Thinking Skills and Creativity 56 (2025) 101699

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: www.elsevier.com/locate/tsc

Understanding the role of virtual mobility on how and what people create in virtual reality

Eugy Han *, Portia Wang, Cyan DeVeaux, Gabriella M. Harari, Jeremy N. Bailenson

Department of Communication, Stanford University, California, United States

ARTICLE INFO

Keywords:
Virtual reality
Creativity
Virtual mobility
Design creations

ABSTRACT

Virtual reality (VR) is considered a compelling tool to foster creativity by allowing its users to create in 3D space. However, the challenge lies in understanding *how* people use these tools and *what* they create, hindering the drawing of meaningful conclusions about VR as a viable tool for creativity. Furthermore, past research has shown that contextual factors shape how people create within VR, suggesting the existence of other factors. Here, we analyze the 3D creations of 137 participants responding to different creativity activities across seven sessions on a social VR platform. Specifically, we evaluate the role of virtual mobility, the capacity to move freely or have restricted movement in virtual space. We additionally present a VR-specific creativity coding scheme that follows recommendations from previous literature. Using dimensions derived from this coding scheme, we examine how these dimensions relate to behaviors and features of the creations in the context of virtual mobility. Results showed the significant role of virtual mobility on the design process, such that participants iterated and revised more by deleting more when their avatars were allowed to teleport and translate freely, compared to when their avatar's movements were restricted to sitting down in virtual chairs. Furthermore, participants built shorter creations and took up less projection space with restricted virtual mobility. Results also showed that participants created more practical, unique, and well-implemented creations the more 3D models they used. Similarly, the more participants deleted, the more well-implemented the creations were. We discuss implications for designers of creation-oriented VR platforms and pedagogy for instructors facilitating activities in educational contexts.

1. Introduction

Digital technologies often serve as tools that support creativity. Previous works have looked at digital technologies of various forms, including but not limited to, web blogs, flipped classrooms, and podcasts, and their positive effects on creativity (Yalcinalp & Avci, 2019). More specifically, digital technologies have been shown to increase autonomy, social belonging, attention, curiosity, and motivation of learning, ultimately increasing their motivation to engage in creativity (Tang et al., 2022). Virtual reality (VR) is one such digital technology that has shown promise given its unique affordances such as immersion and spatial interaction. Such affordances allow VR to be especially beneficial for certain individuals, such as those with high spatial abilities (Sun et al., 2018) or those who learn by doing hands-on activities (Chen et al., 2019). More broadly, the explorative and fun nature of VR has been argued to play

* Corresponding author at: 450 Jane Stanford Way Building 120, Room 110 Stanford, CA 94305-2050 United States.
E-mail addresses: eugyoung@stanford.edu (E. Han), portiaw@stanford.edu (P. Wang), cyanjd@stanford.edu (C. DeVeaux), gharari@stanford.edu (G.M. Harari), bailenso@stanford.edu (J.N. Bailenson).

https://doi.org/10.1016/j.tsc.2024.101699
Received 10 May 2024; Received in revised form 5 November 2024; Accepted 19 November 2024
Available online 20 November 2024

a critical role in the learning experience and help people feel free to suggest, explore, and evaluate ideas (Lau & Lee, 2012).

Despite these positive effects, several gaps in the literature exist. While the affordances of VR make it a potential platform for creativity, there are only a small number of studies that provide empirical results and directly investigate the creative potential for virtual environments (Alahuhta et al., 2014). Specifically, consider virtual mobility, the capacity to move freely or experience restricted movement in virtual space. Different VR platforms offer different types of virtual locomotion (Boletsis, 2017). How virtual mobility can support or hinder creativity and the creation process, demands further research.

The research goals and contributions of the present study are twofold: first, we examine how external factors, such as virtual mobility, influence what and how people create. Second, we present a creativity-evaluation metric specific to the nature of VR, one that fits both the nature of the tasks and the creations in VR. The scheme is informed by dimensions that have previously been identified as the most important aspects of creativity: fluency, originality, flexibility, and elaboration.

In what follows, we review past literature on VR, its affordances, and how these may foster creativity. Second, we discuss the different approaches of measuring creativity, including top-down evaluations from external raters and bottom-up analyzes of features of final creations. Third, we discuss what factors may influence creativity, namely focusing on virtual mobility – the ability to move around inside the virtual environment. Then, we introduce a study in which we use both a top-down and bottom-up approach to analyze participants' actions and outputs (i.e., creation) across seven sessions, and how either having free or restricted virtual mobility influenced these outcomes.

2. Related work

2.1. Virtual reality and creativity

Virtual reality (VR), typically accessed through head-mounted displays (HMDs), is a media that shuts out the physical world and transports users to immersive computer-generated simulations of environments. Compared to other media, VR is unique such that it affords immersion (i.e., generates realistic experiences that foster the feeling of presence), interaction with the surrounding environment (i.e., allows users to change locations, pick up and manipulate objects), and multisensory feedback (i.e., provides visual, auditory, and haptic feedback) (for a history and list of affordances, see Basu, 2019). Such affordances have made VR an attractive media for spatially-oriented tasks such as drawing, dancing, and pointing to, moving around, and scaling objects (Churchill & Snowden, 1998).

In particular, these affordances render VR especially effective for fostering creativity (for a literature review on how VR enhances creativity, see Gong & Georgiev, 2020). By engaging users with multi-sensory experiences, embodied cognition, and 3D spatial cognition, VR allows users to extend their vision and generate ideas (Hu et al., 2021). More specifically, by allowing users to freely observe and interact with objects from multiple perspectives, VR enhances cognitive styles, increasing sensitivity and fluency to induce imagination (Hu et al., 2016). Past research using functional magnetic resonance imaging (fMRI) data showed that sketching in 3D environments such as VR activates more visual-spatial functions compared to sketching in 2D environments (Tung & Chang, 2024). Furthermore, the novelty of being in an immersive environment has shown to boost motivation, inspire creative potential, and encourage the exploration of new ideas (Lau & Lee, 2012).

The effectiveness of VR for enhancing creativity is especially evident when comparing the media to traditional 2D media. For example, Feeman et al. (2018) had 11 different judges rank four different creations by participants, including a chair, maze, truck, and sculpture. Participants used a computer-aided design software either with a standard keyboard and mouse or with VR. In the computer condition, participants tended to create more simple chairs, often considering the task complete as soon as their model resembled a chair. In contrast, models created with VR were more original and complex, likely due to participants being more engaged in the task. The authors suggest that VR promotes creativity across dimensions that media such as a monitor and mouse may limit. Similarly, Yang and Lee (2020) identified the cognitive impact of sketching in VR compared to using paper and pencil. They found that sketching in VR stimulates visual-spatial cognition that transforms initial ideas to generate new concepts.

In a slightly different vein, literature in learning has shed light on the importance of how VR experiences are configured. In comparison to other media such as video, VR surrounds users with visual, auditory, and occasionally haptic information. As a result, VR users have to pay attention to and process multiple streams of information simultaneously (Makransky et al., 2019; Moreno & Mayer, 2002). The highly multisensory nature, interactivity, and high spatial presence introduced by VR experiences may demand processing that exceeds what users' cognitive systems are capable of, resulting in a phenomenon called cognitive overload (Ahn et al., 2022; Mayer, 2002). In particular, poorly designed tasks may lead to even further extraneous processing, causing learners to engage in cognitive processing that is not necessary for making sense of new information, and ultimately wasting cognitive capacity (Moreno & Mayer, 2007). While this present study does not examine cognitive load specifically, it does raise the question of how the lack of virtual mobility – which will likely influence how users engage with material, such as their scale or angle – may have downstream effects on creative outcomes.

Together, these studies highlight VR's unique potential to enhance creativity, but also underscores the need to further investigate the various parameters and factors within VR that may influence these outcomes.

2.2. Measuring creativity

Although what characteristics make up creativity vary across fields (see Amabile, 1996 and Kaufman & Glăveanu, 2021 for an overview of different theories in creativity), creativity is generally described as the internal process of transforming an individual's mental image to an external representation to create ideas, solutions, or products that are novel and valuable (Sarkar et al., 2008). When it comes to evaluating creativity, past studies have evaluated it as a process, an outcome, or both. As a result, there are various tests, schemes, and tools that address the actions performed and the products created. For example, Suwa et al. (1998) came up with a set of action categories to code cognitive actions of designers. According to their scheme, any action, such as what a person is attending to, removing, or inspecting could be encoded into a category that corresponds to the cognitive level of information processing (e.g., physical, perceptual, functional, and conceptual). This scheme would allow researchers to encode what kind of actions people are or are not engaging in, and analyze the relations between each action.

Another example of a tool to evaluate creative output is the Torrance Tests of Creative Thinking (TTCT, Torrance, 1966), one of the most researched and analyzed tools in the field (Kim, 2006). The TTCT includes a series of activities aimed to understand qualities that help people express their creativity. This tool uses game-like, problem solving, rather than testing, to understand the creative process. The output of these activities are evaluated based on several dimensions that are identified as the most important aspects of creativity – specifically creative potential – including fluency, originality, flexibility, and elaboration.

When it comes to understanding and measuring creativity within VR, there have been several different approaches (e.g., Chang et al., 2022). For example, Lee et al. (2019), investigated how fashion students' design cognition and creativity differed while creating a dress concept in either a virtual environment or a 2D one. The researchers used a coding scheme adapted from Suwa et al. (1998) to encode actions performed during a sketching activity in VR. In addition to the encoding, they had design experts evaluate the creations on different categories. Results from this combined approach concluded that VR can stimulate physical and perceptual action and enhance flexible cognitive thinking, leading to more creative outcomes.

In a different study, Lee et al. (2020) set out to investigate how VR could encourage groups of students to expand their thinking in design exploration throughout different phases of the creative process. Students used VR to sketch an idea for a fashion style based on a musical stimulus. The researchers had students evaluate the impact of the VR methods on their creativity based on fluency, flexibility, originality, and elaboration of ideas, which are the dimensions identified in the TTCT tool aforementioned. Results showed that VR expanded their design thinking exploration and representation of coordinated ideas, and enriched the understanding of a given problem context, leading to creative thinking for concept generation.

Similarly, a study by Yang et al. (2018) investigated the effectiveness of VR for creativity and how it compared to other non-immersive tools, such as paper-and-pencil or tablets. This open-ended activity involved designing a wearable device that would function as a smartphone. The researchers collected participants' EEG during and survey data evaluating their experience and state of mind after the creative activity. Researchers also evaluated their creations and rated them using categories related to relevance, novelty, elegance, and creation. Their findings concluded that VR allowed individuals to create higher quality creations, allowed them to see unique 3D points of view that stimulated novel ideas, and stimulated more attention and mental intensity.

In a similar but slightly different vein, recent work has broken down and analyzed 3D VR creations through both users' design behaviors and features of their creation. For instance, work by Wang et al. (2024) maps the processes of 2D sketch shape transformations (Prats et al., 2009), such as modifying the length or width of an element, or replacing an element with a new one, onto VR design behaviors, such as deleting 3D objects or changing viewpoints. Such actions have been studied by scholars across different fields, such as in computer science, where actions such as deleting were interpreted as "undoing" an action that encouraged people to experiment and try different approaches to solving problems through backtracking (Thimbleby, 1990). These actions, which are also often noted through qualitative observations by researchers studying creativity, such as Houzangbe et al. (2022) who commented on the lack of an undo button challenging users who would unintentionally create an undesirable feature, may provide insight into the creativity process.

Additionally, past research has also pointed to the rise of using computational systems to derive meaning and behavioral patterns from sketched shapes through analyzing their properties, such as colors, lines, contours, and texts (Rivard & Leclercq, 2006), which, in the study by Wang et al. (2024), map onto features such as the number of objects used, projection area, height, and volume. Together, using these top-down and bottom-up approaches of analyzing motion- and object-related features of users' behaviors and their creations points towards a holistic and robust approach to understanding creativity processes and creations.

The aforementioned studies highlight the various ways in which creativity – including the process and the final products – have been evaluated, both in general and specifically within the context of VR. Most often, studies use a combination of approaches, such as encoding actions or rating the final product based on dimensions similar to that of the TTCT. This suggests that each approach measures different dimensions of creativity, such as the process signifying different cognitive actions of an individual, or the output signifying observable qualities (e.g., originality, elaboration) expressed across individuals and a 3rd party other (e.g., an expert or a coder). Furthermore, combining multiple approaches is not an uncommon one within the field of creativity research (Long et al., 2014). For instance, work by Cheng (2011), which examined school-based creativity from students' perspectives, used multiple methods in their study and noted how the weakness of one method could offset the strength of the other and vice versa. They were able to validate the study by triangulation of the multiple methods. A similar argument is made by Yang et al. (2018), who highlighted how the incorporation of multiple measurements (e.g., traditional methods such as self-report surveys and expert panels with new measurements such as EEG) could add more data points, potentially increase ecological validity, and help researchers better understand the impact of VR on creativity. More broadly, a review paper on the measurements of creativity by Said-Metwaly et al. (2017) found that all instruments have limitations and that one instrument alone cannot undertake the task of measuring the multifaceted construct of creativity.

2.3. Factors that influence creativity

Past studies have also shown that multiple factors can foster or hinder creative performance, including factors related to the social and physical environment (Amabile, 2012). For example, the visual complexity of an environment may influence performance during design problem solving tasks (Goldschmidt & Smolkov, 2006). In a virtual context, exposure to natural backgrounds during video conferencing can elicit heightened creativity and generate more creative thinking compared to urban or gray backgrounds (Palanica & Fossat, 2022). In a similar domain, the audial environment has also been shown to have effects, such that having music rather than no music positively influences the novelty and design creativity of students engaging in a design task (Xia et al., 2023). Despite the significant role that external contextual factors play on creativity, very few studies examine these factors (Said-Metwaly et al., 2017).

Specifically within VR, Fleury et al. (2021) had participants sketch ideas in either the complete dark, a forest, or an office virtual environment. Creations, which were drawings of a logo representing themselves, were evaluated by two judges on the dimensions of relevance, novelty, elegance, and genesis. Results showed participants were more creative when in natural environments. They suggested that, because exposure to natural elements mobilizes cognitive functions that are inexpensive for mental effort and thus allows for recovery of attentional capacities, participants benefited from more effective attention restoration with the natural environment, giving them more attentional resources to realize their sketches. Moreover, other studies have shown additional factors that shape creativity within VR, such as the nature and duration of the creative activity, whether the activity was done collaboratively or individually, and whether the activity involved brainstorming or sketching (Bourgeois-Bougrine et al., 2022; for a review, see Liu et al., 2023).

Another critical component of creativity is related to movement. Bodily actions heavily shape the perceptual experience, such that information from the head, body, arms, and hands accompany optical information about the environment (see Gibson, 1979; Noë, 2004). Physical movements allow for more interaction with the environment, resulting in more processing of visual, auditory, and haptic information, and consequently, a better understanding of one's surroundings (Dove, 2011). We see evidence of this in Oppezzo and Schwartz's (2014) studies, which show the influence of physical movement on creativity. Participants completed various creative cognitive tasks either sitting or walking. Results showed that the simple act of walking generated more novel and appropriate ideas, and was shown to be beneficial for divergent thinking. They suggest that walking may have taxed the executive function to handle the dual-task nature of walking and thinking, and that this distraction may have allowed for more creativity. They point out that it was not necessarily leg movement that was the cause of these outcomes as there were residual effects of walking, but it could be due to an increase in positive mood. Similarly, another study by Slepian and Ambady (2012), showed that people produced more creative results after fluid versus non-fluid arm movements, highlighting the role of movement in creativity.

2.4. Virtual mobility

In VR, not all movements are physical, but can also be virtual. Virtual mobility, or locomotion, is an interaction component that allows users to navigate within a virtual environment. The ability to move around in virtual environments in unique ways has been shown to minimize cognitive load and time that may be induced by unnecessary travel (Bowman et al., 1998), and allows users to engage with content from various angles and dimensions that might not be possible in the physical world (Dong et al., 2022). Furthermore, mobility has been shown to predict interest in learning and inquisitiveness, and play a critical role in users' sense of self-presence and agency (Kilteni et al., 2012; Markowitz et al., 2018).

Along with technological advances, the ways in which users could move in VR has evolved to become more seamless, user-friendly, and less disorienting (Boletsis, 2017; Bowman et al., 1997). Users can move around in diverse ways, such as through point-and-click teleportation, joystick-based, continuous directional navigation, leaning and moving through the virtual world in the leaned direction, and walking-in-place (Arns, 2002; McVeigh-Schultz et al., 2018). Past research has shown that the type of virtual mobility users engage in can influence factors such as immersion, ease of use, control precision, spatial orientation, tiredness, and simulator sickness (for a review, see Cherni et al., 2020).

Specifically in relation to creativity, virtual visual motion has shown to have significance, such that people performed better at tasks that required divergent creativity when they were in a VR train that was moving, compared to when in a one that was stationary (Fleury et al., 2020). Furthermore, not having virtual mobility in VR may hinder user's willingness during the creativity process. When discussing the improvement of the VR computer-aided design software that was used in their study, Feeman et al. (2018) noted how not having a way to virtually move creations in the virtual environment posed a challenge when users were unwilling to move physically to kneel, lie down, or access parts of the creation. Another study by Fillingim et al. (2012) investigated if movement in VR had comparable effects to physical walking. In their study, they examined creativity outcomes of design tasks after participants either sat still, walked outside, or explored a virtual environment after reading the design prompt. While participants performed at an equal level across all conditions, when considering mood into account, VR was a significantly better incubator for creative outcomes such that participants with more happiness produced more novel designs in the VR condition and less feasible designs in the walking condition.

3. Current study

The field of VR is converging on the notion that environments accessed through VR headsets can foster creativity. However, there exists empirical gaps in the literature. In particular, there are gaps related to measurements in creativity processes and creations. Furthermore, given the spatial aspect of VR plays a critical role in creativity, such that it allows users to move around and see models from unique perspectives, understanding how such factors influence these outcomes is critical.

This study addresses these gaps by analyzing previously unreported data from the Stanford Longitudinal VR Classroom Dataset (SLVRClaD), in which 137 students engaged in various creativity activities in different virtual environments as part of a course (Han & Bailenson, 2024). Whereas Wang et al. (2024) covered the relationship between design behaviors and creations, and investigated how these outcomes are influenced by contextual factors (namely setting and spaciousness), the present paper focuses on a different, unreported variable: virtual mobility.[1]

Given that past literature points toward the value of using multiple approaches, including top-down evaluations from external raters, and bottom-up analyzes of features of final creations (e.g., Long et al., 2014; Wang et al., 2024), we use a combination of these two approaches to evaluate people's actions and outputs (i.e., creation). The goals of this study are to present a creativity-evaluation metric specific to the nature of VR, and provide insight into how and what people create in VR. The central findings of the study include:

- Virtual mobility is a factor that can influence creativity in VR. Participants iterated and revised more by deleting more when their avatars were allowed to teleport and translate freely, compared to when their avatar's movements were restricted to sitting down in virtual chairs.
- Participants built shorter creations that took up less projection area, and built creations that were evaluated as less practical when their virtual movements were restricted compared to when they were allowed to move their avatars freely.
- There were significant, positive correlations between how many times participants deleted and how many 3D models they used. There were also significant, positive correlations between how many times participants deleted and the height, projection area, and volume of their creations.
- Creations were rated as more well-implemented when participants deleted more.
- The more 3D models participants used, the more raters perceived creations as practical, unique, and well-implemented.

4. Methods

The present data is a previously unreported subset of SLVRClaD, which includes data from approximately 500 students across two years, each of whom took one of four separate courses taught using the ENGAGE social VR platform accessed via Meta Quest 2 VR headsets. Students met in groups and consented to have their verbal, nonverbal, and performance data continually tracked during each course, typically for about eight weekly sessions which lasted about 20–30 min per session. In addition, each student self-reported about their experience after each session.

4.1. Participants

The present data is a previously unreported subset of the SLVRClaD dataset and includes data from 137 students. At the beginning of the course, students were invited to participate in an Institutional Review Board-approved (IRB) study of how various exposures to VR influenced their behavior. While all students who were part of the course took part in all the VR activities, only those who consented to participate in the study contributed data for analysis. Safeguards implemented to ensure privacy and consent included review both by the IRB and a second university ethics organization, and third-party oversight of the consent process and data collection, and recurring reminders that they were being recorded at the beginning of every session.

The 137 students who participated in five or more of the eight weekly sessions ($M = 78$, $F = 59$) were between 18 and 49 years old ($M = 20.9$, $SD = 2.78$; $n_{18\sim20} = 62, n_{21\sim23} = 71, n_{24\sim49} = 5$) and identified as Asian or Asian-American ($n = 47$), White ($n = 41$), multiracial ($n = 19$), African, African-American, or Black ($n = 12$), Hispanic or LatinX ($n = 8$), Native Hawaiian or other Pacific Island ($n = 5$), Indigenous/Native American, Alaska Native, First Nations ($n = 2$), declined to or did not respond ($n = 2$), Middle Eastern ($n = 1$), and a racial group not listed ($n = 1$). Participants had varying levels of experience with VR ($n_0 = 50, n_1 = 29, n_2 = 23$, $n_{3\sim10} = 26, n_{20\sim50} = 4, n_{90} = 2, n_{100} = 4$).

4.2. Manipulated variables

4.2.1. Virtual mobility

Virtual mobility was manipulated in terms of whether participants were allowed to move freely or were restricted to virtual seats (active vs. passive). Each week, each group of participants was assigned to one of two mobility conditions via a Latin square randomization scheme that ensured that each group spent 4 weeks in each condition, counterbalanced across weeks and meeting days. In each session, all members of the same group were either *active* or *passive*.

In the *passive* condition, there was a circle of chairs that the participants could teleport to. Sitting in these seats restricted movement and prevented participants from moving and walking around their creation. Participants were instructed to not move their joystick or teleport elsewhere. An instructor would request the participant to go back to their seat if they moved. In the *active* condition,

[1] This independent variable was not investigated in Han et al. (2023) or Wang et al. (2024), as it was concerned with translational motion, which was more relevant to the behavioral outcome variables of interest. Given the scope of this present paper focuses on access to space, being able to spread out, and creativity as a result of this mobility constraint, we include and focus on this independent variable in our analyses.

participants could freely explore the environment. Participants were instructed to move around, explore different areas, and make use of the space around them.

4.2.2. Virtual environment

The virtual environment, which varied each week based on the same Latin square randomization scheme, was also manipulated in terms of setting (outdoors vs. indoors) and spaciousness (panoramic vs. constrained). The setting varied in terms of whether elements of nature were present in the environment (e.g., trees, bodies of water). The spaciousness varied in terms of how much space was visible (e.g., how far and wide the environment was). Each environment was built by research personnel using 3D models. In total, there were 192 uniquely built environments that differed in size of natural elements, moving area and height, and seating arrangements. Further justifications, dimensions, and example images of the environments are presented in Han et al. (2023).

4.3. VR hardware and software

Participants were provided with Meta Quest 2 headsets: standalone head-mounted displays (HMDs) with 1832 × 1920 resolution per eye, 104.00° horizontal FOV, 98.00°FOV, 90 Hz refresh rate, and six-degree-of-freedom inside-out head and hand tracking (503 g) and two hand controllers (126 g) for use in their personal environment.

The software used was ENGAGE, a collaborative social VR platform. On ENGAGE, users are represented as avatars and can interact with one another and the virtual environment. The platform also provides users with various tools that allow them to navigate around the virtual environment, including joystick-based translation and point-and-click teleportation. Other tools allow users to build within the virtual environment, using a 3D pen (allows a user to select a color and pen thickness to freely draw in the air), 3D models (allows a user to choose from a list of immersive effects/3D models to add to the virtual world), whiteboards (allows a user to write/draw on a flat surface), and sticky notes (allows a user to write on small sticky notes and suspend them in space).

4.4. Procedure

Participants selected a discussion group that fit their schedule and availability, resulting in 24 groups that met weekly for eight weeks and varied in size from five to eight members ($M = 6.71$, $SD = 0.81$). Each week, each group was assigned to a set of eight conditions (2 × 2 × 2 design) via a Latin square randomization scheme that ensured each group experienced each condition once and that each condition appeared equally across the weekly schedule. The sessions were led by one of three instructors. Each instructor led the same eight groups every week.

All participants were trained on how to use the ENGAGE interface and the controllers in the first week of the course. During the training, instructors were available to assist via Zoom when students faced technical mishaps in hardware and software.

The first discussion session began in the second week, during which participants completed a series of small-group activities to further familiarize them with the ENGAGE environment and its tools. All discussions, except in the fifth session, had a creativity activity using the tools available on ENGAGE. The 30 min sessions were divided into a 10 min full-group discussion and recap of the course material, a 15 min individual creativity activity based on a prompt, and a 5 min sharing of the creation of the creativity activity portion. The time spent inside the virtual environment was limited to prevent simulator sickness (Han et al., 2022). The activities varied in terms of what the participants were tasked to do, including executing ideas from scratch, reimagining existing ideas or concepts, and ideating on a problem. See Table 1 for all the activity prompts.

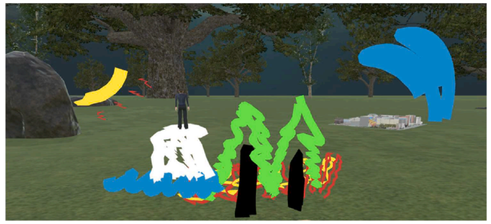
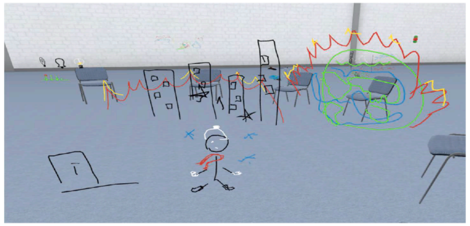
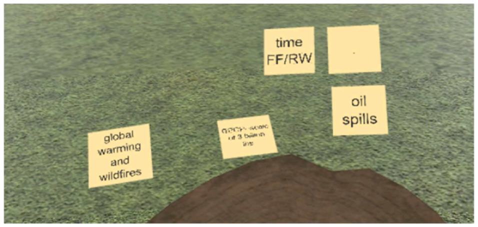
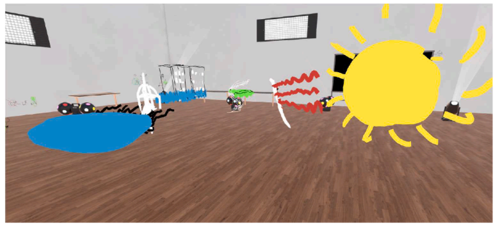
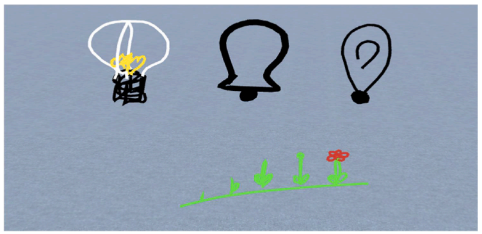
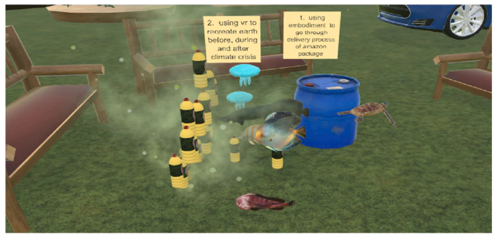
During the individual creativity activity, participants were either limited to remaining in their virtual seat or were allowed to move freely and explore the space around them. Physically, participants reported that they were sitting down for all or most the session (session 1 = 82.0 %, session 2 = 82.3 %, session 3 = 83.6 %, session 4 = 84.0 %, session 6 = 86.1 %, session 7 = 87.4 %, session 8 = 91.0 %). Instructions for participants in the *passive* condition were as follows: “We’ll be holding these discussions in different types of environments, some of which will have seats available, some not. When there are seats available, we’ll be sitting around together. I know it may be tempting to move your joystick around and teleport elsewhere, but to make sure our discussions go smoothly and we can cover everything on time, I will occasionally ask you to teleport back into a seat.”

Instructions for participants in the *active* condition were as follows: “In cases where there aren’t any particular seating arrangements, feel free to move around as you wish, explore different areas, and make use of the space around you.”

Table 1
Weekly creativity activity prompt and type.

Session	Creativity activity prompt
1	Consider the affordances of VR and create a prototype of something that leverages the uniqueness of VR
2	Create something frightening that induces a feeling of high presence
3	Consider the affordances of VR to make a difficult concept easier to understand
4	Create something that reimagines avatars and representations of the self
5	No creativity activity
6	Create a meditation room or “safe-space”
7	Brainstorm an idea of how to communicate a message about climate change
8	Create and playtest a VR-based game

Table 2
Examples of creations and both rater pairs' scores.

<table>
<tr><th></th><th>Dimension and definition</th><th>Low-score example</th><th>High-score example</th></tr>
<tr><td>Idea</td><td>Practicality:
Refers to how well the idea draws on the affordances of VR and is applicable to the physical world</td><td></td><td></td></tr>
<tr><td></td><td>Uniqueness:
Refers to how novel the idea is</td><td></td><td></td></tr>
<tr><td>Final artifact</td><td>Well-implemented:
Refers to how well executed and built the final artifact is</td><td>
</td><td></td></tr>
<tr><td></td><td>Understandability:
Refers to how straightforward and easy it is to understand what the final artifact represents.</td><td></td><td></td></tr>
</table>

Sessions were recorded using ENGAGE's recording feature via a desktop computer. These recorded *myrec* files capture all the tracking and event data from the sessions, including avatars, movements, audio, and objects in 3D space so that interactions can be examined from any angle or distance.

4.5. Measures

4.5.1. Top-down evaluations of the creations

Using the TTCT dimensions, fluency, originality, flexibility, and elaboration as our primary foundation for understanding creativity, we formulated a coding scheme that fit both the nature of the tasks and the creations in VR. Through an iterative process with four raters, the first author and the four raters discussed and modified the coding scheme. The raters then tested the scheme on the creations of a week that was not used in the present analysis. This process was repeated several times until there was a reliable consensus, or intraclass correlation values (ICCs) of at least 0.5 for each dimension, which is considered being moderate reliability, as suggested by Koo and Li (2016). Each dimension was rated on a 7-point Likert scale, with 1 = Strongly disagree to 7 = Strongly agree.

The dimensions were conceptually divided to address the *idea* and *final artifact* of the creations (see Table 2 for dimension definitions and examples). For the final analysis, the raters evaluated the creations for session 7 ($n = 86$). This session was selected because the prompt allowed for the most diverse selection of tools, allowed for participants to create their own original ideas (i.e., in comparison to the reconceptualizing prompts), and would be less abstract for the raters to understand (i.e., in comparison to building prompts, this session had a common theme of climate change).

Two raters made up a pair to evaluate each participant's creation. The step of pairing up the raters was done based on iteration, to allow the raters to discuss each creation together and reach a consensus on a final score for each dimension. The absolute difference between each pair of raters' scores were calculated, and any evaluations that had a difference score equal to or greater than 4-points were dropped due to their large discrepancies ($n_{practicality} = 60$, $n_{uniqueness} = 71$, $n_{understandability} = 78$, $n_{well\text{-}implemented} = 82$). Examples of creations and their ratings are presented in Table 2.

Following this, ICC, a quantitative measure to evaluate the extent to which values in the same group resemble one another, was used to determine whether a pair of two raters reached consensus in assessing each item of the creations. ICC Model 3, k estimates and their 95 % confidence intervals were calculated among complete ratings of the creations using the *psych* package in R, based on a two-way mixed effects, consistency among multiple raters model, following recommendations by Koo and Li (2016) and Shrout and Fleiss (1979).

Despite the iterations, consensus was moderate for *practicality*, ICC(3,k) = 0.690, 95 % CI (0.481, 0.815), *uniqueness*, ICC(3,k) = 0.700, 95 % CI (0.519, 0.813), and *understandability*, ICC(3,k) = 0.721, 95 % CI (0.563, 0.822), and was good for *well-implemented*, ICC (3,k) = 0.762, 95 % CI (0.631, 0.846). Following the ICC evaluation, the mean rating between the two pairs of raters was calculated.

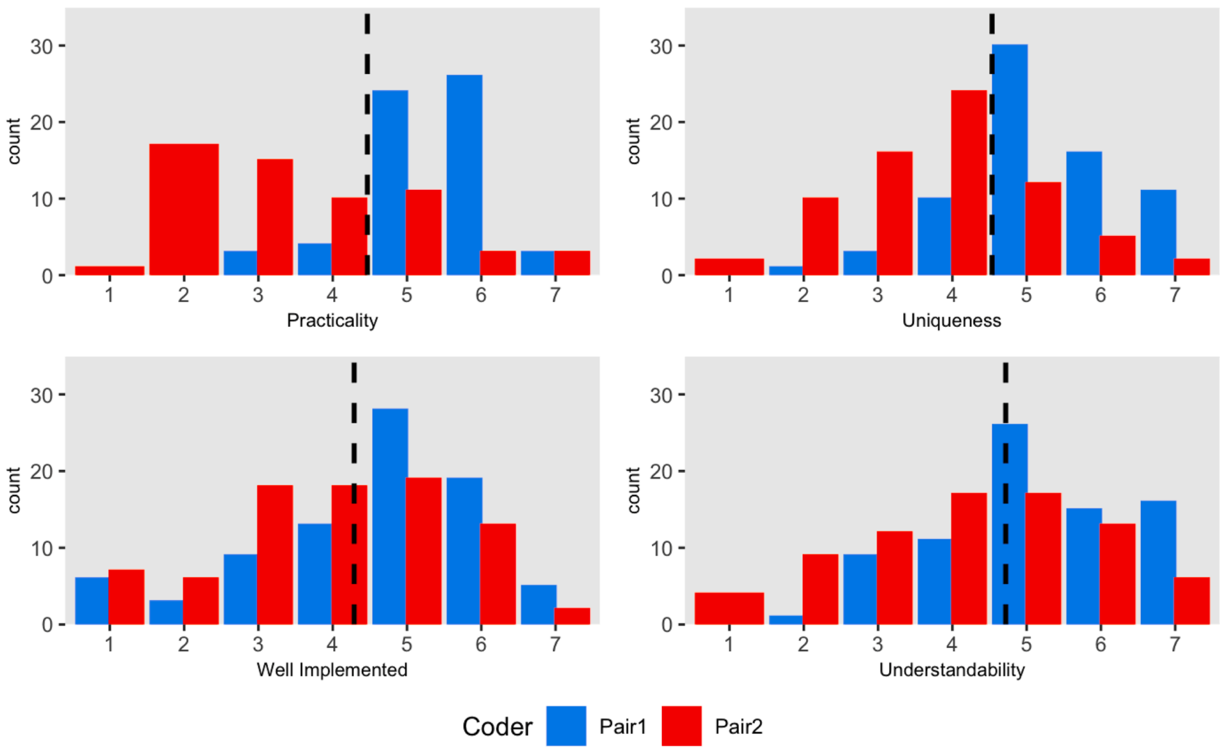


Fig. 1. Distribution of ratings for each pair of coders and the final mean ratings (dashed lines).

Distribution of each pair's ratings and the final mean ratings are presented in Fig. 1.

4.5.2. Bottom-up evaluations of the creations

Various properties of the creations each week were measured, following the procedure by Wang et al. (2024) (see Table 3). These measures were collected at the latest moment when all group members were present in the recording ($n_{week1} \approx 130$, $n_{week2} \approx 130$, $n_{week3} \approx 126$, $n_{week4} \approx 130$, $n_{week6} \approx 130$, $n_{week7} \approx 125$, $n_{week8} \approx 122$). In cases where only sticky notes were used ($n = 24$), these entries were dropped because sticky notes in the ENGAGE recording format could not be traced back to specific participants due to how they were encoded in the data. The measures are conceptually divided into two types: action- and creation-oriented. Action-oriented measures are based on what each participant did throughout a given session whereas creation-oriented measures are based on the dimensions of the creation.

4.6. Data analysis

Following the procedure recommended by Grimm et al. (2016, ch. 3), we examined the relationship between our bottom-up evaluations and our independent variables using linear mixed-effects models that accommodate the nesting of repeated observations within subjects within groups. With group-level variance accounting for between 1.17 % to 18.6 %, person-level variance accounting for between 0 % to 18.01 %, and session-specific variance for between 70.19 % to 94.2 % of total variance, we formalized the analysis of all outcomes using a 3-level multilevel structure with both participant and group random effects.

$$outcome_{tig} = \gamma_{000} + \gamma_{100}\left(week_{tig}\right) + \gamma_{200}\left(virtual\ mobility_{tig}\right) + \gamma_{300}\left(setting_{tig}\right) + \gamma_{400}\left(spaciousness_{tig}\right) + v_{00g} + u_{0ig} + e_{tig}$$

Here, the outcome of interest at occasion t for person i in group g, $outcome_{tig}$ is modeled as a function of a grand intercept, γ_{000}, that indicates the expected level of the outcome on the first week in which students' virtual mobility was restricted (passive) while in a constrained, indoor environment. The other parameters represent a time-related trend, γ_{100}, which represents the prototypical rate of change across weeks; the virtual mobility effect, γ_{200}; the setting effect, γ_{300}; the spaciousness effect, γ_{400}; and residual group-specific, v_{00g}, person-specific, u_{0ig}, and occasion-specific deviations, e_{tig}, that were assumed to be normally distributed with standard deviations σ_{v00}, σ_{u0}, σ_e. Additional models with time-invariant predictors (i.e., individual differences), including the Big 5 personality traits and learning styles which were collected at the beginning of the study (pre-test), and other interaction terms between the manipulated independent variables (i.e., virtual mobility, spaciousness, setting) were tested, but were trimmed due to insignificance. Effect sizes are reported by both the marginal R^2 (R^2m), which is the amount of variance in the outcome variable explained by the fixed effects, and the conditional R^2 (R^2c), which is the amount of variance in the outcome variable explained by the fixed and random effects.

Second, we showed overall effects through correlations. We present exploratory findings on how the top-down evaluations correlate with other outcomes specific to session 7. A 3-level multilevel model similar to the ones aforementioned, with the time-related parameter γ_{100} omitted, was run for the top-down evaluations specific to session 7.

Session 5, which did not have a creativity activity, and any outliers (i.e., cases less or greater than 3 standard deviations) were removed from the analyzes. Incomplete data were treated as missing at random. Statistical significance was evaluated at alpha = 0.05. All models were fit to the data in R using the nlme library and visualized using the ggplot2 library (Wickham, 2011). Conditional and marginal R^2 were calculated using the MuMIn package (Barton, 2009). Confidence intervals (CIs) for marginal R^2 were calculated using the r2glmm package (Jaeger, 2017) using the method suggested by Nakagawa and Schielzeth (2012).

5. Results

5.1. Correlational tests between measures

Correlations between the measures were calculated to identify any patterns.[2] Table 4 shows correlations between the top-down evaluations of the creations. There were significant, positive correlations between practicality and uniqueness, practicality and well-implemented, practicality and understandability, and uniqueness and well-implemented. In particular, the correlation between practicality and well-implemented was strong. And, although not significant, there was a small negative correlation between uniqueness and understandability.

Table 5 shows correlations between the bottom-up evaluations of the creations. There were significant, positive correlations between number of deletions and number of 3D models, number of deletions and height, number of deletions and projection area, number of deletions and volume, number of 3D models and projection area, height and projection area, height and volume, and projection area and volume. In particular, the correlation between height and volume was strong.

Table 6 shows correlations between the top-down and bottom-up evaluations of the creations specifically from session 7. There were significant, positive correlations between practicality and number of 3D models, uniqueness and number of deletions, uniqueness and number of 3D models, well-implemented and number of deletions, well-implemented and number of 3D models. Follow-up

[2] Additional correlations between the evaluations and individual differences (including the Big 5 personality traits and learning styles) were examined but not reported due to insignificance.

Table 3
Description and calculation of each measurement.

Measure	Description
Deletions	This measure represents the number of times a participant deleted a 3D model that they brought into the virtual environment. Units are in frequency of deletions.
Use of 3D models	This measure represents the maximum number of 3D models that a participant brought into the virtual environment. Units are in models.
Height	This measure represents the height of a participant's final creation. This was calculated by taking the height difference (y-plane) between the highest and lowest 3D model or drawing. Units are in meters.
Projection area	This measure represents the amount of space a participant's final creation took up along the x- and z- planes. This was calculated by taking the area of the projection of the created convex hull onto the horizontal plane. Units are in m^2.
Volume	This measure represents the volume of space that a participant's final creation took up. This was calculated by taking the volume of the convex hull that is created based on the positions of each 3D model and drawing. Units are in m^3.

Table 4
Correlation coefficients and their *p*-values between the top-down evaluations of the creations in session 7.

	Practicality	Uniqueness	Well-implemented
Uniqueness	0.437**	–	–
Well-implemented	0.731***	0.4085**	–
Understandability	0.5058***	-0.01049	0.0168

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5
Correlation coefficients and their *p*-values between the measures.

	Deletions	Use of 3D models	Height	Projection area
Deletions	–	–	–	–
Use of 3D models	0.588***	–	–	–
Height	0.282*	0.187	–	–
Projection area	0.271*	0.298*	0.418***	–
Volume	0.376**	0.193	0.758***	0.569***

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6
Correlation coefficients and their *p*-values between the top-down and bottom-up evaluations of the creations in session 7.

	Practicality	Uniqueness	Well-implemented	Understandability
Deletions	0.239	0.2907*	0.385**	-0.1044
Use of 3D models	0.4102**	0.373**	0.6708***	-0.20083
Height	-0.1906	0.0425	0.166	-0.181
Projection area	0.0156	0.0137	0.293	0.00980
Volume	-0.0116	0.138	0.255	-0.02095

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

analyzes using the same 3-level multilevel models aforementioned, with the time-related parameter omitted and bottom-up evaluations set as the predictor (γ_{500}), showed that, in session 7, the more 3D models were used, the more practical ($\gamma_{500} = 0.09803$, $p = 0.0029$), unique ($\gamma_{500} = 0.0720$, $p = 0.0181$), and well-implemented ($\gamma_{500} = 0.182$, $p < 0.000$) the creations were. The more participants deleted, the more well-implemented ($\gamma_{500} = 0.2402$, $p = 0.0085$) the creations were. There was no evidence of effects of the number of deletions on uniqueness ($p > 0.05$).

5.2. *Bottom-up evaluations across time and virtual mobility*

Results from the multilevel models with time-varying predictors week and virtual mobility are presented for the creations.[3] Plots of the raw data overlaid with relevant prototypical trajectories are presented in Fig. 2.

5.2.1. *Deletions*

The prototypical participant's number of deletions decreased from an initial value of $\gamma_{000} = 1.505$, $p < 0.000$ (deletions), though not significantly, at a rate of $\gamma_{100} = -0.000435$, $p = 0.985$ deletions per week. There was a significant effect of virtual mobility such that

[3] Note, we do not report findings of setting or spaciousness in this article, as these findings were discussed by Wang et al. (2024).

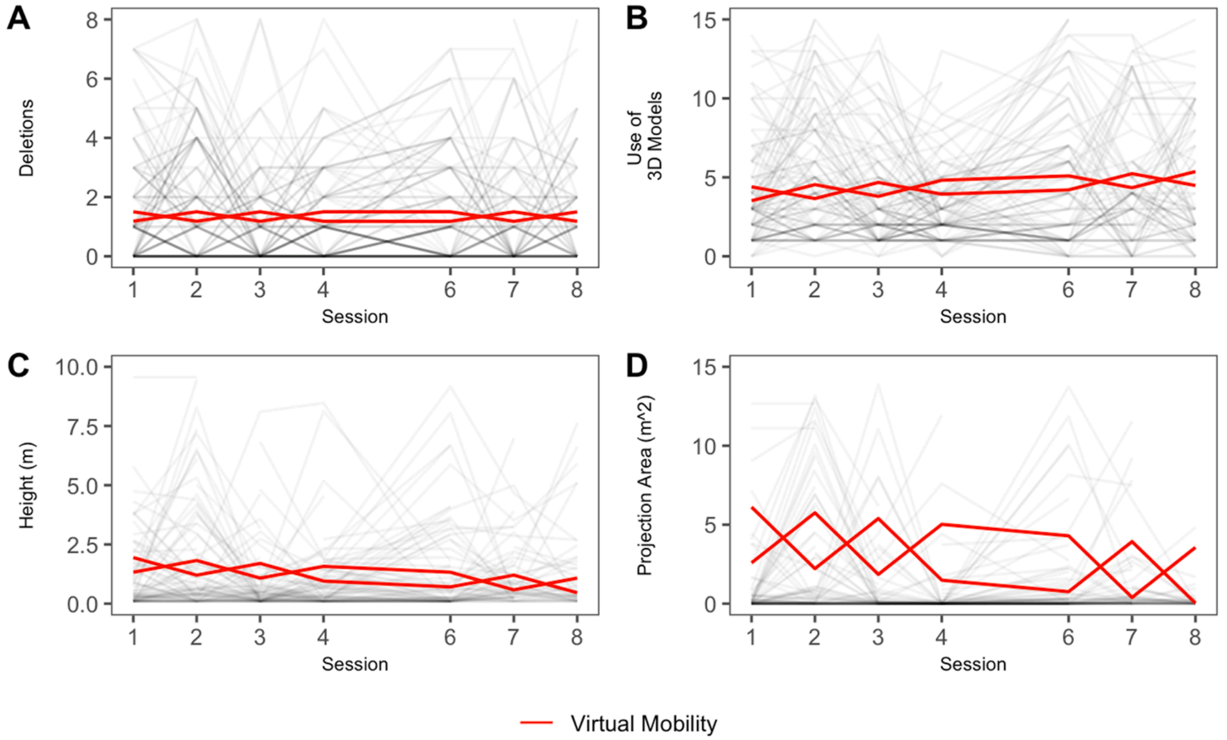


Fig. 2. Raw data and prototypical trajectories of deletions and use of 3D models over time. *Note.* This figure shows the raw data and prototypical trajectories of how deletions and the use of 3D models changed over time for hypothetical participants who alternated weekly between the two virtual mobility conditions (red lines). The gap between sessions 4 and 6 represents session 5, which was omitted from the analysis as it did not have a creativity activity.

participants deleted less when their movements in the virtual environment were restricted, γ_{200} = -0.3202, p = 0.0038 (Panel A). Conditional and marginal R^2 for the model were R^2c = 0.280, R^2m = 0.009802, 95 % CI (0.004, 0.035), respectively.

5.2.2. Use of 3D models

The prototypical participant's number of 3D models used increased from an initial value of γ_{000} = 4.27, $p < 0.000$ (models) at a rate of γ_{100} = 0.138, p = 0.0133 models per week. There was a significant effect of virtual mobility such that participants used less models when their movements in the virtual environment were restricted, γ_{200} = -0.885, p = 0.0010 (Panel B). Conditional and marginal R^2 for the model were R^2c = 0.326, R^2m = 0.02503, 95 % CI (0.014, 0.061), respectively.

5.2.3. Height

The prototypical participant's creation's height decreased from an initial value of γ_{000} = 2.069, p = 0.0001 (m) at a rate of γ_{100} = -0.124, p = 0.0452 m per week. There was a significant effect of virtual mobility such that participants built shorter creations when their movements in the virtual environment were restricted, γ_{200} = -0.6205, p = 0.0351 (Panel C). Conditional and marginal R^2 for the model were R^2c = 0.254, R^2m = 0.0256, 95 % CI (0.012, 0.063), respectively.

5.2.4. Projection area

The prototypical participant's creation's projection area decreased from an initial value of γ_{000} = 6.49, p = 0.0006 (m^2), though not significantly, at a rate of γ_{100} = -0.365, p = 0.176 m^2 per week. There was a significant effect of virtual mobility such that participants built creations with a smaller projection area when their movements in the virtual environment were restricted, γ_{200} = -3.54, p = 0.0059 (Panel D). Conditional and marginal R^2 for the model were R^2c = 0.07505, R^2m = 0.0178, 95 % CI (0.007, 0.049), respectively.

5.2.5. Volume

The prototypical participant's creation's volume decreased from an initial value of γ_{000} = 50.44, p = 0.568 (m^3), though not significantly, at a rate of γ_{100} = -8.98, p = 0.516 m^3 per week. There was no evidence of effects of virtual mobility.

5.3. Top-down evaluations across virtual mobility

Results from the multilevel models with time-varying predictor virtual mobility are presented for the top-down evaluations of the creations.

5.3.1. Practicality

There was a significant effect of virtual mobility on the prototypical participant's creation's practicality $\gamma_{000} = 4.97$, $p < 0.000$ (points, out of 7-points) such that participants built less practical creations when their movements in the virtual environment were restricted, $\gamma_{200} =$ -0.997, $p = 0.0289$. Conditional and marginal R^2 for the model were $R^2c = 0.935$, $R^2m = 0.2038$, 95 % CI (0.041, 0.29), respectively.

5.3.2. Uniqueness, understandability, well-implemented

There was no evidence of effects of virtual mobility on the prototypical participant's creation's uniqueness, understandability, or well-implementation (all p-values > 0.0874).

6. Discussion

In the present study, we examine two approaches of analyzing the creativity process and creations, and how virtual mobility influences these outcomes. Over seven sessions, participants engaged in weekly creativity activities in various virtual environments. These virtual environments varied in terms of setting (outdoors vs. indoors), spaciousness (panoramic vs. constrained), and virtual mobility (active vs. passive). In the top-down approach, two pairs of coders evaluated participants' creations of one of the sessions using a creativity coding scheme that was based on traditional creativity tools, the TTCT, and tailored to a VR context. In the bottom-up approach, participants' actions and dimensions of their creations were calculated following computational methods to analyze properties of creations (Wang et al., 2024). Various measurements related to how (e.g., deleting, using 3D models) and what (e.g., height, projection area, and volume) participants created were collected.

One of the affordances unique and central to VR is its spatiality. Past research has shown that VR is attractive for spatially-oriented tasks where pointing to, moving around, and scaling objects are key (Churchill & Snowden, 1998). Users are able to move around in 3D space, and observe and interact with objects from multiple perspectives. Given this, not being able to move around in the virtual environment may lead to undesirable outcomes (Houzangbe et al., 2022). As past research suggests, creating in VR activates more visual-spatial functions compared to creating in 2D environments. However, how this translates to situations in which a user's movements are limited, such that they are not able to make use of these spatial aspects of VR, remains a question. Notably, the type of virtual mobility made available differs across all platforms, such that some do not allow their users to move around and interact with the surrounding environment at all, while others allow multiple types of mobilities. Not being able to virtually move around freely may have downstream effects on creativity, limiting users' imagination, motivation, and exploration. While past research has focused on the contextual factors that influence design behaviors and dimensions of creative outputs (e.g., Wang et al., 2024), such as amount of visible space, instructor style, and design prompt, this current study narrows down the effects of virtual mobility, as well as top-down, external evaluations of these outputs.

We see evidence of the role of virtual mobility in the current study. We found that participants iterated and revised more by deleting more when their avatars were allowed to teleport and translate freely, compared to when their avatar's movements were restricted to sitting down in virtual chairs. Moreover, participants built shorter creations that took up less projection area and built creations that were evaluated as less practical when their virtual movements were restricted compared to when they were allowed to move their avatars freely.

The number of deletions can be interpreted as "undoing" an action, which has been shown to encourage people to experiment, as it allows people to try out different approaches to solving problems through backtracking (Thimbleby, 1990). This act of deleting suggests engagement in undoing, exploring, and deliberating on a previous action. On a broader level, deleting also suggests reiteration and revision, which, in turn, are signs of more effort, and are good practices for design. Here, this action of deleting suggests that participants explored ideas with more 3D models, took up more virtual space, and deliberated more. This finding is in line with what past studies have found: VR can encourage the exploration of new ideas and transform initial ideas to generate new concepts (Lau & Lee, 2012; Yang & Lee, 2020). However, past studies have also shown that the lack of virtual mobility led to users feeling unwilling to physically move and access parts of their creation (Feeman et al., 2018). This current study aligns with these past findings and provides further evidence of the role of movement in VR and creative outcomes.

Returning to the research on the effects of physical mobility (e.g., moving and walking in the physical world) on idea generation (Oppezzo & Schwartz, 2014), although in our study, the walking was not physical, it was virtual and similarly addressed the ability to move around the surrounding environment and explore different angles. Being able to teleport and translate in their virtual environment may have allowed for more opportunities for distraction, promoting more actions such as deleting and adding (i.e., introducing more 3D models into the creation), compared to those whose experience may have been hindered due to being virtually restricted to chairs.

It may also be possible that having limited virtual mobility restricted how much space participants had, in general. This is in line with findings from Wang et al. (2024), who found that participants moved more in the horizontal space in panoramic virtual environments. Here, the lack of mobility limited horizontal and vertical space, and consequently participants were most likely unable to create as much (i.e., use many 3D models) or reach as high or as far, compared to those who could teleport and translate freely. This potentially explains the findings on participants building shorter creations that took up less projection area and were evaluated as less practical when their virtual movements were restricted. Oppositely, those who were able to move around had access to more space, and consequently more opportunities to deliberate (i.e., delete and add), and take up more horizontal and vertical space.

This is also in line with the findings of significant, positive correlations between how many times participants deleted with how many 3D models they used. There were also significant, positive correlations between how many times participants deleted with the height, projection area, and volume of their creations. Again, while deleting and adding can suggest different meanings, one potential meaning lies in deliberation and exploration, and other actions that suggest that participants iterated more on their creations.

Continuing this thread, there were also significant, positive correlations between how many 3D models were used and how practical, unique, and well-implemented the creation was. Additionally, there were significant, positive correlations between the number of deletions and how unique and well-implemented the creation was. Follow-up analyzes showed that these bottom-up evaluations were significant predictors of the top-down evaluations, such that participants created more practical, unique, and well-implemented creations the more 3D models they used. Similarly, the more participants deleted, the more well-implemented the creations were. However, there was no evidence of effects of the number of deletions on uniqueness. This could suggest that being able to undo and deliberate with what to put into the virtual surroundings, and consequently add other 3D models, potentially ones more relevant to their creation, allows participants to explore more unique ideas and better represent their ideas.

Lastly, other correlational findings include significant, positive correlations with how unique, well-implemented, and understandable a participant's creation was with how practical it was. In particular, the correlation between well-implemented and practicality was strong, suggesting that the more well-presented creations were, though not easily understandable (suggested by a low and insignificant correlation), they were viewed as more unique and practical.

More broadly, we introduced a coding scheme that fit both the nature of the tasks and the creations in VR based on the dimensions that have previously been identified as the most important aspects of creativity: fluency, originality, flexibility, and elaboration. By tying both the bottom-up evaluations to top-down evaluations based on these dimensions, we are able to make sense of and contextualize what design behaviors and dimensions mean in relation to how these outputs are evaluated by a third-party onlooker.

7. Conclusion

7.1. Limitations and future directions

There were several factors, both regarding the design and logistics of the study, that limited the study. First, the study was a field experiment, which serves as both a strength and a limitation. Although field experiments allow researchers to implement interventions and measure outcomes in naturalistic settings that would otherwise be challenging to implement in laboratory settings, there are constraints in how much control the researchers have on sample size, external conditions, and potential confounds. The characteristics of the course and the composition of its students (i.e., college students interested in learning about VR) may have played a role. There are other factors that were out of our control or were difficult to control for, such as where the students were physically located while engaging in the activities in-VR. Given the importance of virtual environmental context, we predict the physical environment also plays a role.

Second, we presented a creativity coding scheme that was tailored to a VR context and based on a task that allowed for a diverse selection of tools and exploration of ideas. We iterated on dimensions from previous literature that have examined creativity, including fluency, originality, flexibility, and elaboration. Through multiple rounds of discussion and refinements, the dimensions of practicality, uniqueness, understandability, and well-implemented were formed. Furthermore, raters paired up together to discuss any discrepancies that might arise in how they evaluated the creation. However, despite the iterations, the consensus was either moderate or good, not reaching excellent. While this challenge may be unique to a platform and its tools, it nonetheless highlights the difficulty in third-person evaluations of creations in 3D space. The results begin to make sense when considered with bottom-up evaluations, such as the action-oriented measures (i.e., what a participant did throughout a session) and creation-oriented measures (i.e., the dimensions of the creations), suggesting that having multiple approaches is central to making sense of what participants build and create in VR. Future research should consider these approaches and how they can be further tailored to the affordances of VR.

Third, regarding the design, we note that all our activities were conducted in a social context, such that all participants were present and could watch one another build. This was not a collaborative task, but it was not necessarily individual, either, as participants could see each other's creations and redo or restart their own creations. Our goal was to broadly examine creation outcomes, as well as in relation to the constraint of virtual mobility. However, we did not explicitly compare outcomes against conditions in which individuals created in isolation or in collaboration. Given the social environment is a factor that influences creativity (Amabile, 2012), future research should account for not only individual and collaborative creative activities, but also those that are done along – but not together – with others.

Fourth, past research shows that the duration spent on creativity activities has ranged from 30 s, 10 min, and close to an hour (e.g., Blanchette et al., 2005; Lee et al., 2020; Tung & Chang, 2024). In the present study, we allotted ~15 min for the creativity activity each session. As a result, how much effect the duration of the creativity activity had on the outcomes remains unclear. Our decision was largely based on simulator sickness, which tends to occur after spending 30 min inside VR (Han et al., 2022). Although simulator sickness did not have a significant effect on any of the outcome variables (all p-values 0.542), we note that, because VR can be used for a wide range of activities and purposes, which require varying amounts of time spent in-headset, how much time should be, or *can* be, spent for creativity demands further research. Furthermore, we underscore that our participants were physically sitting down for most of the time during the activity, which may have influenced their experience with simulator sickness. Again, given VR can be used in a wide variety of contexts that may demand different physical mobilities, how this research translates to cases where physical and virtual mobility interact also remains a question.

Last, a limitation regarding the logistics of the study comes from the bottom-up evaluation of the creations. Due to how the data of certain virtual objects were logged, objects such as sticky notes became encoded as global objects that became challenging to pin to specific creators. This became increasingly challenging when the virtual environment was constrained in space and participants were simultaneously generating objects in proximity. As a result, sticky notes were not accounted for in the creation measures.

7.2. Implications

One of the main goals of this study was to evaluate *how* people use creation tools in VR and *what* they create. We specifically investigated how and what students create in a VR classroom setting. Our first implication, therefore, is directed to instructors and educators. Although digital technologies are core to creativity in education, there is limited work on how the affordances of such technologies intersect with creativity, and how they may play out in classroom settings (Henriksen et al., 2021). As digital technologies like VR are increasingly integrated into both traditional and hybrid forms of classroom instruction (e.g., Han & Bailenson, 2024; McGivney et al., 2022; Solak & Cakır, 2016; Vishwanath et al., 2017), how these tools are used in a creativity context needs more understanding. Our results suggest that, if the goal of using VR for creativity is to encourage students to build something that is practical, unique, and well-implemented, using a platform that has many options for 3D models and other tools may allow students to populate their space more and delete more may be beneficial. Furthermore, our present study highlights the influence of virtual mobility on outcomes. Although virtual mobility can be a constraint placed by the platform (i.e., different platforms provide different ways and freedom in moving around 3D space), it may also be a pedagogical constraint (i.e., dependent on the type of activity). Given past literature has shown association with creativity and positive student outcomes such as positive engagement and ideation (Gajda et al., 2017), choosing a platform and designing activities that allow students to move around freely to fully explore their surrounding space, rather than platforms and activities that may restrict *how* and *how much* students can move, is critical.

Our second implication is directed towards designers of creative platforms in VR, including those that allow users to build, piece, draw, visualize, and listen together. First, consider what tools are made available to users. How many 3D models can users bring into the world? Are primitive objects, such as building blocks, available? What kind of drawing tools are available? The availability of such tools can shape how and what people create. Furthermore, as aforementioned, virtual mobility can be a platform-specific constraint. Some platforms may allow limited virtual mobility, such that they do not allow for free translation or teleportation. How can users move around in such environments, if at all? Are users limited to creating in the immediate space around them, or can they have access to more space?

Our last implication more broadly concerns other technologies that rely on existing in virtual 3D space. As MR HMDs such as the Apple Vision Pro become more prevalent and attractive for use-cases such as collaborating and creating together, this naturally imposes some limitations that a completely virtual environment possible in VR HMDs may not. As the physical and virtual worlds merge, virtual mobility may become more limited, as there is less opportunity for users to explore entirely virtual spaces (i.e., the visual image of the physical world places some restriction). This restriction has implications for virtual mobility and how it can influence the creation process. How will space – both physical and virtual – be used to promote and democratize creativity?

Funding

This research was supported by the National Science Foundation grant (Award 1800922).

CRediT authorship contribution statement

Eugy Han: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Portia Wang:** Writing – review & editing, Data curation. **Cyan DeVeaux:** Writing – review & editing, Data curation. **Gabriella M. Harari:** Writing – review & editing, Writing – original draft, Supervision, Formal analysis. **Jeremy N. Bailenson:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of interests

We have no conflicts of interests to disclose.

Acknowledgments

The authors would like to thank Brian Beams, Daniel Akselrad, Hanseul Jun, Mark Roman Miller, and Mark York for their support with the course and running the study. The authors would also like to thank Benjamin Liao, Casey Manning, Benjamin Martinez, Umar Patel, and Neha Vinjapuri for their help with developing the virtual environments for the study. The authors would also like to thank Arya Marwaha, Caroline Abigail Graham, Francis Santiago, Nattakit Tankongchamruskul, Pedro Jeha Civita, and Vencent Vang for their help developing on and rating using the creativity coding scheme.

Data availability

Data will be made available on request.

References

Ahn, S. J.(Grace), Nowak, K. L., & Bailenson, J. N. (2022). Unintended consequences of spatial presence on learning in virtual reality. *Computers & Education, 186*, Article 104532. https://doi.org/10.1016/j.compedu.2022.104532

Alahuhta, P., Nordbäck, E., Sivunen, A., & Surakka, T. (2014). Fostering team creativity in virtual worlds. *Journal For Virtual Worlds Research, 7*(3). https://doi.org/10.4101/jvwr.v7i3.7062

Amabile, T. M. (1996). *Creativity in Context: Update To The Social Psychology Of Creativity* (1st ed.). Routledge. https://doi.org/10.4324/9780429501234

Amabile, T. M. (2012). *Componential theory of creativity*. Boston, MA: Harvard Business School.

Arns, L. L. (2002). *A new taxonomy for locomotion in virtual environments*. [Iowa State University]. https://doi.org/10.31274/rtd-180813-14259

Barton, K. (2009). MuMIn: Multi-model inference. R package version 1. 0. 0.

Basu, A. (2019). A brief chronology of virtual reality. arXiv preprint arXiv:1911.09605.

Blanchette, D. M., Ramocki, S. P., O'del, J. N., & Casey, M. S. (2005). Aerobic exercise and creative potential: Immediate and residual effects. *Creativity Research Journal, 17*(2–3), 257–264.

Boletsis, C. (2017). The new era of virtual reality locomotion: A systematic literature review of techniques and a proposed typology. *Multimodal Technologies and Interaction, 1*(4), 24. https://doi.org/10.3390/mti1040024

Bourgeois-Bougrine, S., Bonnardel, N., Burkhardt, J.-M., Thornhill-Miller, B., Pahlavan, F., Buisine, S., Guegan, J., Pichot, N., & Lubart, T. (2022). Immersive virtual environments' impact on individual and collective creativity. *European Psychologist, 27*(3), 237–253. https://doi.org/10.1027/1016-9040/a000481

Bowman, D. A., Koller, D., & Hodges, L. F. (1997). Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Proceedings of IEEE 1997 annual international symposium on virtual reality* (pp. 45–52). IEEE.

Bowman, D. A., Koller, D., & Hodges, L. F. (1998). A methodology for the evaluation of travel techniques for immersive virtual environments. *Virtual Reality, 3*(2), 120–131. https://doi.org/10.1007/bf01417673

Chang, Y., Kao, J.-Y., & Wang, Y.-Y. (2022). Influences of virtual reality on design creativity and design thinking. *Thinking Skills and Creativity, 46*, Article 101127. https://doi.org/10.1016/j.tsc.2022.101127

Chen, J., Huang, Y., Lin, K., Chang, Y., Lin, H., Lin, C., & Hsiao, H. (2019). Developing a hands-on activity using virtual reality to help students learn by doing. *Journal of Computer Assisted Learning, 36*(1), 46–60. https://doi.org/10.1111/jcal.12389

Cheng, V. M. (2011). Infusing creativity into eastern classrooms: Evaluations from student perspectives. *Thinking Skills and Creativity, 6*(1), 67–87.

Cherni, H., Métayer, N., & Souliman, N. (2020). Literature review of locomotion techniques in virtual reality. International Journal of Virtual Reality.

Churchill, E. F., & Snowdon, D. (1998). CVE'98: Collaborative virtual environments 1998. *ACM SIGGROUP Bulletin, 19*(2), 21–22.

Dong, Z., Clark, A., Ens, B., & Piumsomboon, T. (2022). Localization in immersive analytics with world-in-miniature, landmarks and constellations. In *Proceedings of the 2022 ACM symposium on spatial user interaction*. https://doi.org/10.1145/3565970.3568186

Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology, 1*. https://doi.org/10.3389/fpsyg.2010.00242

Feeman, S. M., Wright, L. B., & Salmon, J. L. (2018). Exploration and evaluation of CAD modeling in virtual reality. *Computer-Aided Design and Applications, 15*(6), 892–904.

Fillingim, K. B., Shapiro, H., Reichling, C. J., & Fu, K. (2021). Effect of physical activity through virtual reality on design creativity. *AI EDAM, 35*(1), 99–115.

Fleury, S., Agnès, A., Vanukuru, R., Goumillout, E., Delcombel, N., & Richir, S. (2020). Studying the effects of visual movement on creativity. *Thinking Skills and Creativity, 36*, Article 100661. https://doi.org/10.1016/j.tsc.2020.100661

Fleury, S., Blanchard, P., & Richir, S. (2021). A study of the effects of a natural virtual environment on creativity during a product design activity. *Thinking Skills and Creativity, 40*, Article 100828. https://doi.org/10.1016/j.tsc.2021.100828

Gajda, A., Beghetto, R. A., & Karwowski, M. (2017). Exploring creative learning in the classroom: A multi-method approach. *Thinking Skills and Creativity, 24*, 250–267. https://doi.org/10.1016/j.tsc.2017.04.002

Gibson, J. J. (1979). *The ecological approach to visual perception*. Psychology Press.

Goldschmidt, G., & Smolkov, M. (2006). Variances in the impact of visual stimuli on design problem solving performance. *Design Studies, 27*(5), 549–569. https://doi.org/10.1016/j.destud.2006.01.002

Gong, Z., & Georgiev, G. V. (2020). Literature review: Existing methods using VR to enhance creativity. In *Proceedings of the sixth international conference on design creativity (ICDC 2020)* (pp. 117–124).

Grimm, K. J., Ram, N., & Estabrook, R. (2016). *Growth modeling: Structural equation and multilevel modeling approaches*. The Guilford Press.

Han, E., & Bailenson, J. N. (2024). Lessons for/in virtual classrooms: Designing a model for classrooms inside virtual reality. *Communication Education, 73*(2), 234–243.

Han, E., Miller, M. R., DeVeaux, C., Jun, H., Nowak, K. L., Hancock, J. T., … Bailenson, J. N. (2023). People, places, and time: A large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. *Journal of Computer-Mediated Communication, 28*(2), zmac031.

Han, E., Nowak, K. L., & Bailenson, J. N. (2022). Prerequisites for Learning in Networked Immersive Virtual Reality. *Technology, Mind, and Behavior, 3*. https://doi.org/10.1037/tmb0000094 (4: Winter).

Henriksen, D., Creely, E., Henderson, M., & Mishra, P. (2021). Creativity and technology in teaching and learning: A literature review of the uneasy space of implementation. *Educational Technology Research and Development, 69*(4), 2091–2108. https://doi.org/10.1007/s11423-020-09912-z

Houzangbe, S., Masson, D., Fleury, S., Gómez Jáuregui, D. A., Legardeur, J., Richir, S., & Couture, N. (2022). Is virtual reality the solution? A comparison between 3D and 2D creative sketching tools in the early design process. *Frontiers in Virtual Reality, 3*, Article 958223.

Hu, R., Wu, Y. Y., & Shieh, C. J. (2016). Effects of virtual reality integrated creative thinking instruction on students' creative thinking abilities. *Eurasia Journal of Mathematics, Science and Technology Education, 12*(3), 477–486.

Hu, X., Nanjappan, V., & Georgiev, G. V. (2021). Bursting through the blocks in the human mind: Enhancing creativity with extended reality technologies. *Interactions, 28*(3), 57–61.

Jaeger, B. (2017). r2glmm: Computes R squared for mixed (Multilevel) models.

Kaufman, J. C., & Glăveanu, V. P. (2021). An overview of creativity theories (Eds.). In J. C. Kaufman, & R. J. Sternberg (Eds.), *Creativity: An introduction* (pp. 17–30). Cambridge University Press.

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments, 21*(4), 373–387.

Kim, K. H. (2006). Can we trust creativity tests? A review of the torrance tests of creative thinking (TTCT). *Creativity Research Journal, 18*(1), 3–14. https://doi.org/10.1207/s15326934crj1801_2

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lau, K. W., & Lee, P. Y. (2012). The use of virtual reality for creating unusual environmental stimulation to motivate students to explore creative ideas. *Interactive Learning Environments, 23*(1), 3–18. https://doi.org/10.1080/10494820.2012.745426

Lee, J. H., Yang, E. K., Lee, E. J., Min, S. Y., Sun, Z. Y., & Xue, B. J. (2020). The use of VR for collaborative exploration and enhancing creativity in fashion design education. *International Journal of Fashion Design, Technology and Education, 14*(1), 48–57. https://doi.org/10.1080/17543266.2020.1858350

Lee, J. H., Yang, E. K., & Sun, Z. Y. (2019). Design cognitive actions stimulating creativity in the VR design environment. In *Proceedings of the 2019 on creativity and cognition*. https://doi.org/10.1145/3325480.3326575

Liu, J., Burkhardt, J.-M., & Lubart, T. (2023). Boosting creativity through users' avatars and contexts in virtual environments—A systematic review of recent research. *Journal of Intelligence, 11*(7), 144. https://doi.org/10.3390/jintelligence11070144

Long, H. (2014). An empirical review of research methodologies and methods in creativity studies (2003–2012). *Creativity Research Journal, 26*(4), 427–438.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Markowitz, D. M., Laha, R., Perone, B. P., Pea, R. D., & Bailenson, J. N. (2018). Immersive virtual reality field trips facilitate learning about climate change. *Frontiers in psychology, 9*, 2364.

Mayer, R. E. (2002). Multimedia learning. In *Psychology of learning and motivation, 41* pp. 85–139). Academic Press. Vol.

McGivney, E., Tribe, C., & Feng, T. (2022). Remote learning with virtual reality technologies: Student mastery, community and agency. *Educational Innovations and Emerging, Technologies, 2*(1). https://doi.org/10.35745/eiet2022v02.01.0004

McVeigh-Schultz, J., Márquez Segura, E., Merrill, N., & Isbister, K. (2018). What's it mean to "be social" in VR?. In *Proceedings of the 2018 ACM conference companion publication on designing interactive systems*. https://doi.org/10.1145/3197391.3205451

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review, 19*, 309–326.

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of Educational Psychology, 94*(3), 598–610. https://doi.org/10.1037//0022-0663.94.3.598

Nakagawa, S., & Schielzeth, H. (2012). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4* (2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Noë, A. (2004). *Action in perception*. MIT Press.

Oppezzo, M., & Schwartz, D. L. (2014). Give your ideas some legs: The positive effect of walking on creative thinking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(4), 1142–1152. https://doi.org/10.1037/a0036577

Palanica, A., & Fossat, Y. (2022). Effects of nature virtual backgrounds on creativity during videoconferencing. *Thinking Skills and Creativity, 43*, Article 100976. https://doi.org/10.1016/j.tsc.2021.100976

Prats, M., Lim, S., Jowers, I., Garner, S. W., & Chase, S. (2009). Transforming shape in design: Observations from studies of sketching. *Design Studies, 30*(5), 503–520.

Rivard, R. M. H., & Leclercq, R. J. P. (2006). From architectural sketches to feasible structural systems. *Design computing and cognition'06* (pp. 675–694). Netherlands: Springer.

Said-Metwaly, S., Van den Noortgate, W., & Kyndt, E. (2017). Approaches to measuring creativity: A systematic literature review. *Creativity. Theories–Research-Applications,, 4*(2), 238–275.

Sarkar, P., & Chakrabarti, A. (2008). *Studying engineering design creativity*. Workshop on Studying Design Creativity.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. https://doi.org/10.1037//0033-2909.86.2.420

Slepian, M. L., & Ambady, N. (2012). Fluid movement and creativity. *Journal of Experimental Psychology: General, 141*(4), 625.

Solak, E., & Cakır, R. (2016). Investigating the role of augmented reality technology in the language classroom /Istraživanje uloge tehnologije proširene stvarnosti u nastavi jezika. *Croatian Journal of Education - Hrvatski Časopis Za Odgoj i Obrazovanje, 18*(4). https://doi.org/10.15516/cje.v18i4.1729

Sun, R., Wu, Y. J., & Cai, Q. (2018). The effect of a virtual reality learning environment on learners' spatial ability. *Virtual Reality, 23*(4), 385–398. https://doi.org/10.1007/s10055-018-0355-2

Suwa, M., Purcell, T., & Gero, J. (1998). Macroscopic analysis of design processes based on a scheme for coding designers' cognitive actions. *Design Studies, 19*(4), 455–483. https://doi.org/10.1016/s0142-694x(98)00016-7

Tang, C., Mao, S., Naumann, S. E., & Xing, Z. (2022). Improving student creativity through digital technology products: A literature review. *Thinking Skills and Creativity, 44*, Article 101032. https://doi.org/10.1016/j.tsc.2022.101032

Thimbleby, H. (1990). *User interface design*. Association for Computing Machinery.

Torrance, E. P. (1966). *Torrance tests of creative thinking*.

Tung, Y. H., & Chang, C. Y. (2024). How three-dimensional sketching environments affect spatial thinking: A functional magnetic resonance imaging study of virtual reality. *PloS One, 19*(3), Article e0294451.

Vishwanath, A., Kam, M., & Kumar, N. (2017). Examining low-cost virtual reality for learning in low-resource environments. In *Proceedings of the 2017 conference on designing interactive systems*. https://doi.org/10.1145/3064663.3064696

Wang, P., Miller, M. R., Han, E., DeVeaux, C., & Bailenson, J. N. (2024). Understanding virtual design behaviors: A large-scale analysis of the design process in Virtual Reality. *Design Studies, 90*, 101237.

Wickham, H. (2011). ggplot2. *WIREs Computational Statistics, 3*(2), 180–185. https://doi.org/10.1002/wics.147

Xia, T., Sun, Y., An, Y., & Li, L. (2023). The influence of music environment on conceptual design creativity. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1052257

Yalcinalp, S., & Avci, Ü. (2019). Creativity and emerging digital educational technologies: A systematic review. TOJET: The Turkish. *Journal of Educational Technology, 18*(3), 25–45.

Yang, E. K., & Lee, J. H. (2020). Cognitive impact of virtual reality sketching on designers' concept generation. *Digital Creativity, 31*(2), 82–97.

Yang, X., Lin, L., Cheng, P.-Y., Yang, X., Ren, Y., & Huang, Y.-M. (2018). Examining creativity through a virtual reality support system. *Educational Technology Research and Development, 66*(5), 1231–1254. https://doi.org/10.1007/s11423-018-9604-z