypotheses)



Figure 38. Non-STEM virtual mentor's images: game claim (hypotheses)



Figure 39. STEM virtual mentor's images: game evidence (play the game)



Figure 40. Non-STEM virtual mentor's images: game evidence (play the game)



Figure 41. STEM virtual mentor's images: game reasoning (conclusions)



Figure 42. Non-STEM virtual mentor's images: game reasoning (conclusions)



Figure 43. STEM virtual mentor's images: feedback

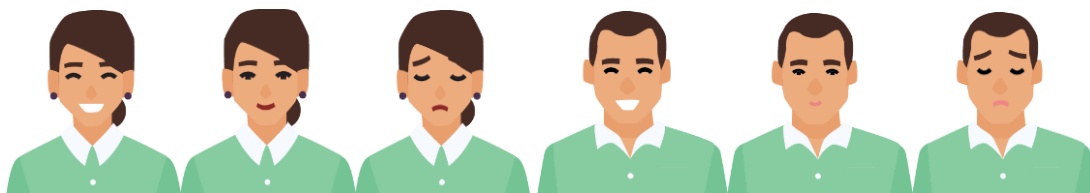


Figure 44. Non-STEM virtual mentor's images: feedback



Figure 45. STEM virtual mentor's images: presenting fun fact



Figure 46. Non-STEM virtual mentor's images: presenting fun fact

APPENDIX B: INTERVIEW IMAGES



Figure 47. Non-STEM virtual mentor: women



Figure 48. Non-STEM virtual mentor: men



Figure 49. STEM worker virtual mentor: women



Figure 50. STEM worker virtual mentor: men

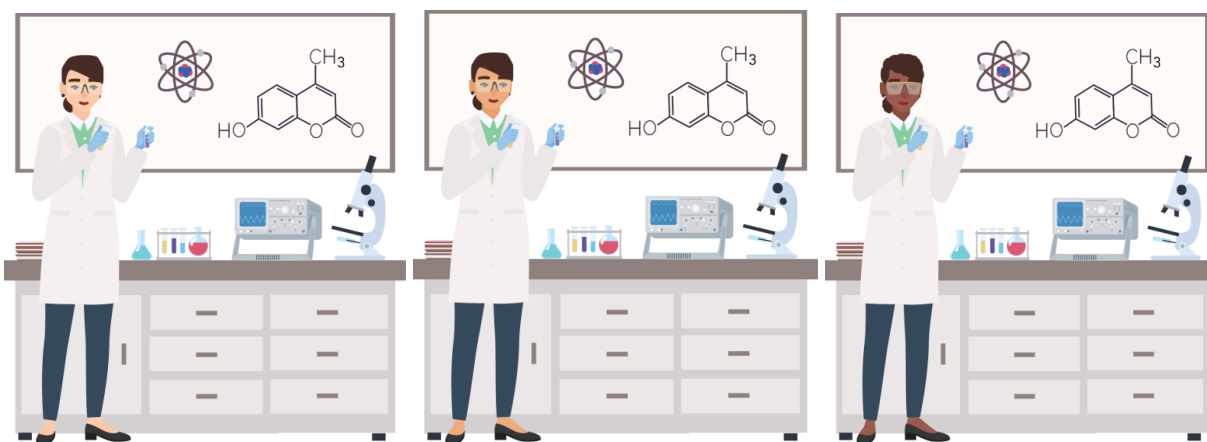


Figure 51. Chemist (Science) virtual mentor: women

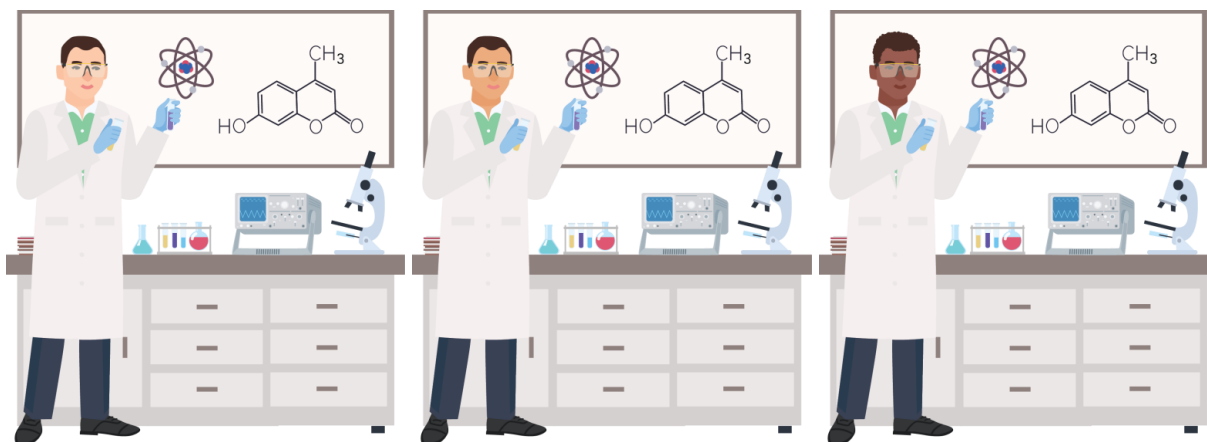


Figure 52. Chemist (Science) virtual mentor: men

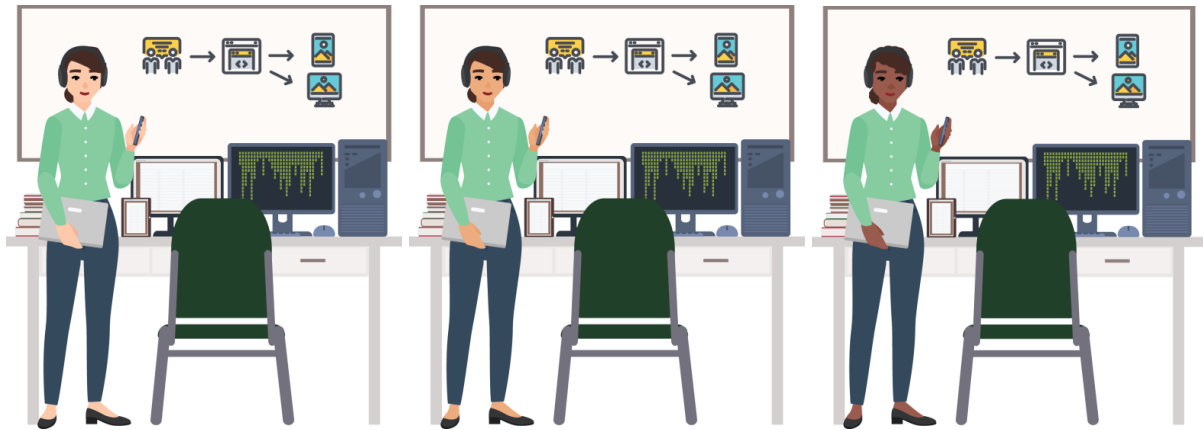


Figure 53. Web developer (Technology) virtual mentor: women

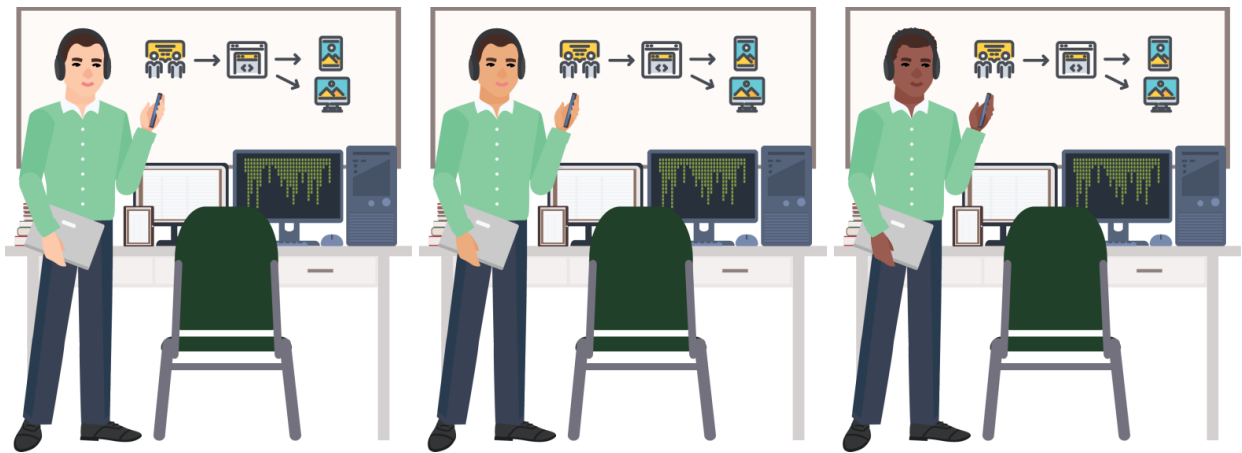


Figure 54. Web developer (Technology) virtual mentor: men



Figure 55. Civil engineer (Engineering) virtual mentor: women



Figure 56. Civil engineer (Engineering) virtual mentor: men



Figure 57. Mathematician (Math) virtual mentor: women



Figure 58. Mathematician (Math) virtual mentor: men

APPENDIX C: WEBSITE WIREFRAME

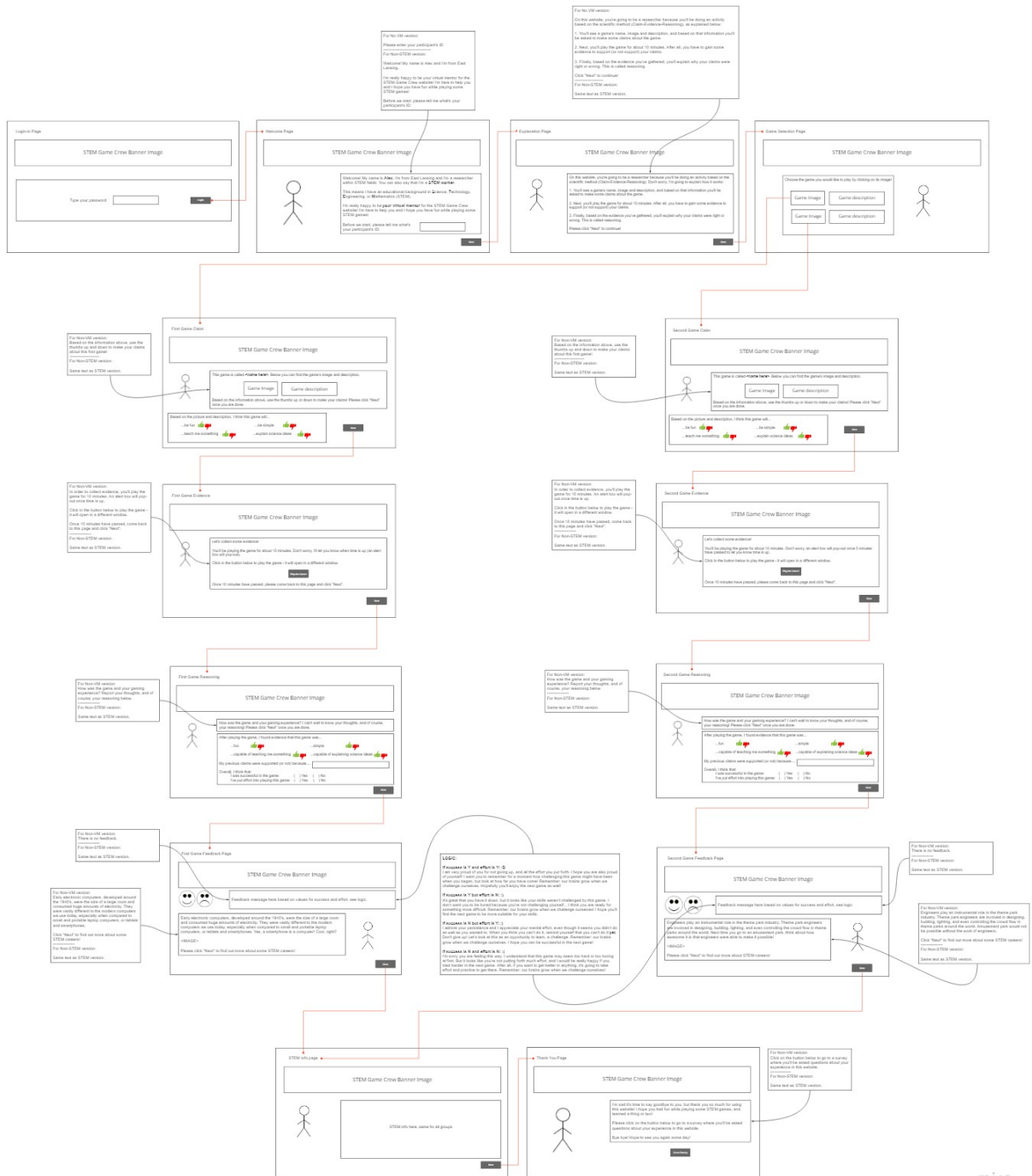


Figure 59. Complete experimental website wireframe

APPENDIX D: SCALES FOR QUANTITATIVE STUDY

Table 11.

Growth Mindset scale.

-
1. No matter how much intelligence you have, you can always change it a good deal. (+)
 2. You can learn new things, but you cannot really change your basic level of intelligence. (-)
 3. I like my work best when it makes me think hard. (+)
 4. I like my work best when I can do it really well without too much trouble. (-)
 5. I like work that I'll learn from even if I make a lot of mistakes. (+)
 6. I like my work best when I can do it perfectly without any mistakes. (-)
 7. When something is hard, it just makes me want to work more on it, not less. (+)
 8. To tell the truth, when I work hard, it makes me feel as though I'm not very smart. (-)
-

Note: items from the 6-point Likert Mindset Assessment Profile scale (Mindset Works, 2007); this scale is composed by items related to Growth Mindset (+) and items related to Fixed Mindset (-). Prompt was: "How much do you agree with the following statements?"

Table 12.

STEM learning self-efficacy (L) and STEM academic self-efficacy (A) scale.

[Science]

1. I am confident in my ability to understand the ideas taught in my Science classes. [L]
2. I can figure out problems and tasks assigned during my Science classes. [L]
3. I am confident in my ability to learn new scientific concepts. [L]
4. I am confident in my ability to do very well in my Science classes. [A]
5. I am able to do well in activities that involve Science. [A]
6. I am confident in my ability to use scientific concepts for class work. [A]

[Technology]

7. I am confident in my ability to understand the ideas taught in my Technology classes. (L)
8. I can figure out problems and tasks assigned during my Technology classes. (L)
9. I am confident in my ability to learn new technologies. (L)
10. I am confident in my ability to do very well in my Technology classes. (A)
11. I am able to do well in activities that involve Technology. (A)
12. I am confident in my ability to use technologies for class work. (A)

[Math]

13. I am confident in my ability to understand the ideas taught in my Math classes. (L)
 14. I can figure out problems and tasks assigned during my Math classes. (L)
 15. I am confident in my ability to learn new mathematical concepts. (L)
 16. I am confident in my ability to do very well in my Math classes. (A)
 17. I am able to do well in activities that involve Math. (A)
 18. I am confident in my ability to use mathematical concepts for class work (A)
-

Note: items adapted from the 5-point Likert Pintrich and De Groot (1990)'s self-efficacy scale in order to fit STEM contexts. Prompt was "When you think about [Science][Technology][Math] and your [Science][Technology] [Math]classes, how much do you agree with the following statements?"

Table 13.

STEM Skills Interest scale.

-
1. Doing experiments in a laboratory.
 2. Going to science museums and/or science fairs.
 3. Thinking about how video games work.
 4. Spending time with photo, video and/or recording technologies.
 5. Visualizing objects in three dimensions from flat drawings.
 6. Creating new games or things to do with your toys, games, etc.
 7. Solving puzzles and/or riddles.
 8. Playing detective and solving mysteries.
 9. Sharing new ideas or things you have created.
 10. Learning about famous inventors and things they have created.
 11. Thinking of new ways to do things.
 12. Finding the answers to questions.
 13. Figuring out how things work and investigating new things.
 14. Working with computers and/or computer programs.
-

Note: 5-point Likert scale built based on the Career Quiz from Washington Career Bridge (2018). Prompt was “How interesting are these activities to you?”.

Table 14.

Virtual Mentor’s ability to facilitate learning (FL) and Virtual Mentor’s credibility (C) scale.

-
1. The virtual mentor made the instruction interesting. (FL)
 2. The virtual mentor encouraged me to reflect what I was learning. (FL)
 3. The virtual mentor presented the material effectively. (FL)
 4. The virtual mentor improved my knowledge of the content. (FL)
 5. The virtual mentor was motivating. (FL)
 6. The virtual mentor was knowledgeable. (C)
 7. The virtual mentor was intelligent. (C)
 8. The virtual mentor was useful. (C)
 9. The virtual mentor was helpful. (C)
 10. The virtual mentor was instructor-like. (C)
-

Note: 5-point Likert scale adapted from the API (Agent Persona Instrument) by Baylor and Ryu (2003). Prompt was “How much do you agree with the following statements?”.

Table 15.

Virtual Mentor Similarity Identification (SI) and Virtual Mentor Wishful Identification (WI) scale.

-
1. My virtual mentor is like me in many ways. (SI)
 2. My virtual mentor resembles me. (SI)
 3. I identify with my virtual mentor. (SI)
 4. My virtual mentor is an extension of myself. (SI)
 5. My virtual mentor is similar to me. (SI)
 6. I resemble my virtual mentor. (SI)
 7. If I could become like my virtual mentor, I would. (WI)
 8. I would like to be more like my virtual mentor. (WI)
 9. My virtual mentor is an example to me. (WI)
 10. My virtual mentor is a better me. (WI)
 11. My virtual mentor has characteristics that I would like to have. (WI)
-

Note: 5-point Likert scale adapted from Van Looy et al. 's (2012) scale for game character identification. Prompt was: "How much do you agree with the following statements?".

APPENDIX E: INTERVIEW QUESTIONS FOR QUALITATIVE STUDY

Table 16.

Questions for the Semi-Structured Interview.

1. Can you talk a bit about your experience when using the website earlier today?
 - Was there anything you liked?
 - Was there anything you disliked?
 2. Do you remember having a virtual mentor named Alex who guided your steps and gave you feedback in the website?
 - Which one of these four images represents how Alex looked like in the website?
 3. If you could pick a virtual mentor, which one of these four would you pick?
 - What do you like about this virtual mentor?
 - Is there anything in this image that catches your attention?
 - (optional) This image is different from the virtual mentor you had earlier today. Why do you think you prefer this one instead of the one you were given?
 4. Do you think any of these four virtual mentors would be better in helping you learn on the website than others?
 - Is there anything in this image that makes you think that?
 5. What if we had more options, for example, more skin colors, or more professions as in these 36 images?
 - Would you like to have more options or fewer options?
 - Which one of these 36 options would you pick as your virtual mentor and why?
 6. Do you think any of these 36 virtual mentors would be better in helping you learn on the website than others?
 - Is there anything in this image that makes you think that?
 7. How would you define success, or someone who is successful?
 - Do you think any of these 36 virtual mentors looks more successful than others?
 - Is there anything in this image that makes you think that?
 8. What if we keep the virtual mentor the same, but change the background?
 - Which one would you pick then and why?
 9. Would you like to have the ability to customize the virtual mentor or the background?
 10. Do you have any suggestions in terms of what could be improved in the design of the virtual mentor or the website?
-

APPENDIX F: SUMMARY OF FINDINGS FOR QUANTITATIVE DATA

Table 17.

Summary of Findings for Quantitative Data organized by STEM Variables of Interest.

<u>Model</u>	<u>STEM Variables of Interest (Pre-Post)</u>			
	<u>Growth Mindset</u>	<u>STEM learning self-efficacy</u>	<u>STEM academic self-efficacy</u>	<u>STEM skill interest</u>
All participants	Increase**	Increase**	Not significant	Not significant
RQ1: Boys compared to girls	No gender interaction	Gender interaction*	Not significant	Gender interaction**
RQ1: Boys only	Not significant	Not significant	Not significant	Not significant
RQ1: Girls only	Increase**	Increase**	Not significant	Increase**
H1: Participants with VM compared to no-VM	No effects of VM existence	No effects of VM existence	No effects of VM existence	No effects of VM existence
RQ2: Participants with STEM VM compared to non-STEM VM; gender included	No gender interaction and/or VM type interaction	Gender interaction**	Gender and VM type interaction*	Gender and VM type interaction**
RQ2: Participants with STEM VM compared to non-STEM VM; gender included; VM metrics included as covariates	No gender interaction and/or VM type interaction	Stronger gender interaction** • Wishful identification** • Similarity identification**	Stronger gender and VM type interaction** • Wishful identification** • Similarity identification**	No gender interaction and/or VM type interaction

* marginally significant finding; ** $p \leq .05$

Table 18.

Summary of Findings for Quantitative Data organized by Virtual Mentor's Variables of Interest.

<u>Virtual Mentor's Variables of Interest (Post)</u>				
<u>Model</u>	<u>VM's credibility</u>	<u>VM's ability to facilitate learning</u>	<u>Similarity identification</u>	<u>Wishful identification</u>
H2: Participants with STEM VM compared to Non-STEM VM	No effects of VM type	No effects of VM type	---	---
RQ3: Participants with STEM VM compared to Non-STEM VM	---	---	No effects of VM type	No effects of VM type
RQ4: Boys with VM compared to girls with VM	No effects of gender	No effects of gender	No effects of gender	No effects of gender

BIBLIOGRAPHY

BIBLIOGRAPHY

American Association of University Women (2010). *Why So Few? Women in Science, Technology, Engineering and Mathematics*. Retrieved from <https://www.aauw.org/files/2013/02/Why-So-Few-Women-in-Science-Technology-Engineering-and-Mathematics.pdf>.

Augusto, J. C., McNair, V., McCullagh, P., & McRoberts, A. (2010). Scoping the potential for anytime-anywhere support through virtual mentors. *Innovation in Teaching and Learning in Information and Computer Sciences*, 9(2), 1-12.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191.

Bandura, A. (1989). Human agency in social cognitive theory. *American psychologist*, 44(9), 1175.

Bandura, A. (2008). An agentic perspective on positive psychology. *Positive psychology*, 1, 167-196.

Baylor, A. L. (2011). The design of motivational agents and avatars. *Educational Technology Research and Development*, 59(2), 291-300.

Baylor, A., & Ryu, J. (2003). The API (Agent Persona Instrument) for assessing pedagogical agent persona. In *EdMedia+ Innovate Learning* (pp. 448-451). Association for the Advancement of Computing in Education (AACE).

Baylor, A., Shen, E., & Huang, X. (2003). Which pedagogical agent do learners choose? The effects of gender and ethnicity. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 1507-1510). Association for the Advancement of Computing in Education (AACE).

Beckem, J. M. (2012). Bringing life to learning: Immersive experiential learning simulations for online and blended courses. *Journal of Asynchronous Learning Networks*, 16(5), 61-70.

Bessière, K., Seay, A. F., & Kiesler, S. (2007). The ideal elf: Identity exploration in World of Warcraft. *Cyberpsychology & behavior*, 10(4), 530-535.

Bian, L., Leslie, S. J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science*, 355(6323), 389-391.

Big Think. (2011, June 8). *Neil deGrasse Tyson on Teaching Science* [Video File]. Retrieved from <https://www.youtube.com/watch?v=tsNrIfDTyWA>.

Big Think. (2013, May 13). *Neil deGrasse Tyson: Want Scientifically Literate Children? Get Out of Their Way* [Video File]. Retrieved from <https://www.youtube.com/watch?v=AIEJjpVlZu0>.

Blanton, H., Crocker, J., & Miller, D. T. (2000). The effects of in-group versus out-group social comparison on self-esteem in the context of a negative stereotype. *Journal of Experimental Social Psychology*, 36(5), 519-530.

Bosworth, D., Lyonette, C., Wilson, R., Bayliss, M., & Fathers, S. (2013). The supply of and demand for high-level STEM skills. *Evidence Report*, 77.

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 43(5), 485-499.

Brush, L. R. (1979). Avoidance of science and stereotypes of scientists. *Journal of Research in Science Teaching*, 16(3), 237-241.

Cantrell, K., Fischer, A., Bouzaher, A., & Bers, M. (2010). The role of E-mentorship in a virtual world for youth transplant recipients. *Journal of Pediatric Oncology Nursing*, 27(6), 344-355.

Carli, L. L., Alawa, L., Lee, Y., Zhao, B., & Kim, E. (2016). Stereotypes about gender and science: Women ≠ scientists. *Psychology of Women Quarterly*, 40(2), 244-260.

Cherchiglia, L., Klug, A., Renius, W., Sanders, H., Oldenburg, S., Stacey, R., Wanek, C., Beyene, M., & Ratan, R. (2016b, August). Proteus Play on a STEM-Game Platform: Examining the role of avatar identity and self-relevance on STEM attitudes and motivation. Poster presented at the Games Learning and Society 12 Conference, Madison, USA.

Cherchiglia, L., Jorae, R., Zhao, Y., Zhang, Q., & Heeter, C. (2017, June). Tech Trek: a serious card-game about STEM careers. Game presented at The Irish Conference on Game-Based Learning 2017, Cork, Ireland.

Cherchiglia, L., & Ratan, R. (2018, October). Avatar Effects in a STEM-Game Website. Paper presented at the Meaningful Play 2018 Conference, East Lansing, USA.

Cherchiglia, L., Ratan, R., Stacey, R., Sanders, H., Oldenburg, S., Renius, W., Klug, A., & Day, T. (2016a, October). Exploring the impact of avatar types and avatar customization on STEM attitudes and motivation to play online STEM games. Poster presented at the Meaningful Play 2016 Conference, East Lansing, USA.

Christenson, J. (2011, November 13). Ramaley coined STEM term now used nationwide. *Winona Daily News*. Retrieved from <http://www.winonadailynews.com/>.

Clarebout, G., Elen, J., Johnson, W. L., & Shaw, E. (2002). Animated pedagogical agents: An opportunity to be grasped?. *Journal of Educational multimedia and hypermedia*, 11(3), 267-286.

Clark, D., Nelson, B., Sengupta, P., & D'Angelo, C. (2009, October). Rethinking science learning through digital games and simulations: Genres, examples, and evidence. In *Learning*

science: Computer games, simulations, and education workshop sponsored by the National Academy of Sciences, Washington, DC. Cote, 2015.

Cote, A. C. (2015). Writing “Gamers”: The Gendered Construction of Gamer Identity in Nintendo Power (1994–1999). *Games and Culture*, Online First, 1-25.

Curran, M. B., & Chatel, R. G. (2013). Virtual Mentors: Embracing Social Media in Teacher. *Pedagogical applications and social effects of mobile technology integration*, 258.

Dabney, K. P., Tai, R. H., Almarode, J. T., Miller-Friedmann, J. L., Sonnert, G., Sadler, P. M., & Hazari, Z. (2012). Out-of-school time science activities and their association with career interest in STEM. *International Journal of Science Education, Part B*, 2(1), 63-79.

Dell, E. M., Verhoeven, Y., Christman, J. W., & Garrick, R. D. (2017). Using Self-Determination Theory to build communities of support to aid in the retention of women in engineering. *European Journal of Engineering Education*, 1-16.

Dehn, D. M., & Van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International journal of human-computer studies*, 52(1), 1-22.

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, 21(8), 1051-1057.

Dudo, A. D., Brossard, D., Shanahan, J., Scheufele, D. A., Morgan, M., & Signorielli, N. (2011). Science on television in the 21st century: Recent trends in portrayals and their contributions to public attitudes toward science. *Communication Research*, 38(6), 754-777.

Dudo, A., Cicchirillo, V., Atkinson, L., & Marx, S. (2014). Portrayals of technoscience in video games: A potential avenue for informal science learning. *Science Communication*, 36(2), 219-247.

Dudo, A., & Besley, J. C. (2016). Scientists’ prioritization of communication objectives for public engagement. *PloS one*, 11(2), e0148867.

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

Dweck, C. S. (2008). *Mindsets and math/science achievement*. New York: Carnegie Corporation of New York, Institute for Advanced Study, Commission on Mathematics and Science Education.

Elsevier (2017). *Gender in the Global Research Landscape*. Retrieved from http://www.elsevier.com/__data/assets/pdf_file/0008/265661/ElsevierGenderReport_final_for-web.pdf

Falloon, G. (2010). Using avatars and virtual environments in learning: What do they have to offer?. *British Journal of Educational Technology*, 41(1), 108-122.

- Flanagan, M. & Kaufman, G. (2016). Shifting Implicit Bias with Games Using Psychology. In *Diversifying Barbie and Mortal Kombat: Intersectional Perspectives and Inclusive Designs in Gaming* (pp. 219-233). Pittsburgh, PA: Carnegie Mellon ETC Press.
- Finson, K. D., Beaver, J. B., & Cramond, B. L. (1995). Development and field test of a checklist for the Draw-A-Scientist Test. *School Science and Mathematics*, 95(4), 195-205.
- Gee, J. P. (2008). Video games and embodiment. *Games and Culture*, 3(3-4), 253-263.
- Guegan, J., Buisine, S., Mantelet, F., Maranzana, N., & Segonds, F. (2016). Avatar-mediated creativity: When embodying inventors makes engineers more creative. *Computers in Human Behavior*, 61, 165-175.
- Hallinen, J. (2015). STEM Education Curriculum. In *Encyclopedia Britannica* [Blog post]. Retrieved from <https://www.britannica.com/topic/STEM-education>.
- Hayes, E. (2016). Girls, Gaming, and Trajectories of IT Expertise. In *Beyond Barbie and Mortal Kombat: New Perspectives on Gender and Gaming* (pp. 216-230). Cambridge, MA: MIT Press.
- Heeter, C., Egidio, R., Mishra, P., Winn, B., & Winn, J. (2009). Alien games: Do girls prefer games designed by girls?. *Games and Culture*, 4(1), 74-100.
- Heeter, C. & Winn, B. (2016). Gender Identity, Play Style, and the Design of Games for Classroom Learning. In *Beyond Barbie and Mortal Kombat: New Perspectives on Gender and Gaming* (pp. 281-300). Cambridge, MA: MIT Press.
- Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning?. *Educational Research Review*, 6(1), 27-54.
- Holman, L., Stuart-Fox, D., & Hauser, C. E. (2018). The gender gap in science: How long until women are equally represented?. *PLoS biology*, 16(4), e2004956.
- Hosein, A. (2019). Girls' video gaming behaviour and undergraduate degree selection: A secondary data analysis approach. *Computers in Human Behavior*, 91, 226-235.
- Hughes, K. (2016). Design to Promote Girls' Agency through Educational Games: the Click! Urban Adventure. In *Beyond Barbie and Mortal Kombat: New Perspectives on Gender and Gaming* (pp. 231-246). Cambridge, MA: MIT Press.
- Hyde, J. S. (2014). Gender similarities and differences. *Annual review of psychology*, 65, 373-398.
- Isbister, K. (2006). *Better game characters by design: A psychological approach*. CRC Press.
- Jenson, J., & De Castell, S. (2010). Gender, simulation, and gaming: Research review and redirections. *Simulation & Gaming*, 41(1), 51-71.

Johnson, A. M., DiDonato, M. D., & Reisslein, M. (2013). Animated agents in K-12 engineering outreach: Preferred agent characteristics across age levels. *Computers in Human Behavior*, 29(4), 1807-1815.

Johnson, I. R., Pietri, E. S., Fullilove, F., & Mowrer, S. (2019). Exploring Identity-Safety Cues and Allyship Among Black Women Students in STEM Environments. *Psychology of Women Quarterly*, doi:0361684319830926.

Kafai, Y. B. (2016). Gender, Games and Computing. In *Diversifying Barbie and Mortal Kombat: Intersectional Perspectives and Inclusive Designs in Gaming* (pp. 261-271). Pittsburgh, PA: Carnegie Mellon ETC Press.

Kahan, D. M., Braman, D., Cohen, G. L., Gastil, J., & Slovic, P. (2010). Who fears the HPV vaccine, who doesn't, and why? An experimental study of the mechanisms of cultural cognition. *Law and human behavior*, 34(6), 501-516.

Kao, D., & Harrell, D. F. (2015, October). Exploring the impact of role model avatars on game experience in educational games. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (pp. 571-576). ACM.

Kao, D., & Harrell, D. F. (2016). Toward Understanding the Impacts of Role Model Avatars on Engagement in Computer Science Learning, 2016 American Educational Research Association Conference, Washington DC, 8-12 April. American Educational Resource Association.

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

Kier, M. W., Blanchard, M. R., Osborne, J. W., & Albert, J. L. (2014). The development of the STEM career interest survey (STEM-CIS). *Research in Science Education*, 44(3), 461-481.

Kim, Y., & Baylor, A. L. (2006). A social-cognitive framework for pedagogical agents as learning companions. *Educational Technology Research and Development*, 54(6), 569-596.

Kim, Y., & Wei, Q. (2011). The impact of learner attributes and learner choice in an agent-based environment. *Computers & Education*, 56(2), 505-514.

Kimball, M., & Smith, N. (2013). *There's one key difference between kids who excel at math and those who don't* [Blog post]. <https://qz.com/139453/theres-one-key-difference-between-kids-who-excel-at-math-and-those-who-dont/>

Klimmt, C., Hefner, D., Vorderer, P., Roth, C., & Blake, C. (2010). Identification with video game characters as automatic shift of self-perceptions. *Media Psychology*, 13(4), 323-338.

Lee, J. E. R., Nass, C. I., & Bailenson, J. N. (2014). Does the mask govern the mind?: Effects of arbitrary gender representation on quantitative task performance in avatar-represented virtual groups. *Cyberpsychology, Behavior, and Social Networking*, 17(4), 248-254.

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.

Loewus, L. H. (2015). *When Did Science Education Become STEM?* [Blog post]. Retrieved from http://blogs.edweek.org/edweek/curriculum/2015/04/when_did_science_education_become_STEM.html.

Losh, S. C. (2010). Stereotypes about scientists over time among US adults: 1983 and 2001. *Public Understanding of Science*, 19(3), 372-382.

MacPhee, D., Farro, S., & Canetto, S. S. (2013). Academic self-efficacy and performance of underrepresented STEM majors: Gender, ethnic, and social class patterns. *Analyses of Social Issues and Public Policy*, 13(1), 347-369.

Marino, M. T., Israel, M., Beecher, C. C., & Basham, J. D. (2013). Students' and teachers' perceptions of using video games to enhance science instruction. *Journal of Science Education and Technology*, 22(5), 667-680.

Martin, C. L., & Ruble, D. N. (2010). Patterns of gender development. *Annual review of psychology*, 61, 353-381.

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.

Maltese, A. V., & Tai, R. H. (2010). Eyeballs in the fridge: Sources of early interest in science. *International Journal of Science Education*, 32(5), 669-685.

Mindset Works (2017). *Mindset assessment*. Retrieved from <http://blog.mindsetworks.com/what-s-my-mindset>

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The Development of Children's Gender-Science Stereotypes: A Meta-analysis of 5 Decades of US Draw-A-Scientist Studies. *Child development*.

Moreno, R., & Flowerday, T. (2006). Students' choice of animated pedagogical agents in science learning: A test of the similarity-attraction hypothesis on gender and ethnicity. *Contemporary educational psychology*, 31(2), 186-207.

National Academies of Sciences. (2016). *Science Literacy: Concepts, and Consequences*. Retrieved from Washington, DC: The National Academies Press.

National Research Council. (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. Washington, DC: The National Academies Press.

National Science Foundation (2014). *Women, Minorities and Persons with Disabilities in Science and Engineering*. Arlington, VA: National Science Foundation.

Oh, Y. J., Jia, Y., Lorentson, M., & LaBanca, F. (2013). Development of the educational and career interest scale in science, technology, and mathematics for high school students. *Journal of science Education and Technology*, 22(5), 780-790.

Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. *Gender differences in mathematics: An integrative psychological approach*, 294-315.

Pearce, C. (2016). Curating for Diversity. In *Diversifying Barbie and Mortal Kombat: Intersectional Perspectives and Inclusive Designs in Gaming* (pp. 200-216). Pittsburgh, PA: Carnegie Mellon ETC Press.

Peña, J. F. (2011). Integrating the influence of perceiving and operating avatars under the automaticity model of priming effects. *Communication Theory*, 21(2), 150-168.

Peña, J., Hancock, J. T., & Merola, N. A. (2009). The priming effects of avatars in virtual settings. *Communication Research*, 36(6), 838-856.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, 82(1), 33.

Ratan, R. (2017). Companions & vehicles: Permutations of avatar-likeness. In J. Banks (Ed.), *Avatars assembled: The social and technical anatomy of digital bodies*. New York: Peter Lang Publishing.

Ratan, R. A., & Dawson, M. (2016). When Mii is me: a psychophysiological examination of avatar self-relevance. *Communication Research*, 43(8), 1065-1093.

Ratan, R., & Sah, Y. L. (2016). Leveling up on stereotype threat: The role of avatar customization and avatar embodiment. *Computers in Human Behavior*, 50, 367-374.

Ratan, R., Rikard, R. V., Wanek, C., McKinley, M., Johnson, L., & Sah, Y. J. (2016, January). *Introducing Avatarification: An experimental examination of how avatars influence student motivation*. In 2016 49th Hawaii International Conference on System Sciences (HICSS) (pp. 51-59). IEEE.

Richard, G. T. (2016). At the Intersections of Play. In *Diversifying Barbie and Mortal Kombat: Intersectional Perspectives and Inclusive Designs in Gaming* (pp. 71-91). Pittsburgh, PA: Carnegie Mellon ETC Press.

Rittmayer, A. D., & Beier, M. E. (2008). Overview: Self-efficacy in STEM. *SWE-AWE CASEE Overviews*.

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research*, 49(1), 1-39.

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In *Development of achievement motivation* (pp. 15-31). Academic Press.

Siekman, G., & Korbel, P. (2016). Defining 'STEM' skills: review and synthesis of the literature. *Adelaide: NCVER*.

Soldner, M., Rowan-Kenyon, H., Inkelas, K. K., Garvey, J., & Robbins, C. (2012). Supporting students' intentions to persist in STEM disciplines: The role of living-learning programs among other social-cognitive factors. *The Journal of Higher Education*, 83(3), 311-336.

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.

Staples, D. S., Hulland, J. S., & Higgins, C. A. (1998). A self-efficacy theory explanation for the management of remote workers in virtual organizations. *Journal of computer-mediated communication*, 3(4), JCMC342.

Steinke, J. (2005). Cultural Representations of Gender and Science: Portrayals of Female Scientists and Engineers in Popular Films. *Science Communication*, 27(1), 27-63.

Steinke, J., Applegate, B., Lapinski, M., Ryan, L., & Long, M. (2012). Gender differences in adolescents' wishful identification with scientist characters on television. *Science Communication*, 34(2), 163-199.

Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., & Kuchibhotla, S. (2007). Assessing media influences on middle school-aged children's perceptions of women in science using the Draw-A-Scientist Test (DAST). *Science Communication*, 29(1), 35-64.

Tai, R. H., Liu, C. Q., Maltese, A.V., & Fan, X. (2006). Planning early for careers in science. *Life sci*, 1, 0-2.

Tyler-Wood, T., Knezek, G., & Christensen, R. (2010). Instruments for assessing interest in STEM content and careers. *Journal of Technology and Teacher Education*, 18(2), 345-368.

U.S. Bureau of Labor Statistics. (2014). *STEM 101: Intro to tomorrow's Jobs*. Retrieved from <http://www.bls.gov/careeroutlook/2014/spring/art01.pdf>.

U.S. Department of Commerce. (2017). *Women in STEM: 2017 Update*. Retrieved from <http://www.esa.doc.gov/sites/default/files/women-in-stem-2017-update.pdf>.

Van Looy, J., Courtois, C., De Vocht, M., & De Marez, L. (2012). Player identification in online games: Validation of a scale for measuring identification in MMOGs. *Media Psychology*, 15(2), 197-221.

Vogt, C. M., Hocevar, D., & Hagedorn, L. S. (2007). A social cognitive construct validation: Determining women's and men's success in engineering programs. *The Journal of higher education*, 78(3), 337-364.

Washington Career Bridge (2018). Career Quiz. Retrieved from http://www.careerbridge.wa.gov/Survey_Cluster.aspx.

Weisgram, E. S. (2016). The cognitive construction of gender stereotypes: Evidence for the dual pathways model of gender differentiation. *Sex Roles*, 75(7-8), 301-313.

Wentling, R. M., & Camacho, C. (2008). Women engineers: Factors and obstacles related to the pursuit of a degree in engineering. *Journal of women and minorities in science and engineering*, 14(1).

Woods-Townsend, K., Christodoulou, A., Rietdijk, W., Byrne, J., Griffiths, J. B., & Grace, M. M. (2015). Meet the Scientist: The Value of Short Interactions Between Scientists and Students. *International Journal of Science Education*, Part B, 1-25.

Yee, N., & Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human communication research*, 33(3), 271-290.

Yee, N., Bailenson, J. N., & Ducheneaut, N. (2009). The Proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research*, 36(2), 285-312

Zahry, N. R., & Besley, J. C. (2017). Students' Perceptions of Agriculture and Natural Resources Majors: Understanding STEM Choice. *Natural Sciences Education*, 46(1).