Virtual Backlash: Nonverbal expression of dominance leads to less liking of dominant female versus male agents

Janet Wessler, Tanja Schneeberger, Leon Christidis, Patrick Gebhard
firstname.lastname@dfki.de
German Research Institute for Artificial Intelligence
Saarland Informatics Campus, Saarbrücken, Germany

Figure 1: Four different agent conditions (from left to right): male neutral, female neutral, male dominant, female dominant.

ABSTRACT

Backlash is a form of social penalty occurring when people behave counter-stereotypically. When promoting themselves, dominant females compared to males are typically liked less and paid worse, because dominance is associated with males, and proscribed for females. Such backlash effects have been shown in human-human interactions, but attempts to replicate them in human-agent interactions have not been successful so far [40]. Here, the goal was to show backlash effects for virtual agents with a nonverbal manipulation of dominance. In an online experiment, N = 223 participants watched the video of a female or male virtual agent presenting themselves as a career coach while using either large or small gestures. They rated the agent's dominance, liking, competence, and made a monetary offer of how much they would pay for the coaching. Agents using large gestures were perceived as more dominant than those using small gestures. Moreover, a backlash effect emerged: Dominant female compared to male agents were liked less. Participants were not penalizing the female dominant agent in monetary offers. Overall, participants rated the female agents as less competent than male ones. The results underline the importance of considering effects of the agent's gender in research on human-agent interaction.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

virtual agents, gender, backlash, dominance, negotiation

1 INTRODUCTION

It is a Catch-22: When women promote themselves, they achieve better business performance, but at the cost of being liked less and perceived as "bossy" — a so-called social backlash [35, 36]. Thus, self-promotion or no self-promotion — both end in unwanted results for females. This dilemma is most prevalent in situations in which self-promotion is a necessary strategy to succeed: Leadership, negotiations, or running a business as a self-employed. Backlash effects can thus have serious consequences on career advancement and the paycheck for females [1].

In the digital age, advances in artificial intelligence research enable authentic interactions with intelligent virtual agents (IVAs). These are virtual body animations that autonomously analyze and adequately respond to social signals of human interaction partners [25, 26, 38]. Humans appear to be similarly susceptible and compliant to IVAs as to humans [30, 34, 37, 38]. Moreover, they perceive IVAs as humanlike and attribute human characteristics to them [29].

But does this perception extend so far that humans also engage in stereotyping non-human IVAs? Do humans transfer their learned and many times implicit associations to a virtual opposite?

Research on social cognition shows that humans perceive females as rather relationship-oriented, likeable and warm, whereas they perceive males as rather self-oriented, agentic and competent (although the gender difference in competence attributions has diminished in the last decades [14]). Such stereotypes can be prescriptive: People expect these characteristics from females and males. Group members who do not behave according to such expectations might face social sanctions. Expressing dominant behaviors might thus be proscribed for women [44]. Indeed, research on human–human interaction has shown that dominance can have more negative outcomes for women than for men, especially in the form of lower likeability ratings [21, 36, 44].

The current research aims at investigating whether such backlash effects also occur for IVAs. We investigated if it is possible (1) to manipulate the perception of an agent's dominance via their nonverbal gesturing behavior and (2) to then find backlash effects against female dominant agents (vs. male dominant agents) in the form of lower likeability and lower monetary outcomes in a business context.

2 BACKGROUND

Dominance generally reflects influence on communication partners [8]. Based on the assumption that people can be influenced by IVAs in a similar way as by humans [30, 34], the agent's expression of dominance should rely on similar parameters as for humans. Next to explicit manipulations of dominance (e.g., via verbal speech), past research has used implicit forms of manipulations, like body postures and gestures [44]. In humans, spreading out and occupying a larger space creates dominant impressions. In contrast, using small gestures and occupying a smaller space creates rather submissive impressions [3, 20, 39, 41]. Apart from bodily expansion [20], dominance can for example be expressed by physical closeness[20], and eye contact, particularly while speaking [13]. Similarly, for IVAs, large vs. small gestures [17], eye contact during speaking [22], and an upright vs. tilted head position [5] increase perceived dominance. Thus, both humans and IVAs are perceived as more dominant through nonverbal behaviors with a larger spatial parameter [7, 10, 17]. For example, in one study, participants were asked to design dominant and submissive agents, respectively, through nonverbal behaviors [33]. When creating dominant agents, participants used rather expansive nonverbal behaviors, while they used smaller nonverbal behaviors when creating submissive agents. Thus, gestures that occupy more space create a more dominant impression, regardless of whether the behavior is exhibited by humans or by IVAs [7, 10, 12, 17]. In the current study, dominant nonverbal behavior refers to using large gestures, while neutral nonverbal behavior refers to using small gestures. Such dominant gesturing might have more negative consequences for female humans and agents than for males.

Gendered expectations are placed on women and men based on their existing gender roles and stereotypes associated with these roles [15]. Gender stereotypes are cognitive structures of socially shared knowledge about characteristic features of women and men [4]. Gender roles include descriptive stereotypes, i.e., how men and women typically act. Such associations develop from observing females and males fulfilling different tasks in society (i.e., child rearing vs. paid labor) [15, 23]. Traits such as being emotional, likeable and caring are more typically attributed to women and often referred to as communal-warm. Traits such as being competent, assertive, and dominant are more associated with men and referred to as agentic-instrumental [31]. However, a recent meta-analysis on US polls showed that the gender difference in agency and competence ratings has diminished over time, while it still exists for communality [14].

Moreover, gender roles include prescriptive stereotypes about how men and women ought to act [15]. In general, study results on expectancy violation show that norm-violating individuals are punished with reduced likeability [32]. Hence, if individuals violate gendered expectations, they will probably experience social backlash: Assertive, self-promoting and dominant women are evaluated as less likable and less hireable, especially in self-promoting situations such as negotiations [2, 6, 35, 36]. A meta-analysis with $k = 63$ studies on the backlash effect found that dominant women were liked less than dominant males and were overall judged as less hireable [44]. Thus, overall, a low dominance expression seems to be required from women [36, 44].

This meta-analysis found no backlash effects for dominant females on competence perceptions: Averaged across the studies, dominant females were rated as competent as dominant males. Dominant targets were rated slightly more competent than non-dominant ones overall. Still, to get hired, employees have to not only be perceived as competent, but even more so as likeable and social [44]. Thus, perceived warmth in the form of high likeability is an important component for forming impressions of others.

Backlash effects can also occur in computer-mediated negotiations: When interacting via text messages only, participants rated an angry female negotiation opposite as more competitive than an angry male one, and tended to like them less [42].

To the best of our knowledge, only one study so far has investigated if assertive female IVAs face a backlash in negotiations [40]. Human participants conducted a salary negotiation in the role of a manager of either a female or male IVA as an employee. During the negotiation, the agent behaved either assertive or non-assertive. The assertive agent made higher offers and used assertive verbal content compared to the non-assertive agent. Moreover, the non-assertive agent had a higher likelihood to accept the participant's last offer. The authors found no significant difference in the final salary outcome for the virtual male and female employees. They found that in general, assertive agents were perceived as less friendly and more arrogant than non-assertive ones, but this pattern was similar for female and male agents. Thus, there was a penalty for being assertive independent of the agent's gender. That study, however, had kept the nonverbal behaviors of the agent's constant in all conditions.

We argue that the nonverbal behavior is an important factor in creating impressions of other humans and virtual agents. This is why we took a different approach: In the present study, the verbal content of the message and value of the offer were kept constant, while the nonverbal behavior differed to create different perceptions of dominance.

3 OVERVIEW AND HYPOTHESES

The present study aims at investigating whether dominant nonverbal behavior by a female virtual agent also leads to more negative reactions in human-agent interactions, thus trying to replicate findings from human-human research. In an online experiment, participants observe a video of an IVA presenting him- or herself as a coach for difficult social situations. The IVA explains that he or she is able to interact authentically with people. We manipulated the agent's gender (male vs. female) and their gesturing behavior (large/dominant vs. small/neutral). They used identical wording to explain their virtual coaching program for situations such as job interviews, and presented a price for this coaching. Participants rated how dominant, likeable and competent they perceived the agent to be. In addition, participants provided a counteroffer for the price of the coaching.

It is expected that IVAs using large gestures are perceived more dominantly than IVAs using smaller gestures. If gender stereotypes apply to IVAs, dominant behavior of a female agent should violate normative expectations, since dominance tends to be associated with masculinity. This norm violation should negatively influence likeability perceptions of the virtual agent and their monetary outcomes. Based on previous findings, though, we expect no differences for dominant female and male on competence perceptions. We formulate the following hypotheses:

H1 (manipulation check): Agents using large gestures are perceived as more dominant than agents using small gestures.

H2: Female agents with dominant behavior are rated as less likable than dominant male agents and neutral female agents (i.e, an interaction effect between agent's gender and gesture type).

H3: Agents with dominant behavior are perceived as more competent than agents with neutral behavior (i.e., main effect of gesture type).

H4: Female agents with dominant behavior receive lower monetary offers than dominant male agents and neutral female agents (i.e, an interaction effect between agent's gender and gesture type).

Furthermore, we asked two exploratory questions: (a) Does the perception of likeability mediate the effect of gender on monetary offers? (if H4 is confirmed) (b) Does the gender of the participants moderate the effects of nonverbal behavior and agent's gender on the monetary offers?

4 METHODS

Following the recommendations by Wessler and colleagues [43], all materials including the online questionnaire, data, and analysis code can be found on osf.io/mkv4f).

4.1 Participants

We aimed for a minimum of $n = 50$ participants in each condition (for recommendations, see [43]. Participants were recruited mainly via groups in social network and had to be fluent in German. Of the total of $N = 275$ participants, $n = 52$ were excluded either due to incomplete responses ($n = 34$), lack of informed consent ($n = 3$), fail to pass the attention checks ($n = 15$), or admitting to giving wrong answers on purpose ($n = 1$). The final sample of $n = 223$ consisted of $n = 102$ males and $n = 121$ females between 18 and 68 years ($M = 27.6$ years, $SD = 9.0$ years). Psychology students ($n = 46$) were rewarded with course credit for participation. All others — students of various other disciplines like economics or biology ($n = 127$) or non-students ($n = 50$) — had the option to participate in a lottery to win one out of five 10€ vouchers for an online store. Regarding political orientation, participants were somewhat normally distributed across a continuous scale from (1) *conservative* to (101) *liberal*. However, the mean ($M = 71.45$, $SD = 18.18$) and median ($Mdn = 74$) indicate that the sample's political orientation was slightly more liberal.

4.2 Design

Participants were randomly assigned to one of four conditions in a 2 (Agent's gender: male vs. female) x 2 (Gesture type: dominant vs. neutral) between-participants design. This led to the following number of participants within each condition: ($n_{m\_dom} = 59$) in male dominant, ($n_{f\_dom} = 53$) in female dominant, ($n_{m\_neut} = 57$) in male neutral, and ($n_{f\_neut} = 54$) in female neutral. Perceptions of dominance, likability, competence, and salary offers were the main dependent variables.

4.3 Procedure

This online experiment was conducted with SoSciSurvey [24] as part of a Master's thesis. First, participants read and agreed to the informed consent form. Participants were then motivated to imagine the following scenario: They are about to graduate and, while looking for a full-time position, they have received feedback from interested companies. In order to prepare for the various job interviews properly, they are looking for a coaching. They come across a video on the internet. The participant starts the video of the virtual agent. A random generator determined whether this agent is female or male and whether he/she expresses nonverbal dominance or not. The virtual agent introduces herself/himself as Luca and explains how an interaction with her/him is possible (verbal content available on osf.io/mkv4f). Luca offers a virtual coaching to prepare for difficult social situations like job interviews. The video ends with Luca stating the price (60€) for five sessions.

Participants answered three attention checks to ensure they had watched the video vigilantly. The following items asked participants regarding their perception of the agent's dominance, likability, competence (see Measures), and their willingness to work with the agent. Participants then imagined that they had had an initial coaching session with the virtual agent: They were convinced by the coaching concept, and they were aware that a coaching with a human is significantly more expensive. They indicated the maximum amount they were willing to pay for coaching, how many coaching session they want, and how much they would pay for each session (see Measures). Finally, participants reported demographic data. A debriefing followed and participants had the possibility to make comments. They could leave their email address to receive a summary of results.

4.4 Dominance Manipulation

We generated the four different versions of the IVAs with the visualization platform VuppetMaster7 [18]. To manipulate the agent's gender, a female and male character was chosen from a given selection with a corresponding female or male voice. Two videos

included a female agent, and two videos a male agent. The verbal content was equal in all four videos. In the two videos manipulating dominance, the agent showed expansive, large gestures, which were fitting to the verbal content. In the other two videos, the agent showed small gestures, which should lead to a neutral perception of the agent. Figure 1 illustrates the manipulation of gender (female vs. male) and nonverbal behavior (dominant vs. neutral) of the virtual agents.

4.5 Measures

Dominance, liking, and competence items were all rated on 7-point Likert scales from (1) *I do not agree at all* to (7) *I totally agree.*
Demographics included gender, age, income, political orientation, highest educational level achieved, field of study, semester, and which degree they aim for.
Attention checks consisted of three items with two answer options each to check if participants had watched the videos attentively (e.g., Which kind of service does the virtual agent offer? with the options "a coaching for difficult social situations like job interviews" or "a coaching for difficult social situations like a sales talk.". Participant who answered one or more items incorrectly were excluded from data analysis.
Dominance was measured with three items which were adapted and translated into German from [36]: "The virtual agent is dominant"; "The virtual agent seems to control others"; "The virtual agents seems arrogant to me", Cronbach's α = .78.
Liking was measured with five items which were adapted and translated into German from previous research on perceptions of virtual agents [19]. Items were "The virtual agent is very likeable.", "Overall, I like the virtual agent.", "The virtual agent seems very kind.", "The virtual agent seems very unfriendly." (reverse), "The virtual agent is very unpleasant." (reverse), Cronbach's α = .89.
Competence was measured with three items adapted and translated from [11], "The virtual agent seems competent to me.", "The virtual agent is intelligent.", and "The virtual agent seems knowledgeable.", Cronbach's α = .84.
*Monetary offers* were counter-offers to the virtual agent's initial offer of 60€ for 5 sessions (300€ total). They were measured with three individual items, "How much would you maximally be willing to pay for the virtual coaching overall? I am willing to maximally pay ....€ for the coaching package.", "How much would you be maximally be willing to pay for one session? I am willing to pay ...€ maximum.", "How many coaching sessions would you be willing to maximally take?" "I would be willing to maximally take ...sessions.". Because the results were similar of each of the three items, here we focus on the willingness to pay for the total coaching package.

5 RESULTS

Data analyses were conducted with R version 4.1.3. Before data analysis, we examined if there were any outlying values with +/- 3 *SD* within each experimental condition. The only dependent variable with outliers was monetary offers from which we excluded *n* = 4 participants from the respective analysis of H4.

5.1 Manipulation Check

If the nonverbal dominance manipulation was successful, participants should perceive the virtual agent in the expansive gesture condition as more dominant than in the neutral condition. In order to test this assumption, a Welch's *t*-test was conducted. Participants who interacted with the dominant virtual agent (*M* = 3.84, *SD* = 1.24) perceived it as significantly more dominant than participants who interacted with the neutral virtual agent (*M* = 3.44, *SD* = 1.44), $t(215.55)$ = 2.26, p = .024, d = .30. Hence, the manipulation of dominance was successful (confirming Hypothesis 1). Figure 2 illustrates the perceived dominance for each experimental condition.
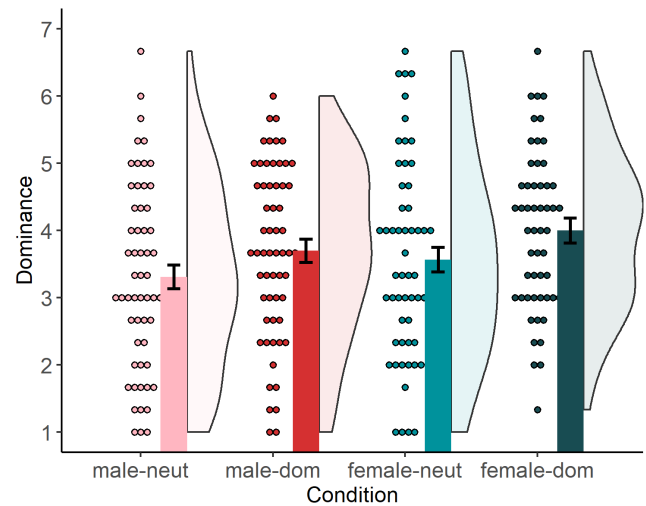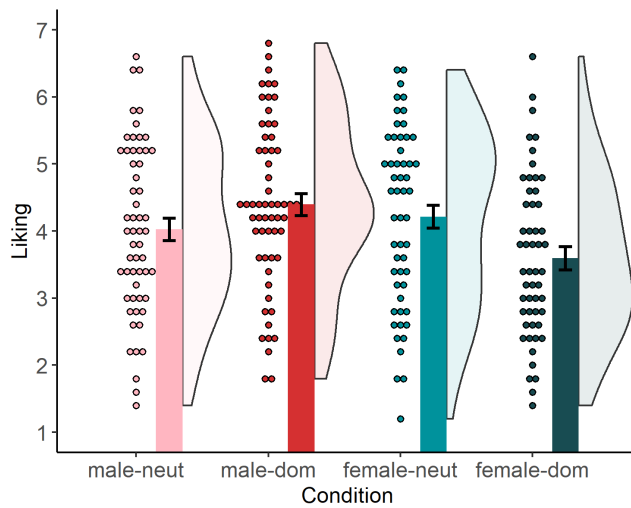

Figure 2: Effects of dominance manipulation on the dominance rating. Dots on the left represent data points of individual participants, violins on the right the distribution within conditions. Error bars represent +/- 1 *SE*.

5.2 Confirmatory Analyses

For Hypotheses 2 to 4 we used 2 (Gender: male vs. female) x 2 (Gestures: dominant vs. neutral) analysis of variances (ANOVA) on the respective measurement, and report main effects and interactions. In case of a significant interaction effect, we disentangled it by looking at simple comparisons between the experimental groups. All means and standard deviations can be found in Table 1.

Hypothesis 2 tested whether female agents with dominant gestures are perceived as less likable than dominant male agents and neutral female agents. The ANOVA revealed a main effect of gender, $F(1, 219)$ = 11.23, p < .001, η_p^2 = .05. This main effect was qualified by a significant interaction of gender and gestures, $F(1, 219)$ = 8.52, p = .003, η_p^2 = .04. Subsequent comparisons showed that dominant female agents were liked less than neutral female agents, $t(219)$ = -2.54, p = .012, d = .49, $CI_{95\%}[0.10, 0.87]$, and also less than dominant male agents, $t(219)$ = -3.35, p < .001, d = .29, $CI_{95\%}[0.27, 1.03]$. Thus, Hypothesis 2 was confirmed: Dominant female agents faced a social backlash, that is, were liked less than their male and female counterparts (see Fig. 3).

Figure 3: Effects of dominance manipulation on liking. Dots on the left represent data points of individual participants, violins on the right the distribution within conditions. Error bars represent +/- 1 *SE*.

Figure 4: Effects of dominance manipulation on competence. Dots on the left represent data points of individual participants, violins on the right the distribution within conditions. Error bars represent +/- 1 *SE*.

Hypothesis 3 states that agents with dominant behavior are perceived more competent than agents with neutral behavior. The ANOVA revealed a main effect of gender, $F(1, 219) = 4.59$, $p < .033$, $\eta_p^2 = .02$, and the predicted main effect of gestures, $F(1, 219) = 7.92$, $p < .005$, $\eta_p^2 = .03$, confirming H3.
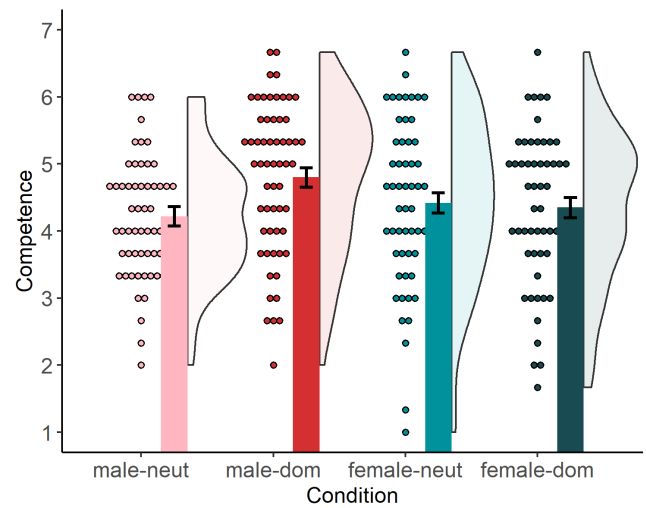
However, surprisingly, these main effects were qualified by an interaction of the two conditions, $F(1, 219) = 4.74$, $p = .031$, $\eta_p^2 = .02$. Simple contrasts showed that dominant male agents were perceived as more competent than neutral male agents, $t(219) = -2.81$, $p = .005$, $d = .57$, $CI_{95\%}[0.20, 0.94]$. Moreover, dominant male agents were perceived as more competent than dominant female agents, $t(219) = -2.14$, $p = .033$, $d = .40$, $CI_{95\%}[0.02, 0.77]$, (see Fig. 4).

Hypothesis 4 stated that female agents with dominant gestures receive lower monetary offers than dominant male agents and neutral female agents. There were neither effects of gender $F(1, 215) = 0.39$, $p = .535$, nor gestures, $F(1, 215) = 0.26$, $p = .613$. Moreover, no interaction could be detected $F(1, 215) = 0.81$, $p = .369$, showing no evidence for Hypothesis 4 (see Fig. 5). [1]

5.3 Exploratory Analyses

Because the interaction effect on monetary offers was non-significant, we did not further explore a possible mediation effect with liking as the mediator. However, we explored whether the participants' gender moderated the effects of agent's gender and nonverbal behavior on monetary outcomes. Multiple regression showed no main effect of gender, $b = 50.54$, $t(211) = 1.64$, $p = .102$, a main effect of gestures, $b = 62.76$, $t(211) = -2.06$, $p = .040$, an interaction of the agent's gender and its gestures, $b = -141.71$, $t(211) = -3.09$, $p = .002$,
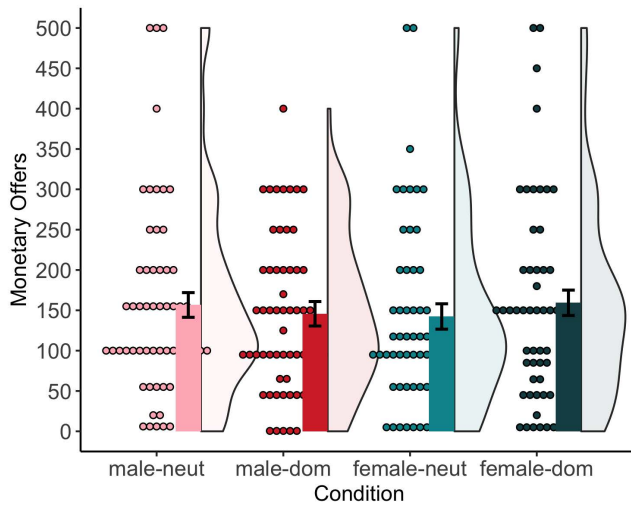
a marginal significant interaction of the agent's and the participant's gender, $b = -72.50$, $t(215) = -1.69$, $p = .093$, an interaction of the agent's gestures and participant's gender, $b = -99.83$, $t(215) = -2.35$, $p = .020$. These interactions were qualified by a significant three-way interaction of gestures, agent's gender and participants' gender, $b = 202.75$, $t(215) = 3.29$, $p = .001$.

To disentangle the interaction, we ran the 2 x 2 ANOVAs separately for male and female participants. For male participants, a significant interaction between agent's gender and gesture type emerged, $F(1, 95) = 7.75$, $p = .006$, while this interaction was not significant for female participants, $F(1, 116) = 2.71$, $p = .102$. All main effects were non-significant, all ps >.067.

For male participants, simple contrasts showed that they were willing to pay equal values to dominant male, $M = 127.6$, $SE = 23.2$, and female agents, $M = 178.2$, $SE = 25.0$, $b = -50.5$ $t(95) = -1.48$, $p = .142$, but more to neutral male agents, $M = 190.4$, $SE = 24.5$, compared to neutral females, $M = 99.2$, $SE = 28.7$, $b = 91.2$, $t(95) = 2.42$, $p = .012$. Moreover, they paid slightly less to the dominant versus neutral male agents $b = -62.8$, $t(95) = -1.86$, $p = .066$. However, they paid the female dominant agents more than the neutral ones, $b = 78.9$, $t(95) = 2.07$, $p = .041$. Thus, it seems as if the male participants were punishing the male agent's for being dominant.

6 DISCUSSION

This paper investigated if IVAs showing expansive, large versus narrow, small hand gestures in a self-promotion video were perceived as more dominant, and if female agents showing such dominance were facing negative consequences compared to male agents. Indeed, participants perceived larger gestures as more dominant than neutral ones (confirming H1). Moreover, a backlash effect for female dominant agents emerged: They were liked less than their male

---

[1]Sixteen participants had values > 300€ on the monetary offer; however, the agent's offer had been only 300€. We did not know if these participants had seriously answered a price higher than the agent's offer or had not paid attention. We left these participants in the sample, but dropping these participants from analyses did not change the pattern of results.

Figure 5: Effects of dominance manipulation on monetary offers. Dots on the left represent data points of individual participants, violins on the right the distribution within conditions. Error bars represent +/- 1 *SE*.

Table 1: Descriptives

	male		female	
	neutral	dominant	neutral	dominant
Variable	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Dominance	3.31 (1.39)	3.70 (1.28)	3.57 (1.50)	4.00 (1.18)
Likeability	4.03 (1.27)	4.40 (1.24)	4.21 (1.33)	3.60 (1.20)
Competence	4.22 (0.91)	4.80 (1.12)	4.42 (1.25)	4.35 (1.14)
Monetary Offer	171.44 (164.47)	157.73 (113.90)	158.24 (163.80)	159.32 (123.79)

Note. N = 223. Dominance was measured on a 7-point scale from 1 (*do not agree*) to 7 (*agree completely*), Likeability was measured using 5 items each consisting of an identical 7-point scale, Competence using 3 items each consisting of an identical 7-point scale (according to Carli et al. (1995)), Monetary Offers were entered manually.

counterparts (H2) and, surprisingly, even judged as less competent. However, they received the same value of monetary offers for the coaching as male agents (not supporting H4).

Women face an impression-management dilemma in situations that require dominant and assertive behavior [35]. Dominant and assertive women are seen as less likable, more unsuitable for a job and are less likely to be promoted than dominant men, and receive worse monetary outcomes in negotiations [1, 6, 21, 35, 36]. In this study, we replicated such backlash effects in a human–agent interaction: Dominant female agents were liked less than dominant male agents and neutral females. This finding is not consistent with a recent study investigating backlash effects in virtual agents. This study found that assertive females were indeed rated as less friendly than non-assertive females, but as equally friendly as the assertive males [40]. They had used a different negotiation scenario and manipulation of assertiveness via verbal content and offer height. This might account for the differences in patterns.

A recent meta-analysis investigated if the backlash effect on liking — statistically speaking the interaction of dominance x target gender — was moderated by explicit versus implicit dominance [44]. Explicit dominance involved directly trying to change the

other's behavior, while implicit dominance was communicated via more subtle modes like nonverbal or paraverbal behaviors — like in the current study. Over the set of *k* = 63 studies, the authors found that the backlash effect was present for explicit, but not for implicit manipulations of dominance. This was due to the fact that the dominance was more detectable in studies using explicit manipulations. In the current study, using an IVA, even the implicit manipulation of dominance by using larger vs. smaller gestures led to a higher dominance perception and a backlash effect for female agents.

For competence, we had predicted that dominant agent would get higher ratings independent of its gender. Although overall participants rated agents using dominant gestures as more competent (confirming H3), this effect depended on the agent's gender. Surprisingly, participants rated the dominant male agents as the most competent, and in general, the female agents as less competent than males. This contradicts findings that the gender gap in competence judgements has diminished in recent decades [14], that backlash effects usually do not emerge on competence [44], and that females under some conditions even have a competence advantage [27]. Moreover, our context was neither specifically male nor female: Although agents were promoting themselves, they offered a coaching for difficult social situations like job interviews. Thus, stereotypical female skills like being communicative and relationship-oriented should be necessary. Given such a task, the gender gap in competence ratings is surprising. Supposedly, some other features might have influenced the competence perception, for example, how the agents were dressed or their hair colour. Future research could replicate this study with agents looking similar and wearing the same clothes.

Like previous studies, this study showed no support for a backlash effect of the monetary offer that agents had received [40, 42]. This is contrary to findings in human–human interactions in which females do not get the outcomes they request, unless advocating for others [2, 28]. One possible explanation could be that participants were not negotiating about real monetary outcomes, and it could have been hard to imagine to actually pay a virtual agent for its service. Participant's gender did moderate the effects on monetary outcomes, but in a way that male participants seemed to rather punish dominant male agents. Indeed, there is evidence that men behave more competitively with other men, but more prosocial with women [9]. Future research could explore this interesting effect further.

The current findings raise the question of how IVAs should be designed. On the one hand, designing gender-neutral or even counter-stereotypical IVAs could be an important step to change and counteract stereotypical associations in the long run (for a similar discussion about gender stereotypes in designing robots, see [16]). One the other hand, if the goal is to create pleasant interactions, designing stereotype-congruent agents might result in better evaluations. In the context of the current study, such evaluations did not influence participant's willingness to pay for the coaching. Also, when exploratorily investigating participant's willingness to work with the agents, no effects of the experimental groups emerged, all *p*s > .14. Although dominant females might be liked less, the behavioral intentions towards them were not different from those towards the stereotype-congruent agents.

The expression of dominance seems to be especially critical for females in self-promoting contexts like a salary negotiation or like in the present study, making an offer as a self-employed coach. If a female agent is showing dominance while promoting somebody else, such backlash effects might diminish — like it is the case for humans [1]. Future research could explore if dominant female agents promoting somebody else rather than themselves face less social costs.

One limitation of this study was that participants were watching a video of an IVA rather than interacting with them. Future research could try to replicate these findings in an interactive scenario. Second, we had manipulated dominance via gesturing, but in an interaction, also other forms of dominance manipulation might be interesting (e.g., eye contact while speaking). Third, possible confounds could have influenced the results of the current study. For example, the female's dominant gestures might have been perceived as less fluid or realistic than the male's dominant gestures. Although we kept the conditions and parameters of the gestures identical, In future research it would be favorable to (1) include a participant question about the perceptions of possible confounds, and (2) to use several different agents from the same gender in a within-participant design in order to rule out agent specific confounds.

7 CONCLUSION

This study replicated the so-called backlash effect known from human–human interaction in a human–virtual agent setting: Participants judged female agents using more spacious and thus more dominant gestures as less likeable and less competent than dominant male and neutral female agents. Nevertheless, dominant females received equally high monetary offers as the other agents. This research shows the importance of taking the agent's gender into account in research on social perception and cognition in human–agent interaction.

ACKNOWLEDGMENTS

REFERENCES

[1] Emily Amanatullah and Catherine Tinsley. 2013. 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* 120 (01 2013), 110–122. https://doi.org/10.1016/j.obhdp.2012.03.006

[2] Emily T Amanatullah and Catherine H Tinsley. 2013. Ask and ye shall receive? how gender and status moderate negotiation success. *Negotiation and Conflict Management Research* 6, 4 (2013), 253–272. https://doi.org/10.1111/ncmr.12014

[3] Elizabeth Aries, Conrad Gold, and Russell Weigel. 1983. Disposition and situational influences on dominance behavior in small groups. *Journal of Personality and Social Psychology* 44 (04 1983), 779–786. https://doi.org/10.1037/0022-3514.44.4.779

[4] R Ashmore and Frances K. Del Boca. 1979. Sex stereotypes and implicit personality theory: Toward a cognitive—Social psychological conceptualization. *Sex Roles* 5 (1979), 219–248. https://doi.org/10.1007/BF00287932

[5] Nikolaus Bee, Colin Pollock, Elisabeth Andre, and Marilyn Walker. 2010. Bossy or wimpy: expressing social dominance by combining gaze and linguistic behaviors, Vol. 6356. 265–271. https://doi.org/10.1007/978-3-642-15892-6_28

[6] Hannah Bowles, Linda Babcock, and Lei Lai. 2007. Social incentives for gender differences in the propensity to initiate negotiations: sometimes it does hurt to ask. *Organizational Behavior and Human Decision Processes* 103 (05 2007), 84–103. https://doi.org/10.1016/j.obhdp.2006.09.001

[7] Judee Burgoon and Norah Dunbar. 2006. *Nonverbal expressions of dominance and power in human relationships.* 279–297.

[8] Judee Burgoon, Michelle Johnson, and Pamela Koch. 1998. The nature and measurement of interpersonal dominance. *Communication Monographs - COMMUN MONOGR* 65 (12 1998), 308–335. https://doi.org/10.1080/03637759809376456

[9] Abraham P Buunk and Karlijn Massar. 2012. Intrasexual competition among males: competitive towards men, prosocial towards women. *Personality and Individual Differences* 52, 7 (2012), 818–821. https://doi.org/10.1016/j.paid.2012.01.010

[10] Zoraida Callejas, Brian Ravenet, Magalie Ochs, and Catherine Pelachaud. 2014. A computational model of social attitudes for a virtual recruiter. *13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014* 1 (05 2014).

[11] Linda L. Carli, Suzanne J. LaFleur, and Christopher C. Loeber. 1995. Nonverbal behavior, gender, and influence. *Journal of Personality and Social Psychology* 68 (1995), 1030–1041. https://doi.org/10.1037/0022-3514.68.6.1030

[12] Dana Carney, Judith Hall, and Lavonia LeBeau. 2005. Beliefs about the nonverbal expression of social power. *Journal of Nonverbal Behavior* 29 (01 2005), 105–123. https://doi.org/10.1007/s10919-005-2743-z

[13] John Dovidio, Clifford Brown, Karen Heltman, Steve Ellyson, and Caroline Keating. 1988. Power displays between women and men in discussions of gender-linked tasks: a multichannel study. *Journal of Personality and Social Psychology* 55 (10 1988), 580–587. https://doi.org/10.1037/0022-3514.55.4.580

[14] Alice H Eagly, Christa Nater, David I Miller, Michèle Kaufmann, and Sabine Sczesny. 2020. Gender stereotypes have changed: a cross-temporal meta-analysis of US public opinion polls from 1946 to 2018. *American psychologist* 75, 3 (2020), 301. https://doi.org/10.1037/amp0000494

[15] Alice H Eagly and Wendy Wood. 2012. Social role theory. *Handbook of theories in social psychology* 2 (2012), 458–476.

[16] Friederike Eyssel and Frank Hegel. 2012. (S)he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology* 42, 9 (2012), 2213–2230. https://doi.org/10.1111/j.1559-1816.2012.00937.x

[17] Patrick Gebhard, Tobias Baur, Ionut Damian, Gregor Mehlmann, Johannes Wagner, and Elisabeth Andre. 2014. Exploring interaction strategies for virtual characters to induce stress in simulated job interviews. *13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014* 1.

[18] Charamel GmbH. 2021. VuppetMaster. https://vuppetmaster.de/

[19] Joanna Hale and Antonia Hamilton. 2016. Cognitive mechanisms for responding to mimicry from others. *Neuroscience Biobehavioral Reviews* 63 (02 2016). https://doi.org/10.1016/j.neubiorev.2016.02.006

[20] Judith A. Hall, Erik J. Coats, and Lavonia Smith LeBeau. 2005. Nonverbal behavior and the vertical dimension of social relations: a meta-analysis. *Psychological bulletin* 131 6 (2005), 898–924. https://doi.org/10.1037/0033-2909.131.6.898

[21] Madeline Heilman and Tyler Okimoto. 2007. Why are women penalized for success at male tasks?: the implied communality deficit. *The Journal of applied psychology* 92 (02 2007), 81–92. https://doi.org/10.1037/0021-9010.92.1.81

[22] Michael Kipp and Patrick Gebhard. 2008. IGaze: studying reactive gaze behavior in semi-immersive human-avatar interactions, Vol. 5208. 191–199. https://doi.org/10.1007/978-3-540-85483-8_19

[23] Anne M Koenig and Alice H Eagly. 2014. Evidence for the social role theory of stereotype content: observations of groups' roles shape stereotypes. *Journal of personality and social psychology* 107, 3 (2014), 371. https://doi.org/10.1037/a0037215

[24] D. J. Leiner. 2021. SoSci survey (version 3.1.6). https://www.soscisurvey.de

[25] Jamy J. Li. 2015. The benefit of being physically present: a survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *Int. J. Hum. Comput. Stud.* 77 (2015), 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

[26] Christine L. Lisetti, Reza Amini, Ugan Yasavur, and Naphtali Rishe. 2013. I can help you change! an empathic virtual agent delivers behavior change health interventions. *ACM Trans. Manag. Inf. Syst.* 4 (2013), 19:1–19:28. https://doi.org/10.1145/2544103

[27] Barbara Masser, Kirsten Grass, and Michelle Nesic. 2007. 'We like you, but we don't want you'—the impact of pregnancy in the workplace. *Sex Roles* 57, 9 (2007), 703–712. https://doi.org/10.1007/s11199-007-9305-2

[28] Jens Mazei, Joachim Hüffmeier, Philipp Alexander Freund, Alice F Stuhlmacher, Lena Bilke, and Guido Hertel. 2015. A meta-analysis on gender differences in negotiation outcomes and their moderators. *Psychological bulletin* 141, 1 (2015), 85. https://doi.org/10.1037/a0038184

[29] Margaret McRorie, Ian Sneddon, Gary McKeown, Elisabetta Bevacqua, Etienne De Sevin, and Catherine Pelachaud. 2011. Evaluation of four designed virtual agent personalities. *IEEE Transactions on Affective Computing* 3, 3 (2011), 311–322. https://doi.org/10.1109/T-AFFC.2011.38

[30] Clifford Nass and Youngme Moon. 2000. Machines and mindlessness: social responses to computers. *Journal of Social Issues* 56 (03 2000), 81–103. https://doi.org/10.1111/0022-4537.00153

[31] Deborah Prentice and Erica Carranza. 2002. What women and en should be, shouldn't be, are allowed to be, and don't have to be: the contents of prescriptive gender stereotypes. *Psychology of Women Quarterly* 26 (12 2002), 269–281. https:

//doi.org/10.1111/1471-6402.t01-1-00066

[32] Deborah A. Prentice and Erica Carranza. 2003. *Sustaining cultural beliefs in the face of their violation: The case of gender stereotypes.* Lawrence Erlbaum Associates, 259–280. https://doi.org/10.4324/9781410608994

[33] Brian Ravenet, Magalie Ochs, and Catherine Pelachaud. 2013. From a user-created corpus of virtual agent's non-verbal behavior to a computational model of interpersonal attitudes. https://doi.org/10.1007/978-3-642-40415-3_23

[34] Byron Reeves and Clifford Nass. 1997. The media equation: how people treat computers, television, and new media like real people and pla. *Bibliovault OAI Repository, the University of Chicago Press* (01 1997).

[35] Laurie Rudman. 1998. Self-promotion as a risk factor for women: the costs and benefits of counterstereotypical impression management. *Journal of personality and social psychology* 74 (03 1998), 629–45. https://doi.org/10.1037//0022-3514.74.3.629

[36] Laurie Rudman, Corinne Moss-Racusin, Julie Phelan, and Sanne Nauts. 2012. Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology* 48 (11 2012). https://doi.org/10.1016/j.jesp.2011.10.008

[37] Tanja Schneeberger, Sofie Ehrhardt, Manuel S Anglet, and Patrick Gebhard. 2019. Would you follow my instructions if I was not human? examining obedience towards virtual agents. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 1–7. https://doi.org/10.1109/ACII.2019.8925501

[38] Tanja Schneeberger, Patrick Gebhard, Tobias Baur, and Elisabeth Andre. 2019. PARLEY: a transparent virtual social agent training interface. 35–36. https:

[39] Larissa Z Tiedens and Alison R Fragale. 2003. Power moves: complementarity in dominant and submissive nonverbal behavior. *Journal of personality and social psychology* 84, 3 (2003), 558. https://doi.org/10.1037/0022-3514.84.3.558

[40] Laura van der Lubbe and Tibor Bosse. 2017. Studying gender bias and social backlash via simulated negotiations with virtual agents. 455–458. https://doi.org/10.1007/978-3-319-67401-8_57

[41] Glenn E. Weisfeld and Jody M. Beresford. 1982. Erectness of posture as an indicator of dominance or success in humans. *Motivation and Emotion* 6 (1982), 113–131. https://doi.org/10.1007/BF00992459

[42] Janet Wessler. 2021. Economic and social consequences of anger and gender in computer-mediated negotiations: is there a backlash against angry females?. In *2021 9th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. 1–6. https://doi.org/10.1109/ACIIW52867.2021.9666337

[43] Janet Wessler, Tanja Schneeberger, Bernhard Hilpert, Alexandra Alles, and Patrick Gebhard. 2021. Empirical research in affective computing: an analysis of research practices and recommendations. In *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 1–8. https://doi.org/10.1109/ACII52823.2021.9597418

[44] Melissa Williams and Larissa Tiedens. 2016. The subtle suspension of backlash: a meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin* 142 (12 2016). https://doi.org/10.1037/bul0000039