

Learning art history through embodied experience with virtual reality

Joseph K. Brown

Lafayette College

Eric Hupe

Lafayette College

Luis F. Schettino

`schettl@lafayette.edu`

Lafayette College

Research Article

Keywords: Virtual Reality, Art History, Embodied Cognition, Embodied Learning

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10817171/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Art History education has traditionally relied on 2D images, limiting the ability of students to fully experience the intent of the artists who created the works. This is especially so in the case of frescoed spaces where a number of elements such as relative position, size and lighting are particularly important. Immersive Virtual Reality (IVR) provides an opportunity for students to experience the original spaces while also providing a sense of presence and enhancing the potential for embodied learning.

Sixteen undergraduate participants with no art history background took part in this study, which mimicked a traditional teaching module consisting of four sessions. The participants were randomly assigned to an IVR condition or a 2D condition. Participants were then tested on their explicit recall but also on non-explicitly mentioned pictorial details. IVR participants demonstrated higher performance on a final test but were also able to locate and elaborate on non-explicitly mentioned information. Our data suggest that IVR facilitated learning of explicitly taught information but also enhanced their perception for spatial relationships, allowing them to better understand the use of architectural elements by Renaissance artists. These results support the potential of IVR as an effective pedagogical tool for art history instruction, particularly for artworks designed to be experienced in three-dimensional space.

Introduction

Immersive virtual reality (IVR) has, in the last decade, emerged as a practicable and advantageous teaching and learning technology [1]. Numerous studies have shown that IVR promotes enhanced learning and student satisfaction due to its ability to present information through novel, interactive means [2, 3, 4]. Similarly, many studies have shown that IVR enhances student enjoyment and motivation, leading to enhanced learning [5]. Nonetheless, beyond its motivational aspects, IVR also improves student comprehension and facilitates ‘sense-making’ under the right conditions. Antonietti & Cantoia [6] have noted that thanks to IVR’s ability to produce immersive three-dimensional experiences, students have the opportunity to experience directly the physical properties of the space in which they find themselves and, by doing so, to recognize the implications of such properties. For example, Chan and colleagues [7] tested the learning of architectural history in both Augmented Reality (AR) and Virtual Reality (VR). They found that, compared to the AR students, VR students not only improved their learning but also were able to gauge and appreciate 3D aspects such as size and proportion, both critical aspects of experiential learning. This ability to move in space at a natural pace while perceiving the environment with the appropriate visual flow provides the opportunity to engage in *embodied learning*.

Embodied learning refers to the idea that bodily action during experience is part of the learning process. Briefly, embodied cognition is the notion that mental processes are fundamentally blended with the interaction of our bodies in the environment [8]. Therefore, the natural process of learning must include our active engagement with surrounding space [9] and the formation of implicit associations of potential actions to the objects around us or “affordances” [10].

Embodied cognition, thus, suggests that learning may include explicit and implicit processes. For example, when learning a new videogame, a person may learn explicitly the rules that govern the game (e.g. number of points obtained for finding a given token), but there are also a number of aspects that are learned implicitly before they become explicit in the mind (e.g. the time that elapses between one event and another or the sequence and rhythm in which certain buttons must be pressed). One of these implicit learning processes involves the ability to integrate spatial information [7, 11] into a 'cognitive map' of space through self-motion cues [12].

In the field of Art History, it is not uncommon to find artworks that are composed not only of a particular sculpture or painting, but also of the architectural space in which such works are presented. This is especially important for works of art that are still *in situ*. In those cases, the precise location of the art within the space is meant to convey information that can only be fully comprehended by wandering in the midst of that space.

Unfortunately, until recently, most Art History instructors and students had to content themselves with 2D images, preventing the perception of the visual and environmental cues that artists carefully chose to employ space (relative location, relative height, relative size, lighting conditions) as a means to transmit specific information or to evoke in the viewer particular emotions. Art History professors often struggle to describe the experience of being in the place itself [13] as 2D images distort the visual cues employed by the artists, forcing the student to 'try to imagine' the original composition.

The advent of IVR technology has given instructors the potential to facilitate the experience of being immersed in a compelling replica of the actual spaces and to provide the opportunity to learn Art History in the embodied way that was conceived by the artists who created them. Hutson & Olsen [14] have called for a closer investigation of the relationship between VR and instructional materials and quantitative measures that provide a link to course learning objectives. Importantly, the ability to provide students with an immersive experience of historically important artistic works regardless of their physical location helps advance two of the United Nations' Sustainable Development Goals, namely, access to quality education and reduced inequalities [15].

In this study, we set out to test the effect of IVR in the learning of Art History versus a traditional (2D) setting. Because we were interested not only in determining whether retention of the materials was improved (an explicit type of memory) but also the extent to which embodied learning was present, we employed virtual simulations of Renaissance era works in which the artists produced large frescoed spaces that aimed to immerse the viewer in a type of early 'virtual reality'. We surmised that an IVR experience of such artworks would provide participants with the opportunity to learn notions in an explicit manner, but that other information would be learned implicitly as the participants moved within the simulated spaces. Furthermore, we modeled our experimental design as a 'learning module' by including lectures followed by brief tests in a series of three sessions occurring within the span of one week with a 'final test' taking place at least 2 weeks but not more than 4 weeks after the third session.

In order to test for these two types of learning, we had participants experience two different frescoed spaces across four sessions either in IVR or in the traditional 2D space used in Art History courses while listening to recorded scripts describing relevant artistic and architectural details. Subsequently, we conducted three types of tests: 1) a traditional written test to assess explicit recall, 2) a verbal test to determine spatial perception processing and 3) a visual search task in which the participants were asked to find particular pictorial details included in the recorded lessons plus details never mentioned in the script. For this last test, we reasoned that the ability to detect non-explicitly mentioned details would be enhanced in those participants who were able to create an integrated cognitive map of the painted spaces and who more fully experienced the spatial information embedded in each space.

Our results indicate that participants' explicit recall was improved in VR relative to 2D space but also that VR participants were able to integrate spatial information to better index the space around them. Moreover, we found evidence that VR participants also employed the available visual cues to support their understanding of the works of art to which they were exposed.

Methods

Participants

Sixteen undergraduate participants were recruited. None of them were Art or Art history majors (Table 1). All participants had normal or corrected-to-normal eyesight. Six identified as male, nine as female and one as non-binary. Mean age was 21.4 years. Videogame experience as reported by the participants was similar across the conditions with 4/8 VR participants reporting extensive experience and 3/8 in the 2D condition reporting similar experience. Participants signed Informed Consent forms approved by Lafayette College's IRB (ref. AY2324-16) and verbally approved to be audio recorded during the tests. Participants were compensated with two \$20 gift cards.

Table 1

Demographics of participants. Videogame lifetime experience: Never = 0, 1 = 1–5 times, 2 = 5–10 times, 3 = 10+ times.

Participant ID	Group	Video game experience	Age	Gender	Major
BR	VR	1	19	Male	Economics
CK	VR	1	22	Female	Psychology
CN	VR	2	19	Non-binary	English
ES	VR	3	22	Male	Biology
MW	VR	1	21	Female	Geology
PF	VR	3	21	Male	Computer Science
TL	VR	3	22	Male	Engineering
TW	VR	3	20	Female	Engineering
AH	2D	1	21	Female	Biology
AS	2D	1	21	Female	Neuroscience
CW	2D	3	21	Male	Engineering
EB	2D	1	22	Female	Engineering
KC	2D	3	24	Male	Chemistry
MD	2D	3	21	Female	Engineering
PW	2D	1	21	Female	Engineering
RS	2D	0	21	Female	Neuroscience

Procedure

Group assignment

Participants were randomly assigned into the two groups. Each participant went through four separate sessions. Participants underwent each of the sessions individually.

Experimental setup

Virtual Reality. Participants were fitted with a Meta Quest (Meta, Inc., Menlo Park, CA) Pro VR HMD and the VR simulations were run on an AlienWare Aurora R15 desktop computer (Dell, Inc., Round Rock, TX). All participants completed two VR tutorials/introductory materials. These were Oculus' *First Steps* and *First Contact* and were run natively on the HMD. The HMD was tethered to the PC using Quest Link.

Audiovisual setup. The 2D group heard the audio recordings through a Bluetooth speaker (Anker Innovations, Hunan, China) and observed the artwork in the form of high resolution 2D photographs (Artstor.org) of the spaces displayed on a 24-inch (diagonally) computer screen (Dell inc., Round Rock, TX). The participant sat at about 50 cm from the computer screen and selected each image by clicking through 3 or 4 browser tabs. Participants had the ability to scroll around the images and zoom in at their discretion using the computer mouse and keyboard.

The VR group was exposed to immersive VR recreations either running natively in the UNITY engine (Bridal Chamber, see below) or run in a web-based social VR platform (Hall of Giants, see below)[1]. Both platforms had the same functionality. VR participants could freely locomote with both physically and with their controllers within a 4.6m X 3.5m space. The cable to the headset was hoisted on the ceiling through a series of pulleys to ensure that the participants did not trip over the cable. Movement sourced from the controllers consisted of both ground-based walking and 'flying' (using the joystick in the game controllers). The participants had the ability to fly and touch the ceiling if so desired.

Voice recordings. Participants' answers to the Scavenger Hunts and Tests 1 and 2 (see below) were recorded via a Blue Yeti microphone (Logitech, Lausanne, Switzerland) using the audio mixer Audacity (www.audacityteam.org). Recordings were then transcribed using automated AI transcription (Otter.ai).

Experimental Design

Participants underwent four sessions (Fig. 1). During **session 1**, participants were asked a series of questions focused on past VR and video game exposure. Gathering information on VR and video game exposure helped elucidate whether participants were used to using joystick and button-based controllers. Information on VR and video game exposure was collected because it may also hint that the participant had a more fluid understanding of how to interact in simulated 3D environments. An independent samples t-test showed that the groups did not differ in their previous experience to videogames ($t_{14}=0.91$, $p = 0.376$).

Training and testing procedures

There were three different types of activities during the sessions: **Lessons**, **Scavenger Hunts** and **Tests**. Learning during each **Lesson** was assessed via a **Scavenger Hunt** and a **Final Test**.

The *Lessons* consisted of pre-recorded audio files containing architectural, historical and artistic descriptions of the different frescoes to which the participants were exposed during the experiment. Each lesson included breaks for participants to view the artwork at their own pace. Topic 1 was the Bridal Chamber (BC, *Camera degli Sposi*, Andrea Mantegna, *Palazzo Ducale*, Mantua), Topic 2 was the Hall of Giants (HG, *Sala dei Giganti*, Giulio Romano, Palazzo Te, Mantua) (Fig. 2).

The Final Test (Session 4) was taken at least two weeks (but not more than four) after completion of the three training sessions in order to test for recall and understanding. The test was divided into two halves. The first half required the participants to analyze previously unseen examples of Pre- and Renaissance era art (see table in appendices) and answer open-ended questions about the art works. The second half had participants answer questions about the two pieces they had lessons on (The Bridal Chamber and the Hall of Giants). This test was completed on a computer and the answers were typed by each participant into a document (see Supplementary_information.pdf file for examples).

During a *Scavenger Hunt* (SH), participants were given free rein to explore the 2D images or VR spaces, according to their group membership. Each one of the two SH consisted of nine timed probes that asked the participant to locate specific details within the paintings. The time to find each item was recorded. Researchers verbally supplied the participants with the objects one at a time and responses were audio recorded and subsequently transcribed for analysis.

The goal of the SH was to gauge perceptual learning outside of the explicit information provided during the *Lessons*. Our hypothesis was that participants in the VR group, by virtue of experiencing the artwork in an immersive environment, would not only acquire explicit knowledge as guided by the narrative, but would also undergo additional implicit learning facilitated by spatial integration of the sensory inputs produced by their own motion within the VR space.

The Spatial Perception Tests consisted a set of three Spatial Perception Questions (SPQ) asked following each SH and were conducted when the participants had (2D group) the images in front of themselves or (VR) were still immersed in the space. The SPQ asked participants to recall information provided explicitly during the lessons. However, they were designed to determine whether the participants had an experiential perception of space while looking at the images. For example, the ceiling of the Bridal Chamber includes a simulated oculus through which a group of people can be seen (Fig. 3). The people are clearly amused at the fact that there is a heavy planter precariously suspended over the viewer. The SPQ for that detail was: *Why are the figures in the oculus giggling?* Even though the lessons include the following description: “a large potted plant extending into the oculus with a thin crossbar support. Seemingly amused, the figures look at the poorly balanced basket and down on the viewer”. We hypothesized that recall of that information would be enhanced in participants who actually saw the planter suspended over their heads. (See Supplementary_information.pdf for SPQs). Participants were not provided with any feedback during the tests other than clarifying the questions.

Statistical Analysis

Verbal answers to **Tests 1** and **2** were transcribed using automated AI transcription. Text files were then analyzed manually and adjusted for accuracy.

Final tests

The final tests were graded by one of us (E.H.) who is a professor of art. The test grader was blind to the identity and condition of the participants throughout the analysis. Grades were assigned as letter grades and these were then converted to numerical values taking the central value of the interval traditionally designated by the letter grade.

Performance on a written test can vary due to a number of factors including interest for the topic, motivation, previous exposure et cetera. Our participants, however, were selected to have no previous leaning towards the topic of Art History. Because we expected participants in the VR condition to exhibit improved retention of the information, our hypothesis was that participants in the VR condition would attain higher grades in the final tests than participants in the traditional (2D) condition. Therefore, we used a directional independent samples t-test with an alpha of 0.05.

Spatial Perception Tests

A researcher blind to group membership (L.S.) analyzed the transcripts and gave one point for each correct answer per participant. Because two of the data sets for participants in the 2D group (AS, CW) for the Hall of Giants were lost due to a technical error, our statistical test was conducted on the remaining six participants. Total correct answers were calculated and the counts entered into a one-tailed Mann-Whitney test for independent samples.

Scavenger Hunts

The purpose of the scavenger hunts (SH) was to assess the integration of non-explicitly provided information in the narrative scripts. We asked participants to locate nine specific details (probes) in each of the two spaces (see [Supplementary_information.pdf](#) for examples of SH probes) while we timed their performance. There were 5 non-explicitly mentioned (NEM) and four explicitly mentioned (EM) probes per space. An experimenter confirmed that each one of the probes was solved by either monitoring the behavior directly on the screen (2D group) or viewing the perspective of the HDM in an associated screen (VR group). Probes with an average solving time exceeding 120 seconds were eliminated from the analysis.

During SH, participants in the 2D group had to toggle between three (BC) and four (HG) tabs to move through the different images while those in the VR group locomoted and directed their gaze towards the walls or ceiling of the simulated space.

Two of the probes (one NEM probe in each space) resulted in average solving times that exceeded 2 minutes and were eliminated from the analysis. The analysis, therefore, included 4 NEM and 4 EM probes per space.

We had a series of hypotheses for this part of the study:

Hypothesis 1

EM defined probes will be found quicker than NEM defined probes by all participants. This is because EM details would have been noticed and 'indexed' as the script was heard.

Hypothesis 2

the temporal difference between groups in the EM probes provides a measure of the 'cost' of using a different form of search. If there was an advantage to searching for items within the VR relative to the 2D setup, the difference between the groups in the EM probes can be used as a measure of that advantage.

Hypothesis 3

the temporal difference for the 2D group across the EM and NEM tasks, provides a measure of the ease/difficulty in attaining implicit learning under traditional learning conditions.

Hypothesis 4

the temporal difference between groups in the NEM probes measures the contribution of VR to the processing of *information not directly mentioned* in the narrative scripts. VR participants, by virtue of their exposure to the seamless experience of space, should be better at transferring not directly attended information into explicitly-guided behavior.

In order to test these hypotheses, we conducted Mixed Factors ANOVAs on the temporal latencies required by the participants to solve the SH probes with **Location** (BC, HG), **Explicitness** (EM, NEM) and **Probe** (P1, P2, P3, P4) as within-group factors and Group (VR, 2D) as between-groups factor. We were mostly interested in the main effect of **Explicitness** (hypothesis 1) and the interaction of **Explicitness * Group** (hypothesis 2) as effects or interactions of Location or Probe depended mostly on the specific details of each probe. Significant interactions of **Explicitness * Group** were followed with Bonferroni-Holm post hoc tests.

Because two of the data sets for the 2D group in the HG space (AS, CW) were lost due to a technical error, we had to drop the data for the same participants from the BC space. In order to check whether this would affect our results, we conducted a Mixed Factors ANOVA on the data for the 2D group in the BC space comparing their results with and without those two data sets. The Mixed Factors ANOVA for the 2D group showed no significant difference between the groups ($F_{1,12}=0.214$, $p = 0.652$), suggesting that the data with and without the two participants are equivalent (see a raincloud plot in Supplementary_information.pdf).

Results

Final tests

In order to test the learning of explicitly mentioned information, participants took a traditional written test at least two-weeks after the training. Table 2 shows the grades and percent equivalences for each

participant.

Table 2
Grades and percent equivalents in the Final Test per participant per group.

Participant	VR	Value	Participant	2D	Value
BR	A	94.5	AH	A	94.5
CK	C+	78	AS	A-	91
CN	A	94.5	CW	D	64.5
ES	A	94.5	EB	C	74.5
MW	B	84.5	KC	D	64.5
PF	A-	91	MD	B-	81
TL	C	74.5	PW	B+	88
TW	A-	91	RS	D	64.5
	Mean	87.8		Mean	77.8
	SD	7.9		SD	12.6

A one-tailed independent samples *t* test was conducted to test the hypothesis that participants in the VR condition would have better recall of the information provided during the lessons relative to those in the 2D condition. The results $t_{14}=1.9$, $p = 0.039$, $d = 0.95$ indicate that the final test grades of participants in the VR condition were significantly higher than those in the 2D condition, suggesting that they were better at recalling and applying the information learned during the training sessions two weeks later (Fig. 4).

Spatial Perception Tests

A one-tailed Mann-Whitney test for independent samples was conducted on the total correct responses per participant across the two tests. A maximum of 6 points could be obtained. The test showed ($U = 40.5$, $p = 0.016$, $r = 0.69$) that participants in the VR group ($M = 4.8$, $SD = 0.83$) were better able to recall the space-related information than participants in the 2D group ($M = 3.0$, $SD = 1.5$). This provides evidence that the participants in VR had processed the spatial information in such a way that their recall was enhanced by virtue of their embodied experience.

Scavenger Hunts

A Mixed Factors ANOVA on the temporal latencies required by our participants to solve the SH probes with **Location** (CP, HG), **Explicitness** (EM, NEM) and **Probe** (P1, P2, P3, P4) as within-group factors and **Group** (VR, 2D) as between-groups factor was conducted to test our hypotheses. The test showed a

significant main effect of Explicitness ($F_{1,12} = 16.4$, $p = 0.002$, $\eta^2_p = 0.58$), supporting hypothesis 1 that EM probes were solved faster than NEM probes by all participants, regardless of group membership (Fig. 5)

There was also a significant interaction of **Explicitness * Group** ($F_{1,12}=12.95$, $p = 0.004$, $\eta^2_p = 0.52$), indicating that the latency difference between EM and NEM probes depends on group membership. A Bonferroni-Holm post hoc test revealed significant differences between the EM and NEM probes in the 2D group (Hypothesis 3, $p = 0.003$) and between the NEM probes for both groups (Hypothesis 4, $p = 0.016$). There was no significant difference ($p = 0.075$) between the groups in the EM probes, suggesting that the cost of working in the 2D setup was relatively low (Hypothesis 2). Figure 6 shows these data.

Discussion and conclusion

In this study, we set out to assess the effect of IVR in the learning of Art History. To that effect, we had two groups of participants listen to recorded lectures while experiencing artworks either in IVR or in a more traditional screen-based format. Because we were interested in modeling the typical learning process at higher education institutions, we had our participants take two lessons followed by a written test at least two weeks after the second lesson.

While we were interested in assessing explicit learning in our participants, as required in the learning objectives of most College-level courses, we hypothesized that experiencing artworks in IVR would also recruit implicit processes associated with embodied cognition. In particular, we expected participants in the IVR group to be able to process spatial aspects of the artworks studied such as relative size and position thanks to the veridical experience of presence rendered possible by that technology.

Importantly, we selected as target artworks two Renaissance era frescoed spaces originally designed to produce convey particular information through the use of both architectural and pictorial details. We expected that the use of space by the artists would be better experienced by the IVR group, promoting implicit learning.

Our results showed that all participants exhibited explicit learning, with those in the IVR group displaying significantly higher scores in the final test relative to those in the 2D group. Furthermore, in the SH tests, explicitly-mentioned (EM) information was found quicker than non-explicitly-mentioned (NEM) information by both groups. On the other hand, participants in the 2D group performed worse than those in the IVR group when locating NEM details. Finally, IVR participants were better able to recall space-related information in the spatial perception questions. These findings provide evidence for enhanced learning in IVR and for a possible role of implicit processes in supporting that learning [18].

It has been suggested that implicit learning may facilitate the scaffolding of explicit processes [19] when the student is able to explore and interact with an environment where information is already embedded. While most of the work in this area has been conducted in computer simulations, we would argue that the immersion of the student in a space where architectural and pictorial information is abundant, if

properly contextualized, is able to foment the learning of explicit knowledge as well as its supporting implicit scaffold (e.g. the symbolic meaning of relative size in a painting).

In our study, following a segment of each lecture, participants were asked to take the time to explore the spaces at their own pace. The VR participants were able to 'fly over' the space while the 2D group could switch between images and zoom in and out of them freely. This methodological component was put in place to allow students to interact with the paintings once they had processed contextual information as guidance.

Limitations

A limitation of this work is its relatively small sample size. However, it must be noted that the resource-intensive nature of our study (4 sessions per participant) put constraints on the final number of participants. Nonetheless, the magnitude of our effect sizes (all medium to large strength) suggests that the methods employed produced a detectable difference between the groups without increasing within-group variability.

While it is possible that as Makransky & Petersen [20] suggest, immersive technologies support learning via affective factors such as enjoyment and interest, the same authors identify the senses of *presence* and *agency* in IVR as the main factors guiding learning. Given that our participants were specifically selected for their relative disinterest in Art History, it is difficult to disentangle the role played by novelty. Disentangling the effect of enhanced attention due to interest would be important in future studies.

Conclusions

Although it may seem intuitive that providing Art History students with an opportunity to experience art in its original setting would lead to enhanced learning, the determination of whether IVR leads to improved learning, the specific aspects of learning that are boosted by immersive technologies and the psychological mechanisms involved has proven to be a challenging process [21, 22, 23]. Nonetheless, diligent experimentation has provided some useful answers [1, 18, 20].

The present work sought to test for learning improvements of Art History through IVR and to obtain evidence for the processes involved. We found evidence for both explicit and implicit learning, suggesting the possibility that the sense of immersive presence allows for implicit scaffolding of explicit knowledge. Having said that, our experiment involved artworks that fuse architectural and pictorial elements to communicate with the viewer, making the sense of presence particularly useful in order to fully comprehend the intended message. It is unclear whether experiencing art in a virtual museum would produce similar gains or involve the same learning mechanisms [24].

As access to VR technology and online coursework becomes more widespread, exposure to simulated visualizations such as the ones described in this study can improve access to high-quality education in

underserved communities improving opportunities for learning and reducing inequalities, two of the Sustainable Development Goals of the U.N. [15]. A plausible way to further increase learning through immersive experimentation in Art History could be to employ modern technologies such as text-to-image Large Language Models to let students create their own painted architectural spaces. By providing students with a set of constraints such as building plans and patron requests, students would need to think like the artists they are studying, helping them recognize the feats of imagination and skill that renaissance artists were able to achieve.

Declarations

Funding

Partial financial support was received from the Sherman Fairchild Foundation and the David M. and Linda Roth Foundation.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author contributions

All authors contributed to the study conception and design. Material preparation and data collection was performed by Joseph K. Brown, analysis was performed by Luis F. Schettino and Eric Hupe. The first draft of the manuscript was written by Luis F. Schettino and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Authorship

The corresponding author confirms they have read the journal policies and is submitting their manuscript in accordance with those policies.

Ethics statement

Our research complies with the Declaration of Helsinki. All participants signed Informed Consent forms approved by Lafayette College's IRB (ref. AY2324-16) and verbally approved to be audio recorded during the tests. Consent to publish: not applicable.

Data availability statement

The data supporting the spatial perception tests and figures 5 and 6 are available in the supplementary materials file.

References

1. Radianti J, Majchrzak TA, Fromm J, Wohlgemant I. A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. *Comput Educ.* 2020;147:103778. <https://doi.org/10.1016/j.compedu.2019.103778>.
2. Di Serio A, Ibanez MB, Kloos CD. Impact of an augmented reality system on students' motivation for a visual art course. *Comput Educ.* 2013;68:586–96. <https://doi.org/10.1016/j.compedu.2012.03.002>.
3. Ho LH, Sun H, Tsai TH. Research on 3D painting in virtual reality to improve students' motivation of 3D animation learning. *Sustainability.* 2019;11(6):1605. <https://doi.org/10.3390/su11061605>.
4. Mulders M. Investigating learners' motivation towards a virtual reality learning environment: a pilot study in vehicle painting. In: Reality V, AIVR). (; 2020 Dec 14–18; Utrecht, Netherlands. Piscataway (NJ): IEEE; 2020. pp. 390-3. <https://doi.org/10.1109/AIVR50618.2020.00081>
5. Serna-Mendiburu GM, Guerra-Tamez CR. Shaping the future of creative education: the transformative power of VR in art and design learning. *Front Educ.* 2024;9:1388483. <https://doi.org/10.3389/educ.2024.1388483>.
6. Antonietti A, Cantoia M. To see a painting versus to walk in a painting: an experiment on sense-making through virtual reality. *Comput Educ.* 2000;34(3–4):213–23. [https://doi.org/10.1016/S0360-1315\(99\)00046-9](https://doi.org/10.1016/S0360-1315(99)00046-9).
7. Chan CS, Bogdanovic J, Kalivarapu V. Applying immersive virtual reality for remote teaching architectural history. *Educ Inf Technol.* 2022;27(3):4365–97. <https://doi.org/10.1007/s10639-021-10786-8>.
8. Shapiro L, Stolz SA. Embodied cognition and its significance for education. *Theory Res Educ.* 2019;17(1):19–39. <https://doi.org/10.1177/1477878518822149>.
9. Fugate JM, Macrine SL, Cipriano C. The role of embodied cognition for transforming learning. *Int J Sch Educ Psychol.* 2019;7(4):274–88. <https://doi.org/10.1080/21683603.2018.1443856>.
10. Gibson JJ. The ecological approach to visual perception. Boston: Houghton Mifflin; 1979.
11. Konig SU, Keshava A, Clay V, Rittershofer K, Kuske N, Konig P. Embodied spatial knowledge acquisition in immersive virtual reality: comparison to map exploration. *Front Virtual Real.* 2021;2:625548. <https://doi.org/10.3389/frvir.2021.625548>.
12. Epstein RA, Patai EZ, Julian JB, Spiers HJ. The cognitive map in humans: spatial navigation and beyond. *Nat Neurosci.* 2017;20(11):1504–13. <https://doi.org/10.1038/nn.4656>.
13. Hupe ER, Carr C. The virtual Renaissance: adopting virtual reality to transform how art history is taught. In: Scherff KD, Sobehrad LJ, editors. *Media technologies and the digital humanities in medieval and early modern studies*. Abingdon: Routledge; 2023. pp. 128–41. <https://doi.org/10.4324/9781003295082-13>.
14. Hutson J, Olsen T. Virtual reality and learning: a case study of experiential pedagogy in art history. *J Intell Learn Syst Appl.* 2022;14:57–70. <https://doi.org/10.4236/jilsa.2022.144005>.
15. United Nations. Sustainable Development Goals [Internet]. New York: United Nations; [cited 2026 Aug 19]. Available from: <https://www.un.org/sustainabledevelopment/development-goals/>

16. Mantegna A. Camera degli Sposi, view of the west and north walls, Palazzo Ducale, Mantua [fresco, 1465–1474]. Wikimedia Commons [cited 2026 Aug 19]. Available from: https://commons.wikimedia.org/wiki/File:Mantegna_-_Ducal_Palace,_View_of_the_west_and_north_walls,_0sposi2.jpg. Public domain (PD-Art).
17. Giulio Romano. Fall of the Giants, Sala dei Giganti, Palazzo Te, Mantua [fresco, c. 1532–1534]. Wikimedia Commons [cited 2026 Aug 19]. Available from: [https://en.wikipedia.org/wiki/Fall_of_the_Giants_\(Romano\)#/media/File:Giulio_Romano_-_View_of_the_Sala_dei_Giganti_\(north_wall\)_-_WGA09553.jpg](https://en.wikipedia.org/wiki/Fall_of_the_Giants_(Romano)#/media/File:Giulio_Romano_-_View_of_the_Sala_dei_Giganti_(north_wall)_-_WGA09553.jpg). Public domain (PD-Art).
18. Slater M. Implicit learning through embodiment in immersive virtual reality. In: Liu D, Dede C, Huang R, Richards J, editors. *Virtual, augmented, and mixed realities in education*. Singapore: Springer; 2017. pp. 19–33. https://doi.org/10.1007/978-981-10-5490-7_2.
19. Podolefsky NS, Moore EB, Perkins KK. Implicit scaffolding in interactive simulations: design strategies to support multiple educational goals. *arXiv:1306.6544 [Preprint]*. 2013 [cited 2026 Aug 19]. <https://doi.org/10.48550/arXiv.1306.6544>
20. Makransky G, Petersen GB. The cognitive affective model of immersive learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. *Educ Psychol Rev*. 2021;33(3):937–58. <https://doi.org/10.1007/s10648-020-09586-2>.
21. Krokos E, Plaisant C, Varshney A. Virtual memory palaces: immersion aids recall. *Virtual Real*. 2019;23(1):1–15. <https://doi.org/10.23919/iLRN47897.2020.9155160>.
22. Kalina E, Johnson-Glenberg MC. Presence and platform: effects of embodiment comparing a 2D computer and 3D VR game. In: 2020 6th International Conference of the Immersive Learning Research Network (iLRN); 2020 Jun 21–25; San Luis Obispo (CA). Piscataway (NJ): IEEE; 2020. pp. 31 – 7. <https://doi.org/10.23919/iLRN47897.2020.9155160>
23. Hurter C, Rogowitz B, Truong G, Andry T, Romat H, Gardy L, et al. Memory recall for data visualizations in mixed reality, virtual reality, 3D and 2D. *IEEE Trans Vis Comput Graph*. 2024;30(10):6691–706. <https://doi.org/10.1109/TVCG.2023.3336588>.
24. Cecotti H, Huisinga L, Pelaez LG. Fully immersive learning with virtual reality for assessing students in art history. *Virtual Real*. 2024;28(1):33. <https://doi.org/10.1007/s10055-023-00920-x>.

Footnotes

1. For a version of the virtual spaces see <https://www.zeuxisvr.com/properties>.

Figures

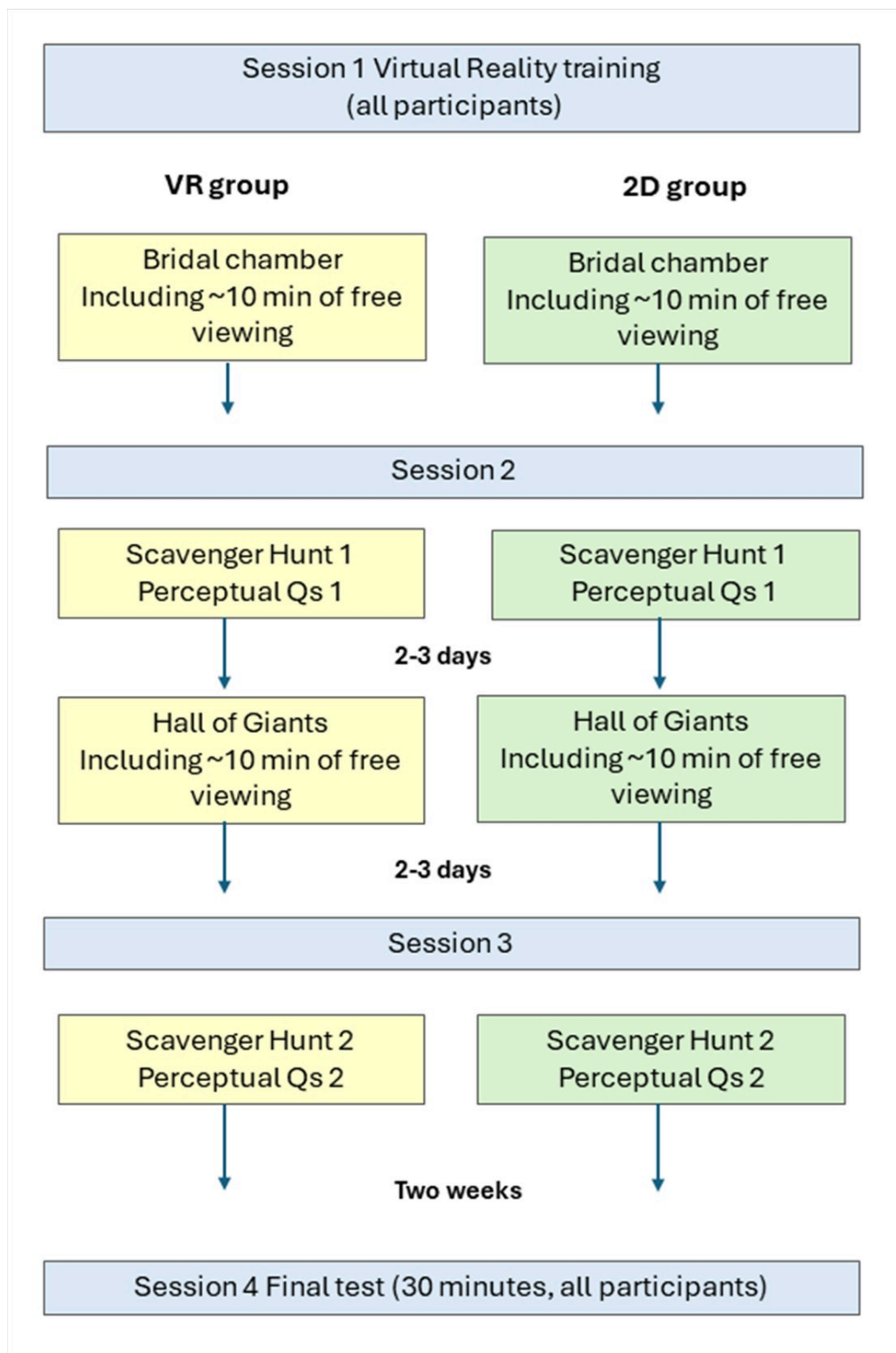


Figure 1

Experimental design



Figure 2

The art works employed in the study. **a** The Bridal Chamber; **b**the Hall of Giants. Images from Wikimedia Commons, public domain [16,17]



Figure 3

The oculus in the Bridal Chamber. Photograph by E.H.

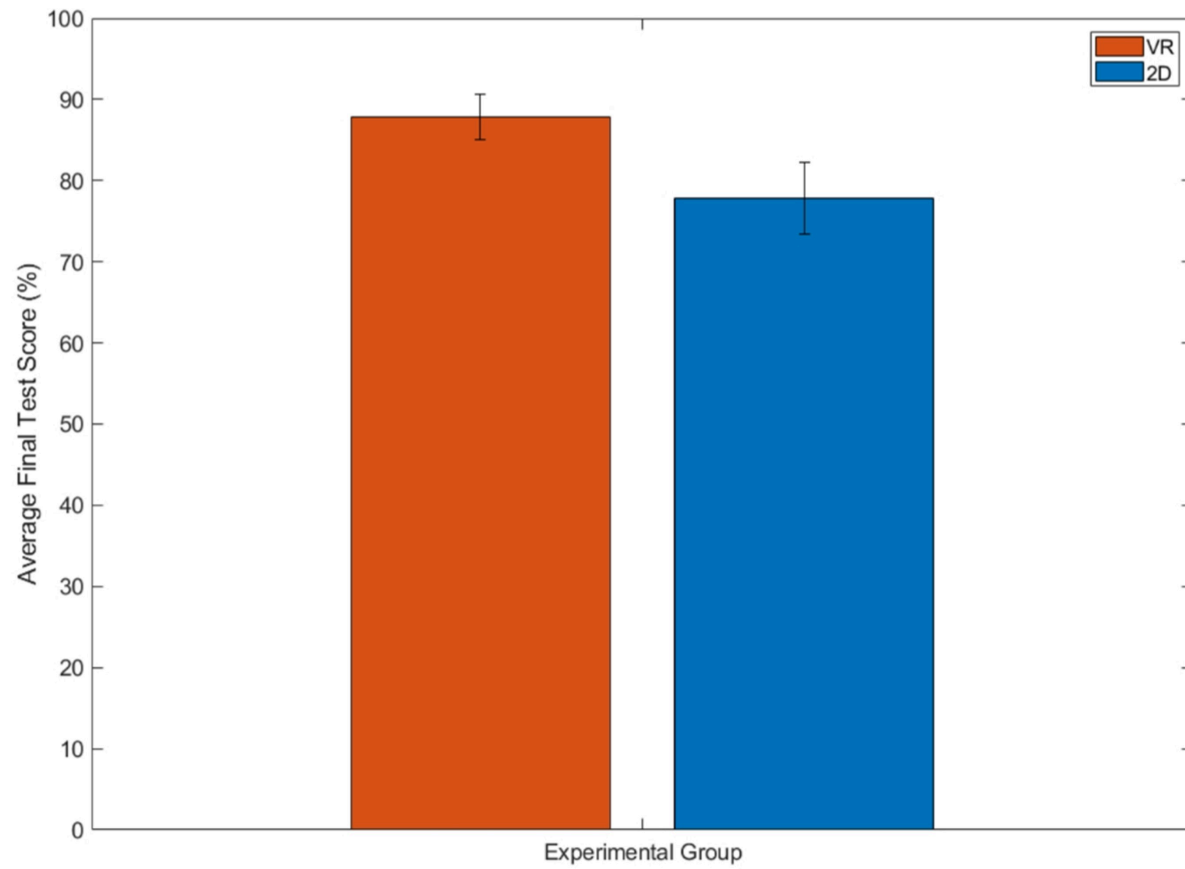


Figure 4

Mean final test scores for the VR and 2D groups. Error bars represent +/- 1SEM

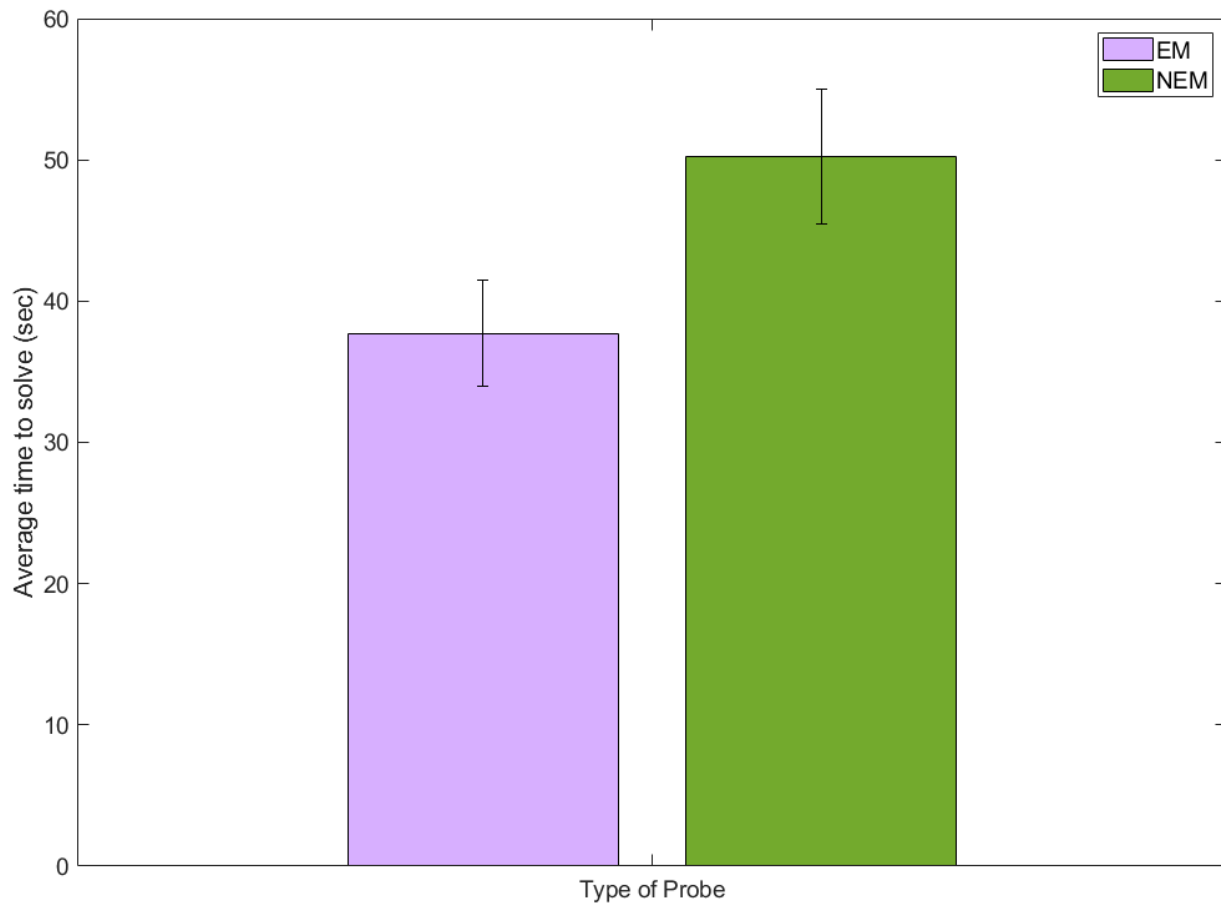


Figure 5

Mean latencies for the EM and NEM probes across groups. Error bars represent +/- 1SEM

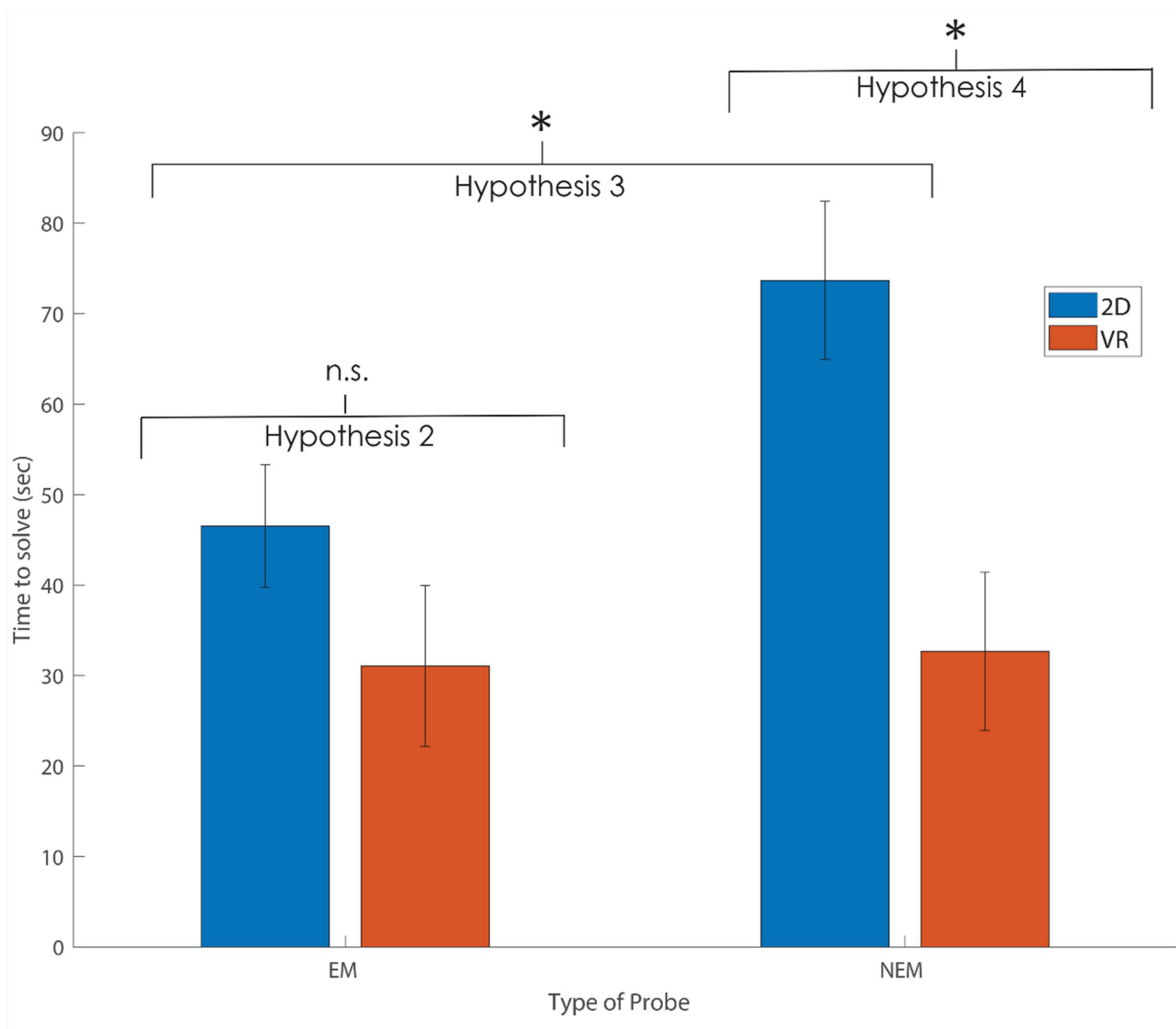


Figure 6

Average latencies per experimental group in the EM and NEM probes. Significant comparisons are denoted by an asterisk. Error bars represent +/- 1SEM

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementarymaterialsArtHistory2.0.pdf](#)