Experimental Evaluation of Implicit and Explicit Learning of Abstract Regularities Following Socio-Emotional Interactions in Mixed Reality

Cristian Pamparău
MintViz Lab, MANSiD Research Center
Ștefan cel Mare University of Suceava
Suceava, Romania
cristian.pamparau@usm.ro

Andrei Costea
Department of Socio-Human Research
Romanian Academy
Cluj-Napoca, Romania
andrei.costea@academia-cj.ro

Răzvan Jurchiș
Department of Psychology
Babeș-Bolyai University
Cluj-Napoca, Romania
razvan.jurchis@ubbcluj.ro

Radu-Daniel Vatavu
MintViz Lab, MANSiD Research Center
Ștefan cel Mare University of Suceava
Suceava, Romania
radu.vatavu@usm.ro

Adrian Opre
Department of Psychology
Babeș-Bolyai University
Cluj-Napoca, Romania
nicolae.opre@ubbcluj.ro

***Abstract*—We present results from a controlled experiment with N=47 participants conducted in a mixed reality environment to assess explicit and implicit learning of cognitive structures instantiated by socio-emotional components. To this end, we implemented a custom version of MR4ISL, the Mixed Reality software tool for Implicit Social Learning, with a task involving colors, numbers, and emotions. Our results show evidence of explicit learning with participants' responses being attributed to conscious response bases, rules, and memory.**

Keywords*— Mixed reality, augmented reality, HoloLens, implicit learning, explicit learning, experiment, holographic avatars

I. Introduction

Mixed Reality (MR) was specified by Milgram *et al.* [1], [2] with the Reality-Virtuality (RV) continuum. More recently, Skarbez *et al.* [3] revised this definition under the consideration that pure VR is challenging to implement because interoceptive senses cannot be controlled with the current level of computer technology. Thus, Skarbez *et al.* put the RV continuum in correspondence with the perception of augmented content, from environments where "the real and virtual world objects are presented within a single display" [1, p.1322] to environments where "real world and virtual world objects and stimuli are presented together within a single percept" [3, p.4]. From this perspective, the RV continuum specifies not just a multitude of possibilities of designing MR systems, but also a multitude of MR experiences.

In this paper, we focus on the specific experience of social learning in a MR environment. We build on MR4ISL [4], a MR HoloLens application designed for psychology experiments, which we customize for a controlled experiment with N=47 participants. We report evidence of explicit social learning, and identify opportunities for future work and further development of the MR4ISL application towards observing implicit learning.

This work was supported by a grant of the Romanian Ministry of Education and Research, CCCDI-UEFISCDI, project number PN-III-P2-2.1-PED-2019-4599 (347PED/2020) within PNCDI III.

II. Related Work

We relate to prior work on Implicit Social Learning (ISL), and discuss MR applications, such as MR4ISL [4], which use holographic avatars for psychology studies.

A. Implicit Learning

We start by introducing the cognitive process of implicit learning (IL) with an example: could an intelligent agent effectively use a complex set of rules in a given domain without knowing that the domain is governed by rules? The answer is a definite yes. For example, a five-year old child can point out when an adult performs a grammatical error in spoken language. However, beyond the fact that the child cannot usually explain the grammatical rule, they do not even know that an entire domain of grammar rules even exists. Over the last decades, cognitive scientists have looked into understanding how such a task can be performed by the human mind. With a relatively general but, by no means unanimous, support, the scientific community considers the human cognitive system capable of unintentionally acquiring information from the environment in the absence of conscious awareness. In his seminal paper, Reber [5] coined the term implicit learning to refer to this family of cognitive processes.

Besides anecdotal evidence, the scientific community has developed standardized paradigms for the investigation of IL. One of the most used methods is the Artificial Grammar Learning (AGL) task [5]. In a prototypical AGL task, participants are informed that they partake in a memory experiment, and are asked to memorize several, apparently meaningless, letter strings between 5 and 9 elements long. After the acquisition phase, participants are informed that the letter strings they just memorized were not constructed at random, but using a very complex set of abstract rules, which are not disclosed. In the second part of the AGL task, i.e., the test phase, participants are presented with a new set of letter strings. They are informed that some of the strings respect the

978-1-6654-8162-5/22/$31.00 ©2022 IEEE

same set of complex rules as the ones they have encountered previously, i.e., grammatical strings, while the others respect a different set of rules, i.e., ungrammatical strings. The task is to indicate, to the best of their ability, whether each of the novel strings is grammatical or not. Evidence of IL appears when the participants classify the grammaticality of the strings above chance level, even when they subjectively feel unaware of the knowledge that guided their responses. In other terms, they perform better than chance even when they feel that they rely on an intuition or simply guess the correct response.

Scientists investigated the functioning of IL in relation to a variety of factors. For instance, Norman and Price [6] asked if the boundary conditions, i.e., the nature of the surface stimuli, influence the nature of learning. They employed AGL and designed a between-groups experiment. Half of the participants completed the acquisition phase with strings composed of letters and the other half with strings of pictures of yoga poses. Results indicated that both groups had acquired knowledge from the task but, when compared with the letters group, the learning achieved by the yoga group was more implicit. Moreover, Eitam et al. [7] showed that participants implicitly learn an artificial grammar when it is instantiated by surface stimuli depicting human faces. However, the same grammar is not learned when instantiated by pictures of buildings. Given that IL seems to be highly sensitive to the perceptual features of the stimuli upon which it operates [7], [8], our chief goal is to develop a research tool to assess the functioning of this cognitive process upon socially relevant surface stimuli in a manner as close as possible to a genuine social interaction.

B. *Mixed Reality Applications for Studies in Psychology*

The scientific literature presents several applications and systems that employ virtual or holographic avatars. For instance, Hatada et al. [9] developed "Double Shellf", a VR application for interacting with virtual avatars. They reported intense eeriness when the virtual avatar was acting autonomously. Putra et al. [10] implemented a holographic avatar for multimodal conversation. Starting from the premise that it is unknown if an avatars' appearance can also influence the user's psychological response to physical exercises, Kocur et al. [11] examined psychological and perceptual responses to athletic avatars while cycling in VR. Shao et al. [12] implemented ASL teaching in an immersive learning environment featuring a virtual avatar. Kocur et al. [11] were also interested in analyzing the effects of self and external perceptions of avatars on cognitive task performance in VR. To the best of our knowledge, there are no MR applications except MR4ISL [4] that employ holographic avatars to facilitate implicit learning. Due to its high relevance to our work, we present MR4ISL in detail in the next subsection.

C. *MR4ISL*

Pamparău et al. [4] introduced MR4ISL, the Mixed Reality tool for Implicit Social Learning, a HoloLens 1st generation application designed to examine the psychological aspects involved by implicit social learning. MR4ISL follows the principles of implicit and explicit learning of socio-emotional information [8], and implements voice and gesture-based input to enable participants from controlled studies to interact with virtual avatars in MR. In a follow-up work, Pamparău et al. [13] introduced XR4ISL by porting MR4ISL to a HTC HMD. They used XR4ISL to discuss differences of conducting experiments in MR and VR; see [13] for more details.

III. Experiment

A. *Objective*

We conducted an experiment to assess learning of cognitive structures instantiated by socio-emotional components in a MR setup with three goals: (1) creating an immersive environment saturated in structural regularities that are likely to be learned through repeated exposure, (2) evaluating whether the environment induces learning, and (3) evaluating whether the environment induces implicit learning.

B. *Participants*

A number of N=47 participants (m_{age}=19.54, SD=0.83), all psychology undergraduate students, underwent this research in exchange for partial course credits. Participants signed an informed consent to participate in the experiment and for their anonymized data to be made publicly available. This research received the approval of the Babeș-Bolyai University's institutional Ethics Committee, and complied with the 1964 Declaration of Helsinki and its later amendments.

C. *Apparatus*

We implemented a custom version of MR4ISL using the second generation Microsoft HoloLens HMD featuring an ARMv8 architecture, 65GB UFS 2.1 flash and 4GB LPDDR4x DRAM memory, and running Windows 10. We used Visual Studio 2019, Unity3D, and the Windows SDK for Windows 10 to implement our custom version of MR4ISL as a C#/C++ Universal Windows Platform (UWP) application. Gesture recognition was implemented with the HoloLens built-in technique for detecting touch gestures.[1] Voice recognition was implemented as a C# script using the HoloLens built-in feature.[2] The experiment was organized in three phases: *training*, *acquisition*, and *testing*. Next, we discuss each phase in detail.

The training phase. The purpose of training was to familiarize participants with the MR environment and interacting with virtual objects. We ensured that participants acquired this skill by presenting them with a structured sequence of events. Audio instructions were provided as guidance during the training phase. For example, the participants were asked to look at their hands in MR and notice the augmentations (Figure 1, left) represented by blue spheres from the top of their index fingers, which were used to touch and interact with MR holograms. Next, the participants followed a brief interaction exercise during which they had to touch/select one cube and then touch/select five cubes in a particular order (Figure 1, middle).

[1] https://docs.microsoft.com/en-us/dynamics365/mixed-reality/guides/authoring-gestures-hl2basic-actions-and-gestures-to-know
[2] https://docs.microsoft.com/en-us/windows/mixed-reality/design/voice-input

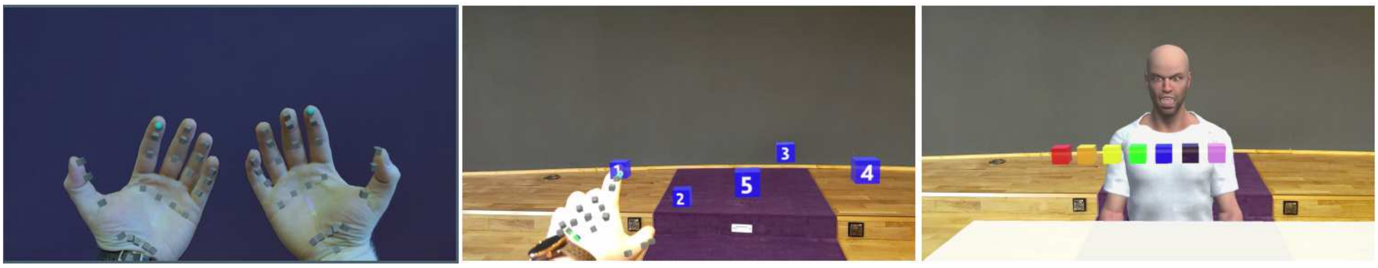

Fig 1. *Left*: The overlayed augmented version of hand; *Middle*: Five blue cubes numbered from 1 to 5 to appear in scene in training phase; *Right*: The answer options for participants.

Fig 2. Emotional facial expressions presented to the participants.

The acquisition phase. The cover story was presented to the participants, i.e., our experiment aims to investigate how colors assist people in adjusting their emotions. Participants were further instructed that they were interacting with Kevin, a virtual avatar that can experience several emotional states. Kevin was initially depicted with a neutral emotion, then displayed the seven emotions and corresponding facial expressions employed in our experiment; see Figure 2. Each facial expression was named to the participant during the audio instructions. The main task followed. Participants were informed that Kevin would change his emotional state only as a reaction to the color he is being shown, and the participant's task was to figure out Kevin's preferences for colors to make him maintain a calm emotional state (neutral facial expression) in as many trials as possible. In the next instructional step, we familiarized participants with the manner in which they could show Kevin colors. Seven colored cubes (Figure 1, right) were displayed in the MR scene, which participants could select with the index finger. Participants completed a trial test to which Kevin reacted with a preprogramed transition from Intense Anger to Low Joy. At this point, participants were reminded that (i) the change in Kevin's expression is a consequence of the selected color, and (ii) Kevin should reach the neutral emotional state in as many trials as possible. In the last instructional step, participants were informed that visual feedback was available over Kevin's right shoulder: (i) a green symbol when Kevin gets in neutral state as a result of the participant's choice, (ii) a red symbol with the text "repeated response" each time they answer with the same choice on consecutive trails, (iii) a red symbol with the text "you have X seconds left" when the participant does not choose an answer within seven seconds from the beginning of the trial, and (iv) a red symbol with the text "slow answer" if they took more than ten seconds to answer. When the participants felt prepared, they initiated the experimental task by pressing a virtual button. The acquisition task consisted in 10 blocks of 30 trials each separated by 30-second breaks. Unknown to the participants, interactions with Kevin were structured by an abstract rule represented by a looped numerical sequence. The starting point of the sequence was set at position 0, the locus where we arbitrarily placed Kevin's facial expression of intense anger and participants' "red" response option. The starting point of the sequence was constant throughout the task. The length of the sequence was determined by a mathematical equation, e.g., to determine the avatar's expression in the 4th trial:

$$S_{t.4} = 0 + [S_{t.3} + (P.Resp._{t.3} - S_{t.3})] \quad (1)$$

where $S_{t.4}$ denotes the avatar's state in the 4th acquisition trial, "0" denotes the starting point of the sequence, $S_{t.3}$ denotes the avatar's state in the 3rd trial, and $P.Resp._{t.3}$ the response given in the 3rd trial. For a detailed description of our implementation of this equation, see Costea *et al.* [8]. Participants were asked to keep the avatar in the neutral emotional state in as many trials as possible. Thus, if participants acquire knowledge from the task, we expect to detect an increase in the number of trials in which the avatar displays the target state.

The testing phase. Participants responded to a task composed of 28 trials. For each trial, one of the seven possible facial expressions was presented, and the participants had to pick a color they thought would regulate Kevin's facial expression in the neutral state. We assessed the implicit/explicit character of learning with subjective measures of awareness. As the psychological literature indicates [14], [15], participants have sometimes a relatively clear idea of what the correct answer is based on a rule or reason they have learned and which they can consciously remember, i.e., a phenomenology that typically occurs in explicit learning. Quite often, however, the participants have just a feeling that a certain answer is correct, but do not know what their feeling is based on. In other cases, participants have no idea of the correct answer, and try to guess it. If participants perform better than expected by chance, this finding is indicative of an implicit learning process.

To prevent learning from the test trials, participants received no feedback about the correctness of their responses. Instead, they were asked to answer a forced-choice question with four options (Guess, Intuition, Rule and Memory) to indicate the basis for their response:

a. *Guess*. Your answer had no basis whatsoever. You could have just as well flipped a coin to decide.

b. *Intuition*. You felt that your answer was correct, but you have no idea why.

c. *Rule*. Your answer was based on a rule (or on a fragment of a rule) that you know consciously and can describe.

d. *Memory*. Your answer was based on the fact that you consciously remember that by responding with that color you were bringing Kevin in the neutral state.

Responses were given via voice input with no time limit.

D. Procedure

The experiment was conducted in a controlled room of the Cognitive Psychology Laboratory, Babeș-Bolyai University. Distracting stimuli, e.g., background noise and ambient light, were kept constant throughout the data collection process. Given the special epidemiological circumstances (SARS-COV-2), we implemented several specific measures for the safety of our participants: only two people were present at any time in the room (i.e., the participant and the experimenter), sanitary masks were worn, the equipment was sanitized, and the space ventilated between the sessions. Participants gave their written informed consent and were assigned anonymized codes. At the end of the experiment, which took approximately 50 minutes, participants were debriefed.

E. Design

We implemented a within-group design with repeated measures. The amount of learning induced by the task was measured with the number of trials in which participants regulated the avatar to the neutral state. Explicit learning was measured with the difference between participants' accuracy in the test phase and the chance level in the trials in which they indicated they relied on explicit decision strategies. Implicit learning was measured with the difference between participants' accuracy in the test phase and the chance level in the trials where they relied on implicit decision strategies.

IV. Results

A. Evidence of Learning

A one-way, repeated measures ANOVA revealed a significant effect of practice on the number of on-target trials, $F_{(9,57)}=1.97$, $p=.041$, $\eta^2_p=.03$, indicating that participants improved their ability to control the avatar's emotional state as the task progressed.

B. Evidence Regarding Implicit and Explicit Learning

Following the practice from the scientific literature [6], [16], [17], we collapsed the test phase responses attributed to Guess and Intuition to create implicit attribution scores. Similarly, we collapsed responses attributed to Rules and Memory to create explicit attribution scores. Conforming to the general pattern of results from the literature, more than half of the responses were attributed to implicit response bases: 57.75% implicit *vs.* 41.56% explicit response attributions. To analyze the type of learning induced by our task, we calculated the chance level to 0.142 (participants had seven response options of which only one was correct). A one-sample *t*-test indicated that responses attributed to conscious response bases (i.e., Rule and Memory) were significantly above chance, $t_{(55)}=5.57$, $p<.001$, $d=-0.744$. This result indicates that participants acquired a significant amount of explicit knowledge from the task. However, a second *t*-test indicated that responses attributed to unconscious structural bases (i.e., Guess and Intuition), were not significantly above chance, $t_{(56)}=-1.95$, $p=.97$, $d=-0.258$, $B_{h(0,.029)}=0.16$, showing that unconscious knowledge was not acquired.

V. Discussion

Different from our expectations, the task generated only explicit and not implicit social knowledge. While explicit knowledge does play a determinant role in our social functioning, the current version of the task did not capture the entire complexity of knowledge acquisition within social interactions [6]. We believe there are two causes behind the exclusively explicit character of learning observed in our task. First, we presented participants with a complex regularity, which could have overloaded the conscious processing capacity and favored IL. In a future experiment, the complex regularity could be broken down in simpler micro-regularities (e.g., if the avatar's expression is "X," then response "Z" brings it to the target state). Second, as opposed to how realistic social situations unfold, the mapping between the avatar's expression and the participants' response was completely arbitrary, which could have made the task feel disfluent or unnatural. A sense of disfluency has been shown to prompt participants to fully mobilize their conscious processing resources (e.g., attention, working memory [18]).

VI. Conclusion

We implemented a custom version of the MR4ISL application for HoloLens to evaluate implicit and explicit social learning in a MR environment. While our experimental task successfully induced learning, learning was exclusively explicit in nature. Future work will address variations of our experimental task and corresponding customizations of the MR4ISL software application to investigate the phenomenon of implicit learning in simulated social contexts instantiated by various socioemotional components. Also, future versions of MR4ISL [4] and XR4ISL [13] will consider design aspects of the user experience [19] delivered to the study participants during their immersion in the mixed/extended reality ISL environments.

Acknowledgment

This work was supported by a grant of the Romanian Ministry of Education and Research, CCCDI-UEFISCDI, project number PN-III-P2-2.1-PED-2019-4599 (347PED/2020) within PNCDI III.

References

[1] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," IEICE Transactions on Information and Systems, vol. E77-D, no. 12, pp. 1321–1329, December 1994.

[2] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum," vol. 2351, 1995.

[3] R. Skarbez, M. Smith, and M. C. Whitton, "Revisiting Milgram and Kishino's reality-virtuality continuum," Frontiers in Virtual Reality, vol. 2, p. 27, 2021.

[4] C. Pamparău, R.-D. Vatavu, A. R. Costea, R. Jurchis,, and A. Opre, "MR4ISL: A Mixed Reality System for Psychological Experiments

Focused on Social Learning and Social Interactions," in Companion of the 2021 ACM SIGCHI Symposium on Engineering Interactive Computing Systems, ser. EICS '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 26–31.

[5] A. S. Reber, "Implicit Learning of Artificial Grammars," Journal of Verbal Learning and Verbal Behavior, vol. 6, no. 6, pp. 855–863, 1967.

[6] E. Norman and M. Price, "Social intuition as a form of implicit learning: Sequences of body movements are learned less explicitly than letter sequences," Advances in cognitive psychology / University of Finance and Management in Warsaw, vol. 8, pp. 121–31, 06 2012.

[7] B. Eitam, R. Glass-Hackel, H. Aviezer, Z. Dienes, R. Shoval, and E. T. Higgins, "Are task irrelevant faces unintentionally processed? implicit learning as a test case," J. Exp. Psychol. Hum. Percept. Perform., vol. 40, no. 5, pp. 1741–1747, oct 2014.

[8] A. Costea, R. Jurchis, , A. Cleeremans, L. Visu-Petra, A. Opre, and E. Norman, "Implicit and Explicit Learning of Socio-Emotional Information in a Dynamic System Control Task," PsyArXiv, 09 2020.

[9] Y. Hatada, S. Yoshida, T. Narumi, and M. Hirose, "Double shellf: What psychological effects can be caused through interaction with a doppelganger?" in Proceedings of the 10th Augmented Human International Conference 2019, ser. AH2019. New York, NY, USA: Association for Computing Machinery, 2019.

[10] R. A. Putra, I. Musyaffa, A. Asfarian, and D. A. Ramadhan, "Curhat: Telling your story to a multimodal conversation bot to alleviating the stress caused by pandemic fatigue," in Asian CHI Symposium 2021, ser. Asian CHI Symposium 2021. New York, NY, USA: Association for Computing Machinery, 2021, p. 177–179.

[11] M. Kocur, F. Habler, V. Schwind, P. W. Wo´zniak, C. Wolff, and N. Henze, "Physiological and perceptual responses to athletic avatars while cycling in virtual reality," in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021.

[12] Q. Shao, A. Sniffen, J. Blanchet, M. E. Hillis, X. Shi, T. K. Haris, J. Liu, J. Lamberton, M. Malzkuhn, L. C. Quandt, J. Mahoney, D. J. M. Kraemer, X. Zhou, and D. Balkcom, "Teaching american sign language in mixed reality," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 4, no. 4, dec 2020.

[13] C. Pamparau, R.-D. Vatavu, A. R. Costea, R. Jurchis, and A. Opre, "XR4ISL: Enabling Psychology Experiments in Extended Reality for Studying the Phenomenon of Implicit Social Learning," ser. MUM 2021. New York, NY, USA: Association for Computing Machinery, 2021, p. 195–197.

[14] Z. Dienes and R. B. Scott, "Measuring unconscious knowledge: distinguishing structural knowledge and judgment knowledge," Psychological Research, vol. 69, pp. 338–351, 2005.

[15] R. B. Scott and Z. Dienes, "The conscious, the unconscious, and familiarity," J. Exp. Psychol. Learn. Mem. Cogn., vol. 34, no. 5, pp. 1264–1288, sep 2008.

[16] A. Costea, "The relationship between implicit learning of cognitive structures with socio-emotional components and subthreshold autistic traits," Journal of Evidence-Based Psychotherapies, vol. 18, pp. 131–142, 09 2018.

[17] Q. Fu, Z. Dienes, and X. Fu, "Can unconscious knowledge allow control in sequence learning?" Consciousness and Cognition, vol. 19, no. 1, pp. 462–474, 2010.

[18] A. Alter, "The benefits of cognitive disfluency," Current Directions in Psychological Science, vol. 22, pp. 437–442, 12 2013.

[19] C. Pamparău and R.-D. Vatavu, "A Research Agenda Is Needed for Designing for the User Experience of Augmented and Mixed Reality: A Position Paper", in 19th International Conference on Mobile and Ubiquitous Multimedia (MUM 2020). Association for Computing Machinery, New York, NY, USA, 323–325.