Using Intelligent Agents to Examine Gender in Negotiations

Emmanuel Johnson
University of Southern California
Playa Vista, CA
ejohnson@ict.usc.edu

Jonathan Gratch
University of Southern California
Playa Vista, CA
gratch@ict.usc.edu

Jill Boberg
University of Southern California
Playa Vista, CA
boberg@ict.usc.edu

David DeVault
Anticipant Speech, Inc.
Irvine, CA
devault@anticipant.ai

Peter Kim
University of Southern California
Los Angeles, CA
kimpeter@marshall.usc.edu

Gale Lucas
University of Southern California
Playa Vista, CA
lucas@ict.usc.edu

ABSTRACT

Women earn less than men in technical fields. Competing theories have been offered to explain this disparity. Some argue that women underperform in negotiating their salary, in-part due to language in job descriptions, called *gender triggers*, which leave women feeling disadvantaged in salary negotiations. Others point to structural and institutional bias: i.e., recruiters make better offers to men even when women exhibit equal negotiation skills. As a final salary is co-constructed though an interaction between employees and recruiters, it is difficult to disentangle these views. Here, we discuss how intelligent virtual agents serve as powerful methodological tools that lend new insight into this psychological debate. We use virtual negotiators to examine the impact of gender triggers on computer science (CS) undergraduates that engaged in a simulated salary negotiation with an automated recruiter. We find that, regardless of gender, CS students are reluctant to negotiate, and this hesitancy likely lowers their starting salary. Even when they negotiate, students show little skill in discovering tradeoffs that could enhance their salary, highlighting the need for negotiation training in technical fields. Most importantly, we find little evidence that gender triggers impact women's negotiated outcomes, at least within the field of CS. We argue that findings that emphasize women's individual deficits may reflect a lack of experimental control, which intelligent agents can help correct, and that structural and institutional explanations of inequity deserve greater attention.

KEYWORDS

Negotiation, gender, bias, methodological tools

WOODSTOCK'18, June, 2018, El Paso, Texas USA

https://doi.org/10.1145/1234567890

1 Introduction

Those that ask will receive: Starting salaries have a strong impact on career earnings and even small initial differences can compound into substantial differences over time. Yet many people accept the first salary they are offered. A 2016 survey by Glassdoor (one of the world's largest recruiting sites) found that 60% of new hires failed to negotiate [1] and a 2019 Swedish national survey found equivalent numbers [2]. This failure to negotiate has consequences: candidates that engage in negotiation achieve higher salaries, particularly when they engaged in open discussions of tradeoffs across different elements of their compensation package [3]. These disparities are concerning as negotiation success varies systematically by discipline and demographics. For example, a recent Academy of Sciences report suggests that student in science, technology, engineering and mathematics (STEM) are especially underprepared to negotiate successfully [4, p. 2] and numerous studies document women's inability to obtain equitable salaries [2, 5, 6]. Understanding the reasons behind these disparities can inform interventions for addressing structural inequities across our society and maximizing the inclusion of underrepresented groups CS.

Research on negotiation disparity focuses on either negotiation inputs or processes. *Inputs* are information and dispositions that might influence candidates in advance of a negotiation. For example, text analysis of job postings shows they use language that can put women on the defensive [7]. Women may also bring different goals to the negotiation, such as greater willingness to accept a low salary in exchange for job security or flexibility [8]. *Processes*, in contrast, refer to actions parties initiate during a negotiation. For example, women may face more aggressive opening offers and greater use of deception by their counterparts [9]. In this paper, we explore the potential for intelligent agents to yield insights into the relationship between negotiation inputs and processes.

Existing research into negotiation processes has adopted one of two experimental methods (dyadic or scripted), each with its own methodological limitations. We argue that intelligent agents offer a third methodological approach that complements the weaknesses of existing methods. In the *dyadic approach*, two participants negotiate with each other (one playing the role of a hiring


BigTech

Company Description:
BigTechCo is a technology company that is driving the information revolution. With widely used products in multiple industries, including consumer hardware, cloud computing, e-commerce, and web search, BigTechCo is revolutionizing how people interact with technology ...

Job Overview:
BigTechCo is seeking a **Software Engineer** to join a team of talented software engineers, data scientists, and machine learning researchers who are trying to solve some of the hardest problems in technology today. **As a software engineer, you will be involved in the design and development of BigTechCo's software platform for a specific new product.** You will work closely with the other software engineers as well as engineers from other groups to design...

Required Characteristics
- Ability to drive your team towards technical perfection as part of our strong engineering culture
- Confidence working in a large application required

Figure 1: Job description (partial) showing female-threatening required characteristics

manager and the other playing the role of the prospective employee). Dyadic studies have found, for example, that women negotiate lower salaries than men when job qualifications are described in stereotypically male ways [10, 11]. The advantage of the dyadic approach is it allows for a rich and natural give-and-take between participants (analogous to what they would face in a real salary negotiation) and facilitates the study of emergent processes such as information exchange and value-creation [12]. The disadvantage is the difficulty in attributing causality to poor outcomes. Do women perform badly because they are poor negotiators or do they perform poorly because their partners treat them differently than men [9]?

To address causality, the *scripted approach* replaces the hiring manager with a completely deterministic computer program that makes the identical sequence of statements and concessions, regardless of the participant's gender or behavior [e.g., 13]. This ensures the hiring manager is truly blind to the participant's gender and other characteristics (thus any differences can be safely attributed to the employee). However, this increase in experimental control comes at a great cost. One of the hallmarks of strong negotiators is the ability to mutually adapt to one's partner: to understand their opponent's interests, communicate their own, and guide the negotiation towards win-win tradeoffs [3, 12]. Such interactive processes and value-creation are precluded by deterministic scripts.

Intelligent negotiation agents offer a way to merge the strengths of these two methodological approaches while avoiding their chief limitations. Research on automated negotiation agents has yielded interactive systems that implement the processes that underlie successful negotiations [14]. For example, agents are able to form accurate models of their opponent's goals, discover opportunities for tradeoffs and propose efficient solutions [15]. A recent focus has been to create agents that can negotiate with human users [16-18] and even model human-like psychological processes. For example, algorithms can simulate or even exploit common negotiation biases such as the fixed-pie bias [19] – a tendency to assume your opponent holds the same preferences as you – or the anchoring bias [20] – a tendency to be influenced by your opponent's opening offer [e.g., see 21]. In this paper, we argue human-like agents realize the strength of the previous two approaches while overcoming their limitations. Like the dyadic approach, they support the study of interactive processes while also providing strong experimental control. Like the scripted approach, they allow for strong experimental control over the factors that might shape outcomes but also enable a level of dynamism that allows the agent to adapt its responses to the user's actions without succumbing to human bias.

In this paper, we use intelligent agents to explore how negotiation inputs – "gender triggers" [22] and gender-differences in negotiation goals [8] – impact the negotiation behavior of CS students. Gender triggers are descriptions about a potential job that appear to trigger gender-divergent behaviors. These include terms that activate gender stereotypes [10] or suggest hostility towards women in the workplace (such as stories or reviews suggesting gender discrimination is rampant).

We bring two novel contributions to this problem. To our knowledge, this is the first work examining gender-bias in negotiation with agent-based technology. It is perhaps the first work to examine salary negotiation processes in the context of undergraduate CS majors (prior literature has focused almost entirely on business students). Our study adds to the nascent literature on using virtual agents as psychological tools [23, 24] but also highlights advances needed in the field of autonomous agents to fully realize the potential of such methods. The next section describes the idea of gender triggers and how they can be manipulated. We then describe the agent-based technology used in this study, before presenting the experiment and discussion.

2 Gender Triggers and Goals

Research on gender in negotiation often accepts the premise that women exhibit poorer salary negotiation skills, but some research attempts to qualify this tendency. Rather than claiming that women are less skilled in general, this work highlights specific situations where negotiation behavior diverges. Bowles and colleagues introduced the term "gender triggers" for the situational factors that elicit gender-related differences in how people negotiate [22]. The most commonly studied trigger is stereotype threat [25]. This is the phenomena that women's performance suffers when a job is described as requiring stereotypically male traits (e.g., requires "strong bias for achievement and confidence when facing risks") but the gender gap can be reversed when the job is described as requiring stereotypically female traits [e.g., requires "good listening skills and good intuition in understanding others, 26].

Figure 1 illustrates a common way to manipulate stereotype threat, which we adopt in the present study. Following Shantz [25] and Bem's Sex Role Inventory [27], we manipulate stereotype threat by adding language to the description of the job participants are negotiating about. Jobs included a "Required Characteristics" section containing five statements crafted to be either *female-threatening* or *female-supportive.* Female-threatening job descriptions included statements that candidates should "be the technical authority in your team", "drive your team toward perfection" and show "confidence working on large applications" whereas female-supportive jobs stated that candidates should "be an asset to your team", "encourage your team to achieve excellence" and show "readiness to collaborate on large applications."

Most research on stereotype threat has utilized business students, where ideal candidates are described as confident, aggressive and self-interested, but stereotypes may play out differently in CS, where teamwork is often emphasized. Thus, we also explore a second potential gender trigger: gender-based hostility in the workplace [28]. Prospective applications often discover a company's culture through reviews at job recruiting sites such as Glassdoor. To manipulate the perceived hostility or openness of a company, we used actual quotes from Glassdoor about real technology companies' work environments. For example, female-threatening jobs included reviews stating "sexism is prevalent", "women have a tough time getting ahead." For female-supportive jobs, we modified these lines to reverse the impressions: "sexism is not an issue", and "women have opportunities to get ahead."

Although gender triggers comprise one input to a negotiation, salary disparity could simply reflect different goals of female versus male negotiators. Some research argues that women's skills are judged unfairly by being compared with metrics developed by largely-male negotiation researchers. Instead, women may simply be trying to achieve other sources of value besides salary. For example, as a group, women tend to be more risk-averse [8] and make greater use of vacation time [29]. Unfortunately, most laboratory studies do not provide an opportunity to examine negotiators' goals. A typical study might use a standard job negotiation exercise, such as *New Recruit* [30]. In this exercise, each side is provided a number of issues (salary, vacation, etc.) and a fixed payoff table on the points they can earn based on what they negotiate for each issue.

To examine gender differences in negotiator goals, we adapt the standard *New Recruit* case to include compensation elements that men and women have been shown to value differently in prior studies. Specifically, stock options and bonuses are seen as riskier issues to which women would presumably assign less value whereas vacation days should be more valued by women, at least according to this prior research. Rather than assigning fixed value to these issues, however, we allow participants to assign their own priorities to these issues to assess if these hypothesized differences are truly present in CS students.

3 Salary Negotiation Exercise

Negotiation Structure: The prospective employee (the participant) and a hiring manger (the agent) negotiate over salary, bonus, stock options and vacation. Each issue has 10 discrete levels (see Figure 2). Players bid by specifying a level for each issue. There are 10^4 possible negotiated agreements.

Following *New Recruit*, participants are provided an explicit payoff function that defines the value of deals as a linear utility function based on the weight they assign to each issue in isolation: their outcome, $u(\omega)$, depends on the level they obtain in issue i, l_i, multiplied by the weight associated with that issue:

$$u(\omega) = \sum_{i=1}^{n} w_i \cdot l_i \quad (1)$$

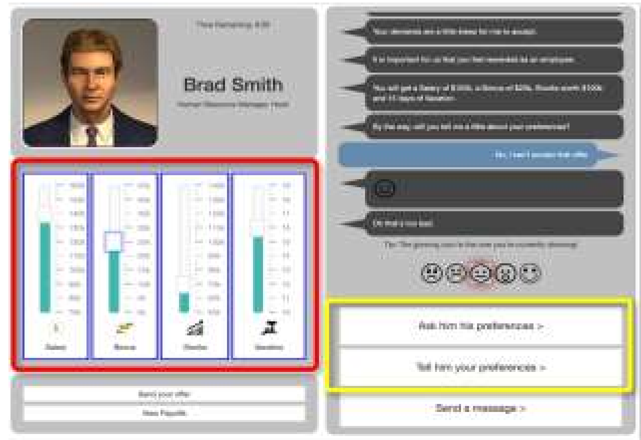

Figure 2: The negotiation interface used in this study. People can exchange offers (highlighted in red), or exchange information about preferences over issues (highlighted in yellow).

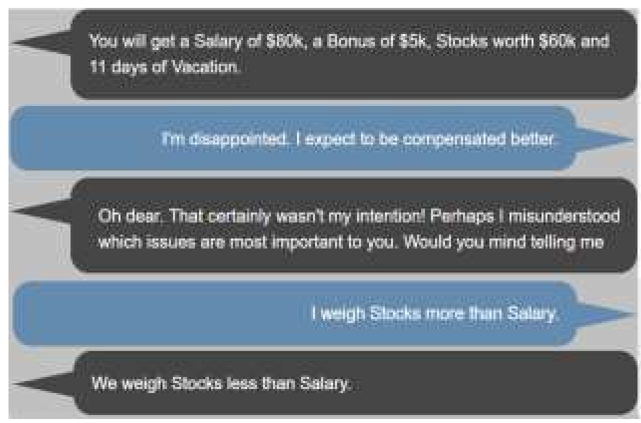

Figure 3: An example dialog involving information exchange.

Different negotiations arise depending on the pattern of these weights. If each party weights issues the same, the negotiation is zero-sum (oft called "fixed-pie"). When opponents hold different weights, there are opportunities to create "win-win" deals by trading-off across issues (e.g. accepting less salary for more vacation). These opportunities are referred to as *integrative potential.* As these weights are typically private, discovering the opponent's preferences is necessary to realize this integrative potential. Negotiations can also be impacted by the general shape of the utility function (e.g., if it is linear), a point we return to in the discussion.

Unfortunately, human negotiators often fail to realize integrative potential. Many simply accept the first offer. When they do negotiate, many fail to realize their counterpart holds different preferences (i.e. they have a "fixed-pie bias" [19], meaning they assume the negotiation is zero-sum). They also fail to exchange information or miss what information has been shared [12]. Even when they understand their opponent's perspective they may fail to exploit this understanding. For example, they may fail to appreciate the importance of trading concessions to realize win-win solutions.

IAGO Negotiation Agent: To simulate how a recruiter might approach this task, we adopt the publicly-available IAGO negotiation framework that serves as the platform for the annual human-agent negotiation competitions at the IJCAI conference [18, see Figure 2]. IAGO supports online negotiation with human participants. Participants exchange offers through sliders. Using menus, participants can assert they prefer stocks over salary, or ask the hiring manager about their relative preferences. Participants can express an emotional attitude towards the negotiation through emojis, or send canned statements such as "I expect to be compensated better" or "We should consider each other's interests."

For this experiment, we adapt the default Pinocchio agent provided with the IAGO framework [see 31 for details]. The model attempts to model how people commonly negotiate, including two tendencies that make it hard for negotiators find win-win-solutions. First, the agent incorporates a "fixed-pie" bias (it assumes a zero-sum orientation to the negotiation). Specifically, it tries to make offers on the Pareto frontier by estimating the human participant's preferences from their preference assertions (e.g., see Figure 3), but this estimate is biased using a fixed-pie prior (i.e., absent information to the contrary, the agent assumes the participant ranks issues the same way as the agent). Second, the agent is reluctant to exchange preference information. The agent responds truthfully to the participant's preference requests but does not actively volunteer this information unless the participant shares first, in which case, it reciprocates each preference revealed. Together, these properties mean that students will fail to find a win-win solution unless they explore the recruiters underlying preferences.

We altered IAGO's text messages to better match the language used in actual salary negotiations. We made one minor adaptation to offer behavior to better align the agent with human performance. Pinocchio aims to be fair by default, always making offers that match the estimated Nash Bargaining Solution [32]. We adjusted this offer behavior to incorporate the aforementioned anchoring bias into the system. As in typical salary negotiations, the agent makes the first offer and tries to induce an anchoring bias in the participant by making a strong initial offer (1,1,1,1). But the agent was also made susceptible to anchoring. Whereas Pinocchio always attempts to offer 50% of the pie, the revised agent initially tries to claim 60% of the pie but this shifts to as low as 40% depending on the strength of the opponent's initial offer.

Job Negotiation Task

Participants were told to imagine they were offered a job and must negotiate their final package. They were provided a cover letter and background materials based on actual job descriptions (see Figure 1). They received background research summarizing the typical salary range for this type of job (and gender triggers that vary with experimental condition). Participants had the opportunity to negotiate over salary, bonus, stocks and vacation. Background materials specified the expected range for this type of job (including historical data stating that salaries range from $70-$120K).

Each participant was asked to express their own priorities over the issues in the negotiation. They assigned weights to issues by distributing 11 discrete points across the four issues (with the understanding that the value of their package is determined by Equation 1). For example, a participant might assign nine points to salary and two points to vacation if they strongly preferred salary over stocks and bonuses. The agent always held the same preferences (as would be expected with any given company). The agent assigned 5 to salary, 3 to vacation, 2 to bonuses and 1 to stock. Participants had ten minutes to complete the negotiation. If they failed to reach agreement, each side received six points.

4 Experiment

We performed an IRB-approved experiment to assess the potential of negotiation agents to give insight into the behavior of CS undergraduates in salary negotiations. Specifically, we examined (1) student's willingness and skill at negotiating (2) if gender triggers explained differences in negotiation processes and outcomes, and (3) if these differences could be explained by gender-specific preferences (e.g., are women risk averse).

Participants: A panel of 440 U.S. undergraduate CS students (308 male) were recruited for this experiment through Qualtrics' panel services. Two declined to report their gender and were excluded from analyses. The strong male-skew (70%) matches the gender imbalance in CS [33]. They were paid for their participation and, further, performance was incentivized with entries into five $25 USD lotteries, one entry for each point earned in the negotiation (based on their stated preferences and Equation 1).

Design, Procedure and Measures: We employed a 2 (Gender: male vs. female) × 2 (Threat: female-threatening vs. female-supportive) × 2 (Manipulation type: Required characteristic vs. Glassdoor comments) between-subjects design. After consenting and completing demographic questions, they were then told to imagine that they were offered a job as a software engineer at a fictitious tech company. Everyone read the same job description but additional language was added to manipulate threat in a female-supportive or female-threatening manner. As described above, threat was manipulated either by a section describing "Required Characteristics" or by a section that included Glassdoor Reviews.

To measure the effectiveness of the threat manipulation, participants next completed a *manipulation check* by rating the masculinity of the job using 3 items ("...reflects masculine characteristics," "men would have an easier time...," "there would be challenges...as a woman") on a scale from 1 (strongly disagree) to 7 (strongly agree). To rule out whether the threat manipulation impacted the perceived job quality, they also rated the *status* of the job using 3 items ("...is a high-status position," "...impressive on a resume," "...people...consider this an attractive opportunity") on a scale from 1 (strongly disagree) to 7 (strongly agree).

Before negotiating, participants reported the minimum salary that they would be willing to accept (*minimally acceptable salary)*. To measure *what issues participants valued*, they were asked to prioritize the relative importance of salary, bonus, stock options, and vacation time by assigning points to each issue (out of eleven total).

Participants then engaged in the negotiation with the virtual agent. The interface allowed the user to: accept or reject the agent's offer(s), make offers themselves, as well as send pre-programed responses and emojis (happy, sad, angry, surprised, neutral). Counts of each of these responses were tracked and served as dependent variables. We focus on: *accepting the agents offer (vs. making offer(s) themselves)* and the *number of times each emoji was used*

In addition to the pattern of offers, the final outcome for participants and the agent, as well as joint points, were derived using Equation 1. Two measures were derived from joint points. *Maximum possible integrative potential* was calculated by finding the maximum joint points of any deal that could be reached between a given participant and the agent (using the participant's stated points); *Realized integrative potential* was measured by the joint points obtained compared to the maximum joint points possible.

Finally, individual differences in motivation were also measured. Specifically, participants rated themselves on 4 items for each type of motivation using a scale from 1 (not at all) to 7 (to a great extent). For *achievement motivation*, they rated to what extent they would 1) get their work tasks done, and 2) get them done well, 3) get a lot of work accomplished, and 4) finish their work in this job. For *communal motivation*, they rated to what extent they would 1) get along with coworkers, 2) be a team player, 3) be liked by coworkers, and 4) build relationships in this job.

5 Results

Manipulation Check: The manipulation was successful; participants in the female-threatening condition reported that the position was significantly more masculine (M = 4.89, SE = 0.12) than those in the female-supportive condition (M = 4.21, SE = 0.11; F(1, 430) = 18.66, p < .001). The manipulation type also affected the perception of masculinity such that participants in the Glassdoor comments condition reported that the position was significantly more masculine (M = 4.72, SE = 0.11) than those in the Required Characteristics condition (M = 4.38, SE = 0.11; F(1, 430) = 4.59, p = .03. Moreover, there was a significant interaction such that these effects were driven by the female-threatening Glassdoor comments condition (M = 5.27, SE = 0.16; F(1, 430) = 6.70, p = .01), which was perceived the most masculine, followed by female-threatening Characteristics condition (M = 4.52, SE = 0.17); ratings in the female-supportive conditions were lower and comparable to each other (M = 4.17, SE = 0.16 vs M = 4.24, SE = 0.15, respectively). Although it was only a trend (F(1, 430) = 2.60, p = .11), the effect of the manipulation was slightly stronger for female participants (M = 4.05 vs. M = 4.99) than male participants (M = 4.37 vs. M = 4.79). All other effects failed to approach significance (Fs < 1.3, ps > .26).

The gender-trigger manipulation did not impact the perceived desirability of the job. The job was uniformly seen as having high status (M = 5.90, SD = 1.00, on scale from 1 to 7). All effects for perceived status failed to reach significance (Fs < 2.9, ps > .09).

In summary, we successfully manipulated threat but, for CS students, this was most effectively achieved through the Glassdoor reviews rather than Required Characteristics. As intended, this manipulation did not make jobs seem more or less desirable. As we did see a main-effect of threat regardless of how it was conveyed, the remainder of the paper will focus on the impact of gender and threat (ignoring if threat was conveyed through Glassdoor comments or Required Characteristics).

Minimally Acceptable Salary: Participants' minimum acceptable salary fell near the bottom of the stated salary range for this job (which was $70K). There was a borderline significant interaction (F(1, 414) = 3.81, p = .052) showing that women's bottom line was influenced by threat. As can be seen in Figure 4, while men did not significantly differ across threat conditions (t(418) = -0.88, p = .38), women marginally raised their bottom line when the job was described in a female-supportive way (t(418) = 1.87, p = .06). All other effects failed to reach significance (Fs < 1.6, ps > .21).

In summary, consistent with prior research on business students, women were particularly impacted by gender triggers. When jobs advertisements contained female-threatening language, women lowered their salary minimums compared to when jobs were described in a female-supportive way, whereas men were unaffected.

Willingness to Negotiate: If women have lower aspirations, they may feel less motivation to negotiate and simply accept the first offer. To examine this, we created a dichotomous variable distinguishing participants that accepted the agent's first offer versus those who made at least one offer to the agent. Overall, 43% did not negotiate. However, willingness to negotiate was unaffected by gender or threat (χ^2s < 0.58, ps > .44); e.g., women were not less likely to counter the first offer (54%) compared to men (56%).

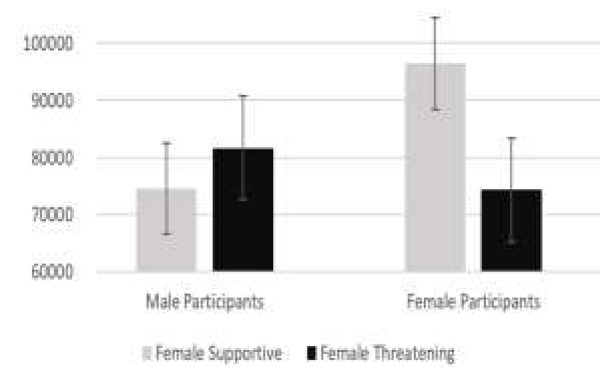

Figure 4: Minimally acceptable salary (USD) as reported before the negotiation began.

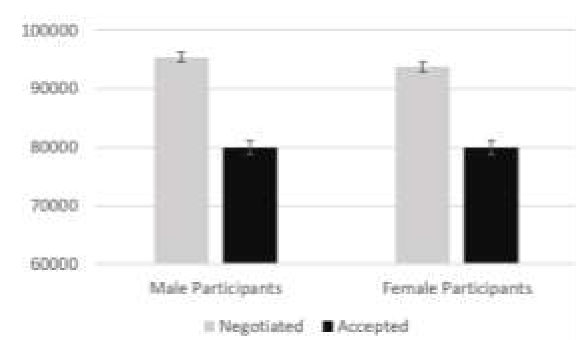

Figure 5: Final negotiated salary (USD) for those willing to negotiate versus those simply accepting the first offer.

Motivations did influence willingness to negotiate. Negotiators with higher achievement motivation were significantly more likely to make at least one offer (vs. none at all; B = .23, Wald(1) = 5.10, p = .02), though communal motivations had no effect (B = .02, Wald(1) = 0.05, p = .82).

In summary, there was no evidence that gender triggers lead women to be more reluctant to negotiate their salary.

Final Negotiated Salary: To examine the outcome of the negotiation, we excluded 13 participants that failed to reach agreement in the negotiation (neither men nor women were more likely to fail to reach an agreement). We find that final outcome was not significantly influenced by gender or threat (Fs < 1.01, ps > .31).

Final salary was strongly impacted by willingness to negotiate: participants who made at least one offer did significantly better in terms of overall points (M = 25.13, SE = 0.79) than those who just accepted the agent's first offer (M = 11.29, SE = 0.25; t(419) = -14.76, p < .001) and final salary (see Figure 5): participants who made at least one offer received a significantly higher salary (M = $92,946, SE = 1106.44) than those who just accepted the agent's first offer (M = $80,000, SE = 0.00; t(419) = -10.11, p < .001).

Note, regardless of willingness to negotiate, the vast majority (85.5%) achieved a salary that exceeded their minimally-acceptable salary. There was no significant difference, based on gender or threat, in the proportion of participants that settled for a salary below their stated bottom line.

In summary, in contrast to prior work [e.g., 2], there were no gender differences in in the salary obtained, nor an impact of gender triggers. Rather, salary was determined by participants willingness to counter the hiring manager's initial offer (which was also unaffected by gender or threat). Those that did negotiate obtained $13,000 USD more per year (a difference roughly comparable with what has been previously reported).

Integrative Potential: Although we did not see differences in outcome by gender or gender triggers, it is possible that differences could emerge when we control for the goals people bring to the negotiation. If men and women systematically weight issues differently, they might face quite different integrative potential.

Participants expressed a wide range of priorities and this led to a wide distribution of possible joint outcomes, and priorities differed by gender. Women valued stock options significantly less (M = 1.70, SE = 0.12) than men (M = 1.99, SE = 0.08; F(1, 430) = 4.55, p = .03). This is consistent with the existing literature which suggests that women are less tolerant of risk. All other effects failed to reach significance (Fs < 0.8, ps > .13).

Gender-differences in preferences did not, however, alter the distribution of points that participants had the potential to receive. To examine this, we again excluded the participants that failed to reach agreement in the negotiation (n = 13), and then examined the effect of condition on maximum integrative potential. Men and women showed no significant differences. All other effects failed to reach significance (Fs < 1.8, ps > .18).

In summary, CS undergraduates brought varying preferences to the negotiation. Women valued stocks less than men, but this gender difference did not substantially shift the space of possible deals that could be reached given the agent's preferences.

Realizing Integrative Potential: Participants often failed to realize the integrative potential available to them. Overall, participants realized 80% of the possible joint value (they left 20% of the value sitting on the table) as measured by the joint points obtained compared to the maximum joint points possible. Participants did worse on tasks with greater integrative potential: maximum integrative potential was negatively correlated with percent of this obtained, r(N=409) = -.90, p < .001.

There were no gender differences in the ability to realize integrative potential. The difference between women and men in percent of integrative potential obtained failed to reach significance (1, 417) = 1.76, p = .19). All other effects failed to reach significance for realizing integrative potential (Fs < 2.0, ps > .16).

We also tested whether individual differences in motivation predicted realizing integrative potential. There was no significant relationship between communal motivations and ability to realize integrative potential (r(N=408) = -.01, p = .81). Nor was there was a significant relationship between achievement motivations and realizing integrative potential (r(N=409) = .08, p = .11).

In summary, CS undergraduates failed to realize the integrative potential in their negotiations. However, this effect was not influenced by gender, gender triggers or motivation.

Emoji Use: Emotional expressions are sometimes an indicator of people's attitudes towards a negotiation or their partner. Overall, happy emojis were most commonly used (M = 1.45, SE = 0.20) and negative expressions – angry (M = 0.11, SE =0.03) and sad (M = 0.17, SE = 0.03) – were rare. Neutral emoji were the second most common emotional signal (M = 0.70, SE = 0.07). Prior work suggests neutral expression are interpreted as negative signals [34].

There were gender differences in what emojis participants used. Women used sad emojis significantly less (M = 0.06, SE = 0.05) than men (M = 0.21, SE = 0.03; F(1, 430) = 5.34, p = .02). For use of neutral emojis, there was a significant interaction (F(1, 430) = 4.91, p = .03) showing that each gender used more neutral emojis when their specific gender was threatened. Men tended to use more neutral emojis in the female-supportive condition than in the female-threatening condition (t(434) = -1.54, p = .13). For women the pattern was reversed. Women tended to use more neutral emojis in the female-threatening condition than in the female-supportive condition (t(434) = 1.83, p = .07).

In summary, CS undergraduates generally signaled positive emotions during the negotiation. However, their use of neutral emotions suggests less satisfaction when the job description was less supportive of their gender.

6 Discussion and Next Steps

This experiment explored the potential for agent-based technology to give insights into issues impacting the underrepresentation of women in STEM. We used a gender-blind automated negotiation agent to examine the salary negotiation behavior CS undergraduates, particularly with regard to how gender bias might shape outcomes. Prior findings suggest that women will fare poorly when the job description contains gender triggers (such as the work environment is hostile to women). Our results partially support these findings but failed to find evidence that women's negotiation behavior, in and of itself, leads to lower salaries.

As expected from prior work, gender triggers did impact women's salary *expectations*. Women set lower minimum salary expectations when the job description suggested a more threatening environment, whereas men were unaffected. Yet these initial expectations failed to impact the *process* by which they negotiated or *final outcomes*. Women and men negotiated equivalent salaries (if they negotiated at all), and women and men showed equal proclivity for negotiating, regardless of threat.

Some research has suggested that disparities in women's outcomes may reflect different goals and we did find support for this. Women were more risk-averse – they assigned less value to stock options – but this did not systematically shape outcomes in our study. It should be noted that the potential for individual and joint games will depend on the preferences each party holds. In our experiment, the hiring manager's preferences were fixed but participants' preferences varied. In that women's preferences differed from men's, it should be possible to devise a set of preferences for the hiring manager that would systematically alter the space of possible deals in ways that might benefit either men or women. Future work should explore this possibility.

Overall, CS students were poor negotiators. Almost half simply accepted hiring manager's first offer, replicating recent survey findings [1, 2]. When they negotiated, students were poor at discovering win-win solutions. This reinforces recommendations by the National Academy that STEM fields are lacking in important interpersonal skills such as negotiation.

We have argued that negotiation agents eliminate some important weakness in the scripted agents used in negotiation research, but some limitations qualify our results and readers should take caution in extrapolating these findings to real-world salary negotiations. Characteristics of the IAGO agent might serve to minimize gender differences. Assuming a linear utility function (Eq. 1) may mismatch participants true preferences, though this assumption shared by most prior psychological studies on gender and negotiation. Participants may have felt less threat because they were negotiating with a computer program. For example, prior agent research suggests that threat perceptions are reduced when people interact with computer agents, compared with human-human interaction [35]. Prior research also suggests that gender differences are smaller when the structure of the negotiation task is more explicit [22]. By explicitly listing the set of issues under negotiation and their allowable range, IAGO reduces the structural ambiguity, perhaps minimizing gender differences. Our use of an explicit manipulation check before the negotiation, though it helped ensure our manipulation was effective, could have enhanced the salience of gender effects.

Of course, outcomes were also influenced by the agent's negotiation behavior or even the mere fact that it was an agent [36]. Although the agent incorporates human-like elements, gender differences could emerge if the agent used different negotiation tactics or adopted more realistic assumptions about participants' utility [37, 38]. Further studies must also verify the robustness of these findings with respect to elements of agent design, and if these findings translate to interactions with human negotiators.

More broadly, the present study follows the common practice in the negotiation literature of focusing on economic outcomes and stylized economic games that strip out face-to-face communication in service of experimental control. However, this approach can overlook important strengths that underrepresented groups bring to the negotiation process. For instance, research suggests that women place greater emphasis than men on creating relational capital and that this can result in economic benefits that are not captured by traditional measures of negotiation success [39]. To enhance the methodological value of agent-based approaches, more research is needed into natural language interfaces [16, 40], rapport [41], and behavioral (rather than classical) economic models of decision-making that incorporate the role of emotions, fairness and non-monetary sources of value.

The current study examined differences in women's negotiation behavior, but negotiation research highlights that salary disparities reflect bias in hiring managers. For example, female negotiators face systemic bias: they are perceived as less competent then men, regardless of skill [42], and this results in their receiving more aggressive opening offers and a greater use of deception by their counterparts, especially when they adopt an assertive negotiation style [9]. This research suggests that even if women adopt the same negotiation tactics as men, they will be perceived very differently and have different consequences for outcomes. An obvious next study would be to reverse the situation, have participants play the role of hiring managers, and examine the impact of employee gender on their behavior.

In conclusion, this study highlights the potential for agent-based technology to yield insights into and important societal problem, namely gender-bias in salary negotiations. Our approach provides little support for the idea that women's negotiation ability is undermined by gender threat (claims that have been made and have been widely-accepted in the gender and negotiation literature). This lack of support, furthermore, can be observed despite our ability to manipulate both gender and threat quite effectively. And though one might choose to dwell on how this new methodology might be tweaked to increase the chances of finding gender effects, the results also provide the basis for a more troubling question about the gender-related findings in the negotiation literature. If it is so difficult to uncover meaningful gender differences in a study that maintains experimental control while also increasing realism, and provides so many opportunities for these kinds of gender differences to emerge (e.g., in response to triggers, based on differences in preferences, and differences in motivations), then are the experimental findings from prior studies the result of such experiments being designed to uncover gender differences researchers were hoping to find? Are such gender differences simply less robust in settings beyond those that are specifically tailored to make a case for such effects? And if robust gender differences in negotiation do exist in more naturalistic settings, then perhaps these researchers need to look beyond individual dispositions and skills to the role of social structures and institutions (cf. [43]).

7 ACKNOWLEDGMENTS

This work was supported the National Science Foundation under grant 1822876 and the Army Research Office, accomplished under Cooperative Agreement Number W911NF-20-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. We appreciate the feedback of the anonymous reviewers.

8 REFERENCES

[1] Glassdoor Team, "3 in 5 employees did not negotiate salary [Blog Post]," in *In the News*, ed: Glassdoor, 2016.

[2] J. Säve-Söderbergh, "Gender gaps in salary negotiations: Salary requests and starting salaries in the field," *Journal of Economic Behavior & Organization,* vol. 161, pp. 35-51, 2019/05/01/ 2019.

[3] M. Marks and C. Harold, "Who asks and who receives in salary negotiation," *Journal of Organizational Behavior,* vol. 32, pp. 371-394, 2011.

[4] National Academies of Sciences, Engineering and Medicine,, *Promising Practices for Strengthening the Regional STEM Workforce Development Ecosystem*: National Academies Press, 2016.

[5] K. Gray, A. Neville, A. H. Kaji, M. Wolfe, K. Calhoun, F. Amersi, *et al.*, "Career Goals, Salary Expectations, and Salary Negotiation Among Male and Female General Surgery Residents," *JAMA Surgery,* 2019.

[6] A. Leibbrandt and J. A. List, "Do Women Avoid Salary Negotiations? Evidence from a Large-Scale Natural Field Experiment," *Management Science,* vol. 61, pp. 2016-2024, 2015/09/01 2014.

[7] S. Tang, X. Zhang, J. Cryan, M. J. Metzger, H. Zheng, and B. Y. Zhao, "Gender Bias in the Job Market: A Longitudinal Analysis," *Proc. ACM Hum.-Comput. Interact.,* vol. 1, p. Article 99, 2017.

[8] R. Croson and U. Gneezy, "Gender Differences in Preferences," *Journal of Economic Literature,* vol. 47, pp. 448-74, 2009.

[9] E. T. Amanatullah and C. H. Tinsley, "Punishing female negotiators for asserting too much…or not enough: Exploring why advocacy moderates backlash against assertive female negotiators," *Organizational Behavior and Human Decision Processes,* vol. 120, pp. 110-122, 2013/01/01/ 2013.

[10] L. J. Kray, L. Thompson, and A. Galinsky, "Battle of the sexes: Gender stereotype confirmation and reactance in negotiations," *Journal of Personality and Social Psychology,* vol. 80, pp. 942-958, 2001.

[11] U. Tellhed and F. Björklund, "Stereotype threat in salary negotiations is mediated by reservation salary," *Scandinavian Journal of Psychology,* vol. 52, pp. 185-195, 2011.

[12] L. L. Thompson, "Information exchange in negotiation," *Journal of Experimental Social Psychology,* vol. 27, pp. 161-179, 3// 1991.

[13] E. T. Amanatullah and M. W. Morris, "Negotiating gender roles: Gender differences in assertive negotiating are mediated by women's fear of backlash and attenuated when negotiating on behalf of others," *Journal of personality and social psychology,* vol. 98, p. 256, 2010.

[14] N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, C. Sierra, and M. Wooldridge, "Automated Negotiation: Prospects, Methods and Challenges," *International Journal of Group Decision and Negotiation,* vol. 10, 2001.

[15] T. Baarslag, M. Hendrikx, K. Hindriks, and C. Jonker, "Measuring the performance of online opponent models in automated bilateral negotiation," in *AI 2012: Advances in Artificial Intelligence*, ed: Springer, 2012, pp. 1-14.

[16] A. Rosenfeld, I. Zuckerman, E. Segal-Halevi, O. Drein, and S. Kraus, "NegoChat: a chat-based negotiation agent," presented at the 13th International Conference on Autonomous Agents and Multi-agent Systems, Paris, France, 2014.

[17] R. Lin, Y. Oshrat, and S. Kraus, "Investigating the benefits of automated negotiations in enhancing people's negotiation skills," in *8th International Conference on Autonomous Agents and Multiagent Systems*, 2009, pp. 345-352.

[18] J. Mell, J. Gratch, T. Baarslag, R. Aydrogan, and C. Jonker, "Results of the First Annual Human-Agent League of the Automated Negotiating Agents Competition," presented at the 18th International Conference on Intelligent Virtual Agents, Sydney, Australia, 2018.

[19] F. Harinck, C. de Dreu, and A. E. M. van Vianen, "The Impact of Conflict Issues on Fixed-Pie Perceptions, Problem Solving, and Integrative Outcomes in Negotiation," *Organizational Behavior and Human Decision Processes,* vol. 81, pp. 329-358, 3// 2000.

[20] A. D. Galinsky and T. Mussweiler, "First offers as anchors: The role of perspective-taking and negotiator focus," *Journal of Personality and Social Psychology,* vol. 81, pp. 657-669, 2001.

[21] Z. Nazari, G. Lucas, and J. Gratch, "Fixed-pie Lie in Action," presented at the 17th International Conference on Intelligent Virtual Agents, Stockholm, Sweden, 2017.

[22] H. R. Bowles, L. Babcock, and K. L. McGinn, "Constraints and triggers: Situational mechanics of gender in negotiation," *Journal of personality and social psychology,* vol. 89, p. 951, 2005.

[23] J. Blascovich, J. Loomis, A. Beall, K. Swinth, C. Hoyt, and J. N. Bailenson, "Immersive virtual environment technology as a methodological tool for social psychology," *Psychological Inquiry,* vol. 13, pp. 103-124, 2002.

[24] C. de Melo, P. J. Carnevale, and J. Gratch, "Using virtual condederates to research intergroup bias and conflict," presented at the 74th Annual Meeting of the Academy of Management, Philadelphia, PA, 2014.

[25] A. Shantz, "The effect of stereotype threat on the interview performance of women," *Advancing Women in Leadership,* vol. 32, pp. 92-106, 2012.

[26] L. J. Kray, A. D. Galinsky, and L. Thompson, "Reversing the Gender Gap in Negotiations: An Exploration of Stereotype Regeneration," *Organizational Behavior and Human Decision Processes,* vol. 87, pp. 386-409, 2002/03/01 2002.

[27] S. L. Bem, "Bem sex role inventory," *Journal of personality and social psychology,* 1981.

[28] H. R. Bowles, L. Babcock, and L. Lai, "Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask," *Organizational Behavior and Human Decision Processes,* vol. 103, pp. 84-103, 5// 2007.

[29] D. J. Maume, "Gender Differences in Taking Vacation Time," *Work and Occupations,* vol. 33, pp. 161-190, 2006.

[30] M. A. Neale, "New recruit," *Teaching materials for negotiations and decision making. Evanston, IL: Northwestern University, Dispute Resolution Research Center,* 1997.

[31] J. Mell and J. Gratch, "Grumpy and Pinocchio: The effect of language and strategy in human-agent negotiation," presented at the 16th International Conference on Autonomous Agents and Multiagent Systems, Sao Paulo, Brazil, 2017.

[32] K. Fujita, T. Ito, and M. Klein, "Approximately fair and secure protocols for multiple interdependent issues negotiation," in *AAMAS (2)*, 2009, pp. 1287-1288.

[33] T. Camp, "Women in computer sciences: Reversing the trend," *Syllabus Magazine,* pp. 24-26, 2001.

[34] M. Weiß, J. Gutzeit, J. Rodrigues, P. Mussel, and J. Hewig, "Do emojis influence social interactions? Neural and behavioral responses to affective emojis in bargaining situations," *Psychophysiology,* vol. 56, p. e13321, 2019.

[35] G. Lucas, J. Gratch, A. King, and L.-P. Morency, "It's only a computer: Virtual humans increase willingness to disclose," *Computers in Human Behavior,* vol. 37, pp. 94-100, 2014.

[36] C. de Melo, J. Gratch, and P. J. Carnevale, "The Effect of Agency on the Impact of Emotion Expressions on People's Decision Making," presented at the International Conference on Affective Computing and Intelligent Interaction, Geneva Switzerland, 2013.

[37] T. Ito, M. Klein, and H. Hattori, "A multi-issue negotiation protocol among agents with nonlinear utility functions," *Multiagent and Grid Systems,* vol. 4, pp. 67-83, 2008.

[38] G. B. Northcraft, J. N. Preston, M. A. Neale, P. H. Kim, and M. C. Thomas-Hunt, "Non-linear Preference Functions and Negotiated Outcomes," *Organizational Behavior and Human Decision Processes,* vol. 73, pp. 54-75, 1998/01/01/ 1998.

[39] M. J. Gelfand, V. S. Major, J. L. Raver, L. H. Nishii, and K. O'Brien, "Negotiating relationally: The dynamics of the relational self in negotiations," *Academy of Management Review,* vol. 31, pp. 427-451, 2006.

[40] D. Traum, J. Gratch, S. Marsella, J. Lee, and A. Hartholt, "Multi-party, Multi-issue, Multi-strategy Negotiation for Multi-modal Virtual Agents," presented at the 8th International Conference on Intelligent Virtual Agents, Tokyo, Japan, 2008.

[41] R. Zhao, A. Papangelis, and J. Cassell, "Towards a dyadic computational model of rapport management for human-virtual agent interaction," in *Intelligent Virtual Agents*, 2014, pp. 514-527.

[42] I. Ayres and P. Siegelman, "Race and gender discrimination in bargaining for a new car," *The American Economic Review,* pp. 304-321, 1995.

[43] S. Haslanger, "What is a (social) structural explanation?," *Philosophical Studies,* vol. 173, pp. 113-130, 2016.