ACTA UNIVERSITATIS LODZIENSIS

FOLIA SOCIOLOGICA 86, 2023

https://doi.org/10.18778/0208-600X.86.01

Member since 2019
JM14479

Aleksander Walas*

https://orcid.org/0009-0004-3768-380X

IMMERSIVE VIRTUAL REALITY AS A SOCIAL RESEARCH ENVIRONMENT. THE IMPACT OF THE PRESENCE OF AN EMBODIED AGENT ON THE STROOP TASK PERFORMANCE BASED ON OWN RESEARCH[1]

Abstract. This empirical study examined the effects of the presence of others (physical observers and embodied agents) on the performance of the Stroop task within an immersive virtual reality (VR) environment. Drawing from the research conducted by Pascal Huguet and his colleagues, who explored the social facilitation and inhibition effects in traditional lab conditions using the Stroop test, the current study sought to test these effects in VR conditions. Participants were divided into four groups, each experiencing different conditions – performing the task alone, under the awareness of being observed, or in the presence of embodied agents (formally dressed as Dr. Piotr or casually as Piotrek). A one-way analysis of variance (ANOVA) for independent samples was conducted to evaluate the results. Findings showed a significant reduction in Stroop interference when participants were aware of being observed by the experimenter. However, the presence of embodied agents in the VR environment, regardless of their regardless of their dress or manner of presentation, did not significantly influence the results. These results not only confirmed the occurrence of the Stroop interference effect in VR conditions, comparable to traditional lab conditions, but they also suggested that the social presence of the experimenter could enhance task performance in the VR environment.

* MA, Doctoral School of Social Sciences, University of Lodz, ul. Matejki 21/23, pok. 114 (I piętro), 90-237 Łódź, e-mail: aleksander.walas@edu.uni.lodz.pl, alex@walas.net.

[1] The article is based on an MA thesis written in the Department of the Sociology of Politics and Morality, Faculty of Economics and Sociology of the University of Lodz, under the supervision of Professor Alicja Łaska-Formejster. The thesis was awarded in the Krystyna and Jan Lutyńscy competition (under the auspices of the University of Lodz) and the Florian Znaniecki competition (under the auspices of the Polish Sociological Association).

Received: 20.06.2023. Verified: 10.07.2023. Revised: 25.07.2023. Accepted: 25.07.2023.
© by the author, licensee University of Lodz – Lodz University Press, Lodz, Poland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/)

There was no significant influence observed from the presence of embodied agents, though. These observations can greatly contribute to the development of VR. However, further research is required to expand these findings, considering factors such as sample size and realism of embodied agents.

Keywords: Virtual reality, VR, immersive virtual reality, social facilitation, social inhibition, Stroop task, Stroop interference, embodied agent, social research.

IMMERSYJNA RZECZYWISTOŚĆ WIRTUALNA JAKO ŚRODOWISKO BADAŃ SPOŁECZNYCH. WPŁYW OBECNOŚCI UCIELEŚNIONEGO AGENTA NA POZIOM WYKONANIA ZADANIA STROOPA NA PODSTAWIE BADAŃ WŁASNYCH

Abstrakt: Badaniom poddany został wpływ obecności innych osób (fizycznego obserwatora i ucieleśnionych agentów) na wykonanie zadania Stroopa w immersyjnym środowisku wirtualnej rzeczywistości (VR). Opierając się na pracach przeprowadzonych przez Pascala Hugueta i współpracowników, którzy badali efekty facylitacji i hamowania społecznego w tradycyjnych warunkach laboratoryjnych za pomocą testu Stroopa, obecne badanie miało na celu przetestowanie tych efektów w warunkach VR. Uczestnicy zostali podzieleni na cztery grupy, z których każda przebadana była w innych warunkach – wykonując zadanie samodzielnie, ze świadomością bycia obserwowanym lub w obecności wcielonych agentów (formalnie ubranego dr Piotra lub nieformalnie Piotrka). W celu oceny wyników przeprowadzono jednoczynnikową analizę wariancji (ANOVA) dla niezależnych próbek. Wyniki wykazały znaczną redukcję interferencji Stroopa, gdy uczestnicy byli świadomi tego, że są obserwowani przez eksperymentatora. Jednak obecność ucieleśnionych agentów w środowisku VR, niezależnie od ich ubioru czy sposobu przedstawienia, nie wpłynęła znacząco na poziom wykonania zadania. Wyniki nie tylko potwierdziły występowanie efektu interferencji Stroopa w warunkach VR, porównaniu z tradycyjnymi warunkami laboratoryjnymi, ale także zasugerowały, że obecność społeczna eksperymentatora może poprawić wydajność wykonywania zadań w środowisku VR. Nie zaobserwowano jednak znaczącego wpływu obecności ucieleśnionego agenta. Obserwacje te mogą znacznie przyczynić się do rozwoju aplikacji VR. Jednak potrzebne są dalsze badania, w których uwzględnione zostaną takie czynniki, jak wielkość próby i realizm ucieleśnionych agentów.

Słowa kluczowe: rzeczywistość wirtualna, VR, immersyjna rzeczywistość wirtualna, facylitacja społeczna, hamowanie społeczne, zadanie Stroopa, interferencja Stroopa, ucieleśniony agent, badania społeczne.

1. Introduction

Virtual reality (VR) has been envisaged as an innovative communication medium since its inception, with the potential to radically alter the way humans interact (Oh et al. 2018: 1). Some VR pioneers envisioned it as the ultimate form of interaction between humans and machines, a medium that could liberate the human spirit (Biocca et al. 1995: 7). Others saw it as the first step towards the creation of an optimal communication medium. William Gibson described cyberspace as a consensual hallucination experienced daily by billions of authorized users (Gibson 1984: 51).

Despite conceptualizing VR as a social medium where individuals can coexist and interact, the earliest VR research focused on single-user systems using head-mounted displays (HMDs), generally restricted to laboratory settings. However, in recent years, VR technology has migrated, in various forms, to home environments. This increased accessibility, coupled with improved technical specifications and lower costs, has reignited interest in its social applications, evidenced by the launch of platforms such as AltSpace VR, Facebook Spaces (now Meta Horizon), and VR Chat (Oh et al. 2018: 1).

VR is believed to offer a significantly higher level of social presence compared to other technology-mediated communication forms – the subjective experience of being with a real person, gaining access to their thoughts and emotions (Biocca 1997). Through visual, auditory, tactile, and to a lesser extent, olfactory information, VR systems can provide a wide range of social cues. Consequently, it has been argued that it's critical to understand how various technology-related aspects of VR affect the perception and experience of social presence (Oh et al. 2013: 1).

The phenomenon of social facilitation in immersive virtual reality environments forms the core of this study. Social facilitation, or more broadly social facilitation and inhibition (SFI), is the simplest aspect of group activity, associated with the mere presence of others (Wojciszke 2002: 388). As demonstrated in the meta-analysis by F.C. Bond and L.J. Titus (Bond, Titus 1983: 265), social presence increases the speed of performing simple, well-learned tasks but slows the performance of poorly learned tasks.

This study aims to investigate the phenomenon of social facilitation and inhibition in immersive virtual reality. Two goals were established: 1) Determine the similarity between the impact of another human's presence and the influence of an embodied agent on task efficiency in immersive VR, and 2) Demonstrate the potential of immersive virtual reality systems as an environment for researching selected social behaviors.

2. Virtual reality as a laboratory for social science research

In the academic sphere, it is increasingly recognized that beyond the established position of Virtual Reality (VR) in diagnostics (psychiatric, psychological), or therapy, this technology can be successfully used as a tool for experimental research. The methodology based on VR solutions exhibits a multitude of advantages, especially in studies of psychopathological phenomena. These advantages include the realism of the environment, the ability to control a large number of variables while simultaneously providing precise and comprehensive response recording (neurological, physiological, self-reports, etc.), and the absence of the need to sacrifice external validity to achieve greater internal validity. As individuals are immersed in an environment that mirrors reality, simulating the environment and tasks performed in the real world, the generalizability of results is enhanced (Baños et al. 1999: 283).

An important characteristic of VR is its capability to control social environments, which is regarded as one of the key variables, for example, in understanding psychosis. Freeman and colleagues (2010: 85) employed VR to study paranoid delusions, examining whether virtual avatars (computer-generated representations of people in virtual environments) are perceived as hostile. The work of Freeman and his team yielded satisfactory results and demonstrated that using VR as a research environment is feasible and safe for patients with psychotic disorders (Freeman et al. 2010: 89).

Botella et al. (2017: 575), citing Dunsmoor et al. and Shibana et al., note that using VR and virtual embodied agents in research involving social conditioning stimuli carries many benefits. VR-based environments can simulate social interaction much more effectively than traditional computer systems with a standard monitor. Moreover, the use of space and movement in VR enhances ecological validity, and participants' attention is concentrated on the experiment, not distracted by the laboratory or experimenter.

VR technology may also provide a useful way to conduct research that would be challenging to carry out in a real environment due to ethical considerations. For example, Slater and colleagues suggested that VR can be beneficial in any social and psychological study where, for safety or ethical reasons, it is not possible to place participants in the actual phenomenon to be studied. To illustrate this, the researchers employed VR to replicate Stanley Milgram's classic experiment on obedience to authority (Slater et al. 2006: 2).

3. Social facilitation and inhibition

Social facilitation and inhibition are phenomena related to the influence of others' presence on our behavior. These studies began in the 19th century with observations by Triplett (1983: 514), who noticed that cyclists racing against others achieved better results than cycling alone.

In the 1930s, Pessin and Husband demonstrated the existence of social inhibition or deterioration of performance in the presence of others. However, due to conflicting results, this problem was abandoned for nearly 30 years (Wojciszke 2002: 338).

The issue was revisited in the 1960s when Zajonc (1965: 270) tried to resolve these contradictions. He noticed that the presence of others facilitated the performance of simple tasks, but hindered the performance of complex tasks. He suggested that the mere presence of others was sufficient to cause social facilitation or inhibition effects (Zajonc 1965: 274).

According to Zajonc, the presence of others enhances arousal, which strengthens the tendency to perform the dominant reaction (Wojciszke 2002: 340). The presence of others increases the likelihood of errors in difficult tasks and reduces it to easy ones, confirmed by a meta-analysis by Bond and Titus (Bond, Titus 1983: 265).

Alternative explanations of the social facilitation and inhibition effects are connected with the interpretation of the observer's behavior by the observed. Cottella's model of evaluation apprehension (1972, cited in Bond 1982: 1042) suggests that the effect of social facilitation is evoked by the presence of an evaluative observer, but the majority of empirical studies did not confirm this (Wojciszke 2002: 340).

Baron's attention distraction model postulates that the mere presence of others attracts attention and causes a conflict between focusing on the other person and the task. Baron (1986) indicates this conflict as a mediator of facilitation and inhibition effects.

4. The Stroop task in measuring social facilitation effects

As previously mentioned, models explaining the influence of the presence of others on task performance – Robert Zajonc's drive model and Robert Baron's attention model – predict that the presence of others enhances performance in easy tasks (social facilitation) and reduces it in challenging tasks (social inhibition). The differences between these models lie in the underlying psychological mechanisms that explain this relationship. Zajonc's model suggests that the differentiating factor between easy and difficult tasks is the accuracy of the dominant response to task stimuli. In contrast, Baron's model suggests the level of task performance depends on the amount of stimuli required to solve the task. The presence of others causes the cognitive load, which, for easy tasks, leads to an improvement in task performance, and for difficult tasks, to a decrease. As Wojciszke (2002: 391) observes, in nearly all previous research, the validity of both theoretical models has been confirmed. Easy tasks are characterized by both a small number of signals needed for their execution, and the correct response being the dominant one. On the other hand, difficult tasks involve a high amount of stimuli, and the dominant response is incorrect.

Charles F. Bond Jr. (1982: 1042) sought to disrupt this pattern in studies of the self-presentation model, explaining the social facilitation phenomenon. The author embedded easy elements in a difficult task and difficult elements in an easy task. According to the adopted model, the observer's presence hindered the learning of the simple elements present in the difficult task and facilitated the learning of difficult elements placed in the easy task (Bond 1982: 1048).

The Stroop test, however, is an example of a task where the incorrect response is dominant but requires taking into account a very small number of signals (Huguet et al. 1999: 1012). This test is a frequently used method for evaluating executive control through the so-called interference effect (Okruszek, Rutkowska 2013:

216). In the standard version of the Stroop task, subjects are presented with words denoting colors (e.g., green, red) printed in different colors. The sequence of words and colors is determined randomly, and the subjects are asked to name the color in which a given word is presented as quickly as possible (Vakil et al. 1996: 314). Successful performance of the Stroop test requires the ability to inhibit habitual reactions, which is associated with the activation of attention processes, cognitive plasticity (mainly selective functions), and effective control, allowing flexible adjustment to changing task rules and conditions (Tomaszewska et al. 2010: 35). The Stroop test is believed to be a measure of cognitive control over the disruptive influence of the automated reading reaction, hence it is also used to measure inhibition control in conflict situations (Jodzio 2008: 263). The interference effect is understood as the time measured by the prolongation of the reaction time to the color of a color-inconsistent word compared with the same reaction to the color of a color-consistent word (Śpiewak 1999: 130).

Relating the Stroop task to the phenomena of social facilitation and inhibition, the dominant response in this task is processing the word content, which causes the interference effect, understood as a decrease in the speed of color recognition due to the word's content (which indicates a different color). However, this is an incorrect response, as the task involves identifying the color, not recognizing the word's content.

According to Zajonc's theory, it is a difficult task, as the presence of others should decrease performance. Conversely, according to Baron's theory, it is an easy task, as there are only two signals, one of which (the word's meaning) is irrelevant. From the perspective of this theory, the presence of others should enhance performance, as the narrowing of attention leads to ignoring the irrelevant signal (Wojciszke 2002: 392).

5. Methods

The research methodology that was used in this work, similar to the studies of Pascal Huguet and colleagues (Huguet et al. 1999: 1011), was based on the Stroop task. These researchers showed that the presence of others leads to an increase in the Stroop task performance level (decrease in interference values), meaning that ignoring word content was facilitated. These results are consistent with the model adopted by Baron but are difficult to explain by Zajonc's model, even though the latter is widely accepted (Huguet et al. 1999: 1011).

Similar conclusions (with some limitations) to those of Huguet and colleagues (Huguet et al., 1999: 1019) were reached by the author of this work, transferring the study of social facilitation to the virtual reality environment.

The study conducted as part of this research aimed to determine the level of social facilitation occurring in the immersive virtual reality experimental conditions using the Stroop test. The role of the facilitator was performed by the experimenter and a virtual, embodied agent. Simultaneously, an attempt was made to relate the obtained results to the findings of previous studies by other authors carried out in real (non-virtual) conditions. This was aimed at identifying potential prospects for the application of immersive virtual reality technology as an environment for selected social research.

The study included 64 participants (42 women), comprised of employees from a manufacturing plant (40 individuals) and acquaintances of the study's author. The participant sample spanned a wide age range from 19 to 65 years and a variety of professional roles. A precondition for participation in the study was the absence of vision disorders, such as color blindness. Each participant was randomly assigned to one of four groups, each corresponding to a different type of social presence (levels of the study). The participants provided written consent to participate in the study using VR technology.

Following the study's assumptions, participants were divided into 4 groups. To achieve an equal number of women and men in each group, randomization was performed separately for each gender.

For the study, in Unreal Engine 4.7 was created a computer program was installed and run on Oculus Quest 2 virtual reality goggles. Figure 1 shows an examination using goggles Oculus 2.

Figure 1. Examination using the Oculus Quest 2 VR device

Source: own study

The independent variables – factors – measured on the nominal scale (Grzeszkiewicz-Radulska et al. 2020: 104) that were present in the study belonged to the following groups:

The first group. The type of social presence simulated during the study. This factor had the following values (levels):

- Level 1 – the task was performed alone, and there were no characters in the virtual environment.

- Level 2 – there were no characters in the virtual environment, but the participant was informed that the researcher would observe the study on a monitor duplicating the participant's point of view.
- Level 3 – an avatar dressed in sportswear, whose gaze was directed at the participant, was present in the participant's field of view in the three-dimensional virtual reality environment.
- Level 4 – an avatar dressed in a white shirt, whose gaze was directed at the participant, was present in the participant's field of view in the three-dimensional virtual reality environment.

The second group. The color inconsistency of the displayed images. In this group of factors, there were two levels: 1) A colored control pattern in the form of a "+" character string of varying length was displayed (consistent displays); 2) A word was displayed, the meaning of which was not consistent with the color in which it was presented (e.g., the word BLUE displayed in red, or SUN in green).

During the study, three dependent variables were measured. The first dependent variable was the reaction time to the stimulus – the delay in milliseconds between the start of displaying the colored word or the control string of the Stroop test and the moment the participant pressed any virtual button (quantitative variable). The difference between the average reaction time to color-inconsistent words and the average reaction time to control signs was the Stroop interference – this was the second dependent variable (also quantitative).

The third dependent variable was the correctness of the response, i.e., the consistency of the displayed color with the pressed corresponding button. The number of correct answers given by the participant divided by the number of displays represented the correctness coefficient of the answers.

Task performance was understood as (determined independently): A smaller Stroop interference effect – a higher level of task performance. The level of response correctness – expressed by the correctness coefficient – a higher coefficient indicated a higher level of task performance.

Three versions of the program were developed, differing from each other in the presence/absence of a virtual agent, as well as its appearance and representation. Each group was tested using a specific version of the program. Thus, Group 1 and Group 2 performed the primary test without the presence of an agent, group 3 with the presence of the agent 'Piotrek', and Group 4 with the presence of the agent 'Dr. Piotr'.

Every participant in the study, regardless of the group they belonged to, went through three levels (parts) of the program.

Introduction Level. The purpose of the first level was to enable the test subject to adapt to the new situation of experiencing immersive virtual reality, as well as to present basic information regarding VR technology and the hand motion tracking system. The virtual scenery of the introduction level was organized in a minimalist way, allowing however to capture the spatial relations characteristic of virtual reality.

At this stage, information regarding the purpose and course of the study, as well as the possibility of interrupting it at any time, was also conveyed to the participant. The readiness to proceed to the next level had to be confirmed by the participant in an interactive way – by pressing a virtual button. All information at the introduction level, as well as on the next two levels, was delivered verbally. Figure 1 presents the scenery of the first (introduction) level.

Training Level. The intent of this stage was to uniformly and thoroughly explain the role assigned to the study participant. Simultaneously, this level served as practical preparation for performing the Stroop test in a virtual environment. An assumed, hypothetical effect of the participant going through the training level was to compensate for any psychomotor differences (e.g., the time needed to acquire specific manual skills in a VR environment) that could disrupt the study results.

Upon completing the introductory level, the participant was transferred to the training level. The starting point, which is the place in the virtual space from where the participant begins to observe the scenery, was located in a way that allowed observation of a significant part of the room. The height of the observation point was determined automatically and reflected the current distance between the participant's eye level and the real floor level. The location of the interactive elements (buttons and board) was determined in a similar way – it considered the participant's line of sight and the reach of their hands.

The training level was divided into two parts: informative and practical. In the first part, virtual buttons and a board were introduced, on which the words and control signs of the Stroop test were displayed.

The participant was informed through verbal messages about the task at hand (they are supposed to press one of the buttons "R", "G", "B", "Y" – corresponding to the color of the displayed image). They were then informed about the three practical training series that awaited them. During each series, 20 so-called frames (color words or sequences of "x" characters) were displayed, and each subsequent cycle had an increased level of difficulty. At the beginning of the training cycle, the displayed image only disappeared after pressing a button, while incorrect responses were corrected with a voice message. In the second training series, verbal cues were retained, but the display time was shortened to 700ms, and the rhythm of frame changes was set to 3500ms. The third series had the same time parameters as the proper test at the third level, i.e., the frame display time was 300ms, rhythmically every 2000ms, and voice messages were also disabled. In the training module, only color-neutral words (e.g. "bench", "idea") were displayed on a gray background in one of the four colors used in the Stroop task (blue, green, red, yellow).

After going through the three training series, the participant was notified that they would soon be redirected to the main task. At this point, if the version for group 3 or group 4 is launched, verbal information appears about the presence of a virtual agent during the test.

Depending on which group a given person was assigned to – one of the program variants was launched during the study:

- Version A – no embodied agent or avatar was present in the virtual room while performing the Stroop task (Group 1 and 2).
- Version B – an embodied agent was present in the room, watching the participant – the agent was introduced as Doctor Piotr and was dressed in a white shirt (Group 3).
- Version C – an embodied agent was present in the room, watching the participant – agent was introduced as Peter and was dressed in sportswear (Group 4).

After completing the second level, the **third level** was launched. When the participant confirmed their readiness to begin the task by pressing any button, the display of the test frames began. Exposed at a tempo (300ms display, rhythm 2000ms) were words associated with colors (e.g., “blood”) or denoting colors (blue, green, red, yellow) in a color inconsistent with the word’s meaning. Control positions were displayed with sequences of “+” characters of various lengths, in one of the aforementioned colors. Each element was displayed twice, equating to 80 test displays (frames) (2 x 20 words and 2 x control sequences). During the actual test, the time between the start of the display of a colored word or sequence of colored “+” characters and the pressing of the virtual button (reaction time) was measured. The recording of the reaction time took place automatically during the task in a text file, along with a timestamp and the parameters of the given word or control sequence. Figure 2 shows the third level with embodied agent Dr. Piotr.

Figure 2. Scenery of the third level (agent Dr. Piotr)

Source: own study

To answer the research questions posed, the data obtained were subjected to statistical analysis.

The average values of dependent variables (reaction time to control signs, reaction time to color-inconsistent words, Stroop interference, accuracy rate) obtained during the study for different groups are presented in Table 1.

Table 1. Average of dependent variables obtained during the study

	Variables			
	Display time – control sign	**Display time – incoherent words**	**Stroop Interference**	**Accuracy rate**
The task performed in solitude	1101	1199	98,3	0,983
The examiner monitors the screen	1156	1171	15,5	0,853
In the field of view is Agent Dr. Piotr	1090	1146	56,6	0,875
In the field of view is Agent Piotrek	1116	1195	79,4	0,898

Source: own study

The data obtained in all groups were characterized by a longer reaction time to the display of color-inconsistent words compared to the reaction time to the display of control signs, indicating the occurrence of the Stroop interference effect. This effect was greatest in Group 2 (members were informed that the experimenter was observing their actions in virtual reality on a cloned monitor during the Stroop test), and smallest in Group 1 (the task was performed in solitude, i.e., the experimenter ensured that he was not observing the subjects in virtual reality and there was no virtual figure in the VR field of view).

In Group 3, where there was a virtual agent dressed in a white shirt and introduced as Dr. Piotr in the VR field of view during the study, the interference effect was larger than in Group 4, where the virtual agent present was dressed in a sporty sweatshirt and introduced as Piotrek.

At the same time, Group 2 observed a longer reaction time (compared to the other groups) to the display of control signs with a similar reaction time to the display of color-inconsistent words as in other groups.

The average values of the accuracy rate (a higher value of the rate indicates fewer mistakes made in color designation during the test) fell within a relatively narrow range – between 0.983 (Group 1) and 0.853 (Group 2). Fewer mistakes were also made in Group 4 (0.898) compared to Group 3 (0.875).

Using statistical methods: mixed ANOVA, one-way ANOVA, and contrast analysis (Grzeszkiewicz-Radulska et al. 2020: 104; Sosnowski 2004: 370), hypotheses were tested and verified.

To test Hypothesis 1, asserting the occurrence of interference effects (measured in milliseconds of statistically significant differences in reaction times to the control signs displays and color-inconsistent words) in individuals performing the Stroop task in immersive virtual reality conditions, a mixed ANOVA method was used.

The within-subject factor was the reaction time measurement (control displays, color-inconsistent words displays), while the between-subject factor was the group (4 levels: solitary task performance, experimenter watching the monitor screen, presence of an agent introduced as Dr. Piotr agent, presence of an agent introduced as Piotrek).

The analysis revealed significant differences between the measurements – the main effect of measurement was statistically significant: $F(1.60) = 89.86$; $p < 0.01$; $\eta\ 2p = 0.6$. Bonferroni post-hoc tests showed statistically significant differences between the reaction time measurement for color-inconsistent word displays and control sign displays: mean difference = 62.00; $p < 0.01$. Bonferroni post-hoc tests considering time and groups demonstrated statistically significant differences between reaction time measurement for color-inconsistent word displays and control sign displays in Group 1: mean difference = 97.8; $p = 0.02$, in Group 3: mean difference = 56.0; $p < 0.01$, and in Group 4: mean difference = 79.3; $p < 0.01$. For Group 4, no statistically significant differences were demonstrated. There were also no statistically significant inter-group differences: $F(3,60) = 0.560$; $p = 0.643$. However, the interaction effect of factors proved significant: $F(3.60) = 7.43$; $p < 0.01$; $\eta\ 2p = 0.271$.

To test Hypotheses 2, 3, and 4, which postulated a significant effect of others' presence (the awareness of being observed, the presence of embodied agent Dr. Piotr or agent Piotrek in the visual field) on decreasing the Stroop interference (expressed in ms) during the Stroop task performance in immersive virtual reality, a one-way analysis of variance (ANOVA) for independent samples was conducted. The analysis showed that the compared groups significantly differed from each other. This indicates that the presence of others (either a person observing the task performance on a monitor or an embodied agent) influences the Stroop interference during the test in an immersive virtual reality environment, $F(3.60) = 7.35$; $p < 0.001$; $\eta^2 = 0.26$.

The result indicated that at least one group statistically significantly differed from the others.

Subsequently, under Hypotheses 2–4, a procedure involving the analysis of selected planned contrasts (a priori contrasts) was applied. Three contrasts were defined.

The analysis of contrast 1 showed that groups 1 and 2 differed significantly from each other: $F(1.60) = 17.018$; $p < 0.001$. The results showed that Group 2, which performed the Stroop task with the awareness that the experimenter was observing the test process from the perspective of the subject (on a cloned monitor), had a lower level of Stroop interference than the group performing the task in solitude (the difference in averages was 93.9ms). This allowed for the acceptance of the hypothesis about the influence of the awareness of being observed during the Stroop task in immersive virtual reality on the reduction of interference value.

The results obtained for contrast 2 and contrast 3 did not confirm the existence of statistically significant differences (respectively: $p = 0.401$, $p = 0.225$), therefore it was necessary to reject Hypothesis 3, which postulated that the presence of an embodied agent in the subject's visual field reduces the interference value during the Stroop task in immersive virtual reality compared to the situation when the task was performed in solitude. Hypothesis 4, which postulated that the appearance and presentation manner of the virtual agent (formal attire, presented as Dr. Piotr vs. sportswear, presented as Piotrek) influence the reduction of Stroop interference, also did not receive confirmation.

Testing Hypothesis 5, postulating the effect of others' presence on the number of errors in color recognition during the Stroop task in virtual reality, using a one-way ANOVA, showed no significant differences between groups: $F(3.60) = 1.27$; $p < 0.291$. This suggests the lack of influence of others' presence (either an observer or an embodied agent) on correct color recognition in the Stroop test.

The statistical analysis conducted in this study confirmed Hypothesis 1 and Hypothesis 2. However, Hypotheses 3, 4, and 5 did not receive confirmation. Therefore, the obtained results allow us to assume that during the performance of the Stroop task in the immersive virtual reality environment, the interference effect occurred, just like in classic (non-virtual) laboratory conditions.

There was also a statistically significant reduction in the interference value in the situation where the subject performed the Stroop test with the awareness of being observed in virtual reality by the examiner. No such influence was observed when an embodied agent was introduced into the subject's field of vision. The manner of dressing and the introduction of the agent also did not have a significant impact.

The statistically significant influence of the presence of an agent or the awareness of being observed on the degree of correct color recognition in the Stroop test in virtual reality was not confirmed either.

6. Conclusions and discussion

The following conclusions emerge from the empirical research conducted in this study. First, during the performance of the Stroop task in immersive virtual reality, the interference effect occurred, just as it happens in real (non-virtual) laboratory conditions. Statistically significant prolongation of the reaction time, which consisted of determining the color of the displayed word whose color was inconsistent with its meaning, was observed in the studied groups compared to the time of a similar reaction to the display of color control signs (+++++). This means that this study constitutes a significant premise for designing and further testing tools based on immersive VR technology – both implementing the Stroop task and other research techniques.

Secondly, the influence of social presence on improving the effectiveness of the Stroop task performance was observed in the immersive virtual reality environment. A statistically significant reduction in inference values was obtained in the group of people performing the Stroop test with the awareness of being observed in virtual reality by the examiner, compared to the group performing the task in solitude (the examiner did not observe the actions of the subject in VR on the monitor, no computer-generated character is in the field of view).

However, no significant reduction in interference values was observed in groups where an embodied agent was present in the field of vision during the task. The agent's dress and manner of presentation also had no impact. This means that only one of the detailed hypotheses, stating the improvement in the effectiveness of performing the Stroop task in immersive virtual reality due to the presence of others, was confirmed. This hypothesis also referred to the presence of a real person (the examiner).

Therefore, the main hypothesis stating that the presence of an embodied agent affects the level of Stroop task performance in an immersive virtual reality environment in the same way as the real presence of the experimenter in the study conducted under laboratory conditions by Pascal Huguet and his colleagues was not confirmed.

However, the analysis of the obtained results at the descriptive level may shed some light. The average Stroop interference values obtained in each group are presented in Chart 1 (error bars represent 95% confidence intervals).

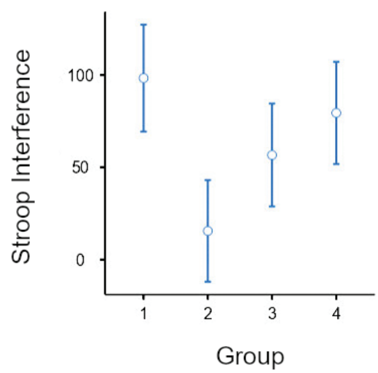

Chart 1. Average Stroop interference values obtained in each group

Source: own study

In addition to the statistically significant difference in average interference values, observed between Groups 1 and 2, there is a decrease in these values in Groups 3 and 4 compared to Group 1. At the same time, the average interference value in Group 3 is lower than in Group 4. This means that there were trends (not statistically significant) consistent with the assumptions underlying the research questions and hypotheses formulated in this study.

Similarly, when analyzing the average correctness coefficient values, it can be noticed that the most mistakes in determining the displayed color were made in

Group 2 and the fewest in Group 1 and Group 4. These averages are presented in Chart 2 (error bars represent 95% confidence intervals).

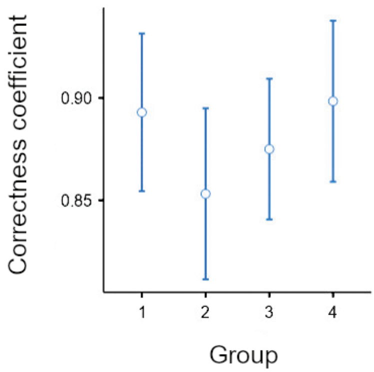

Chart 2. Average correctness coefficients obtained in each group

Source: own study

The observed trend was consistent with the findings of Huguet and colleagues (Huguet et al. 1999: 1016), who also noted a descriptive level decrease in the correctness coefficient in the presence of another person compared to the situation when the Stroop task was performed in solitude.

The lack of statistically significant results confirming Hypotheses 3 and 4 could be due to many reasons. One of the main issues was the sample size. As per the assumptions, the study was to be conducted on 4 levels, which required the division of the subjects into 4 groups, significantly reducing the size of each group (16 people). The characteristics of the tool itself could also have influenced the study results – primarily the realism of the embodied agent's appearance and behavior. The extent to which the agent resembles a real figure is largely limited by the device's technical parameters. The naturalness of the computer character's behavior is a derivative of the workshop engaged to prepare the program (individually developed animation, AI-driven responsiveness). In this study, the agent's behavior realism was not high (simulation "free posture" used in computer games) – it can therefore be cautiously speculated that more natural and responsive agent behavior would affect the facilitation effect's strength.

An interesting issue that emerged after the results analysis is the answer to the question: what impact would presenting a computer character as an avatar and informing the subject that from this character's perspective, another real user is observing their action, have on the Stroop task performance level? The subject would be visible in virtual reality as an avatar, as would the observer. Therefore, it can be assumed that there would be an interaction between real people hidden behind avatars. Should we expect a stronger social facilitation effect in such a situation than the embodied agent's effect? The study conducted within this work indirectly points to this possibility – the examiner in Group 2 served as an observer in virtual reality and although not visible, the mere awareness of this presence had a significant impact on the task performance level.

The continuous presence of VR technology in daily life, education, medicine, and research is a given. Despite current ergonomic challenges, it's expected that future advancements will enhance user comfort. VR's attractiveness extends beyond user experience, as it provides a vast data pool for network solutions. However, there are potential risks concerning data use for coercion, behavioral modeling, and surveillance (Brzezińska 2020: 33).

VR already plays a significant role in fields such as education, medicine, and research. It's important to note that research using VR also significantly contributes to the technology's development. Understanding how embodied agents or avatars affect task performance in VR can be valuable for designers of VR applications, including games, social services, education, and medicine.

This study, while not yielding unequivocal results, provides useful insights for future research and VR application design. It confirmed the possibility of creating a VR diagnostic tool equivalent to the traditional Stroop task. Moreover, it showed the Stroop interference effect occurs in immersive VR conditions. This can open up numerous research opportunities that would be impossible or difficult in conventional laboratory conditions.

The study also examined social facilitation and inhibition in VR environments, finding that awareness of being observed reduced the interference effect. The presence of an embodied agent only resulted in a non-significant descriptive reduction in interference values. Further research is needed in this area.

VR technology is still in its early stages of application in science and other non-entertainment areas. As VR technology advances, its potential applications will undoubtedly become apparent. Its ultimate use and development will rely heavily on multidisciplinary researchers working with this technology, including social scientists.

Bibliography

Baños R.M., Botella C., Perpiñá C. (1999), *Virtual reality and psychopathology*, "Cyberpsychology & Behavior", vol. 2(4), pp. 283–292.

Baron R.S. (1986), *Distraction-conflict theory: Progress and problems*, [in:] L. Berkowitz (ed.), *Advances in experimental social psychology*, Academic Press, New York.

Biocca F. (1997), *The cyborg's dilemma: progressive embodiment in virtual environments*, "Journal of Computer-Mediated Communication", 3:JCMC324, https://doi.org/10.1111/j.1083-6101.1997.tb00070.x (accessed: 17.08.2022).

Bond C.F. Jr. (1982), *Social facilitation: A self-presentational view*, "Journal of Personality and Social Psychology", vol. 42(6), pp. 1042–1050.

Bond C.F. Jr., Titus L.J. (1983), *Social facilitation: A metanalysis of 241 studies*, "Psychological Bulletin", vol. 94(2), pp. 265–292.

Botella C., Baños R.M., Garcia-Palacios A., Quero S. (2017), *Virtual Reality and Other Realities*, [in:] S.G. Hofmann, G. Asmundson (eds.), *The science of cognitive behavioral therapy*, Boston.

Brzezinska A. (2020), *Od mediów do rzeczywistości. Jakie ryzyka i szanse niesie ze sobą popularyzacja VR?*, Fundacja Centrum Cyfrowe, Warszawa.

Cottrell N.B. (1972), *Social facilitation*, [in:] C.G. McClintock (ed.), *Experimental social psychology*, Holt, Rinehart & Winston, New York.

Frankfort-Nachmias C., Nachmias D. (2001) *Metody badawcze w naukach społecznych*, Zysk i S-ka, Poznań.

Freeman D., Pugh K., Vorontsova N. (2010), *Testing the Continuum of Delusional Beliefs: An Experimental Study Using Virtual Reality*, "Journal of Abnormal Psychology", vol. 1, pp. 83–89.

Grzeszkiewicz-Radulska K., Wiktorowicz J., Grzelak M.M. (2020), *Analiza statystyczna z IBM SPSS Statistics*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

Huguet P., Galvaing M.P., Jean M., Monteil J.M., Dumas F. (1999), *Social Presence Effects in the Stroop Task: Further Evidence for an Attentional View of Social Facilitation*, "Journal of Personality and Social Psychology", vol. 77(5), pp. 1011–1025.

Milgram P., Kishino F. (1994), *A Taxonomy Of Mixed Reality Visual Displays*, "IEICE Transactions on Information Systems", vol. E77-D, no. 12, December.

Okruszek E., Rutkowska A. (2013), *Badanie kontroli wykonawczej za pomocą testu interferencji strupa u chorych na schizofrenię i osób z uszkodzeniami płatów czołowych*, "Polskie Forum Psychologiczne", vol. 18, no. 2.

Slater M., Antley A., Davison A,. Swapp D., Guger C., Barker Ch., Pistrang N., Sanchez--Vives M.V. (2006), *A Virtual Reprise of the Stanley Milgram Obedience Experiments*, "PLOS ONE", vol. 1(1): e39, https://doi.org/10.1371/journal.pone.0000039

Smelser N. (1995), *Social science: an imperfect art*, "La Revue Tocqueville", vol. 16, no. 1, pp. 3–24.

Sosnowski T., (2004), *Analiza kontrastów: między eksploracją a testowaniem hipotez*, "Przegląd Psychologiczny", vol. 47, no. 4.

Śpiewak S, (1999), *Wpływ bodźców supra i subliminalnych na efekt interferencji w zadaniu Stroopa*, "Czasopisma Psychologiczne", vol. 5, no. 2, pp. 129–139.

Tomaszewska M., Markowska, A., Borkowska A. (2010), *Test Stroopa – wartość diagnostyczna w psychiatrii*, "Neuropsychiatria i Neuropsychologia", vol. 5, no. 1, pp. 35–41.

Triplett N. (1898), *The Dynamogenic Factors in Pacemaking and Competition*, "The American Journal of Psychology", vol. 9(4), pp. 507–533.

Vakil E., Manovivh R., Ramati E., Blachstein H. (1996), *The Stroop Color-Word Task as a Measure of Selective Attension: Efficiency*, "Developmental Neuropsychology", vol. 12(3), pp. 313–315.

Wojciszke B. (2002), *Człowiek wśród ludzi. Zarys psychologii społecznej*, Wydawnictwo Naukowe Scholar, Warszawa.

Zajonc R.B. (1965), *Socjal Facilitation*, "Science, New Series", vol. 149, no. 3681, pp. 269–274.