



HAL
open science

Toward a Tool Against Stereotype Threat in Math: Children's Perceptions of Virtual Role Models

Marjorie Armando, Isabelle Régner, Magalie Ochs

► **To cite this version:**

Marjorie Armando, Isabelle Régner, Magalie Ochs. Toward a Tool Against Stereotype Threat in Math: Children's Perceptions of Virtual Role Models. ICMI 2023 International Workshop on Multimodal, Interactive Interfaces for Education, Oct 2023, Paris, France. 10.1145/3610661.3616179 . hal-04195168v1

HAL Id: hal-04195168

<https://hal.science/hal-04195168v1>

Submitted on 6 Sep 2023 (v1), last revised 23 Jan 2024 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Toward a Tool Against Stereotype Threat in Math: Children’s Perceptions of Virtual Role Models

Marjorie Armando
Aix Marseille Univ, Université de
Toulon, CNRS, LIS, Marseille, France
Aix Marseille Univ, CNRS, LPC,
Marseille, France
marjorie.armando@univ-amu.fr

Isabelle Régner
Aix Marseille Univ, CNRS, LPC,
Marseille, France
isabelle.regner@univ-amu.fr

Magalie Ochs
Aix Marseille Univ, Université de
Toulon, CNRS, LIS, Marseille, France
magalie.ochs@lis-lab.fr

ABSTRACT

Users have the propensity to apply real-life gender stereotypes on virtual agents. Such stereotypes could impair users’ learning in virtual environments with pedagogical agents. In particular, girls and women are negatively affected by these stereotypes in math tests. The phenomenon called Stereotype Threat refers to the fear of being negatively stereotyped in a skill area that reduces learning and performance. We wonder if virtual agents can be used to reduce Stereotype Threat effects in math. For this purpose, a virtual agent has to be perceived as a social role model. We present in this paper two pilot studies with children to explore their perception of a male agent and a female agent, both child agents, whether they can establish rapport or not. We also investigate if children apply gender stereotypes when they interact with child virtual agents, and if such agents could embody an efficient role model. Results show children found the male agent as competent as the female agent, which is usually not the case in other research involving adult participants. More importantly, the child virtual agents are perceived as social role models, opening the possibility to use them to improve children performance in STEM fields.

KEYWORDS

Pedagogical Virtual Agents, Perceptions, Learning Environments, Virtual Role Models, Gender Stereotypes, Stereotype Threat, Children, Education

1 INTRODUCTION

Pedagogical agents refer to virtual characters designed to enhance learning outcomes within educational environments [19, 34, 35]. Certain traits of pedagogical agents have the potential to influence the learning process [11, 26]. However, few studies have evaluated the impact of pedagogical agents as social role model [1]. Research in Social Cognition has documented the positive impact of social role models on performance and learning. In particular, studies have shown that *Stereotype Threat* (ST), the fear of confirming a negative stereotype in a skill area, impair learning and performance [33, 38]. ST applies to girls and women in math tests [30]. ST effects can be reduced using different strategies, such as presenting a successful role model before taking a test [2, 3]. *Presenting pedagogical agents as role models could be a potential solution for reducing the effects of ST for girls in math.* Few studies have used virtual role models to change women’s perceptions on Science, Technology, Engineering, and Mathematics (STEM) [31]. However, as found by Buck et al. [9], female participants did not immediately perceive female scientists

as role models solely through interaction, but after establishing personal connections with them. In Social Psychology, rapport is a way to establish these kinds of connections through a positive attitude (e.g., acquiescence, smiles), mutual attention (e.g., mutual gaze) and coordination of behaviors (e.g., synchrony, mimicry) [40]. In the field of virtual agents, virtual rapport has been studied and induced to improve the global human-agent interaction [4, 12, 16]. Given the importance of rapport to create a role model, in this article, we aim at investigating children’s perception of rapport during an interaction with a child virtual agent. More generally, we study if a virtual child can be perceived by a child as a role model by trying to answer to the following questions: (1) Can children identify to a virtual role model? (2) Can children perceive virtual rapport? (3) How do children perceive virtual agents? As far as we know, there is no research on how children can identify to virtual agents in order to perceive them as role models, nor on how children perceive virtual rapport. In this article, we have two main objectives: to verify that we can induce a virtual rapport through a virtual child agent towards children users, and verify children can perceive those virtual children agents as role models. As we are interested in math stereotype that opposes skills and abilities of girls/women and boys/men, participants were girls and boys, and agents’ appearances were feminine or masculine.

2 RELATED WORK

2.1 Theoretical Background

Stereotype Threat. Several research in the field of social cognition has shown that the fear of confirming a negative stereotype in a skill domain produces negative thoughts, which in turn reduces working memory capacity of individuals and impair learning and performance [33]. This phenomenon, called *Stereotype Threat* (ST), applies to different stereotypes and social groups, such as girls and women in math [36]. ST may explain why girls tend to perform less well in math than boys on standardized tests. The apprehension of reinforcing the negative stereotype of women’s abilities in math can affect their performance on a task [13]. This may also account for why women are less inclined than men to pursue in scientific professions [8], as when ST becomes a chronic condition, it can lead to demotivation and disinterest [37].

Reducing ST Effects With Social Role Models. Providing a female social role model who succeed in math [2] can lead to improve performance of girls and women in math, as shown by Bagès and Martinot [2] when young learners read a story of a female social role model in math just before a math test. Furthermore, when the

role model explained how he/she get to succeed in math, learners' performance improved even more, regardless of the role model's and learner's gender. To be effective, the social role model must be slightly older than the target student and portrayed as someone who is working to succeed, as it is a way to counteract the belief in innate math abilities [37], and therefore, it is a way to counteract ST.

2.2 Virtual Agents' gendered appearance

Impact on individuals' perceptions. As individuals communicate with virtual agents by applying social rules and expectations as social beings [21], it is not surprising that they also apply gender stereotypes to virtual agents and their interactions with them [1]. Female virtual agents are usually seen as less expert, less knowledgeable, and less powerful than male virtual agents [7, 23], and more likable and attractive than male virtual agents [18, 23].

Those perception differences can affect people's decisions, as individuals tend to listen more to virtual agents when their gender stereotypically matched the context: participants followed more advices from a female agent when the topic was about cosmetics, and they followed more advices from a male agent when the topic was about sports [17]. Moreover, participants tend to attribute communion traits (e.g., affectionate) to robots perceived as women [27], and they attribute conventional male roles to male robots and conventional female roles to female robots. [25]. It is important to note participants and virtual agents' appearance in those studies were adults. In the pilot studies presented in this article, we focus on children: participants and virtual agents' appearance were children. **Virtual Agents as Role Models.** Using pedagogical agents as role models may offer a promising approach to reduce ST effects. While research in this area is still limited, some studies have investigated the effects of gendered appearances of virtual agents on learners' interest in mathematics and their performance, showing positive outcomes. For example, Rosenberg-Kima et al. [31] showed the efficacy of a female virtual agent engineer, who narrated a story about accomplished female role models in STEM, in increasing the interest of women in STEM. Female participants were less likely to adhere to conventional STEM stereotypes after interacting with the female agent than those who interacted with the male agent, and were more inclined to believe in the potential for women to succeed in STEM. Despite these positive outcomes, some gender stereotypes still persisted: they were slightly more inclined to consider STEM as useful with a male agent engineer. In a similar study by Plant et al. [28], after engaging with a female agent, all participants demonstrated enhanced performance, self-efficacy, interest and usefulness of STEM. Interestingly, male participants were less likely to subscribe to conventional STEM stereotypes regardless of the agent's gender, while female participants were less likely to endorse traditional STEM stereotypes when interacting with a female agent or in the absence of an agent, as opposed to a male agent. Another similar study by Rosenberg-Kima et al. [32] showed that Black virtual agents were more effective in increasing STEM interest and reducing STEM gender stereotypes for Black women, while female virtual agents (Black or White) were more effective for White women on the same criteria. These results underscore the significance of other factors, such as the virtual agents' color skin.

This observation aligns with Bandura's social cognitive learning theory, which suggests that people frequently learn by emulating those whom they perceive as similar or superior to themselves and who, as a result, are accepted as role models [5].

3 PILOT STUDIES OF PEDAGOGICAL VIRTUAL CHILDREN

In the pilot studies presented in this article, we analyzed the impact of virtual rapport and agents' gender on students' identification and other aspects that make an effective role model through two kind of agents: the Rapport Agent and the Neutral Agent (both either male or female). We specifically designed the Rapport Agents to create a virtual rapport between the user and the agent. Their behaviors are described below. To our knowledge, these analyses have never been done with children participants. We hypothesized that: (H1) Children perceive a stronger feeling of rapport when they interact with a Rapport Agent compared to a Neutral Agent, (H2) Children perceive the Rapport Agents as more likable than the Neutral Agents, (H3) Children identify better with the Rapport Agents compared to the Neutral Agents, (H4) Children perceive the Female Agents as more likable than the Male Agents, (H5) Children perceive the Female Agents as less successful in math than the Male Agents. In this article, we propose to evaluate these hypotheses through two pilot studies. In the first study, the virtual children only introduced themselves and said they were successful in math because they worked hard. In the second one, in addition, the agents explained a math lesson. It allows us to be as close as possible to our future research on using these agents on ST-induced condition, to analyze their actual impact on children's performance in math.

3.1 The Virtual Children

Agents' appearance. Two virtual agents, a boy and a girl, were made on Character Creator 3 and animated on Unity. Following Bandura's social cognitive learning theory [5] and the studies of Bagès and Martinot [2] and Bagès et al. [3], we designed the agents to be slightly older than participants, and looking like the majority of participants in term of ethnicity. All agents wore the same clothes. The agents were sitting in a classroom like a student. The classroom was designed by following a study by Cheryan et al. [10] where no STEM-stereotypical items decorate the place to avoid girls to feel they do not belong there, and it looks like a typical French classroom. The female virtual agent was made first thanks to a free photo of a real girl. The male virtual agent was then made by masculinizing the female agent following the study of Nag and Yalçın [20]. The agents are illustrated in Table 1 in the supplementary material.

Agents' behavior. To induce rapport, we created the *Rapport Agents* with particular communication strategies [14, 15, 22, 29, 41, 42]: (1) *self-disclosure*: head nods, gazing at the user, high usage of first person singular pronoun to speak about personal concerns, (2) *intimate self-disclosure*: describing more intimate personal concerns, (3) *negative self-disclosure*: validating the user's negative personal concerns so they won't feel disconnected from the agent, (4) *acknowledgment*: head nods and smiles, (5) *praising* specific behavior, (6) *referring to shared experience*: remembering what the user said, (7) *conveying intimacy*: non-dominance, forward body lean and open body posture, (8) *back-channeling*: signifying the agent's

attention or agreement though head nods, (9) *enthusiasm*: quick, demonstrative gestures, many different facial expressions, lot of eye contact and a highly descriptive vocabulary, (10) *teaching*: deictic and iconic gestures, jabbing one's arm in the air with an index finger extended to underscore a lecture point being verbally presented. As opposed to the Rapport Agents, we created the *Neutral Agents*, who only have idle animations, blinking, and position shift. Facial animations (blinks, labial animation, eyebrows movements) were computed with Salsa LipSync Suite V2. Children voices were generated with Text-to-Speech method on Acapela-Box.

Agents' intervention. In both pilot studies, the Neutral and Rapport Agents introduced themselves and stated they were here to help the student in math, because they have good grades as they study for it. This is particularly important to be an effective role model and potentially reduce ST effects (see Section 2.1). In the second pilot study, they explained a math lesson. Unlike the Neutral Agents, the Rapport Agents asked the student's name and how they feel about math. They also coordinated their answers to students' answers. Participants could answer all agents' questions through several fixed set of answers. When explaining the math lesson, the Rapport Agents made sure students were understanding everything so far by asking them. Also, the Rapport Agents directly spoke to students when showing examples, unlike the Neutral Agents.

3.2 Measures

All scales in this article were evaluating through Likert-scale items, from 1="I totally disagree" to 5="I completely agree". We adapted and translated in French the Godspeed questionnaire [6] to measure *agent's math success* (5 items, e.g., competent in math), *agent's likability* (5 items, e.g., nice), and *agent's level of anthropomorphism* (7 items, e.g., move like a robot). In addition, we measured the feeling of *rapport* (17 items, e.g., the agent appreciated talking to me) by adapting the Virtual Intimacy Scale (VIS) [29]. To measure the potential of agents to be role models, we adapted and translated in French a *role model identification* scale (4 items, e.g., lot of commons with the agent) [24]. We created another scale, *success identification*, to measure how much children identify with the agents' success in math (4 items, e.g., I am capable to succeed in math just like the agent). It is important to note two classes did not evaluate the *success identification* scale in the second pilot study (Section 5). Internal consistency of each scale has been calculated with Cronbach's alpha for each pilot study (see the Results paragraph in Section 4 and 5). Measures' score was calculated by averaging their items' score.

4 PILOT STUDY 1: MIXED DESIGN

Experimental design. This pilot study was a 2 (*agents' gender*: female/male) \times 2 (*agents' behavior*: rapport/neutral) mixed design. Only the last variable was within-subject. Order of agents' behavior was counterbalanced.

Participants. Two French classes participated in this pilot study. A total of N=49 participants were recruited (26 boys, 23 girls, $M_{age}=11.39$, $SD=.09$). Twenty-five children interacted with the Male Agents, 24 with the Female Agents.

Experimental task. Participants were at their school with their math teacher. A female experimenter explained they would watch the same virtual agent representing a real high school student twice,

but behaving differently. Participants have been notified their answers were anonymous. They were all equipped with tablets and headphones. They first answered some questions about them (age, gender, how often they play video games), watched an agent (either the Rapport Agent or the Neutral Agent), answered the questionnaire described in Section 3.2, watched the second agent (behaving differently than the first one) and answered the questionnaire again. Agents introduced themselves and stated they were successful in math because they worked hard for it. Participants had one hour to complete the task. Participants interacted with each agent for about 3 minutes.

Results. We first calculated the internal consistency of each scale using Cronbach's alpha: agent's math success ($\alpha=.64$ on Rapport Agent's questionnaire, $\alpha=.78$ on Neutral Agent questionnaire), agent's likability ($\alpha=.85$ on Rapport Agent's questionnaire, $\alpha=.88$ on Neutral Agent's questionnaire), agent's level of anthropomorphism ($\alpha=.79$ on both questionnaires), feeling of rapport ($\alpha=.87$ on Rapport Agent's questionnaire, $\alpha=.86$ on Neutral Agent's questionnaire), role model identification ($\alpha=.72$ on Rapport Agent's questionnaire, $\alpha=.81$ on Neutral Agent's questionnaire), and agent's math success identification ($\alpha=.78$ on Rapport Agent's questionnaire, $\alpha=.76$ on Neutral Agent's questionnaire). Results from Skewness and Kurtosis tests supported the assumption of normality. The optimal analysis with all independent variables would have been 2 (*participants' gender*: girl/boy) \times 2 (*agents' gender*: female/male) \times 2 (*agents' behavior*: rapport/neutral) mixed ANOVA. Given the small number of participants (9 girls saw the Female Agents), we decided to perform a two-way mixed ANOVA of agents' behavior (rapport/neutral) and agents' gender (female/male) to test our hypothesis. Children perceived a stronger feeling of rapport in presence of the Rapport Agents ($M=4.04$, $SD=.08$) compared to the Neutral Agents ($M=3.51$, $SD=.09$), $p<.001$, $\eta_p^2=.42$, validating H1. Also, the Rapport Agents ($M=4.56$, $SD=.08$) were perceived as more likable than the Neutral Agents ($M=4.14$, $SD=.11$), $p<.001$, $\eta_p^2=.27$, validating H2. Children had the tendency to find the Rapport Agents ($M=3.04$, $SD=.11$) as more human than the Neutral Agents ($M=2.88$, $SD=.11$), $p=.061$, $\eta_p^2=.07$. Children did not identify more to the Rapport Agents than the Neutral Agents, neither they did on Rapport Agents' math success, invalidating H3. Nonetheless, a test of Student on the mid point of Likert-scale revealed children identified to both type of agents, the Rapport Agents ($M=3.49$, $SD=.89$, $t(48)=3.84$, $p<.001$) and the Neutral Agents ($M=3.28$, $SD=.95$, $t(48)=2.08$, $p=.043$). Children also identified to the Rapport Agents' math success ($M=4.47$, $SD=.75$, $t(48)=13.64$, $p<.001$) and the Neutral Agents' math success ($M=4.52$, $SD=.64$, $t(48)=16.62$, $p<.001$). No effect of agents' gender was found, invalidating H4 and H5. In the supplementary material: figure 1 shows scores of each scale depending on agents' behavior. Figure 2 shows scores of each scale depending on agents' gender. Table 2 details descriptive results.

5 PILOT STUDY 2: BETWEEN-SUBJECT DESIGN

Experimental design. This pilot study was a 2 (*agents' gender*: female/male) \times 2 (*agents' behavior*: rapport/neutral) between-subject.

Participants. Eight French classes participated to this pilot study. A total of N=95 participants were recruited (51 boys, 44 girls, $M_{age}=11.73$,

$SD=.09$). Twenty-four children interacted with the Female Rapport Agent, 22 with the Female Neutral Agent, 19 with the Male Rapport Agent, 30 with the Male Neutral Agent.

Experimental task. Participants were at their school with their math teacher. A female experimenter explained they would watch a virtual agent representing a real high school student. This agent would introduce itself and explain a math lesson. The experimenter explained that their opinions were very important, as we were trying to design the best virtual agent to help students improve in math. Participants have been notified their answers were anonymous. They were all equipped with tablets and headphones. They first answered some questions about them (age, gender, how often they play video games), watched an agent (either the Rapport Agent or the Neutral Agent, male or female), then answered the questionnaire described in Section 3.2. Participants interacted with an agent for about 6 minutes.

Results. We first calculated the internal consistency of each scale using Cronbach's alpha: agent's math success ($\alpha=.72$), agent's likability ($\alpha=.82$), agent's level of anthropomorphism ($\alpha=.75$), feeling of rapport ($\alpha=.86$), role model identification ($\alpha=.81$), and agent's math success identification ($\alpha=.67$). Results from Skewness and Kurtosis tests supported the assumption of normality. Just like in the first pilot study, the optimal analysis to answer all questions with all independent variables would have been a 2 (*participants' gender*: girl/boy) \times 2 (*agents' gender*: female/male) \times 2 (*agents' behavior*: rapport/neutral) ANOVA. Given the small number of participants in some conditions (9 girls saw the male Neutral Agent), we decided to perform a 2 (*agents' behavior*) \times 2 (*agents' gender*) ANOVA. Children had the tendency to perceive a stronger feeling of rapport when interacting with the Rapport Agents ($M=3.67$, $SD=.10$) compared to the Neutral Agents ($M=3.41$, $SD=.09$), $p=.057$, $\eta_p^2=.04$. Interestingly, this main effect was moderated by the agent's gender: children significantly perceived a stronger feeling of rapport with the Female Agents ($M=3.69$, $SD=.10$) compared to the Male Agents ($M=3.39$, $SD=.10$), $p=.029$, $\eta_p^2=.05$. H1 is partially validated: the feeling of rapport depends on the agent's gender. Moreover, children had the tendency to perceive the Female Agents ($M=4.38$, $SD=.10$) as more likable than the Male Agents ($M=4.10$, $SD=.10$), $p=.054$, $\eta_p^2=.04$. H4 is partially validated. A significant interaction effect occurred between agents' behavior and agents' gender on the level of anthropomorphism, $p=.001$, $\eta_p^2=.12$: when interacting with the Rapport Agents, children found the Female Agent ($M=3.13$, $SD=.14$) as more human than the Male Agent ($M=2.37$, $SD=.16$), $p<.001$. Moreover, the Female Rapport Agent was evaluated as more human than the Female Neutral Agents ($M=2.59$, $SD=.15$), $p=.009$. Interestingly, it was the contrary for the Male Agents: the Male Rapport Agent ($M=2.37$, $SD=.16$) was perceived as less human than the Male Neutral Agent ($M=2.83$, $SD=.13$), $p=.024$. Concerning identification, children identified more to the Rapport Agents' math success ($M=4.58$, $SD=.14$) compared to the Neutral Agents ($M=4.22$, $SD=.10$), $p=.036$, $\eta_p^2=.07$, validating H3. They also identified more to the Female Agents ($M=3.42$, $SD=.15$) compared to the Male Agents ($M=2.97$, $SD=.14$), $p=.031$, $\eta_p^2=.05$. H2 and H5 were not validated. In the supplementary material: figure 3 shows scores of each scale depending on agents' behavior. Figure 4 shows scores of each scale

depending on agents' gender. Both figures display significant results of tests of Student on the mid point of Likert-scale. Table 3 details descriptive results.

6 DISCUSSION AND PERSPECTIVES

The present pilot studies aimed at evaluating virtual agents as role models for children to reduce ST effects in future studies, depending on whether they exhibit a neutral or a rapport behavior. Both pilot studies showed children perceived virtual rapport and were able to identify to the children virtual agents' success in math. They identified better with the Rapport Agents compared to the Neutral Agents, highlighting the importance of rapport. They perceived virtual children as successful in math and likable. Contrary to adults in most studies, children did not perceive Male Agents as more competent in math than Female Agents. However, we think it would be hasty to conclude children do not match agents' gender with stereotypical skills based solely on our results. A study by Nag and Yalçın [20] showed adult participants applied competent traits the same way to male agents and female agents, contradicting other studies so far. However, their agents were very similar in appearance, which is also the case for our agents in this article. Going beyond gender, while important, the level of perceived masculinity can also play a role, as shown by Świdrak et al. [39] where male participants accepted more offers from masculine agents (whatever agents' gender) in a negotiation game. In our pilot studies, participants were not asked to rate the level of agents' masculinity/femininity, which could have been interesting given the results. However, gender stereotypes mildly persisted: in the second pilot study, children had the tendency to perceive the Female Agents as more likable than the Male Agents. All of these results are important for future research in virtual learning environments and pedagogical virtual agents, opening the possibility of a rapport-female virtual role model to not only be effective to improve students' math performance, but also to encourage more girls in STEM by having more female representation. The next step in this research is to use these agents in ST-induced condition to analyze their actual impact on children's math performance.

REFERENCES

- [1] Marjorie Armando, Magalie Ochs, and Isabelle Régner. 2022. The Impact of Pedagogical Agents' Gender on Academic Learning: A Systematic Review. *Frontiers in Artificial Intelligence* 5 (2022).
- [2] Céline Bagès and Delphine Martinot. 2011. What is the best model for girls and boys faced with a standardized mathematics evaluation situation: A hardworking role model or a gifted role model? *British Journal of Social Psychology* 50, 3 (2011), 536–543.
- [3] Céline Bagès, Catherine Verniers, and Delphine Martinot. 2016. Virtues of a hardworking role model to improve girls' mathematics performance. *Psychology of Women Quarterly* 40, 1 (2016), 55–64.
- [4] Jeremy N Bailenson and Nick Yee. 2005. Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological science* 16, 10 (2005), 814–819.
- [5] Albert Bandura. 1986. Social foundations of thought and action: A social cognitive theory.
- [6] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1 (2009), 71–81.
- [7] Amy L Baylor and Yanghee Kim. 2004. Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. In *International Conference on Intelligent Tutoring Systems*. Springer, 592–603.
- [8] Talia Ben-Zeev, Seth Duncan, and Chad Forbes. 2005. Stereotypes and math performance. (2005).

- [9] Gayle A Buck, Vicki L Plano Clark, Diandra Leslie-Pelecky, Yun Lu, and Patricia Cerda-Lizarraga. 2008. Examining the cognitive processes used by adolescent girls and women scientists in identifying science role models: A feminist approach. *Science Education* 92, 4 (2008), 688–707.
- [10] Sapna Cheryan, Andrew N Meltzoff, and Saenam Kim. 2011. Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education* 57, 2 (2011), 1825–1835.
- [11] Jonathan Gratch, Ning Wang, Jillian Gerten, Edward Fast, and Robin Duffy. 2007. Creating rapport with virtual agents. In *International Workshop on Intelligent Virtual Agents*. Springer, 125–138.
- [12] Lixing Huang, Louis-Philippe Morency, and Jonathan Gratch. 2011. Virtual Rapport 2.0. In *Intelligent Virtual Agents: 10th International Conference, IVA 2011, Reykjavik, Iceland, September 15–17, 2011. Proceedings 11*. Springer, 68–79.
- [13] Pascal Huguet and Isabelle Regner. 2007. Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of educational psychology* 99, 3 (2007), 545.
- [14] Sin-Hwa Kang and Jonathan Gratch. 2010. Virtual humans elicit socially anxious interactants' verbal self-disclosure. *Computer Animation and Virtual Worlds* 21, 3-4 (2010), 473–482.
- [15] Sin-Hwa Kang, Jonathan Gratch, Candy Sidner, Ron Artstein, Lixing Huang, and Louis-Philippe Morency. 2012. Towards building a virtual counselor: modeling nonverbal behavior during intimate self-disclosure. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. 63–70.
- [16] Nicole C Krämer, Bilge Karacora, Gale Lucas, Morteza Dehghani, Gina Rütter, and Jonathan Gratch. 2016. Closing the gender gap in with friendly male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers & Education* 99 (2016), 1–13.
- [17] Eun-Ju Lee. 2003. Effects of "gender" of the computer on informational social influence: The moderating role of task type. *International Journal of Human-Computer Studies* 58, 4 (2003), 347–362.
- [18] Renaud Lunardo, Gregory Bressolles, et al. 2016. The interacting effect of virtual agents' gender and dressing style on attractiveness and subsequent consumer online behavior. *Journal of Retailing and Consumer Services* 30, C (2016), 59–66.
- [19] Mohamed Taib Mohtadi, Abdelmajid Hajami, and Hakim Allali. 2014. Pedagogical agent for metacognitive scaffolding in interactive learning environments. In *2014 International Conference on Multimedia Computing and Systems (ICMCS)*. IEEE, 652–656.
- [20] Procheta Nag and Özge Nilay Yalçın. 2020. Gender Stereotypes in Virtual Agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*. 1–8.
- [21] Clifford Nass and Youngme Moon. 2000. Machines and mindlessness: Social responses to computers. *Journal of Social Issues* 56, 1 (2000), 81–103.
- [22] Hanh Thi Nguyen. 2007. Rapport building in language instruction: A microanalysis of the multiple resources in teacher talk. *Language and Education* 21, 4 (2007), 284–303.
- [23] Jay F Nunamaker, Douglas C Derrick, Aaron C Elkins, Judee K Burgoon, and Mark W Patton. 2011. Embodied conversational agent-based kiosk for automated interviewing. *Journal of Management Information Sys* 28, 1 (2011), 17–48.
- [24] Laurie T O'Brien, Aline Hitti, Emily Shaffer, Amanda R Van Camp, Donata Henry, and Patricia N Gilbert. 2017. Improving girls' sense of fit in science: Increasing the impact of role models. *Social Psychological and Personality Science* 8, 3 (2017), 301–309.
- [25] Oronzo Parlangeli, Paola Palmitesta, Margherita Bracci, Enrica Marchigiani, and Stefano Guidi. 2023. Gender role stereotypes at work in humanoid robots. *Behaviour & Information Technology* 42, 3 (2023), 316–327.
- [26] Florian Pecune, Angelo Cafaro, Magalie Ochs, and Catherine Pelachaud. 2016. Evaluating social attitudes of a virtual tutor. In *International Conference on Intelligent Virtual Agents*. Springer, 245–255.
- [27] Giulia Perugia, Latisha Boor, Laura van der Bij, Okke Rikmenspoel, Robin Foppen, and Stefano Guidi. 2023. Models of (Often) Ambivalent Robot Stereotypes: Content, Structure, and Predictors of Robots' Age and Gender Stereotypes. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 428–436.
- [28] E Ashby Plant, Amy L Baylor, Celeste E Doerr, and Rinat B Rosenberg-Kima. 2009. Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education* 53, 2 (2009), 209–215.
- [29] Delphine Potdevin. 2020. *Vers des agents conversationnels animés sociaux: Quelle influence de l'intimité virtuelle sur l'expérience utilisateur et la relation-client?* Ph. D. Dissertation. Université Paris-Saclay.
- [30] Isabelle Régner, Jennifer R Steele, Nalini Ambady, Catherine Thinus-Blanc, and Pascal Huguet. 2014. Our future scientists: A review of stereotype threat in girls from early elementary school to middle school. *Revue internationale de psychologie sociale* 27, 3 (2014), 13–51.
- [31] R.B. Rosenberg-Kima, A.L. Baylor, E.A. Plant, and C.E. Doerr. 2008. Interface agents as social models for female students: The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior* 24, 6 (2008), 2741–2756.
- [32] Rinat B Rosenberg-Kima, E Ashby Plant, Celeste E Doerr, and Amy L Baylor. 2010. The influence of computer-based model's race and gender on female students' attitudes and beliefs towards engineering. *Journal of Engineering Education* 99, 1 (2010), 35–44.
- [33] Toni Schmader and Michael Johns. 2003. Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology* 85, 3 (2003), 440.
- [34] Noah L Schroeder, Olusola O Adesope, and Rachel Barouch Gilbert. 2013. How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research* 49, 1 (2013), 1–39.
- [35] Noah L Schroeder, William L Romine, and Scotty D Craig. 2017. Measuring pedagogical agent persona and the influence of agent persona on learning. *Computers & Education* 109 (2017), 176–186.
- [36] Steven J Spencer, Claude M Steele, and Diane M Quinn. 1999. Stereotype threat and women's math performance. *Journal of Experimental Social Psychology* 35, 1 (1999), 4–28.
- [37] Claude M Steele. 1997. A threat in the air: How stereotypes shape intellectual identity and performance. *American psychologist* 52, 6 (1997), 613.
- [38] Claude M Steele and Joshua Aronson. 1995. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* 69, 5 (1995), 797.
- [39] Justyna Świdrak, Grzegorz Pochwatko, and Andrea Insabato. 2021. Does an agent's touch always matter? Study on virtual Midas touch, masculinity, social status, and compliance in Polish men. *Journal on Multimodal User Interfaces* 15, 2 (2021), 163–174.
- [40] Linda Tickle-Degnen and Robert Rosenthal. 1990. The nature of rapport and its nonverbal correlates. *Psychological Inquiry* 1, 4 (1990), 285–293.
- [41] Ran Zhao, Alexandros Papangelis, and Justine Cassell. 2014. Towards a dyadic computational model of rapport management for human-virtual agent interaction. In *Intelligent Virtual Agents: 14th International Conference, IVA 2014, Boston, MA, USA, August 27–29, 2014. Proceedings 14*. Springer, 514–527.
- [42] Ran Zhao, Tanmay Sinha, Alan W Black, and Justine Cassell. 2016. Automatic recognition of conversational strategies in the service of a socially-aware dialog system. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 381–392.