Assessing Human Reactions in a Virtual Crowd Based on Crowd Disposition, Perceived Agency, and User Traits

BENJAMIN BENDIKSEN, University of Massachusetts, Boston, USA
NANA LIN, University of Massachusetts, Boston, USA
JIEHYUN KIM, University of Massachusetts, Boston, USA
FUNDA DURUPINAR, University of Massachusetts, Boston, USA

Immersive virtual environments populated by real and virtual humans provide valuable insights into human decision-making processes under controlled conditions. Existing literature indicates elevated comfort, higher presence, and a more positive user experience when virtual humans exhibit rich behaviors. Based on this knowledge, we conducted a web-based, interactive study, in which participants were embodied within a virtual crowd with complex behaviors driven by an underlying psychological model. While participants interacted with a group of autonomous humanoid agents in a shopping scenario similar to Black Friday, the platform recorded their non-verbal behaviors. In this independent-subjects study, we investigated behavioral and emotional variances across participants with diverse backgrounds focusing on two conditions: perceived agency and the crowd's emotional disposition. For perceived agency, one group of participants was told that the other crowd members were avatars controlled by humans, while another group was told that they were artificial agents. For emotional disposition, the crowd behaved either in a docile or hostile manner. The results suggest that the crowd's disposition and specific participant traits significantly affected certain emotions and behaviors. For instance, participants collected fewer items and reported a higher increase of negative emotions when placed in a hostile crowd. However, perceived agency did not yield any statistically significant effects.

CCS Concepts: • **Computing methodologies** → **Animation**.

Additional Key Words and Phrases: crowd simulation

ACM Reference Format:
Benjamin Bendiksen, Nana Lin, JieHyun Kim, and Funda Durupinar. 2023. Assessing Human Reactions in a Virtual Crowd Based on Crowd Disposition, Perceived Agency, and User Traits. *ACM Transactions on Applied Perception* 1, 1 (October 2023), 21 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The growing integration of digital technologies into our daily lives necessitates an in-depth understanding of our interactions within virtual environments. Exploring our relationship with virtual humans and crowds that populate these environments presents opportunities and challenges across various domains such as social psychology, human-computer interaction, emergency management, and gaming. One such opportunity involves

Authors' addresses: Benjamin Bendiksen, p.bendiksen001@umb.edu, Department of Computer Science, University of Massachusetts, Boston, 100 Morrissey Boulevard, Boston, MA, 02125, USA; Nana Lin, nana.lin002@umb.edu, Department of Computer Science, University of Massachusetts, Boston, 100 Morrissey Boulevard, Boston, MA, 02125, USA; JieHyun Kim, jiehyun.kim001@umb.edu, Department of Computer Science, University of Massachusetts, Boston, 100 Morrissey Boulevard, Boston, MA, 02125, USA; Funda Durupinar, funda.durupinarbabur@umb.edu, Department of Computer Science, University of Massachusetts, Boston, 100 Morrissey Boulevard, Boston, MA, 02125, USA.

XXXX-XXXX/2023/10-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

the ability to conduct controlled studies that would be unattainable or non-replicable in real-life settings. For instance, when examining events that involve crowds, such as those characterized by emergent behaviors like panic, stampedes, or riots, the consequences of irrational behavior are discernible; yet the underlying mechanisms can only be approximated through retrospective analyses or field observations. Given the realistic responses of people in immersive virtual worlds [Slater et al. 2020; Slater and Sanchez-Vives 2016] and the high ecological validity of virtual social interactions [Bombari et al. 2015], we can gain critical insights into human decision-making processes through controlled virtual crowd experiments in a low-risk and efficient manner.

Existing approaches that integrate humans into virtual crowds predominantly focus on crowd or human movement characteristics [Kim et al. 2016; Moussaïd et al. 2016; Nelson et al. 2020], or consider a limited set of other parameters such as crowd density [Dickinson et al. 2019], eye gazes of agents [Narang et al. 2016], or character appearance [Nelson et al. 2020]. However, a holistic approach needs to consider nuanced behaviors rooted in the complexities of human psychology. Understanding the reactions of a user when embodied in a crowd simulation system that portrays the psychological nuances of its agents remains a promising area for exploration. Such an understanding can elucidate human psychology and decision-making patterns in virtual or real crowds, depending on the level of immersion.

In this paper, we explore the impact of a virtual crowd's emotional disposition and perceived agency of crowd members on human emotions and non-verbal behaviors. We assess emotional disposition on two levels: hostile versus docile crowds. Perceived agency also involves two levels that indicate whether users believe crowd members to be human-controlled avatars or autonomous agents. The literature does not have consensus on the effects of perceived agency [Oh et al. 2018]. While some works suggest that avatars and agents elicit similar responses in users [Von der Pütten et al. 2010], others report stronger emotional responses toward avatars [Fox et al. 2015; Kothgassner et al. 2017]. However, existing studies only involve interactions with individual virtual humans, and do not extend to crowds.

We introduce an independent-subjects study where users were embodied in a web-based 3D platform involving emotionally expressive virtual crowds exhibiting complex behaviors. The web-based nature of the study allowed us to efficiently recruit a diverse set of participants. We designed a scenario mirroring a shopping event resembling Black Friday. The participants were told to collect as many items as they could from a virtual store. We introduced a reward strategy directly mapped from real life, where participants earned additional compensation based on the number of items they collected. While each person engaged with the environment as a crowd member, the platform collected data regarding their simulation behaviors and affective states.

In addition to crowd disposition and perceived agency, we explored the impact of individual differences across participants on their emotional and behavioral responses. Specifically, we considered their age, gender, personality traits, and familiarity with 3D virtual environments. This work poses the following research questions:

- **RQ1**. How does the virtual crowd's emotional disposition towards aggression affect user emotions and behavior?
- **RQ2**. How does perceived agency affect user emotions and behavior? Specifically, do user behaviors and emotions vary when they perceive the other crowd members as avatars controlled by real individuals as opposed to artificial agents?
- **RQ3**. How do users' demographics affect their emotions and behavior when interacting with virtual crowds?
- **RQ4**. Are there any correlations between users' personality traits and their emotions and behaviors in the simulation?

Our work, expanding upon traditional approaches to crowd simulation research, holds the potential to advance the understanding of human behavior in social environments while contributing to the development of more effective and adaptable crowd simulation strategies.

2 RELATED WORK

2.1 Interactions with Crowds in Virtual Environments

The promise of immersive virtual environment technologies to support social experiments has long been established [Blascovich et al. 2002; Loomis et al. 1999]. This can be attributed to the high ecological validity of virtual social interactions [Bombari et al. 2015]. Multiple studies have shown that people treat and communicate with virtual humans naturally as if they were real humans [de Borst and de Gelder 2015; Nass and Reeves 2003]. Virtual humans increase the sense of immersion in virtual worlds [Llobera et al. 2010; Pelechano et al. 2008], provide standardization for social experiments, and offer low-cost solutions for training applications [Bombari et al. 2015].

Besides individual exchanges with virtual humans, understanding human interactions in multi-agent scenarios carries a significant potential to advance both our knowledge of human psychology and the state of the art in crowd simulation research [Pelechano and Allbeck 2016]. Since Pelechano et al.'s seminal work that evaluated user presence in a virtual crowd [Pelechano et al. 2008], many studies have explored human behaviors when embodied in virtual crowds, especially in VR environments. These studies have predominantly examined the steering or walking behaviors of human participants in crowds. Nelson et al. [Nelson et al. 2019], evaluated walking within a virtual crowd where the participant was actually walking in a motion capture studio. They found that an extremely dense virtual crowd significantly altered the movement behavior of participants in terms of their speed and walking time. This finding was also supported by Koilias et al. [Koilias et al. 2020a], who reported a high impact of density, low speed, and diagonal direction on the speed, deviation, and trajectory lengths of participants. In a separate work, Nelson et al. found participant movement speed, deviation, and interpersonal distance to be significantly affected by the appearance of the virtual characters [Nelson et al. 2020]. Bruneau et al. [Bruneau et al. 2015] studied the interactions between individual walkers and groups by applying the Principle of Minimum Energy (PME) and found that humans behaved as predicted by the PME. For instance, participants preferred to go through large and sparse groups and go around small and dense groups. The study took place in a CAVE environment where the participants navigated with a joystick. The work demonstrated the promise of VR experiments to guide the improvement of crowd simulations by directly designing an algorithm to imitate human behavior. Olivier et al. [Olivier et al. 2017] also performed CAVE experiments to show that collision avoidance trajectories and levels of perception in VR were sufficiently realistic for human locomotion studies.

Jiang et al. [Jiang et al. 2018] investigated the coordination of joint actions on road-crossing in VR, which provides a safe environment for such tests. They observed that participants' impulses to move in coordination when crossing a virtual road were consistent with real-world studies. Koilias et al. [Koilias et al. 2020b] evaluated the movement coordination of participants with virtual crowd members and found significant differences between the movements of participants and crowd agents. Particularly, the participants moved slower, followed longer paths, performed less smooth motions, and had higher interpersonal distance. Although they found moderate associations, the crowd's influence was not enough to make the participant become a part of it.

Studies show proxemics to be an influential factor in user experience. In an early work by Llobera et al. [Llobera et al. 2010], virtual characters approached a stationary participant in groups of one to four agents. The authors measured electrodermal activity to evaluate participant arousal and found that skin conductance levels increased more the closer the virtual characters approached the participants. Christou et al. [Christou et al. 2015] also evaluated proxemics and found similar results by measuring electrodermal activity, again with a stationary participant and groups of agents. Additionally, they showed a declined cognitive performance under close proximity. Different from earlier works, Dickinson et al. [Dickinson et al. 2019] designed an experiment where the participant was allowed to navigate naturally in the environment with dynamically tracked handheld controllers and a head-mounted display. They found high density to increase the difficulty of participants' carrying out the given task and negatively impact their affective states as measured by PANAS self-reports.

Instead of employing crowds of virtual humans, Moussaïd et al. [Moussaïd et al. 2016] conducted a study with multiple human participants, demonstrating the efficacy of virtual platforms for social experiments. They replicated a high-stress emergency situation from a real-life experiment on a screen-based study involving the simultaneous participation of multiple human subjects. An important observation they report is that herding resulted from the high density and not from a change in the individual tendency to imitate neighbors.

Gaze and gestures positively affect user experience by increasing perceived realism and feeling of comfort. Narang et al. [Narang et al. 2016] introduced PedVR, a VR framework that couples crowd simulations with realistic 3D body movements and gestures of a large number of virtual characters. Their within-subjects experiments indicated that the gaze behavior of agents through eye contact with the user improved perceived believability. Kyrikou et al. [Kyriakou et al. 2017] suggested that in addition to collision avoidance with the participant, interactions such as verbal salutations, gaze, waving, and other gestures should be employed to enhance the sense of presence.

Volonte et al. [Volonte et al. 2020] developed an agent-based crowd model with rich behaviors including eye gaze, facial expressions, body motion, verbal, and non-verbal behaviors. They evaluated the impact of a crowd of emotional virtual humans on users' affective and non-verbal behaviors in a VR setting and found that participants interacted with the positive emotional crowds more than the negative ones. Our work shares similarities with this study in the exploration of complex virtual human behaviors and the emphasis on affective and non-verbal behaviors. Differently, our crowd simulation system is controlled by a multi-layered parametric psychological model that incorporates personality and emotions to control low and high-level agent behaviors. Additionally, the experimental settings, scenarios, and research questions diverge significantly, which demonstrates the applicability and breadth of virtual social interactions for understanding human reactions.

Although undoubtedly more immersive, VR with head-mounted displays or in CAVE is more restrictive than desktop environments regarding participant recruitment. The aforementioned studies typically involved a small number of participants, ranging from 13 [Bruneau et al. 2015] to 18 [Nelson et al. 2019, 2020]. Given our primary emphasis on high-level behavioral parameters and emotional outcomes rather than steering behaviors, precise locomotion control was not a pressing requirement for our work. Thus, we preferred a browser-based environment, which facilitated access to a broader participant base.

2.2 Perception of Aggression in Crowds and Virtual Humans

Several studies have examined the perception of emotions in crowds. For example, Hansen and Hansen [Hansen and Hansen 1988] showed an anger superiority effect, where angry faces were more easily detected in crowds of happy faces compared to happy faces in crowds of angry ones. Bucher and Voss [Bucher and Voss 2019] asked participants to rate the overall mood of the crowd and found that happy faces were more likely to be attended to and their predominance was assessed more accurately than the predominance of angry faces. Mihalache evaluated anger bias towards crowds with varying ratios of angry to happy faces and found that anger bias emerges particularly in the context of perceptual uncertainty, i.e. with low intensities of expressions [Mihalache et al. 2021]. It is important to note that in these studies, crowds were composed of static images of real or computer-generated faces, not real or simulated crowds. Also, the participants were merely observers.

Studies involving individual virtual humans also confirm anger bias. For instance, participants took longer times to recognize patterns when presented with angry faces [Rapuano et al. 2023], tended to increase spatial distances, and reacted with increased emotional arousal [Ruggiero et al. 2021].

2.3 Perceived Agency

Perceived agency, or "agency belief", is a concept for which there is no clear consensus in the literature regarding its effects. There are many studies that did not find any differences towards agents vs. avatars in social judgments [Nowak 2004; Von der Pütten et al. 2010]. This can be explained by the "ethopoeia" concept, introduced by Nass and Moon, which indicates that systems elicit social responses as long as they provide social cues [Nass and Moon 2000]. Von der Pütten et al. later discussed that the strength of reactions would depend on the amount or strength of the social cues [Von der Pütten et al. 2010] and demonstrated that higher behavioral realism led to stronger social effects.

Contrary to these works, in a study by Guadagno et al., subjects assessed a virtual human's behaviors as more realistic and reported higher levels of social presence when they believed that it was controlled by a human [Guadagno et al. 2007]. The authors also found interaction effects of agency with the gender of the virtual human, indicating in-group favoritism and stereotypical effects. Other studies also support the relationship between social presence and agency in a similar way, reporting stronger effects when others were perceived as avatars instead of agents [Appel et al. 2012; Fox et al. 2015; Gajadhar et al. 2008; Weibel et al. 2008]. Poinsot et al. discussed the interaction effects of emotional communication and perceived agency, reporting that emotional communication led computer-controlled opponents to elicit a stronger sense of co-presence than human-controlled ones [Poinsot et al. 2022].

In the opposite direction, Williams and Clippinger found higher levels of aggression toward computer opponents than humans [Williams and Clippinger 2002]. Lim and Reeves [Lim and Reeves 2010] later showed interaction effects of agency with the type of game activity (cooperative vs. competitive). Thus, the apparent contradictions in research results toward agency may be a result of other factors such as visual and behavioral fidelity, type of interaction, and in-group effects.

3 METHOD

3.1 Stimuli Creation

We developed an interactive, browser-based platform using Unity[1] and WebGL, building upon an existing crowd simulation system that models crowd behavior using social psychology theories and features autonomous humanoid agents with controllable psychological states [Durupinar et al. 2016a]. The system allows authoring scenarios by initializing agents with social roles, cognition, and personalities, which determine the agents' emotions, decision-making, and actions. Agents interact with each other and the environment; have emotional facial expressions and body postures; and exhibit complex behaviors. The system uses the Five-Factor model of personality [Goldberg 1990a] to indirectly influence all these features and directly control low-level dynamics parameters such as velocities, forces, and local steering choices. The system also employs an epidemiological emotional contagion model, simulating a phenomenon that causes emotions to propagate among the crowd, which potentially yields irrational emergent behavior, turning regular crowds into emotional mobs. Our extended platform incorporates the human user as a crowd member with a 3D body who can navigate the environment with a first-person perspective.

We launched the platform's Unity WebGL build on a public-facing server that employs Express [2], a Node.js application framework. The collected user data was managed on a MySQL server. Our platform's web support facilitates deployment on crowdsourcing websites like MTurk. Users accessed the study's website, which was embedded in an HTML inline frame (iFrame) as part of the MTurk Human Intelligence Task (HIT). Worker IDs and HIT information were automatically transmitted to our server (Figure 1). As a quality assurance measure, we hid the "submit" button until a study completion token was sent from the server.

[1]https://unity.com/
[2]https://expressjs.com/

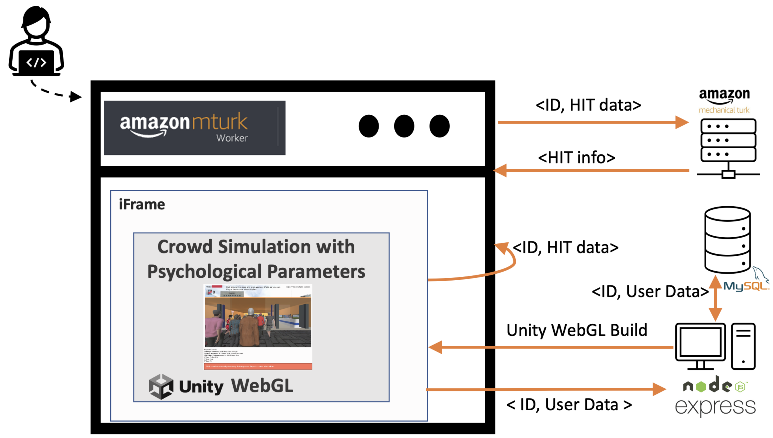

Fig. 1. Platform architecture for the study

The scene involved a big sale event in a store offering discounts, with the goal of purchasing iPads. There were 20 virtual agents and a human user. Participants were told that they would receive bonuses based on the number of iPads they collected, incentivizing them with a direct mapping to the scenario's theme. Both the human user and the computer agents could navigate the store, collecting iPads from the shelves and paying at the counter. Users were given the option to start fights with virtual humans, with the winner collecting the loser's iPads. Users could interact with the environment through keyboard controls using the arrows for navigation, the 'C' key for collecting iPads, 'F' for fighting, and 'P' for paying at the counter (Figure 2). We preferred keyboard controls over mouse or touchpad for accessibility.

Each press of 'F' animated a punching clip on the human's avatar and included a punching sound. To enhance immersion, we also assigned footstep sounds to the participant's walking. We used freely available Mixamo [3] and Adobe Fuse [4] virtual human models with facial blendshapes.

3.2 Study Design

We conducted a study with a *2x2* factorial design model with independent subjects on MTurk. Each task consisted of six scenes: a warm-up scenario, demographic data survey, participant personality assessment, pre-study participant emotional state assessment, the main scenario, and post-study participant emotional state assessment (Figure 3). The warm-up scene displayed the environment without any virtual humans, where the participants were instructed to collect iPads. The study started with a warmup so that the workers could decide whether to continue the study or not. We assessed personality by a brief Five-Factor personality measure, the Ten Item Personality Inventory (TIPI) [Gosling et al. 2003]. For the emotional state assessment, we used a short version of PANAS [Watson et al. 1988] for brevity: the International Positive and Negative Affect Schedule Short-Form (I-PANAS-SF) [Karim et al. 2011]. At the end of the study, we presented a questionnaire with the following questions:

- Please rate your overall experience during the iPAD collection task on a scale of 1 to 7, where 1 indicates "Not at all like interacting with real persons" and 7 indicates "Exactly like interacting with real persons."

[3] https://www.mixamo.com/
[4] https://www.adobe.com/tr/products/fuse.html

Fig. 2. A first-person view of the environment at the beginning of the study.

- Did you feel that the behaviors of others during the task were consistent and predictable? Please explain your answer.
- How would you describe the personalities of others based on your interactions? (e.g., friendly, helpful, competitive, cooperative)
- Did you notice any emotional responses from other characters during the task? If yes, please describe the emotions you observed.
- To what extent did you feel that other characters understood your perspective and communicated effectively? Rate on a scale of 1 to 7, where 1 indicates "Not at all" and 7 indicates "Completely."
- Were there any moments during the task that made you doubt or suspect that other characters might not be other MTurkers but artificial characters? If yes, please describe those moments. (Question reworded as "Were there any moments during the task that made you doubt or suspect that other characters might not be artificial characters but other MTurkers? for the "Agent" condition.)
- Please provide any additional comments or observations about your experience during the task, including any thoughts or feelings you had.

The study conditions included the emotional disposition of the crowd (*docile* vs. *hostile*) and whether the participants were told the others in the crowd were *virtual agents* or *avatars controlled by other MTurk workers*. In the *docile* scenario, crowd agents were assigned neutral personalities (with all the five factors set as 0, e.g., neither introvert nor extrovert), so they only walked around the store, collecting iPads and not showing strong emotions or fighting (Fig. 4(a)). In the *hostile* scenario, the personalities were assigned as unconscientious, extroverted, disagreeable, and neurotic ($openness = 0$, $conscientiousness = -1$, $extroversion = 1$, $agreeableness = -1$, $neuroticism = 1$), so that the agents would be more assertive and prone to starting fights with the human participant and other agents, displaying negative emotions such as angry facial expressions, and exiting the store without paying [Durupinar et al. 2016b](Fig. 4(b)). During each scenario, we recorded the total time, average speed, the number of fights, punches, iPads grabbed from others, and the total number iPads (collected from

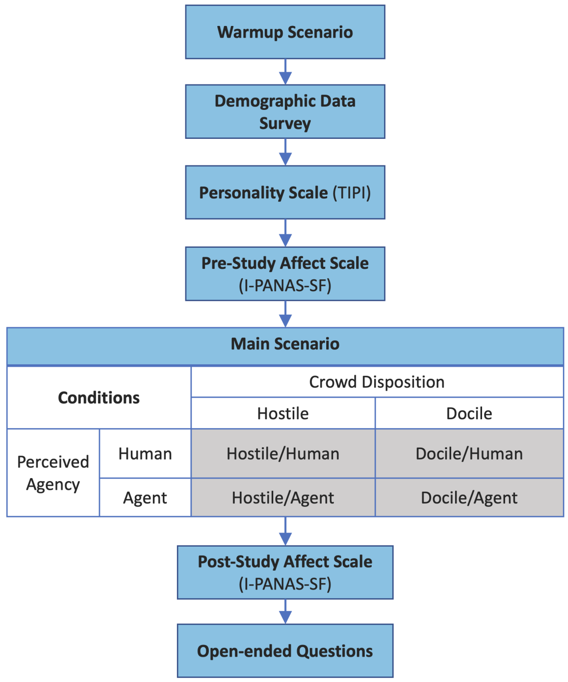

Fig. 3. Independent-subjects study procedure

Fig. 4. (a) A docile crowd; (b) a hostile crowd. Participant fights with an agent.

shelves and grabbed in a fight). The University Institutional Review Board approved the experiment protocol. The participants were paid \$1 per task and an additional \$0.02 bonus per iPad they collected.

3.3 Participants

We set participation requirements as having an acceptance rate of $> 95\%$ and experience of more than 100 HITs and Masters qualifications. To ensure quality, we placed attention-checking questions in the questionnaires and discarded the responses of participants who did not pass these tests. For a medium effect size (Cohen's $d = 0.25$)

for both main effects and their interaction and power of 0.80 at a significance level of 0.05, we collected 30 participant responses per group (120 in total).

Of the 120 unique participants (83M/37F/0 other), the average age was = 37.083 ± 11.41. The ethnic distribution was 81 White, 25 Asian, 6 Black, and 7 Hispanic/LatinX, and 1 other. We also asked about familiarity with first-person view video games on a scale of 0 ("not familiar at all") to 5 ("highly familiar"). The mean familiarity level was 1.533 with a standard deviation of 1.66.

3.4 Analysis

3.4.1 The Effects of Study Conditions on Participant Emotions and Behaviors. We collected the emotional states of participants before and after the study using a short and international version of PANAS. We collected scores for ten emotions, five positive and five negative, on a 5-point Likert scale. The positive emotion items were *active*, *determined*, *attentive*, *inspired*, *alert*, and the negative emotion items were *afraid*, *nervous*, *upset*, *hostile*, and *ashamed*. To test for differential effects, we subtracted the pre-study scores from post-study scores. Figure 5 shows the frequency histograms of the positive and negative affect differences scores, indicating approximately normal distributions.

As behavioral parameters, we collected the number of fights (*fightCnt*) each participant involved, punches administered (*punchCnt*), iPads collected from the shelves (*collectedItemCnt*), and iPads stolen from others (*stolenItemCnt*), as well as the time spent (*timeSpent*), average speed (*avgSpeed*), and the total distance covered during the simulation (*totalDist*). Figure 6 shows the frequency histograms of the behavior parameters for the whole study.

To analyze the effects of the perceived agency (whether participants believe others are humans vs. agents) and crowd disposition (docile vs. hostile) on emotional state changes, we designated these two factors as independent variables with two levels each, and the affect score differences as the response variables. With a lack of evidence for unequal variances across conditions (with Levene's test) we employed a two-way independent subjects Analysis of Variance (ANOVA) to test the effects of the study conditions on affective differences.

The ANOVA test returned a statistically significant main effect of crowd disposition on *afraid* $F(1,116) = 4.545, p = 0.035$, *upset* $F(1,116) = 4.474, p = 0.037$, *hostile* $F(1,116) = 12.865, p = 0.0005$, and *ashamed* $F(1,116) = 9.901, p = 0.002$. We did not find any main effects of agency or interaction effects. To control for potential inflation of the Type I error rate, we employed Benjamini-Hochberg (BH) procedure for False Discovery Rate (FDR) detection. After FDR correction, only *hostile* and *ashamed* remained statistically significant. Figure 7 illustrates the comparative distributions of affect score differences between post and pre-study responses across the study conditions. We also calculated the composite positive and negative affect scores as the means of the positive and negative items respectively as dictated by PANAS. ANOVA yielded a significant main effect of post and pre-study negative affect score difference with $F(1,116) = 13.954, p = 0.0003$. The effect remained statistically significant after the BH procedure. Table 1 summarizes the statistically significant results after ANOVA and FDR correction.

To measure simulation conditions on participant behaviors, we again took participant belief and crowd disposition as independent variables. Because of the ANOVA model's robustness given independent and sufficiently large data (typically $n = 30$) we repeated a two-way independent subjects ANOVA to test the effects of the study conditions on behavior parameters. The ANOVA test yielded statistically significant main effects of crowd type on *fightCnt* with $F(1,116) = 9.405, p = 0.003$, *collectedItemCnt* with $F(1,116) = 15.917, p = 0.0001$, *timeSpent* with $F(1,116) = 5.299, p = 0.023$, *totalDist* with $F(1,116) = 4.696, p = 0.032$. We did not find any main effects of agency or interaction effects. To control for potential inflation of the Type I error rate, we employed Benjamini-Hochberg procedure. After FDR correction, only *fightCnt* and *collectedItemCnt* remained statistically significant.

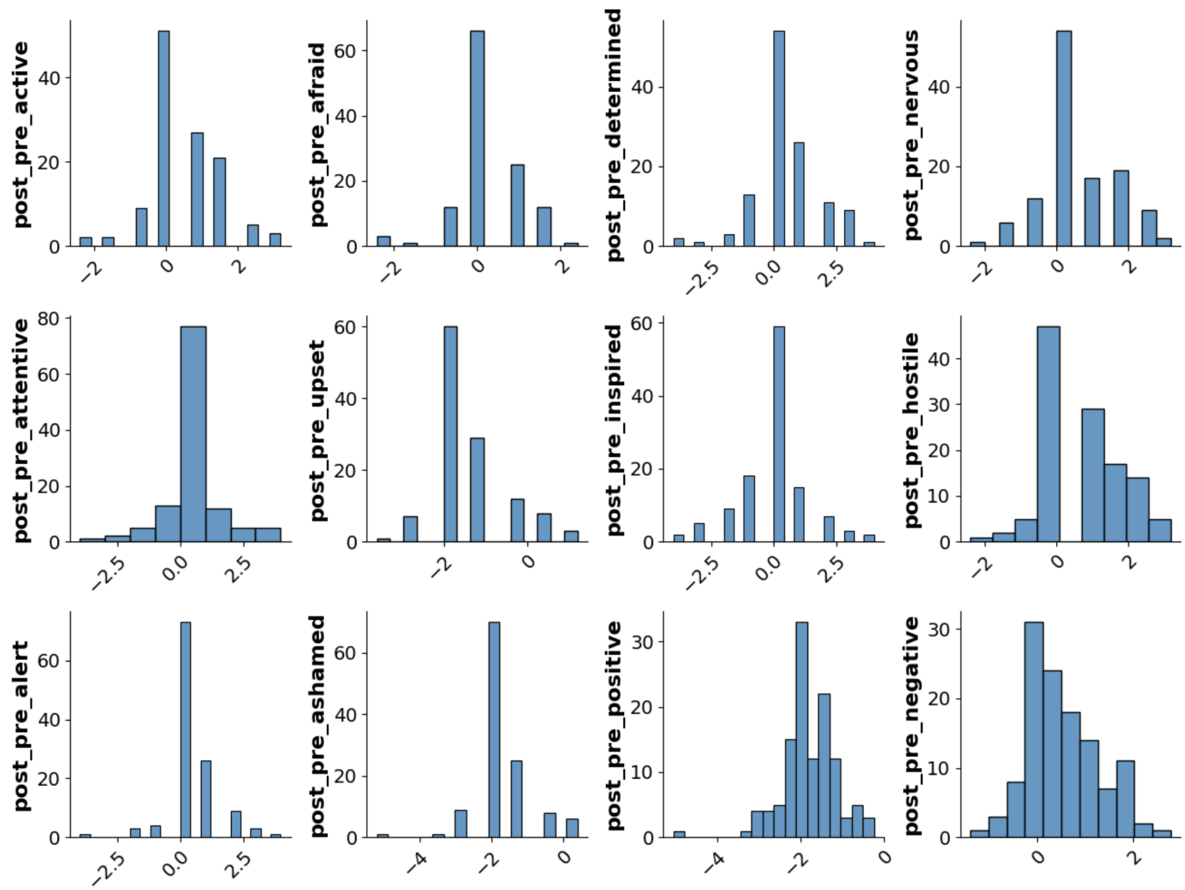

Fig. 5. Frequency histograms for the differences of the post and pre-study affect scores

Table 1. Two-way ANOVA results for crowd disposition and perceived agency on post and pre-study participant affect differences. Statistically significant factors are indicated by **, with $p < 0.01$ and * with $p < 0.5$.

Parameter	Model term	n	F value	p-value	Adjusted p-value
post_pre_afraid	crowd disposition	120	4.545	0.035	0.274
post_pre_upset	crowd disposition	120	4.474	0.037	0.274
post_pre_hostile*	crowd disposition	120	12.865	0.0005	0.015
post_pre_ashamed*	crowd disposition	120	9.901	0.002	0.031
post_pre_negative**	crowd disposition	120	13.954	0.0003	0.002

Table 2 summarizes the significant results after ANOVA and FDR correction. Figure 8 illustrates the comparative distributions of behavior parameters across the study conditions.

3.4.2 The Effects of Demographics on Participant Emotions and Behaviors. From the demographic parameters, we discarded ethnicity due to its disproportionate distribution. Additionally, a least-squares linear regression to find

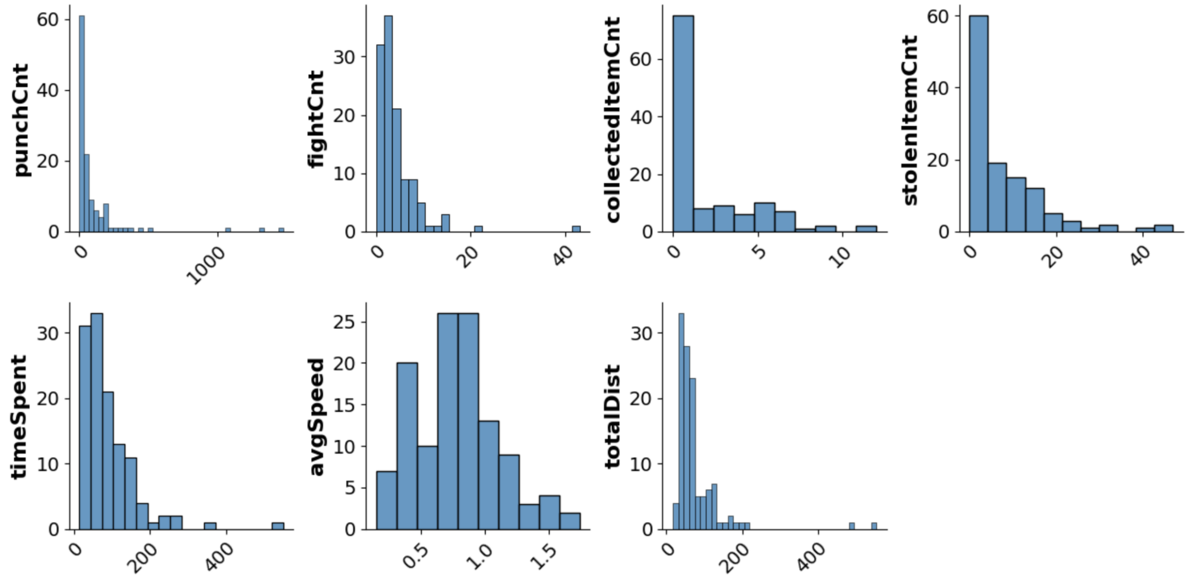

Fig. 6. Frequency histograms of participant behavior parameters

Table 2. Two-way ANOVA results for crowd disposition and perceived agency on participant behavior parameters. Statistically significant factors are indicated by **, with $p < 0.01$ and * with $p < 0.5$.

Parameter	Model Term	n	F value	p-value	Adjusted p-value
fightCnt*	crowd disposition	120	9.405	0.003	0.028
collectedItemCnt**	crowd disposition	120	15.917	0.0001	0.002
timeSpent	crowd disposition	120	5.299	0.023	0.162
totalDist	crowd disposition	120	4.696	0.032	0.169

the effects of age, a continuous variable, on emotional changes did not return any statistically significant effects. As nominal and ordinal variables respectively, we assessed the effect of gender and familiarity on participant emotions using a two-way ANOVA. Familiarity was recorded on a 7-point Likert scale ranging from “not at all” to “highly familiar”. We binned familiarity into two groups low (n = 32) and high (n = 88), corresponding to the ranges [-3,0] and [1, 3]. Thus, we treated it as a binary variable with sufficient sample sizes per category and explored its interaction effects with gender.

We ran a 2x2 ANOVA to test the effects of gender and familiarity on the differences in participants’ emotions after and before the study. The ANOVA returned a statistically significant main effect of familiarity on the differential affect scores of *afraid* with $F(1, 116) = 9.158, p = 0.003$ and *ashamed* with $F(1, 116) = 6.810, p = 0.01$. There were interaction effects of gender and familiarity on the differential affect scores of *attentive* with $F(1, 116) = 6.194, p = 0.014$. However, we did not find any statistically significant effects after BH correction. We also tested the composite affect scores, where familiarity yielded a significant main effect on differential negative emotions $F = (1, 116) = 5.088, p = 0.026$; familiarity and gender yielded a significant interaction effect on differential positive emotions $F = (1, 116) = 4.495, p = 0.036$. Again, after correcting for FDR, we did not find any significant factors. Table 3 summarizes the significant results after ANOVA and FDR correction.

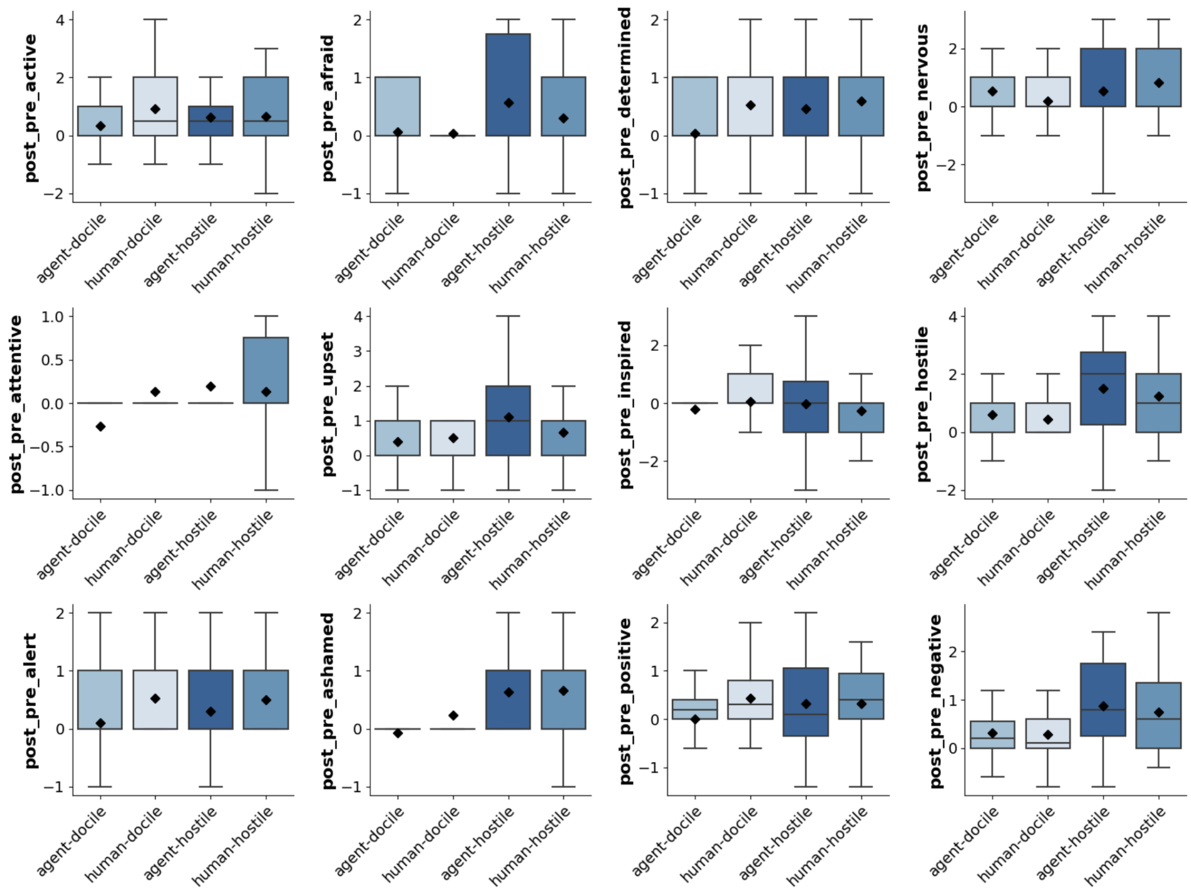

Fig. 7. Box plot diagrams for the differences between the post and pre-study affect scores by study condition

Figure 9 shows the distributions of affect score differences for gender and familiarity, respectively.

Table 3. Two-way ANOVA results for gender and familiarity on participant post and pre-study participant affect differences.

Parameter	Model Term	F value	p-value	Adjusted p-value
post_pre_afraid	familiarity	9.158	0.003	0.092
post_pre_attentive	gender: familiarity	6.194	0.014	0.142
post_pre_ashamed	familiarity	6.810	0.010	0.142
post_pre_positive	gender:familiarity	4.495	0.036	0.108
post_pre_negative	familiarity	5.088	0.026	0.108

We ran a 2x2 independent subjects ANOVA to test the effects of gender and familiarity on participant behavior parameters. ANOVA returned a statistically significant main effect of familiarity on *fightCnt* with $F(1, 116) = 8.677, p = 0.004$, *timeSpent* with $F(1, 116) = 13.423, p = 0.0004$, and *avgSpeed* with $F(1, 116) = 9.886, p = 0.002$. We also found an interaction effect of gender and familiarity on *totalDist* with $F(1, 116) = 3.961, p = 0.049$. BH

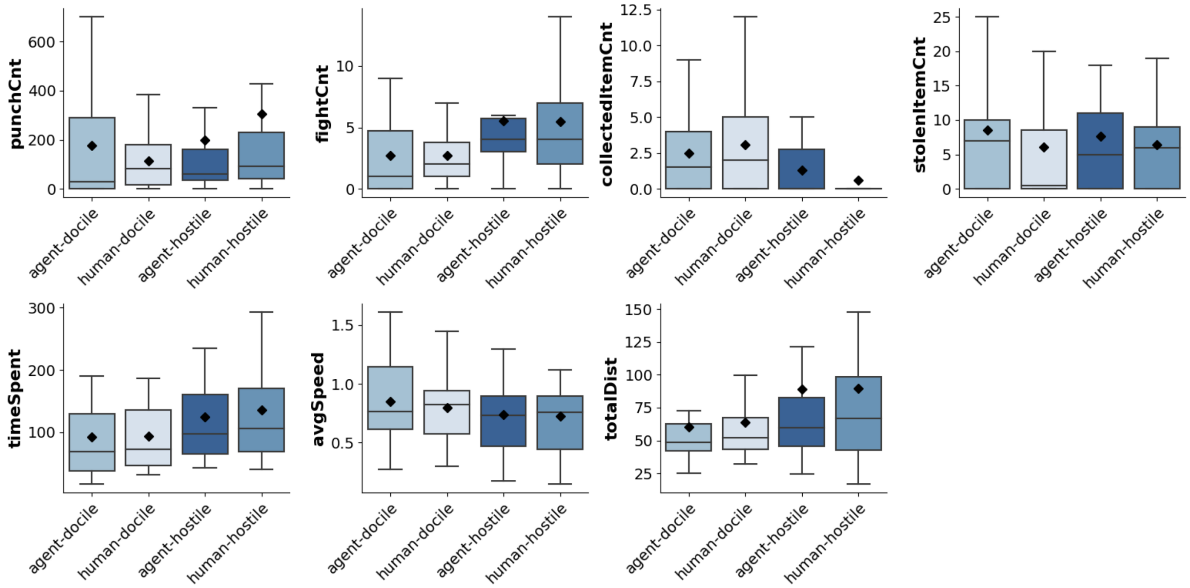

Fig. 8. Box plot diagrams for the participant behaviors by study condition

correction yielded only the main effects to be statistically significant. Table 4 summarizes the significant results after ANOVA and FDR correction. Figure 10 depicts the distributions of behavior parameter values for gender and familiarity, respectively.

Table 4. Two-way ANOVA results for gender and familiarity on participant behaviors. Statistically significant factors are indicated by **, with $p < 0.01$ and * with $p < 0.5$

Parameter	Model Term	F value	p-value	Adjusted p-value
fightCnt*	familiarity	8.677	0.004	0.027
timeSpent**	familiarity	13.423	0.0004	0.008
avgSpeed*	familiarity	9.886	0.002	0.022
totalDist	gender:familiarity	3.961	0.049	0.257

3.4.3 Correlations of Participant Personality Factors with Their Emotions and Behaviors . We computed the participant personality scores for the Five-Factor personality model, which defines personality on five orthogonal dimensions of (O)openness, (C)conscientiousness, (E)xtroversion, (A)greeableness, and (N)euroticism. Data was collected for each personality dimension on a 5-point scale where zero indicated “strongly disagree”, and four indicated “strongly agree”. We computed the responses per participant according to the TIPI guidelines. For instance, the agreeableness score was computed by taking the average of the responses for the question "I see myself as *sympathetic, warm*." and the score for "I see myself as *critical, quarrelsome*" subtracted from four.

We performed Spearman rank-order correlations between personality factors and participants’ affect score differences between post and pre-study responses. To control for Type I errors, we performed FDR correction with BH procedure. We evaluated composite positive and negative affect scores separately due to their dependence on

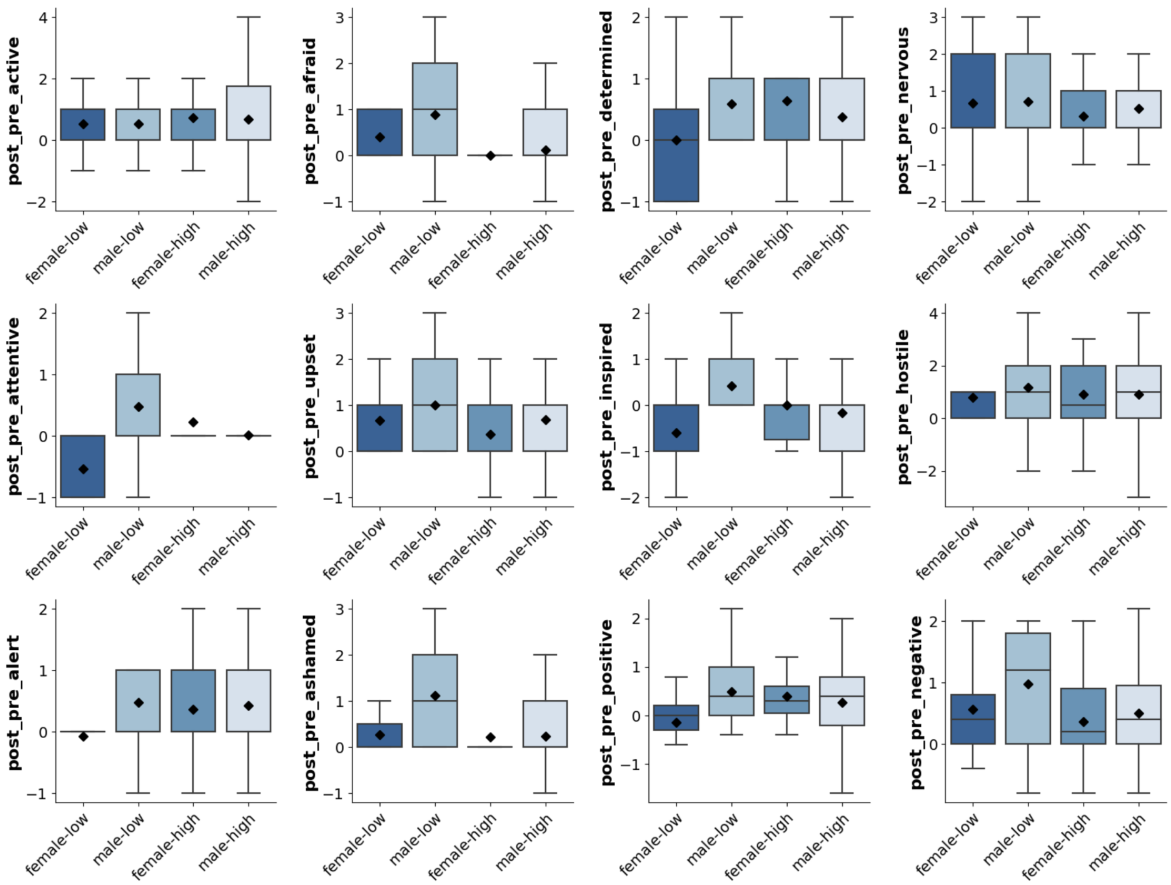

Fig. 9. Box plot diagrams for the differences between the post and pre-study affect scores by gender and familiarity

other factors, again correcting for FDR. After FDR correction, the results indicate a moderate negative correlation between extroversion and the differential score of *active* emotion ($\rho = -0.456, p < 0.001$), a low positive correlation between neuroticism and the change in *active* emotion ($\rho = -0.305, p < 0.001$), a low negative correlation between extroversion and the change in overall positive affect score ($\rho = -0.29, p = 0.0013$), and a low negative correlation between conscientiousness and the change in overall positive affect score ($\rho = -0.25, p = 0.006$). Figure 11 (a) shows the correlations between participants' differential affect scores and their personality traits, where the statistically significant correlations are highlighted with their corresponding adjusted p-values.

To further explore the relationship between initial emotions and personality scores, we also calculated Spearman rank-order correlations between personality factors and participants' pre-study affect scores and applied FDR correction with BH procedure. We observe slight to moderate positive correlations between positive affect scores (individual and composite) and the personality factors of openness, conscientiousness, extroversion, and agreeableness. In contrast, we find negative correlations between negative affect scores and neuroticism. Figure 11 (b) shows the correlations between pre-study affect scores and personality traits. The statistically significant correlations are highlighted with their corresponding adjusted p-values.

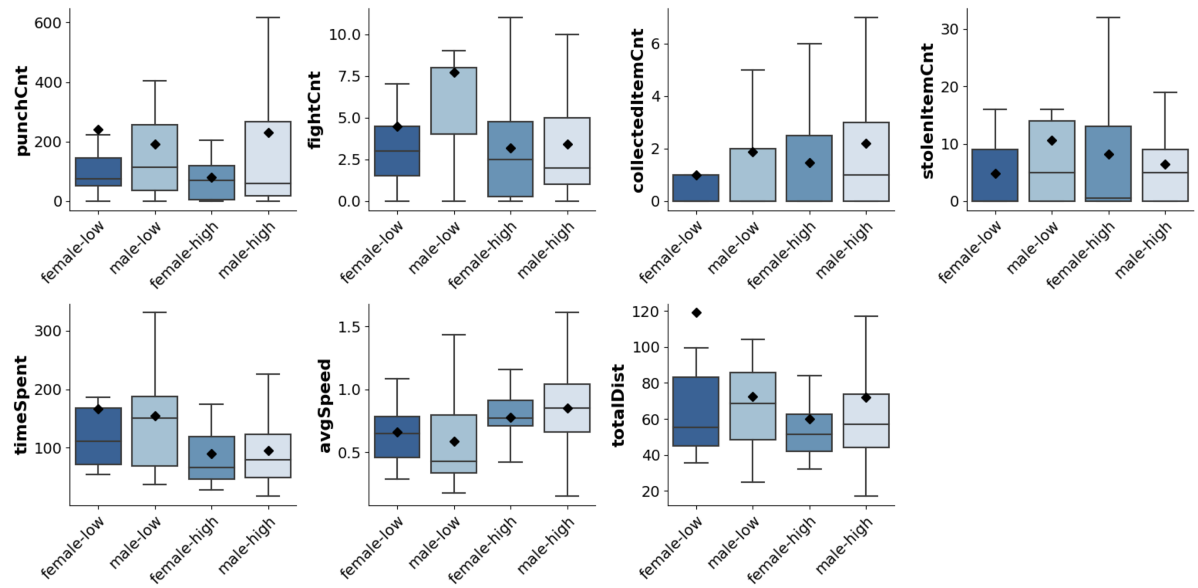

Fig. 10. Box plot diagrams for the participant behaviors by gender and familiarity

Although Spearman rank-order correlations between personality factors and participant behavior parameters returned a low positive correlation between neuroticism and *punchCnt* ($\rho = 0.202, p = 0.027$), openness and *timeSpent* ($\rho = 0.187, p = 0.041$), after FDR correction using the BH, none of these values remained statistically significant. Figure 12 shows the correlation matrix of personality factors and behaviors, where statistically significant correlations are highlighted with their corresponding adjusted p-value.

3.4.4 Post-Study Survey. The mean ratings for q1 and q5 were $\mu = 3.306 \pm 1.952$ and $\mu = 2.953 \pm 1.731$, respectively. These indicate that the participants had neutral experiences and found the character communication to be neutral. Most participants found the characters consistent and predictable.

All participants stated that they never suspected the agency when they were told that others were artificial agents.

More people believed the characters to be genuinely human-controlled in the hostile condition than the docile. Around 40% explicitly stated that they never suspected them to be artificial and 16% stated they were suspicious when agents did not attack others with more iPads or when they quickly gathered so many iPads. Others who did not believe they were avatar-controlled, mentioned robotic movements, the improbability of having the simulation start simultaneously on MTurk, and the game-like nature of the study. In the docile-human condition, only 25% believed others were humans.

When the crowd was hostile, some of the adjectives to describe the personality of the crowd included 'greedy', 'rude', 'competitive', 'aggressive', 'hostile', 'primordial', and 'selfish'. In the docile condition, some also described the crowd as 'competitive', 'greedy', and 'rude,' but there were also mentions of '(very) friendly', 'stoic', 'passive', 'uninterested', and 'cooperative'.

An interesting response for the docile-human condition, where the participant believed others were humans was "I felt pretty ashamed because I just started swinging, trying to get what I needed. I felt like I didn't have

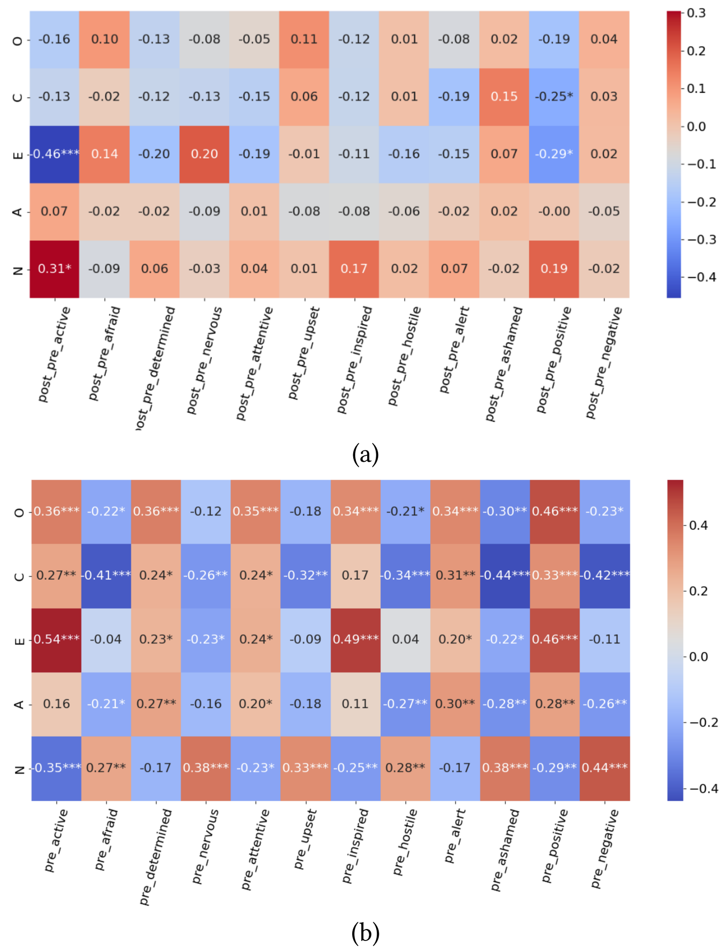

Fig. 11. Spearman correlations between participant personality and emotional changes. Statistically significant factors after FDR correction are indicated by **, with $p < 0.01$ and * with $p < 0.5$.

much of a choice because of the other players. It seemed like one of those Black Friday videos but happening in a virtual world. It was crazy."

4 DISCUSSION

RQ1 explores the effect of the crowd's emotional disposition on participant emotions and behaviors. We calculated the differences between the post and pre-study affect scores of participants to understand the emotional changes. Thus, the analyses compare these differential scores rather than the absolute emotions before or after the study. In the hostile crowd condition, there was a higher increase in the scores of *hostility*, *shame*, and *overall negative emotions* after the study, compared to the docile crowd condition. Additionally, the hostile condition yielded a slightly higher increase in *fear* and *upset* although the results did not remain statistically significant after the FDR correction. However, all these negative emotions have likely contributed to the combined negative affect score. There were no statistically significant differences between conditions for the changes in positive emotions.

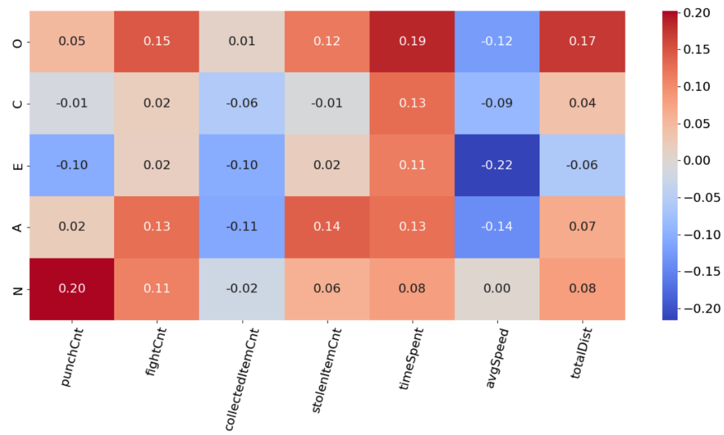

Fig. 12. Spearman correlations between participant personality and behavior parameters.

The results for emotions are in line with the participant behavior differences, showing a higher fight count in the hostile crowd than in the docile crowd condition. Yet, there was no significant difference in the number of punches they administered. Participants were attacked by crowd agents in the hostile setting. So, they had to engage in fights even when they did not initiate them. The lack of a significant difference in the number of punches between the two conditions indicates that participants generally responded to fights only when needed, which is also supported by their responses to open-ended questions. The number of collected iPads was lower in the hostile condition. This is also an expected result of the simulation setting as the hostile agents were able to attack the participants and steal their iPads. The participants spent slightly longer time and traveled longer distances in the store in the hostile setting. Regardless, reaching a definitive conclusion is challenging because these results did not retain statistical significance after the FDR correction. We can speculate that some participants might have extended their stay to recover the iPads that were stolen from them.

The results indicate that by imbuing agents with emotional behaviors dictated by personality factors, we can control users' behaviors and affective states. The increase in negative emotions, especially hostility, affirms the effectiveness of the simulation conditions in eliciting emotional responses.

In response to RQ2, we did not find any statistically significant effects of agency on human emotions or behaviors. The answers to open-ended questions hint at the main reason being the failure of the instructions and simulation to convince the participants that the agents were human-controlled. Although a considerable number of participants believed that narrative, others were skeptical, as a multi-user study that starts simultaneously on all computers is not usual on a crowdsourcing platform. Considering that the participants were selected from "Masters", i.e., people with high experience, deceiving them was especially challenging. Participants in the hostile crowd condition had a higher belief in the stated agency, possibly because the agents' actions were less predictable. The literature has mixed findings on the impact of perceived agency. Therefore, other factors than the participants' beliefs may have played a role in the lack of a difference between the two conditions of avatar vs. agent. An explanation might be the effect of the anonymity on MTurk. Because the experiment was short and did not involve further interaction other than the simulation, the participants may have distanced themselves from others, depersonalizing them, regardless of their agency [Chen and Dang 2022]. Another explanation may simply be ethopoeia, as suggested by Nass and Moon [Nass and Moon 2000].

Regarding RQ3, we did not find any significant effects of age and gender on participants' emotions or behaviors. Although high familiarity suggested higher increases in *fear*, *shame*, and *negative emotions*, the effects were not found to be statistically significant after controlling for Type I errors. Regarding behaviors, familiarity was found to be related to a lower number of fights and the time spent, and an increased average speed with statistically significant effects after FDR correction. Having more experienced participants perform the study in a shorter time and with higher speed is an expected outcome. The lower number of fights was likely a result of the shorter time spent in the simulation. The lack of a significant effect of familiarity on punch count can be attributed to more effective fights for more familiar participants.

RQ4 explored potential correlations between the dependent variables and participants' personality traits. Although no significant correlations were observed after applying the FDR correction, we found slight positive correlations between neuroticism and punch count, and openness and time spent. Neuroticism is associated with anger, nervousness, and emotionality [Goldberg 1992] which could explain the higher punch count. Similarly, openness is described with curiosity and inquisitiveness [Goldberg 1990b], which could indicate the increased time in the environment. As these results are in line with the literature [Durupinar et al. 2011], we postulate that significant effects might be achieved with a larger sample size.

Although our main focus was the change in emotional states, we examined correlations between the participants' emotional states before the study and their personality traits to test the coherence of the collected data with the literature. Openness, conscientiousness, extroversion, and agreeableness, which can be considered the personality factors with "positive" connotations, were correlated with the participants' initial positive affect scores, i.e. being *alert*, *determined*, *attentive*, and *inspired*. In contrast, neuroticism, which is described as a tendency to experience negative affect, was correlated with the initial scores of *afraid*, *nervous*, *upset*, *hostile*, and *ashamed*. Those affinities are in alignment with the trait descriptions of the personality factors in the literature [Watson and Clark 1992; Watson et al. 1992].

We found statistically significant correlations between some differential affect scores and personality traits. For instance, extroversion was negatively correlated and neuroticism was positively correlated with the increase of *active* emotion. In other words, higher extroversion and emotional stability (low neuroticism) scores were linked to higher decreases in feeling active after the study. In addition, higher extroversion and conscientiousness were found to be correlated with higher decreases in positive affect scores. An explanation could be about the computer-based nature of the study, leading to lower activity and less excitement for extroverted and emotionally stable individuals. The unpredictable nature of the simulation might have caused higher distress in the more conscientious individuals. Although the correlations between individual positive emotions and conscientiousness were not statistically significant, their combination was. In addition, all the positive emotions, i.e., *activity*, *determination*, *attentiveness*, *inspiration*, and *alertness* were negatively correlated with conscientiousness.

5 CONCLUSION AND FUTURE WORK

This paper presents a web-based study to explore the effects of perceived agency and emotional disposition of virtual crowds on a human participant embodied as a crowd member. The system recorded each participant's actions during the simulation. We also collected data about their personalities and emotional states before and after the study, in addition to their demographics. We found statistically significant effects of crowd disposition and familiarity with first-person 3D games on certain behavioral parameters and differential emotion scores but did not observe any effects of perceived agency.

The findings provide a foundation and direction for the next phase of crowd simulation research. By incorporating the insights gained from these results, researchers can work towards more detailed, realistic, and adaptive virtual environments that mirror the intricacies of human behavior and decision-making in crowds. The clear impact of crowd disposition on individual emotions and behaviors emphasizes the promise of incorporating emotional temperaments in simulations. The absence of significant effects of perceived agency suggests that the distinction between avatars and agents might not be crucial in certain contexts. However, it also opens the door for more detailed research, perhaps exploring subtler nuances or different scenarios where perceived agency might have a more pronounced impact. Because not all the participants believed that the crowd was human-controlled, a more convincing study needs to be performed to reach a conclusion about the lack of significant effects for perceived agency. More diverse animations, perhaps with deliberately erratic behaviors, and a more convincing platform for multi-user studies can provide different results.

The findings about the effects of participant familiarity and personality on their behaviors and emotions can guide researchers and animators to design personalized simulations based on user profiles. This can increase the accuracy of predicted simulation effects and provide customized virtual experiences.

Our future plans include repeating the study in a VR environment. Although VR headsets provide higher immersion, their usage has limitations in terms of efficiency, diversity of the participant pool, and participant comfort. Thus, we preferred a browser-based environment instead of a lab-based VR study for this exploratory study. We expect to gather more detailed data and pronounced differences in a more immersive setting. Such a setting will also permit recording physiological responses in addition to self-reported measures. Additionally, we will explore different scenarios that analyze the attitudes of people toward virtual agents with diverse personalities and behaviors. Our study involved a single user but the platform can easily be extended to incorporate multiple participants, allowing the design of more variable scenarios.

REFERENCES

Jana Appel, Astrid von der Pütten, Nicole C Krämer, and Jonathan Gratch. 2012. Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. *Advances in Human-Computer Interaction* 2012 (2012), 13–13.

Jim Blascovich, Jack Loomis, Andrew C Beall, Kimberly R Swinth, Crystal L Hoyt, and Jeremy N Bailenson. 2002. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological inquiry* 13, 2 (2002), 103–124.

Dario Bombari, Marianne Schmid Mast, Elena Canadas, and Manuel Bachmann. 2015. Studying social interactions through immersive virtual environment technology: virtues, pitfalls, and future challenges. *Frontiers in psychology* 6 (2015), 869.

Julien Bruneau, Anne-Helene Olivier, and Julien Pettre. 2015. Going through, going around: A study on individual avoidance of groups. *IEEE transactions on visualization and computer graphics* 21, 4 (2015), 520–528.

Alica Bucher and Andreas Voss. 2019. Judging the mood of the crowd: Attention is focused on happy faces. *Emotion* 19, 6 (2019), 1044.

Gewei Chen and Jianning Dang. 2022. Crowds' malice behind the screen: The normative influences of online dehumanization on discrimination against foreigners. *Group Processes & Intergroup Relations* (2022), 13684302221127228.

Chris Christou, Kyriakos Herakleous, Aimilia Tzanavari, and Charalambos Poullis. 2015. Psychophysiological responses to virtual crowds: implications for wearable computing. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 35–41.

Aline W. de Borst and Beatrice de Gelder. 2015. Is it the real deal? Perception of virtual characters versus humans: an affective cognitive neuroscience perspective. *Frontiers in Psychology* 6 (may 2015). https://doi.org/10.3389/fpsyg.2015.00576

Patrick Dickinson, Kathrin Gerling, Kieran Hicks, John Murray, John Shearer, and Jacob Greenwood. 2019. Virtual reality crowd simulation: effects of agent density on user experience and behaviour. *Virtual Reality* 23 (2019), 19–32.

Funda Durupinar, Uğur Güdükbay, Aytek Aman, and Norman I Badler. 2016a. Psychological Parameters for Crowd Simulation: From Audiences to Mobs. *IEEE Trans. Vis. Comp. Graph.* 22, 9 (2016), 2145–2159.

F Durupinar, N Pelechano, J M Allbeck, U Gudukbay, and N I Badler. 2011. How the Ocean personality model affects the perception of crowds. *IEEE Computer Graphics and Applications* 31, 3 (2011), 22–31.

Funda Durupinar, Kuan Wang, Ani Nenkova, and Norman I. Badler. 2016b. An Environment for Transforming Game Character Animations Based on Nationality and Profession Personality Stereotypes. In *The Twelfth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-16)*.

Jesse Fox, Sun Joo Ahn, Joris H Janssen, Leo Yeykelis, Kathryn Y Segovia, and Jeremy N Bailenson. 2015. Avatars versus agents: a meta-analysis quantifying the effect of agency on social influence. *Human–Computer Interaction* 30, 5 (2015), 401–432.

Brian J Gajadhar, Yvonne AW De Kort, and Wijnand A Ijsselsteijn. 2008. Shared fun is doubled fun: player enjoyment as a function of social setting. In *Fun and Games: Second International Conference, Eindhoven, The Netherlands, October 20-21, 2008. Proceedings*. Springer, 106–117.

Lewis R. Goldberg. 1990a. An Alternative "Description of Personality": The Big-Five Factor Structure. *Journal of Personality and Social Psychology* 59 (1990), 1216–1229.

Lewis R Goldberg. 1990b. An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology* 59 (1990), 1216–1229.

Lewis R Goldberg. 1992. The development of markers for the Big-Five factor structure. *Psychological assessment* 4, 1 (1992), 26.

Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann. 2003. A very brief measure of the big-five personality. *Journal of Research in Personality* 36, 6 (2003), 504–528.

Rosanna E Guadagno, Jim Blascovich, Jeremy N Bailenson, and Cade McCall. 2007. Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology* 10, 1 (2007), 1–22.

Christine H Hansen and Ranald D Hansen. 1988. Finding the face in the crowd: an anger superiority effect. *Journal of personality and social psychology* 54, 6 (1988), 917.

Yuanyuan Jiang, Elizabeth E O'neal, Junghum Paul Yon, Luke Franzen, Pooya Rahimian, Jodie M Plumert, and Joseph K Kearney. 2018. Acting together: Joint pedestrian road crossing in an immersive virtual environment. *ACM Transactions on Applied Perception (TAP)* 15, 2 (2018), 1–13.

Jahanvash Karim, Robert Weisz, and Shafiq Ur Rehman. 2011. International positive and negative affect schedule short-form (I-PANAS-SF): Testing for factorial invariance across cultures. *Procedia - Social and Behavioral Sciences* 15 (2011), 2016–2022.

Sujeong Kim, Aniket Bera, Andrew Best, Rohan Chabra, and Dinesh Manocha. 2016. Interactive and adaptive data-driven crowd simulation. In *2016 IEEE Virtual Reality (VR)*. IEEE, 29–38.

Alexandros Koilias, Michael Nelson, Sahana Gubbi, Christos Mousas, and Christos-Nikolaos Anagnostopoulos. 2020b. Evaluating human movement coordination during immersive walking in a virtual crowd. *Behavioral Sciences* 10, 9 (2020), 130.

Alexandros Koilias, Michael G Nelson, Christos-Nikolaos Anagnostopoulos, and Christos Mousas. 2020a. Immersive walking in a virtual crowd: The effects of the density, speed, and direction of a virtual crowd on human movement behavior. *Computer Animation and Virtual Worlds* 31, 6 (2020), e1928.

Oswald D Kothgassner, Mirjam Griesinger, Kathrin Kettner, Katja Wayan, Sabine Völkl-Kernstock, Helmut Hlavacs, Leon Beutl, and Anna Felnhofer. 2017. Real-life prosocial behavior decreases after being socially excluded by avatars, not agents. *Computers in human behavior* 70 (2017), 261–269.

Marios Kyriakou, Xueni Pan, and Yiorgos Chrysanthou. 2017. Interaction with virtual crowd in Immersive and semi-Immersive Virtual Reality systems. *Computer Animation and Virtual Worlds* 28, 5 (2017), e1729.

Sohye Lim and Byron Reeves. 2010. Computer agents versus avatars: Responses to interactive game characters controlled by a computer or other player. *International Journal of Human-Computer Studies* 68, 1-2 (2010), 57–68.

Joan Llobera, Bernhard Spanlang, Giulio Ruffini, and Mel Slater. 2010. Proxemics with multiple dynamic characters in an immersive virtual environment. *ACM Transactions on Applied Perception (TAP)* 8, 1 (2010), 1–12.

Jack M Loomis, James J Blascovich, and Andrew C Beall. 1999. Immersive virtual environment technology as a basic research tool in psychology. *Behavior research methods, instruments, & computers* 31, 4 (1999), 557–564.

Diana Mihalache, Sarah Ariel Lamer, Josh Allen, Mackenzie Maher, and Timothy D Sweeny. 2021. Anger bias in the evaluation of crowds. *Journal of Experimental Psychology: General* 150, 9 (2021), 1870.

Mehdi Moussaïd, Mubbasir Kapadia, Tyler Thrash, Robert W Sumner, Markus Gross, Dirk Helbing, and Christoph Hölscher. 2016. Crowd behaviour during high-stress evacuations in an immersive virtual environment. *Journal of The Royal Society Interface* 13, 122 (2016), 20160414.

Sahil Narang, Andrew Best, Tanmay Randhavane, Ari Shapiro, and Dinesh Manocha. 2016. PedVR: Simulating Gaze-Based Interactions between a Real User and Virtual Crowds. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (Munich, Germany) *(VRST '16)*. Association for Computing Machinery, New York, NY, USA, 91–100. https://doi.org/10.1145/2993369.2993378

Clifford Nass and Youngme Moon. 2000. Machines and mindlessness: Social responses to computers. *Journal of social issues* 56, 1 (2000), 81–103.

Clifford Nass and Byron Reeves. 2003. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.* CTR FOR STUDY OF LANG & INFO. https://www.ebook.de/de/product/3743492/clifford_nass_byron_reeves_the_media_equation_how_people_treat_computers_television_and_new_media_like_real_people_and_places.html

Michael Nelson, Alexadros Koilias, Sahana Gubbi, and Christos Mousas. 2019. Within a virtual crowd: Exploring human movement behavior during immersive virtual crowd interaction. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*. 1–10.

Michael G Nelson, Angshuman Mazumdar, Saad Jamal, Yingjie Chen, and Christos Mousas. 2020. Walking in a crowd full of virtual characters: Effects of virtual character appearance on human movement behavior. In *Advances in Visual Computing: 15th International Symposium, ISVC 2020, San Diego, CA, USA, October 5–7, 2020, Proceedings, Part I 15*. Springer, 617–629.

Kristine L Nowak. 2004. The influence of anthropomorphism and agency on social judgment in virtual environments. *Journal of Computer-Mediated Communication* 9, 2 (2004), JCMC925.

Catherine S Oh, Jeremy N Bailenson, and Gregory F Welch. 2018. A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI* 5 (2018), 409295.

Anne-Hélène Olivier, Julien Bruneau, Richard Kulpa, and Julien Pettré. 2017. Walking with virtual people: Evaluation of locomotion interfaces in dynamic environments. *IEEE transactions on visualization and computer graphics* 24, 7 (2017), 2251–2263.

N. Pelechano and J. M. Allbeck. 2016. Feeling crowded yet?: crowd simulations for VR. In *2016 IEEE Virtual Humans and Crowds for Immersive Environments (VHCIE)*. 17–21.

Nuria Pelechano, Catherine Stocker, Jan Allbeck, and Norman Badler. 2008. Being a Part of the Crowd: Towards Validating VR Crowds Using Presence. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1* (Estoril, Portugal) *(AAMAS '08)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 136–142.

Killian Poinsot, Geoffrey Gorisse, Olivier Christmann, Sylvain Fleury, Simon Richir, et al. 2022. Effect of Social Actors Perceived Agency on Social Presence in Computer-Mediated Communication. *Advances in Human-Computer Interaction* 2022 (2022).

Mariachiara Rapuano, Tina Iachini, and Gennaro Ruggiero. 2023. Interaction with virtual humans and effect of emotional expressions: anger matters! *Journal of clinical medicine* 12, 4 (2023), 1339.

G Ruggiero, M Rapuano, Alice Cartaud, Yann Coello, and Tina Iachini. 2021. Defensive functions provoke similar psychophysiological reactions in reaching and comfort spaces. *Scientific reports* 11, 1 (2021), 5170.

Mel Slater, Cristina Gonzalez-Liencres, Patrick Haggard, Charlotte Vinkers, Rebecca Gregory-Clarke, Steve Jelley, Zillah Watson, Graham Breen, Raz Schwarz, William Steptoe, et al. 2020. The ethics of realism in virtual and augmented reality. *Frontiers in Virtual Reality* 1 (2020), 1.

Mel Slater and Maria V Sanchez-Vives. 2016. Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI* 3 (2016), 74.

Matias Volonte, Yu-Chun Hsu, Kuan-Yu Liu, Joe P Mazer, Sai-Keung Wong, and Sabarish V Babu. 2020. Effects of interacting with a crowd of emotional virtual humans on users' affective and non-verbal behaviors. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 293–302.

Astrid M Von der Pütten, Nicole C Krämer, Jonathan Gratch, and Sin-Hwa Kang. 2010. "It doesn't matter what you are!" Explaining social effects of agents and avatars. *Computers in Human Behavior* 26, 6 (2010), 1641–1650.

David Watson and Lee Anna Clark. 1992. On traits and temperament: General and specific factors of emotional experience and their relation to the five-factor model. *Journal of personality* 60, 2 (1992), 441–476.

David Watson, Lee Anna Clark, Curtis W McIntyre, and Stacy Hamaker. 1992. Affect, personality, and social activity. *Journal of personality and social psychology* 63, 6 (1992), 1011.

David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54, 6 (1988), 1063.

David Weibel, Bartholomäus Wissmath, Stephan Habegger, Yves Steiner, and Rudolf Groner. 2008. Playing online games against computer-vs. human-controlled opponents: Effects on presence, flow, and enjoyment. *Computers in human behavior* 24, 5 (2008), 2274–2291.

Russell B Williams and Caryl A Clippinger. 2002. Aggression, competition and computer games: computer and human opponents. *Computers in human behavior* 18, 5 (2002), 495–506.