Social Presence as a Moderator of the Effect of Agent Behavior on Emotional Experience in Social Interactions in Virtual Reality

Michael Pfaller[1,], Leon O. H. Kroczek[1], Bastian Lange[2], Raymund Fülöp[3], Mathias Müller[2] and Andreas Mühlberger[1]*

[1]Department of Psychology, Clinical Psychology and Psychotherapy, University of Regensburg, Regensburg, Germany, [2]Department of Research and Development, VTplus GmbH, Würzburg, Germany, [3]Institute for Information and Media, Language and Culture, Chair of Information Science, University of Regensburg, Regensburg, Germany

Background: Exposure therapy involves exposure to feared stimuli and is considered to be the gold-standard treatment for anxiety disorders. While its application in Virtual Reality (VR) has been very successful for phobic disorders, the effects of exposure to virtual social stimuli in Social Anxiety Disorder are heterogeneous. This difference has been linked to demands on realism and presence, particularly social presence, as a pre-requisite in evoking emotional experiences in virtual social interactions. So far, however, the influence of social presence on emotional experience in social interactions with virtual agents remains unknown.

Objective: We investigated the relationship between realism and social presence and the moderating effect of social presence on the relationship between agent behavior and experienced emotions in virtual social interaction.

Methods: Healthy participants (N = 51) faced virtual agents showing supportive and dismissive behaviors in two virtual environments (short interactions and oral presentations). At first, participants performed five blocks of short one-on-one interactions with virtual agents (two male and two female agents per block). Secondly, participants gave five presentations in front of an audience of 16 agents. In each scenario, agent behavior was a within subjects factor, resulting in one block of neutral, two blocks of negative, and two blocks of positive agent behavior. Ratings of agent behavior (valence and realism), experience (valence and arousal), and presence (physical and social) were collected after every block. Moderator effects were investigated using mixed linear models with random intercepts. Correlations were analyzed via repeated measures correlations.

Results: Ratings of valence of agent behaviors showed reliable relationships with experienced valence and less reliable relationships with experienced arousal. These relationships were moderated by social presence in the presentation scenario. Results for the interaction scenario were weaker but potentially promising for experimental studies. Variations in social presence and realism over time were correlated but social presence proved a more reliable moderator.

Conclusion: Our findings emphasize the role of social presence for emotional experience in response to specific agent behaviors in virtual social interactions. While these findings should be replicated with experimental designs and in clinical samples, variability in social presence might account for heterogeneity in efficacy of virtual exposure to treat social anxiety disorder.

1 INTRODUCTION

1.1 Theoretical Background

Exposure Therapy is considered the gold-standard treatment for anxiety disorders, especially phobic disorders. It involves "the systematic confrontation of feared stimuli" (Steinman et al., 2016). Whereas the confrontation *in vivo* (in the real world) is highly effective for a variety of disorders, there are a number of problems which hinder its application in routine care, among them the considerable effort involved in preparing and conducting the exercises, often outside the common therapy setting (Neudeck and Einsle, 2012). These problems can be countered by the application of exposure therapy *in virtuo*, i.e., in Virtual Reality (Virtual Reality Exposure Therapy, VRET; Rothbaum et al., 1997). VRET has been proven to be as effective as *in vivo* Exposure Therapy for multiple anxiety disorders (Carl et al., 2019). With respect to Social Anxiety Disorder (SAD), however, the results have been heterogeneous. Some studies found an advantage of *in vivo* exposure (Kampmann et al., 2016), others a comparable effectiveness between *in vivo* and in virtuo exposure therapy (Chesham et al., 2018), and others still, a superiority of virtual exposure (Bouchard et al., 2017). On a meta-analytic level, Carl et al. (2019) concluded that VRET is as effective as *in vivo* exposure for SAD, whereas a recent meta-analysis, with more stringent criteria on what constitutes *in vivo* exposure, concluded that *in vivo* exposure is more effective compared to VRET for Social Phobia (Wechsler et al., 2019). Given these contrasting results, a better understanding of the underlying working mechanisms of virtual reality exposure in the case of social anxiety is clearly required.

An important mechanism underlying VRET is presence. Presence can be commonly defined along the lines of a "sense of being there in a mediated environment" (Lombard and Jones, 2015). The precise nature of the bi-directional relationship between presence in VR and experienced emotions is not fully understood and the focus of ongoing research (Bouchard et al., 2008; Diemer et al., 2015). It has been suggested that a certain amount of presence is only a precondition to experience emotion but does not influence its intensity (Felnhofer et al., 2014, 2015). Yet a common finding in VRET is a correlation between sense of presence and experienced anxiety, strongest for participants fulfilling criteria for anxiety disorders (Ling et al., 2014). Importantly, this correlation could not be established for Social Phobia (Ling et al., 2014). A specific form of presence, social presence, has commonly been defined in distinction to physical presence, among others as "the feeling of being together (and communicating) with someone" (Ijsselsteijn et al., 2000) or as "a psychological state in which virtual [...] social actors are

experienced as natural social actors" (Lee, 2004). Social presence has been rarely reported in studies on VRET for Social Phobia until recently. But social presence appears to be more strongly related to the anxiety response in social phobia than physical presence (Felnhofer et al., 2019).

A different conceptualization of presence by Slater (2009) distinguishes between place illusion (PI) and plausibility illusion (Psi). PI, the feeling of being in the virtual environment, is related to most conceptualizations of physical presence and could be sufficient to trigger fear related to places, objects, and animals. But to trigger the core fear of social anxiety (i.e., evaluation by other humans), plausibility of the displayed behaviors may be essential. According to Slater (2009), Psi "includes the notion of the credibility of events in comparison with what would be expected in reality in similar circumstances." In the Multimodal Presence Scale, a questionnaire designed to measure the three dimensions of presence defined by Lee (2004), the subscale *social presence* contains but is not limited to "human realism," which includes experiencing virtual agents as "credible" (Makransky et al., 2017). Slater (2009) states that "when Psi (plausibility illusion) breaks, it is unlikely to recover." This implies that the effect of evaluative behaviors by these virtual characters could have diminished or ceased, i.e., this perceived "reality" of the characters (or the related experiential construct social presence) would moderate between the evaluative behaviors and the experience in response to them.

The goal of the study is to show in a correlative setting that an effect of the perception of behaviors with evaluative content (by agents in VR) on corresponding experience (in valence and arousal) exists and is moderated by the perceived realism of these behaviors and the experienced social presence in these situations.

1.2 Primary Hypotheses

It is hypothesized that the perceived valence of the behavior of agents in virtual social situations evokes an experience congruent to the behavior (i.e., the lower/more negatively valenced the behavior of agents is rated, the lower/more negative are ratings of experienced valence and the higher are ratings of experienced arousal and vice versa; hypothesis 1a valence and 1b arousal). The perception of realism of agent behavior over the different behavior variants correlates with the experience of social presence (hypothesis 2). The effect of valence of agent behavior on emotional experience (valence and arousal) is moderated by perceived realism of agent behavior (hypothesis 3a valence and 3b arousal) and social presence (hypothesis 4a valence and 4b arousal).

1.3 Secondary Hypotheses - Evaluation of Scenarios

We assume that the non-neutral behavior variants were successful in depicting behaviors that were distinguishable from neutral in evoking corresponding experiences. Indicators of a successful implementation are ratings for valence of agent behavior and experienced valence and arousal in all four non-neutral conditions that differ from neutral in the expected direction (larger/more positive for supportive and interested, smaller/more negative for uninterested and dismissive). Perceptions of differences of agent behavior have to be considered as the most basic indicator, differences in experienced valence as intermediate and differences in arousal as the strongest indicator of a successful paradigm. Ideally, no differences are apparent between behavior variants in physical presence, realism of agent behavior, and social presence.

2 METHODS

2.1 Participants

In total, 52 healthy participants of both genders were recruited *via* bulletin boards and social media. Participants received 10€ per hour or course credit. Exclusion criteria were self-reported neurological or mental illnesses. One person was excluded from analysis post hoc due to self-reported mental disease. The remaining 51 participants were included in the analysis (33 female, aged between 18 and 47, $M = 23.16$, SD = 4.83). Seven participants had prior experience with giving presentations in VR. Experimental procedures were in line with the Declaration of Helsinki and the study was approved by the Ethics Committee of the University of Regensburg. All participants gave written informed consent.

2.2 Design

The study consisted of two experimental scenarios, i.e., an interaction scenario and an oral presentation scenario. Within each scenario, we manipulated the within-subject factor agent behavior by presenting five distinct blocks which differed only in the behavior that the virtual agents showed towards the participant. The implemented agent behavior variants aimed at representing *supportive*, *interested*, *neutral*, *uninterested*, and *dismissive* behavior, leading to five factor levels (the operationalization of the behavior variants are described in the procedure section of the respective scenario). All agent behavior variants were presented for every participant. To ensure comparability for evaluation of behavior variants in contrast to neutral behavior, neutral was the first variant for every participant. The order of the remaining variants was counterbalanced across participants.

2.3 Questionnaires

Physical and social presence were assessed separately for both scenarios using the Multimodal Presence Scale (MPS; Volkmann et al., 2018). Social anxiety was measured using the Social Phobia Inventory (SPIN; Stangier and Steffens, 2002), fear of negative

evaluation using the Brief Fear of Negative Evaluation questionnaire (BFNE; Leary, 1983), fear of public speaking using the Personal Report of Confidence as a Speaker questionnaire (PRCS; Paul, 1966) and submissive behavior using the Submissive Behavior Scale (SBS; Allan and Gilbert, 1997). Symptoms of VR sickness were assessed with the Virtual reality sickness questionnaire (VRSQ; Kim et al., 2018) separately for both scenarios. In the sample, SPIN ranged from 5 to 35 ($M = 17.12$, SD = 7.98), BFNE from 11 to 43 ($M = 24.65$, SD = 7.62), PRCS from 4 to 23 ($M = 13.2$, SD = 5.1) and SBS from 11 to 44 ($M = 24.4$, SD = 6.17). Results of MPS and VRSQ are displayed separately for both scenarios in **Table 1**. Custom questionnaires were used to assess demographic information (age and gender), previous experience with presentations in VR (yes/no; if answered yes, follow up questions on number of studies and topics of presentation for these studies) and self reported presence of mental illness (yes/no; if answered yes, follow up questions on diagnosis, psychological and medical treatment as yes/no).

2.4 Apparatus

The virtual environments were created by VTplus (Würzburg, Germany) using Unreal Engine (Version 4.25, Epic Games, inhouse Interface VrSessionModUDK 1.0.16) and were presented *via* the HTC Vive Pro Eye head mounted display (HTC Corporation, Taoyuan, Taiwan). Participants wore the inbuilt headphones for presentation of auditory stimuli. The sequence of statements comprising the interactions and control of behavior of virtual agents constituting the different conditions was handled by software and a graphical user interface developed in the context of the project OPTAPEB by VTplus, the Chair of Information science at the University of Regensburg and the Zentrum für Telemedizin (Bad Kissingen, Germany). Agent responses during the interactions were presented automatically using output from speech recognition software (Povey et al., 2011).
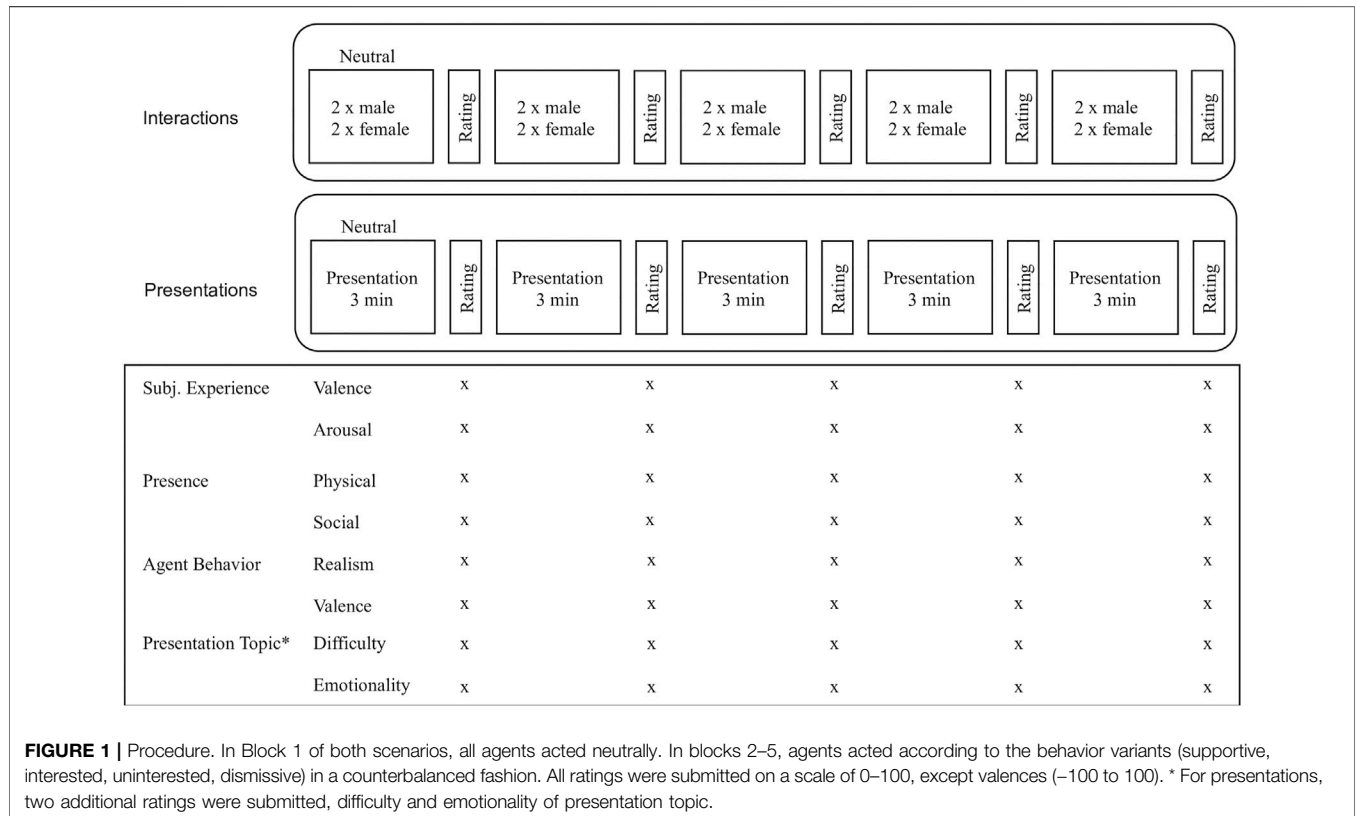
2.5 Measures

After every block of interactions and every presentation, participants rated their experience on the dimensions of valence and arousal, their sense of physical and social presence, and the realism and valence of agent behavior (see **Figure 1**). Ratings were given on a scale of 0–100 (0 representing very low or completely absent, 100 representing very high or completely), ratings of valences were coded on a scale of −100 (extremely unpleasant/extremely dismissive) to 100 (extremely pleasant/extremely supportive). The direction of valence (positive or negative) was determined by a prior question of "rather pleasant or unpleasant" or "rather supportive or dismissive" after which the corresponding numeric value was given on a scale of 0–100 and appropriately coded by the experimenter. Physical presence was phrased as "sense of being in VR" (similar to MPS question 4), social presence as "sense of interacting with interaction partners/listeners instead of with a computer simulation" (based on MPS question 10). Realism was phrased as "how realistic did you find the behavior" of agents in VR. Valence of agent behavior was phrased as "how would an objective observer probably rate the behavior" on a scale from

	Interactions		Presentations		—			
	M	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*
VRSQ	4.78	3.70	4.33	3.22	0.65	96.57	0.519	0.13
MPS Physical	16.28	4.22	17.84	3.73	−1.97	97.02	0.051	0.39
MPS Social	14.68	4.42	16.49	4.15	−2.12	98.32	0.036	0.42

VRSQ, virtual reality sickness questionnaire; MPS, multimodal presence scale.

FIGURE 1 | Procedure. In Block 1 of both scenarios, all agents acted neutrally. In blocks 2–5, agents acted according to the behavior variants (supportive, interested, uninterested, dismissive) in a counterbalanced fashion. All ratings were submitted on a scale of 0–100, except valences (−100 to 100). * For presentations, two additional ratings were submitted, difficulty and emotionality of presentation topic.

dismissive to supportive, experienced valence was phrased as "how did you experience the situation" on a scale from unpleasant to pleasant. After every presentation, additional ratings of difficulty and the emotionality of the presentation topic were given on a scale of 0–100. Measures of physiology (ECG and EDA) and behavior (Gaze, Distances in VR, Voice) were recorded but are not reported here.

2.6 Procedure

Participants signed informed consent and filled out questionnaires for demographic information and trait measures. Subsequently, electrodes for psychophysiological measurements (ECG and EDA) were attached and participants entered the VR environment. The interaction scenario was then presented (constituting five blocks of four interactions each), followed by the presentation scenario (five presentations of 3 min length plus 1 min preparation).

2.6.1 Interaction Scenario

In VR, participants found themselves in a room with eight tables and eight distinct agents (four male, four female), one agent standing at each table. The agents and their position were constant over the whole scenario. The agents at the table appeared to be preoccupied by looking down at their phones. The experimenter instructed the participant to select two agents of the same sex in succession and approach them, greet them, ask for the time, and to say goodbye. Agents reacted to the approach by looking up at a remaining distance of 1.5 m, a value determined as optimal for such a setting by a previous study (Kroczek et al., 2020). The subsequent behavior of the agent reflected one of the five behavior variants, expressed as facial expression, posture, verbal response and tone of voice. The behavior variants ranged from orienting fully towards the participant after approach, showing lots of eye contact and a clear smile

expression, statements and a tone of voice expressing enjoyment of the conversation (*supportive*) to merely turning the head initially, little eye contact, facial expressions of slight anger and contempt and short statements in a tone of voice expressing annoyance (*dismissive*). Between these two poles lay the variants *interested* (less pronounced smile), *neutral* (an absence of facial expressions or tone of voice of a particular valence, a medium level of orientation towards the participant and eye contact) and *uninterested* (slightly weaker expressions of the dismissive behaviors). After each approach block (with two men and two women each, order counterbalanced) for a specific behavior variant, the screen went dark and participants were asked to rate their experience, presence, and impressions of agent behavior. Subsequently, participants continued from the previous location in the room. In the first of the five approach blocks, behavior of all agents was neutral. The order of the remaining behavior variants was counterbalanced across all participants. See **Figure 1** for an overview of blocks and ratings during the scenario. After the last block and removal of VR equipment, participants rated their presence (MPS) and VR sickness symptoms (VRSQ) during the interaction scenario.

2.6.2 Presentation Scenario
In VR, participants found themselves in a room in front of a laptop. The topic of the next presentation was displayed on the laptop. The topics were "personal introduction" and opinion on the following topics relating to Germany: "stronger promotion of alternative energy sources," "stronger promotion of local public transport," "prohibition of plastic," and "autobahn speed limit." Introduction was always the first topic, the remaining topics were counterbalanced to be presented in the given order for one half of the participants and in reversed order for the other half. Participants were given 1 minute to prepare a 3 minute presentation on the instructed topic. At the end of the preparation period, they were guided into the presentation room by an agent. In case of technical issues they were teleported into the presentation room for all remaining presentations in order to prevent negative effects on presence. The presentation room consisted of four rows of tables with four agents each. The audience consisted of 16 distinct agents (nine male, seven female), including the eight agents from the interaction scenario. At this point, and in the first 10 s of the presentation, the audience acted neutrally, i.e., no discernible facial expressions, and some idling behavior, like twitching or hair stroking. The topic was presented on a computer screen on a table positioned slightly to the right of the center position in front of the audience and on the wall behind the position of the presenter (like a video projection). The participant was instructed to begin their 3 min presentation on the given topic. After approximately 10 s, listeners gradually started displaying the facial expressions, postures and behaviors defined by the behavior variant for this presentation. For *neutral*, agents showed no specific facial expressions or movements indicating interest or lack of interest. For *supportive*, listeners attended to the participant with a friendly expression and nodded occasionally. The

variant *interested* was similar, but the expressions were weaker and behaviors less frequent. For *uninterested*, the listeners displayed no specific facial expressions but showed behavioral expressions of boredom such as slouching and looking around the room, at their neighbors or at their phone. For the variant *dismissive*, listeners showed facial expressions of anger, contempt and disgust while looking at the participant, and behavioral expressions of discontent, such as shaking their head. After the presentation, the screen went dark and participants were asked to rate their experience, presence, impressions of agent behavior and difficulty and emotionality of the presentation topic. Subsequently, they were teleported back into the into the preparation room in front of the laptop with the topic of the next presentation. The first presentation always constituted neutral behavior, the order of the remaining behavior variants was counterbalanced across participants. See **Figure 1** for an overview of presentations and ratings. After the last block and removal of VR equipment, participants rated their presence (MPS) and VR sickness symptoms (VRSQ) during the presentation scenario. After removing electrodes, participants were debriefed. The total duration of the study was approximately 2 hours.

2.7 Data Processing
Questionnaire and rating data entry and documentation was performed in SPSS (version 25, IBM Corp, 2017). Statistical analysis and report generation were performed using R (R Core Team, 2020) and knitr (Xie, 2015), graphics were created with ggplot2 (Wickham, 2016). For moderator analyses, the investigated moderators were dichotomized at the median of the respective rating over all conditions (including neutral) for this scenario (medians for interaction scenario are: agent realism 50, social presence 50, physical presence 60; medians for presentation scenario are: agent realism 60, social presence 60, physical presence 66).

2.8 Statistical Analysis
All described analyses were performed separately for interactions and presentations. All significance tests were conducted with $\alpha = 0.05$. All ratings are assumed to represent continuous data at an interval scale. Relationships between ratings over the four non-neutral conditions were investigated using repeated measures correlations (Rmcorr; Bakdash and Marusich, 2017). Rmcorr accounts for dependence among observations by using analysis of covariance (ANCOVA) to remove variance between participants with varying intercepts. Moderator effects, i.e., how the relationship between two variables is influenced by a third variable (moderator), were tested by analyzing the interaction of the two variables in a linear regression (Baron and Kenny, 1986). The examined models are outlined in **Figure 2**. To improve interpretability of the results, the investigated moderators were dichotomized using a median split for all ratings of all participants in the given category across the four non-neutral conditions used for analysis. To account for non-independent measurement due to repeated measurements, the interactions were analyzed in a multilevel model with fixed
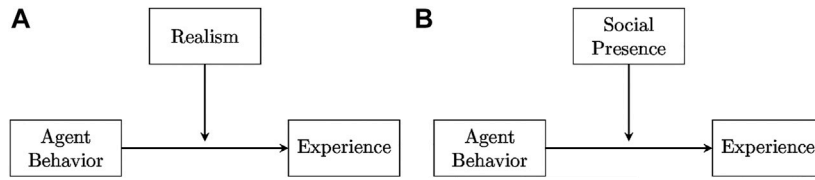
FIGURE 2 | Graphical representation of moderator analyses for **(A)** realism and **(B)** social presence. Analyses were conducted separately for both scenarios and for both aspects of experience, i.e., valence and arousal.

TABLE 2 | Interactions: Repeated measures correlations among ratings.

	Experience		Presence		Agent Beh.
	Valence	Arousal	Physical	Social	Realism
EX Arousal	−0.31***	—	—	—	—
PR Physical	0.04	0.00	—	—	—
PR Social	0.21*	0.06	0.02	—	—
AG Realism	0.41***	−0.18*	−0.05	0.26**	—
AG Valence	0.82***	−0.21*	−0.01	0.18*	0.40***

*EX, experience; PR, presence; AG, agent behavior. *p < 0.05, **p < 0.01, ***p < 0.001.*

slope but random intercept per participant. Evaluation of the scenarios was carried out through comparisons of all behavior variants to neutral condition using Welch's t-tests (Delacre et al., 2017).

3 RESULTS

3.1 Primary Hypotheses

During blocks of interactions with non-neutral behaviors of interaction partners, ratings of valence of agent behavior correlate with ratings of experienced valence, $r(149) = 0.82$, $p < 0.001$, and experienced arousal, $r(149) = −0.21$, $p = 0.011$. Ratings of realism of agent behavior correlate with Social Presence, $r(149) = 0.26$, $p = 0.001$ (full correlation matrix for interactions is presented in **Table 2**). Concerning the hypothesized moderators (realism and social presence) between valence of agent behavior and experience (valence and arousal), realism is not a significant moderator of experienced valence, $b = 0.05$, $t(147) = 0.59$, $p = 0.556$, or arousal, $b = 0.03$, $t(147) = 1.17$, $p = 0.246$. Social presence is not a significant moderator for experienced valence, $b = 0.05$,
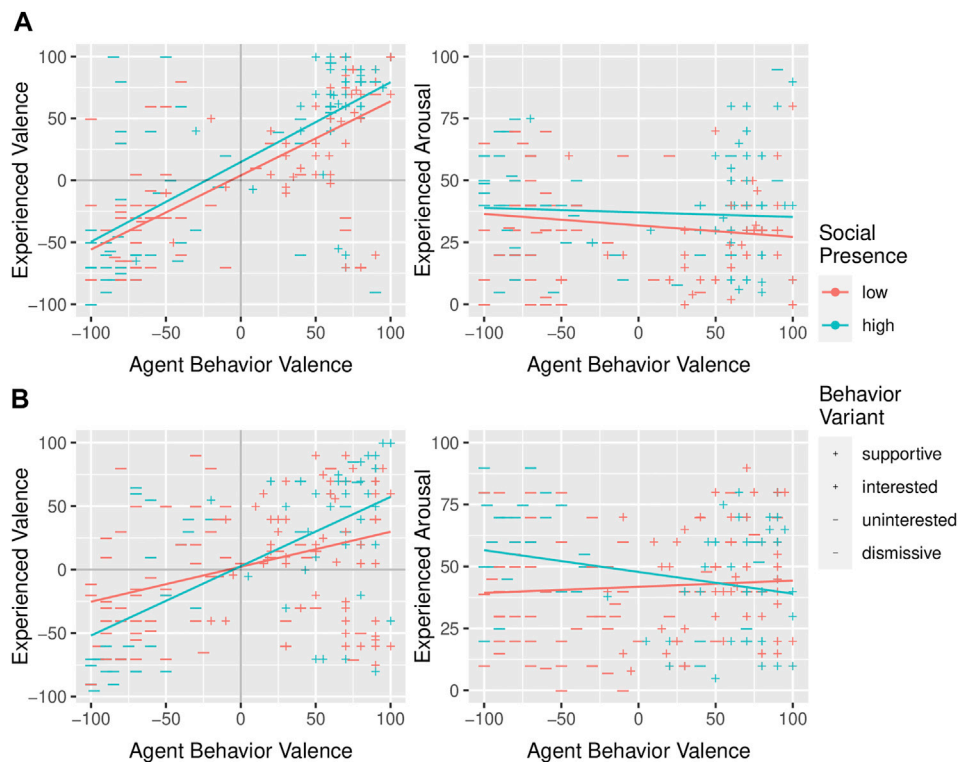

FIGURE 3 | Moderator effects of social presence: Not significant during **(A)** interactions, significant for valence and arousal during **(B)** presentations. Agent behavior variants are displayed grouped by their targeted valence as positive (+) or negative (−), neutral was omitted for moderator analysis.

TABLE 3 | Moderators between agent behavior and experience.

	b	SE	t	p
Interactions-DV: Experienced valence				
Agent realism	0.05	0.08	0.59	0.556
Presence social	0.05	0.08	0.60	0.547
Presence physical	−0.11	0.08	−1.45	0.148
Interactions-DV: Experienced arousal				
Agent realism	0.03	0.03	1.17	0.246
Presence social	0.03	0.02	1.14	0.257
Presence physical	−0.00	0.03	−0.16	0.876
Presentations-DV: Experienced valence				
Agent realism	0.21	0.09	2.29	0.024
Presence social	0.27	0.09	3.09	0.002
Presence physical	0.17	0.09	1.95	0.053
Presentations-DV: Experienced arousal				
Agent Realism	−0.04	0.03	−1.29	0.200
Presence Social	−0.11	0.03	−3.92	0.001
Presence Physical	−0.02	0.03	−0.54	0.589

Each row represents the linear mixed model involving the specified dichotomized moderator. Only the interaction term is reported for each model.

$t(147) = 0.6$, $p = 0.547$, or arousal, $b = 0.03$, $t(147) = 1.14$, $p = 0.257$. The absence of a moderator effect of social presence on experienced valence and arousal during interactions, i.e., similar slopes for high and low values of the moderator social presence, are illustrated in **Figure 3A** (see **Table 3** for full results of moderator analysis).

During presentations with non-neutral audience, ratings of valence of agent behavior correlate with ratings of experienced valence, $r(152) = 0.59$, $p < 0.001$, but not experienced arousal, $r(152) = −0.13$, $p = 0.101$. Ratings of realism of agent behavior correlate with social presence, $r(152) = 0.38$, $p < 0.001$; full correlation matrix for presentations is presented in **Table 4**). Concerning the moderators between valence of agent behavior and experienced valence, realism is a significant moderator for experienced valence, $b = 0.21$, $t(150) = 2.29$, $p = 0.024$, but not for arousal, $b = −0.04$, $t(150) = −1.29$, $p = 0.200$. Social presence is a significant moderator, both for experienced valence, $b = 0.27$, $t(150) = 3.09$, $p = 0.002$, and arousal, $b = -0.11$, $t(150) = -3.92$, $p < 0.001$. The moderator effect of social presence on experienced valence and arousal during presentations, i.e., the different slopes for high and low values of social presence, are illustrated in **Figure 3B** (see **Table 3** for full results of moderator analysis).

3.2 Secondary Hypotheses - Evaluation of Scenarios

Ratings for valence of all agent behaviors differ significantly from the neutral condition in the expected directions, i.e., larger/more positive valence for *supportive* and *interested* and smaller/more negative valence for *uninterested* and *dismissive*, for both scenarios (all ps. <0.05). No difference to neutral condition in physical presence is significant in both scenarios. The corresponding plots are displayed in **Figure 4** (interactions) and **Figure 5** (presentations). Ratings for experienced valence differ from the neutral condition for all conditions in the expected directions (see above) except *supportive* and *interested* in the interaction scenario and all except *dismissive* in the presentation scenario. Ratings of arousal do not differ from the neutral condition in any condition in both scenarios. Realism of agent behavior and social presence was rated higher than neutral in both positive behavior conditions in the interaction scenario whereas in the presentation scenario no differences were found for any condition.
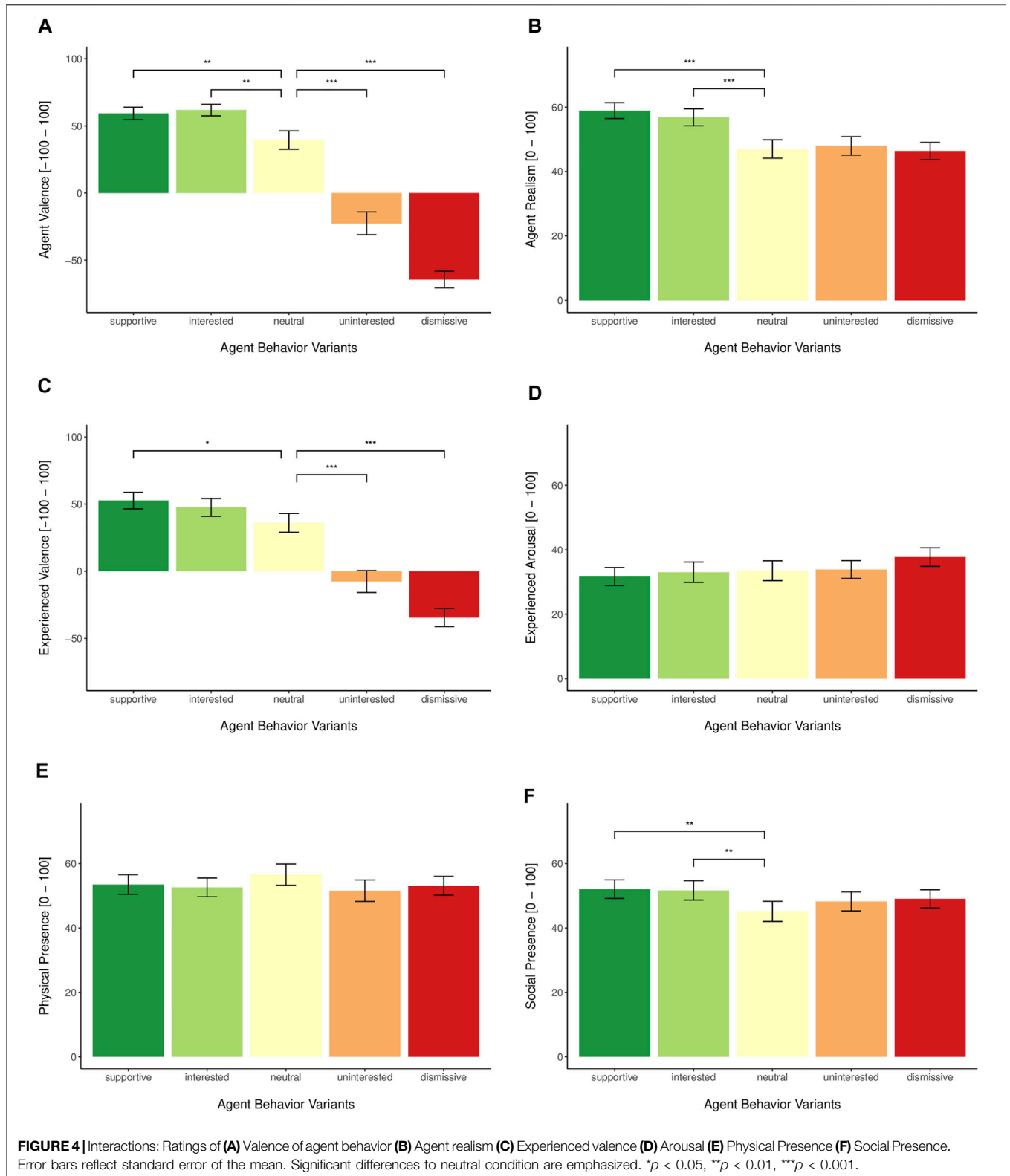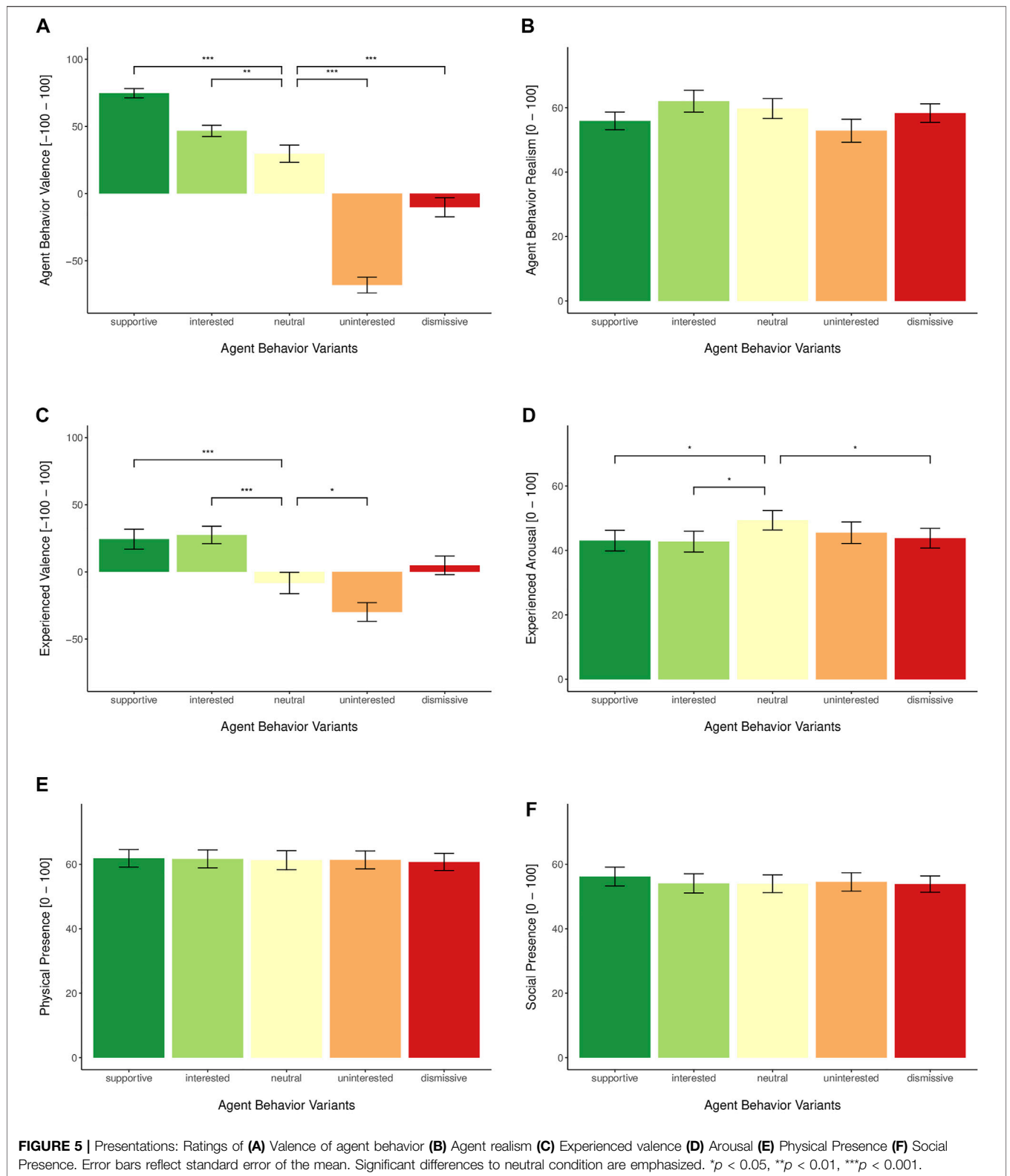
3.3 Exploratory Analysis

In some instances, the direction of valence of the agent behavior does not correspond to the direction of valence of the experience (i.e., recognizing a certain valence in the displayed behavior but experiencing the opposite valence; upper left and lower right quadrant in left plots of **Figure 3A,B**). These data points may represent errors in assessment or measurement or other factors influencing emotional experience, e.g., anxiety in the presentation scenario (despite supportive agent behavior) or amusement over dismissive agent behaviors in the interaction scenario. In the context of the current correlational study, these factors could be considered noise, i.e., qualitatively different influences affecting the quantification of the relationships in question and their hypothesized moderators. Therefore, to secure that such factors have not biased our results, we removed ratings of opposite polarity for agent behavior valence and experienced valence in a supplemental analysis (see **Supplementary Figure S1** for remaining data points). Under these conditions, social presence is still a moderator for experienced valence and arousal in the presentation scenario ($b = 0.22$, $t(93) = 4$, $p < 0.001$ and $b = −0.12$, $t(93) = −4.07$, $p < 0.001$), but also a significant moderator for experienced valence in the interaction scenario, $b = 0.1$, $t(115) = 2.25$, $p = 0.026$ (see **Supplementary Table S1** for full results).

TABLE 4 | Presentations: Repeated measures correlations among ratings.

	Experience		Presence		Agent Beh.		Pres. topic
	Valence	Arousal	Physical	Social	Realism	Valence	Challenge
EX Arousal	−0.17*	—	—	—	—	—	—
PR Physical	0.00	0.17*	—	—	—	—	—
PR Social	0.10	0.03	0.37***	—	—	—	—
AG Realism	0.16*	0.00	0.35***	0.38***	—	—	—
AG Valence	0.59***	−0.13	0.08	0.15	0.23**	—	—
PT Challenge	−0.39***	0.20*	0.11	0.08	−0.05	−0.24**	—
PT Emotion	0.23**	0.24**	−0.05	0.00	0.00	0.22**	−0.32***

*EX, experience; PR, presence; AG, agent behavior; PT, presentation topic. * <0.05, **p < 0.01, ***p < 0.001.*

FIGURE 4 | Interactions: Ratings of **(A)** Valence of agent behavior **(B)** Agent realism **(C)** Experienced valence **(D)** Arousal **(E)** Physical Presence **(F)** Social Presence. Error bars reflect standard error of the mean. Significant differences to neutral condition are emphasized. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

FIGURE 5 | Presentations: Ratings of **(A)** Valence of agent behavior **(B)** Agent realism **(C)** Experienced valence **(D)** Arousal **(E)** Physical Presence **(F)** Social Presence. Error bars reflect standard error of the mean. Significant differences to neutral condition are emphasized. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

4 DISCUSSION

The present study investigated the interplay of virtual agent behavior, social presence, and experienced emotions. In two virtual scenarios, an interaction scenario and a presentation scenario, participants were confronted with agents displaying different levels of agent behavior, i.e., negative, neutral, or positive behavior.

4.1 Primary Hypotheses

Valence of agent behavior correlated with experienced valence for both virtual scenarios (hypothesis 1a fully confirmed) whereas it only correlated with experienced arousal during interactions, not presentations (hypothesis 1b partially confirmed). Realism of agent behavior correlated with social presence in both scenarios (hypothesis 2 fully confirmed). Realism of agent behavior was a significant moderator of the effect of agent behavior for experienced valence only in the presentation scenario (hypothesis 3a partially confirmed) but not for arousal in any scenario (hypothesis 3b not confirmed). Social presence was a significant moderator of the effect of agent behavior on experienced valence only during presentations, not interactions (hypothesis 4a partially confirmed) and on arousal only during presentations, not interactions (hypothesis 4b partially confirmed). In summary, we were able to demonstrate the feasibility of investigating the moderating role of constructs like realism and social presence between agent behavior and evoked emotional experience but could not demonstrate moderator effects reliably in both scenarios.

The analysis of moderating effects on arousal is complicated by the fact that data points show a u-shaped pattern that could potentially be modeled using a quadratic term. But if a difference based on a potential moderator is present mostly for negatively valenced agent behavior (leading to high arousal) whereas no systematic differences are present for positively valenced agent behavior (but a large variation in whether this constitutes low or high arousal is present), then a linear relationship seems more suitable to capture this effect.

4.2 Secondary Hypotheses - Evaluation of Scenarios

The representation of different kinds of behaviors was successful for both scenarios. For valence of agent behavior, all conditions for interactions and presentations differed from neutral and for experienced valence, only the interested condition in the interaction scenario and dismissive condition in the presentation scenario showed no difference to neutral. For the interaction scenario, the similarity between positive conditions and neutral is congruent with feedback from speakers who recorded the statements that reported trouble in creating two different kinds of positive sentiments and a difference to "neutral" in the short statements comprising the interactions. More importantly, given the context of social anxiety, the negative conditions in the interaction scenario were successful in evoking a negative experience, one in evoking increased arousal. For the presentations, the uninterested condition was successful in evoking the targeted experienced valence. The finding that for

dismissive, valence of agent behavior clearly differed from neutral but experienced valence did not, could be an indication that displaying dismissive behaviors in an audience not necessarily translates to negative experiences.

Unexpectedly, for presentations, an effect of listening audience behaviors on arousal was practically absent in the given sample and overshadowed by the sequence effect (neutral condition first) and effects of the presentation topic (as indicated by the Rmcorrs between arousal and presentation topic in **Table 4**). It is possible that a sample with more pronounced social anxiety or SAD patients would show an effect on arousal. It is likewise conceivable that arousal was higher as a function of related traits than situational factors like audience behavior (a position taken by e.g., Ayres, 1990). Contrary to this position, Hsu (2009) showed that, in presentations with real audiences trained to give positive or negative feedback, audience behavior did have an substantial impact on state anxiety. But the necessary levels of realism of agent behavior and/or physical and social presence when depicting these behaviors in VR to achieve a similar effect are unknown.

For physical presence, no differences were apparent between any conditions in both scenarios. For social presence during interactions, the conditions supportive and interested differed from neutral. This potentially reflects that realism was rated higher for these conditions (than any other). It is unclear whether this reflects differences in the quality of the implementation between conditions or a genuine effect-potentially related to social expectations of polite responses. Such a finding, if replicated, would make it necessary to "package" negative behaviors of virtual agents as more realistic, potentially introducing it gradually in longer interactions.

4.3 Interpretation

The current study can be interpreted as preliminary evidence that social presence plays an important role in how evaluative behaviors of agents in VR are experienced. In research on the conditions that enable the processing of social stimuli in VR in a similar fashion to outside VR, Strojny et al. (2020) found that co-presence and realism were moderators on social facilitation effects i.e., the effect of virtual agents being present (analogous to humans outside VR being present) was dependent on the agents' perceived social realism. Despite the differences in context and methodology, the similarity in findings and our finding of a more reliable effect for social presence further pronounces the importance of focusing on social presence when intending to increase effects of social settings in VR.

4.4 Limitations and Strengths

A clear limitation of the current study is its correlational nature without systematic variation of realism of agent behavior or factors influencing social presence. Therefore, no causal inferences can be made. On the other hand, the study allowed us to investigate the effects of naturally occurring differences in perception and experience of VR which has its own merit. The fact that removing data points with opposing valence leaves the moderating effects in the presentation scenario intact and

introduces a moderating effect for experienced valence in the interaction scenario suggests that an experimental study, with less noise, could show more reliable moderator effects in multiple settings.

4.5 Implications

It is conceivable that social presence is one of the key mechanisms that underlies the effectiveness of VRET for social anxiety. Understanding its role in "translating" the social events in VR into an emotional experience (and potential peculiarities of SAD patients in this regard) could enable more precise measurements and early estimates of probable effectiveness of scenarios when used with patients. Future studies could systematically vary realism by counterbalancing two pairs of a positive and a negative condition (attempting high similarity in effects among the conditions of the same valence) and variations of realism (e.g., longer delay until agents utter statements during the interactions) or social presence (e.g., robotic vs. natural sounding voice). A similar analysis strategy could be used, i.e., different slopes between perception of agent behavior and its effect on experience for different degrees of realism, potentially investigating specific characteristics of high socially anxious participants or SAD patients.

An important question is whether the assumption that social presence is non-recoverable (Slater, 2009) is correct-even in SAD patients- and if so, how generalized this phenomenon is, i.e., does a breakdown of social presence negatively influence the complete VR session, just the current scenario or the current agent. Another open question is the nature of the moderation of social presence. It is conceivable that non-linear effects could be necessary to adequately model the moderation, maybe reflecting the notion of presence as "precondition" for the experience of emotion as a "jump" around a certain level of social presence. From an applied standpoint, finding the threshold of social presence necessary to evoke adequate responses in SAD patients could be a main goal in the development and refinement of social scenarios for VRET.

In conclusion, our findings underscore the importance of the construct social presence in evoking emotional responses to specific agent behaviors in virtual social interactions. Replication of the results is called for, especially in experimental designs using systematic variations of realism and other variables relevant to social presence and in clinically anxious samples, as well as investigations of the nature of the moderating relationship of social presence.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Committee of the University of Regensburg. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MP, LK, and AM designed research. BL and MM created the virtual environment. RF programmed scripts for automatized presentation. MM supervised virtual environment creation and experiment paradigm programming. MP supervised data acquisition and analyzed data. MP, LK, and AM wrote the paper. BL, RF, and MM commented on the paper.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2021.741138/full#supplementary-material

REFERENCES

Allan, S., and Gilbert, P. (1997). Submissive Behaviour and Psychopathology. Br. J. Clin. Psychol. 36, 467–488. doi:10.1111/j.2044-8260.1997.tb01255.x

Ayres, J. (1990). Situational Factors and Audience Anxiety. Commun. Educ. 39, 283–291. doi:10.1080/03634529009378810

Bakdash, J. Z., and Marusich, L. R. (2017). Repeated Measures Correlation. Front. Psychol. 8, 456. doi:10.3389/fpsyg.2017.00456

Baron, R. M., and Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. J. Personal. Soc. Psychol. 51, 1173–1182. doi:10.1037/0022-3514.51.6.1173

Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, É., Forget, H., et al. (2017). Virtual Reality Compared within Vivoexposure in the Treatment of Social Anxiety Disorder: A Three-Arm Randomised Controlled Trial. Br. J. Psychiatry 210, 276–283. doi:10.1192/bjp.bp.116.184234

Bouchard, S., St-Jacques, J., Robillard, G., and Renaud, P. (2008). Anxiety Increases the Feeling of Presence in Virtual Reality. Presence Teleop. Virtual Environ. 17, 376–391. doi:10.1162/pres.17.4.376

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., et al. (2019). Virtual Reality Exposure Therapy for Anxiety and Related

Disorders: A Meta-Analysis of Randomized Controlled Trials. *J. Anxiety Disord.* 61, 27–36. doi:10.1016/j.janxdis.2018.08.003

Chesham, R. K., Malouff, J. M., and Schutte, N. S. (2018). Meta-analysis of the Efficacy of Virtual Reality Exposure Therapy for Social Anxiety. *Behav. Change* 35, 152–166. doi:10.1017/bec.2018.15

Delacre, M., Lakens, D., and Leys, C. (2017). Why Psychologists Should by Default Use Welch's T-Test Instead of Student's T-Test. *RIPS* 30, 92. doi:10.5334/irsp.82

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., and Mühlberger, A. (2015). The Impact of Perception and Presence on Emotional Reactions: A Review of Research in Virtual Reality. *Front. Psychol.* 6, 26. doi:10.3389/fpsyg.2015.00026

Felnhofer, A., Hlavacs, H., Beutl, L., Kryspin-Exner, I., and Kothgassner, O. D. (2019). Physical Presence, Social Presence, and Anxiety in Participants with Social Anxiety Disorder during Virtual Cue Exposure. *Cyberpsychol. Behav. Soc. Netw.* 22, 46–50. doi:10.1089/cyber.2018.0221

Felnhofer, A., Kothgassner, O. D., Hetterle, T., Beutl, L., Hlavacs, H., and Kryspin-Exner, I. (2014). Afraid to Be There? Evaluating the Relation between Presence, Self-Reported Anxiety, and Heart Rate in a Virtual Public Speaking Task. *Cyberpsychol. Behav. Soc. Netw.* 17, 310–316. doi:10.1089/cyber.2013.0472

Felnhofer, A., Kothgassner, O. D., Schmidt, M., Heinzle, A.-K., Beutl, L., Hlavacs, H., et al. (2015). Is Virtual Reality Emotionally Arousing? Investigating Five Emotion Inducing Virtual Park Scenarios. *Int. J. Human Compu. Stud.* 82, 48–56. doi:10.1016/j.ijhcs.2015.05.004

Hsu, C.-F. (2009). The Relationships of Trait Anxiety, Audience Nonverbal Feedback, and Attributions to Public Speaking State Anxiety. *Commun. Res. Rep.* 26, 237–246. doi:10.1080/08824090903074407

IBM Corp (2017). *IBM SPSS Statistics for Windows, Version 25.0.* Armonk, NY: IBM Corp.

Ijsselsteijn, W. A., De Ridder, H., Freeman, J., and Avons, S. E. (2000). "Presence: Concept, Determinants, and Measurement," in *Proceedings of the SPIE* (San Jose, CA: Human Vision and Electronic Imaging V), 3959, 520–529. doi:10.1117/12.387188

Kampmann, I. L., Emmelkamp, P. M. G., Hartanto, D., Brinkman, W.-P., Zijlstra, B. J. H., and Morina, N. (2016). Exposure to Virtual Social Interactions in the Treatment of Social Anxiety Disorder: A Randomized Controlled Trial. *Behav. Res. Ther.* 77, 147–156. doi:10.1016/j.brat.2015.12.016

Kim, H. K., Park, J., Choi, Y., and Choe, M. (2018). Virtual Reality Sickness Questionnaire (VRSQ): Motion Sickness Measurement index in a Virtual Reality Environment. *Appl. Ergon.* 69, 66–73. doi:10.1016/j.apergo.2017.12.016

Kroczek, L. O. H., Pfaller, M., Lange, B., Müller, M., and Mühlberger, A. (2020). Interpersonal Distance during Real-Time Social Interaction: Insights from Subjective Experience, Behavior, and Physiology. *Front. Psychiatry* 11, 561. doi:10.3389/fpsyt.2020.00561

Leary, M. R. (1983). A Brief Version of the Fear of Negative Evaluation Scale. *Pers Soc. Psychol. Bull.* 9, 371–375. doi:10.1177/0146167283093007

Lee, K. M. (2004). Presence, Explicated. *Commun. Theor.* 14, 27–50. doi:10.1111/j.1468-2885.2004.tb00302.x

Ling, Y., Nefs, H. T., Morina, N., Heynderickx, I., and Brinkman, W.-P. (2014). A Meta-Analysis on the Relationship between Self-Reported Presence and Anxiety in Virtual Reality Exposure Therapy for Anxiety Disorders. *PloS one* 9, e96144. doi:10.1371/journal.pone.0096144

Lombard, M., and Jones, M. T. (2015). "Defining Presence," in *Immersed in Media.* Editors M. Lombard, F. Biocca, W. Freeman, R. IJsselsteijn, and R. J Schaevitz (Springer International Publishing), 13–34. doi:10.1007/978-3-319-10190-3_2

Makransky, G., Lilleholt, L., and Aaby, A. (2017). Development and Validation of the Multimodal Presence Scale for Virtual Reality Environments: A Confirmatory Factor Analysis and Item Response Theory Approach. *Comput. Hum. Behav.* 72, 276–285. doi:10.1016/j.chb.2017.02.066

Neudeck, P., and Einsle, F. (2012). "Dissemination of Exposure Therapy in Clinical Practice: How to Handle the Barriers?," in *Exposure Therapy: Rethinking the Model - Refining the Method.* Editors P. Neudeck and H.-U Wittchen (New York, NY: Springer), 23–34. doi:10.1007/978-1-4614-3342-2_3

Paul, G. L. (1966). *Insight vs. Desensitization in Psychotherapy: An experiment in Anxiety Reduction.* Stanford, CA: Stanford University Press.

Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., et al. (2011). "The Kaldi Speech Recognition Toolkit," in IEEE 2011 Workshop on Automatic Speech Recognition and Understanding (IEEE Signal Processing Society) (Hilton Waikoloa Village, Big Island: IEEE). Catalog No: CFP11SRW-USB.

R Core Team (2020). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rothbaum, B. O., Hodges, L., and Kooper, R. (1997). Virtual Reality Exposure Therapy. *J. Psychotherapy Pract. Res* 6, 219–226.

Slater, M. (2009). Place Illusion and Plausibility Can lead to Realistic Behaviour in Immersive Virtual Environments. *Phil. Trans. R. Soc. B* 364, 3549–3557. doi:10.1098/rstb.2009.0138

Stangier, U., and Steffens, M. (2002). *Social Phobia Inventory (SPIN)–Deutsche Fassung.* Frankfurt am Main: Psychologisches Institut der Universität Frankfurt am Main.

Steinman, S. A., Wootton, B. M., and Tolin, D. F. (2016). "Exposure Therapy for Anxiety Disorders," in *Encyclopedia of Mental Health.* Editors H. S. Friedman (Oxford: Elsevier Academic Press), 186–191. doi:10.1016/B978-0-12-397045-9.00266-4

Strojny, P. M., Dużmańska-Misiarczyk, N., Lipp, N., and Strojny, A. (2020). Moderators of Social Facilitation Effect in Virtual Reality: Co-presence and Realism of Virtual Agents. *Front. Psychol.* 11, 1252. doi:10.3389/fpsyg.2020.01252

Volkmann, T., Wessel, D., Jochems, N., and Franke, T. (2018). "German Translation of the Multimodal Presence Scale," in Mensch und Computer 2018-Tagungsband. Editors G. Dachsel and R. Weber (Bonn, Germany: Gesellschaft für Informatik e.V). doi:10.18420/muc2018-mci-0428

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag.

Xie, Y. (2015). *Dynamic Documents with R and knitr.* Boca Raton, FL: Chapman and Hall/CRC.

Wechsler, T. F., Mühlberger, A., and Küumpers, F. (2019). Inferiority or even Superiority of Virtual Reality Exposure Therapy in Phobias? - A Systematic Review and Quantitative Meta-Analysis on Randomized Controlled Trials Specifically Comparing the Efficacy of Virtual Reality Exposure to Gold Standard *in vivo* Exposure in Agoraphobia, Specific Phobia and Social Phobia. *Front. Psychol.* 10, 1758. doi:10.3389/fpsyg.2019.01758