Spatial Hand Actions: Exploring the Hand Actions used to Represent Spatial Thinking for 3D Assembling Tasks

Prashant Rawat
Dalhousie University
Halifax, Canada
prashant.rawat@dal.ca

Mayra Donaji Barrera Machuca
Dalhousie University
Halifax, Canada
mbarrera@dal.ca

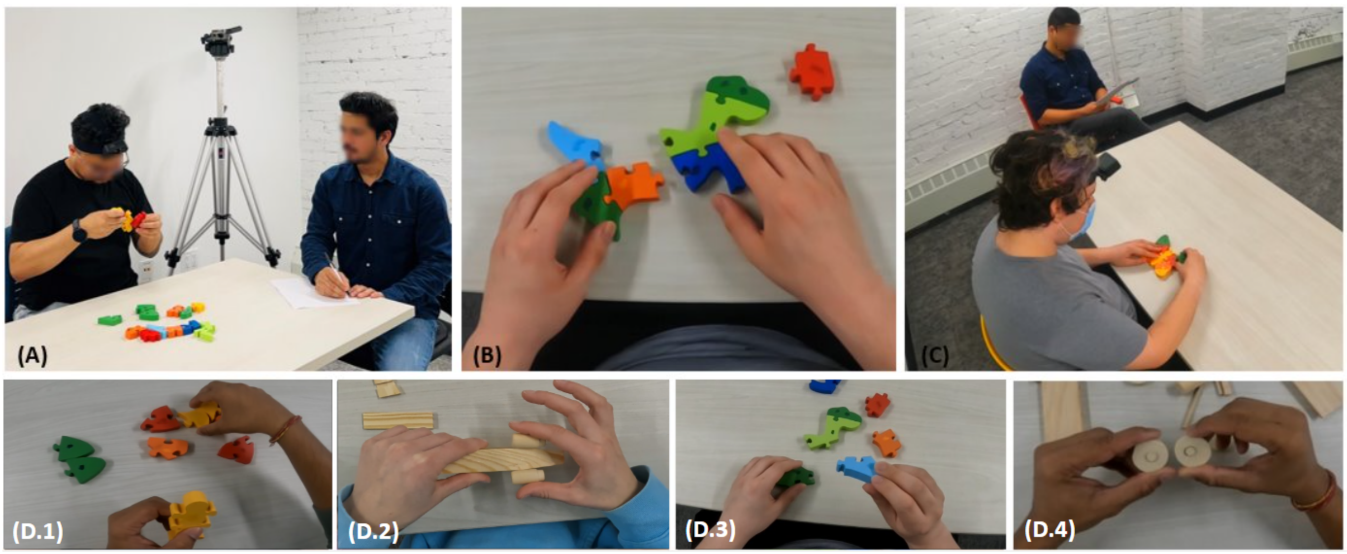


Figure 1: A) Experimental Setting, B) Egocentric View of the Experiment, C) Top View of the Experiment, and D) Examples of Spatial Hand Actions (D.1: Grabbing Object, D.2: Placement of Object, D.3: Orientation, D.4: Comparison)

Abstract

When designing 3D objects in 3D virtual environments using naturalistic 3D user interfaces, people use their hands to manipulate the environment and objects inside it. At the same time, people utilize their spatial thinking to understand the spatial relationship of the objects in the scene. Yet, the relationship between spatial thinking and hand actions remains unclear. Here, we present a user study with 18 participants that examines the association between 3D assembling tasks and reflective hand movements that allow people to enhance their spatial thinking. Utilizing a mixed-methods protocol, we identified nine Spatial Hand Actions and three Spatial Themes people use when designing 3D objects. Then, we analyzed a subset of the participants to understand the relationship between Spatial Hand Actions and spatial abilities. Our results will help develop better hand-based naturalistic 3DUI that considers the spatial thinking abilities of the users.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; **Gestural input**; Virtual reality.

Keywords

Virtual Reality, 3D Design, Hand Gestures, Spatial Thinking, Observational Study

ACM Reference Format:
Prashant Rawat and Mayra Donaji Barrera Machuca. 2025. Spatial Hand Actions: Exploring the Hand Actions used to Represent Spatial Thinking for 3D Assembling Tasks. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 21 pages. https://doi.org/10.1145/3706598.3713424

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713424

1 Introduction

Current technological advances have made it possible to create naturalistic hand-based 3D user interfaces (3DUI) that seamlessly bridge the physical and virtual worlds, facilitating more natural and embodied user experiences [18]. Examples of naturalistic 3DUI include grabbing a 3D object with the hand and rotating the object by moving the hand. These naturalistic 3DUI have high expressiveness [117] and spatial precision to replicate the real-world interaction [104]. Moreover, incorporating hand-based interactions into 3DUI creates intuitive applications by aligning with users' innate behaviors, making them more intuitive and easier to learn [108, 114]. Current hand-based 3DUI for immersive technologies like virtual reality (VR) use communicative gestures, e.g., symbolic gestures, as input methods [90, 100, 103]. Communicative gestures like pinch, swipe, thumbs up or hand wave are representational and are often used to communicate thoughts and ideas to other people or with the system [3, 84]. Yet, using communicative gestures in 3DUI removes some of the benefits of natural interactions as they provide limited support for self-understanding critical relationships in 3D spaces [59, 67, 80]. Plus, users need to memorize their meaning [110, 150].

This paper focuses on naturalistic 3DUI that utilize reflective hand movements as inputs for 3D design [75, 109, 149]. Building 3D objects in 3D space is closely intertwined with the human capability to envision objects in 3D [114], perceive their spatial arrangement [47], and imagine the eventual appearance of the object [39]. This mental process relies on people's spatial thinking abilities [118]. Yet, when engaging in spatial thinking, people encounter challenges that place a substantial cognitive load on their spatial skills [141], which can impact the performance of individuals designing 3D objects. One way people solve this challenge is to use reflective hand movements to communicate ideas [31]. Reflective hand movements occur spontaneously during self-explanation, helping with cognitive processing [3, 135], which aid in activating spatial information, maintaining it in memory [126] and highlighting spatial information in reasoning and problem solving [5]. Past research has shown that spatial information carried in by reflective hand movements enhances spatial learning, thinking and understanding across a wide range of tasks, including word learning [101], problem solving [14, 136] and designing [25, 108, 114]. Thus, it is important to understand how people use their hands while building 3D objects, as their movements allow them to understand better the spatial relationships within 3D space [5, 126].

One challenge of studying reflective hand movements for immersive technologies is the limitations in tracking and hand recognition, as these actions are non-symbolic and can be challenging to detect. Here, we follow a long list of previous work that use elicitation studies to understand people's thoughts about space [30, 31, 40, 59, 93, 120, 141] or how their spatial abilities affect their performance while creating new objects in VR [10, 12, 46, 63, 128]. For example, Cortesa et al. [40] used a block construction task to gain insight into the participants' cognitive processes and spatial knowledge. Yet, the findings across these studies are about spatial thinking and its relationship to single design activities, e.g., 3D drawing [10, 81], VR interactions [71], and block construction [93, 114]. Thus, this paper aims to identify the reflective and naturalistic hand-actions that people use to aid their spatial thinking during 3D assembling tasks to create a database of hand actions that help improve hand action-based design tools for immersive technologies.

To achieve our goal, we studied how people utilize reflective gestures to help them think in 3D, focusing on 3D shapes and space representations with their hands. In a user study, 18 participants solved four 3D assembling tasks while following the "think-out-loud" method [113]. Then, four domain experts used a mixed-methods protocol that included factor analysis, structural and protocol analysis, and thematic analysis to understand the relationship between reflective hand actions and spatial thinking. We found nine unique Spatial Hand Actions, e.g., hand movements people use to aid their spatial thinking when solving a 3D design problem, and three Spatial Themes, representing the relationship between these Spatial Hand Actions and the person's spatial thinking, as shown in Figure 6. We also analyzed a subset of our participants with high or low spatial abilities to understand the effect of spatial abilities on Spatial Hand Actions. Our result shows that the frequencies of some Spatial Hand Actions correspond with the participants' spatial abilities, e.g., individuals with higher spatial abilities tend to employ comparatively more hand actions.

In summary, this paper presents a new classification of hand movements called Spatial Hand Actions, which helps people with their spatial thinking when designing 3D objects. Spatial Hand Actions address a gap in the spatial thinking-action-object relationships research, as they merge a cognitive process (spatial thinking) with hand movements. These findings extend previous work on the utility of hand actions when designing new objects [39] and spatial thinking when designing 3D objects [92]. We also extend previous work that analyzes actions for design [59] and tries to understand spatial thinking while designing 3D objects [35, 61]. Finally, we provide insights to the designers of naturalistic 3DUI that plan to use hand actions as input methods in their design applications. Our contributions are the following

- **Identifying Spatial Hand Actions and Spatial Themes**: We found nine unique hand actions related to three specific spatial thinking abilities by analyzing the user's actions and explanations during a 3D assembling task.
- **Identifying the relationship between Spatial Hand Actions and spatial abilities**: We found that participants with high spatial abilities utilize more Spatial Hand Actions than participants with low spatial abilities.
- **Based on Empirical Findings, proposing Actionable Insights for Designers of naturalistic 3DUI**: We propose recommendations for designers of naturalistic 3DUI, that help them implement the identified Spatial Hand Actions into their systems. Moreover, to advance research in this direction, we release our extensive Spatial Hand Actions dataset here: https://git.cs.dal.ca/research1/vertexlab/spatial-hand-actions-dataset.git, comprising 2431 hand actions collected to aid spatial thinking in 3D assembling tasks.

2 Related Work

This section discusses past work on naturalistic hand-based 3DUI for 3D design, spatial thinking for 3D design, and reflective hand actions for spatial thinking.

2.1 Naturalistic Hand-based 3DUI for 3D Design

Researchers have focused on using hand gestures to interact with computers for more than 30 years [53, 147]. This paper distinguishes between reflective hand actions and communicative gestures. Communicative gestures are symbolic movements, e.g., waving, that convey meaning without causing physical environmental change and influence communication, often displaying cognitive processes and facilitating interaction with others [29]. In contrast, reflective hand actions are linked to self-understanding of spatial thinking and manipulating 3D objects in space rather than symbolic communication [3, 135].

For naturalistic hand-based 3DUI, gestures emulate real-world physical interactions [69, 155], allowing users to manipulate and explore virtual environments more naturally [78, 79]. These hand-based 3DUI provide an intuitive and immersive interaction method [6, 19, 121]. The advantages of using naturalistic hand-based 3DUI include a seamless experience without requiring extensive learning or adaptation [44]. Moreover, several studies have highlighted the importance of naturalistic hand actions integration in improving user performance in task manipulation, design, and data visualization [94, 153]. For example, naturalistic hand actions enhance spatial understanding, increase immersion, and reduce the learning curve for interacting with complex 3D environments [38, 125, 127]. Past work has highlighted challenges users face in VR environments when manipulating and merging virtual objects using controllers or author-defined hand gestures, which can feel less natural [43, 54, 129] and more difficult to learn [111]. For example, Alem et al. [2] evaluated expressive gesture representations, such as overlaying hands and cursor pointers, finding limited performance improvements. Few other studies explored user-defined gestures to understand better their impact on the structure of collaboration and discourse in 3D design tasks [83, 148]. While gesture-based interaction has a long research story, few past work focus on understanding how people use reflective hand movements to help with spatial thinking while solving a 3D assembling task.

2.2 Spatial Thinking and 3D Design

Generally speaking, spatial ability is the capacity to understand, reason, and remember the visual and spatial relations among objects or space [37, 66, 95, 96, 118]. Spatial abilities consist of diverse skills that impact performance in tasks involving spatial-visual development processes [10, 31, 93, 120]. Still, no consensus exists about the name, or the types of abilities involved. For example, Kimura et al. [82] identified six spatial factors: spatial orientation, spatial location memory, targeting, spatial visualization, disembedding, and spatial perception. Burton et al. [23] indicated that spatial abilities encompass spatial visualization and orientation; Cho et al. [33] asserted that they contain spatial perception, rotation, and imagination; Ho et al. [68] described spatial abilities as converting 3D structures into 2D images, envisioning 3D structures of 2D objects [102], and identifying changes in object structural direction [115]. According to Hayatpur et al. [63], spatial abilities involve converting 2D and 3D images, requiring object generation, retention, and manipulation. Finally, Hegarty et al. [66] argued that spatial relations, orientation, and visualization are the most recognized factors in understanding spatial thinking within design tasks.

For 3D design, past work has focused on understanding how spatial thinking affects design performance [8, 40, 42, 72, 81, 114] and found a strong association between design behavior and individuals' spatial abilities [10, 31, 63]. For example, Barrera Machuca et al. [10] found that spatial abilities impact the user's 3D sketching task performance regarding shape accuracy. Another example is Chandrasegaran et al. [31], who studied VR sketching tools' innovative approach, enabling designers to express and evaluate their ideas visually. Past work also found that the spatial information embedded within objects' geometric and material properties, such as size, shape, location, and orientation, also affect user interaction [47, 120]. Moreover, past work shows the importance of spatial abilities for design tasks by showing the importance of understanding spatial factors [118] and their impact on design development [93, 132], emphasizing the need for tailored design interfaces that accommodate diverse spatial skills [10, 26, 130]. For example, Liao et al. [93] study how people's spatial ability affects their 3D sketching and found that the ability of spatial visualization is interlinked with the acquisition and understanding of architectural plan development, which requires the skills of creating structural 3D images.

Past work also found that people who lack abilities or experience in using spatial thinking may encounter difficulties in designing tasks [31, 92, 98, 122]. One limitation of all these past works is that they focus on identifying the relationship between spatial thinking and specific design activities like sketching [20, 137], 3D design performance [33], data visualization [134] and architecture [33, 77]. See Table 1 for a summary of past work in the area. We extends this past work by focusing on 3D assembling tasks, which are general 3D design task involving spatial language, directional gestures, symmetry recognition, map reading, tangram puzzles, Lego building, and block construction. Jones et al. [74] proposed a theoretical model for 3D design understanding, comprising three phases: Visualization (perceiving spatial relations), Construction (creating and mentally rotating images), and Reasoning (solving design tasks), as these tasks require individuals to visualize spatial relationships, mentally transform objects across scales, rotate objects to view different perspectives, and generate new viewing angles. Additionally, 3D assembling tasks involve recalling spatial configurations and representations.

2.3 Reflective hand actions for spatial thinking

Our work focused on identifying reflective hand-actions, which serve as a medium for externalizing spatial thinking cognitive processes, providing a tangible and intuitive way to engage with and manipulate spatial information. Yet, it is challenging to identify reflective hand actions as these movements are integrated into building the object and are non-symbolic, and there is an absence of techniques for testing user experiences [154]. Moreover, most past work on spatial thinking measures spatial abilities using geometric representations [65] or standardized metrics derived from tasks like paper-folding and card-rotation [49].

Despite these challenges, past work has used elicitation studies for understanding user preferences and designing natural, spatially aware hand interaction input methods[151]. For example, Vatavu et al. [140] ran an elicitation study to compare mid-air gestures with remote control inputs in home entertainment systems, finding that users preferred familiar interaction models like point-and-click and drag-and-drop. Jahani et al. [70] introduced descriptive mid-air

Table 1: Relevant work of Spatial Thinking for enhancing user experience in 3D Design

Paper Id	Design Task Type	Population	Analysis Approach	Spatial ability test	Spatial Ability Evaluated	Focus	Main finding
[23]	Imagery task	Psychology Students	Qualitative: Factor Analysis	Paper Folding (Vz-2), Speed Rotation (SR), Closure Speed (CS), Flexibility of Closure (CF), and Perceptual Speed (P)	Visual Imaginary	Hypothetical imagery vs Experimental imagery	High Hypothetical Imagery in Spatial Ability concepts
[33]	Picture selection	Architecture Students	Quantitative: Two-way ANOVA	Paper Folding (Vz-2), Spatial Orientation Test (SOT), Architectural Spatial Ability Test (ASAT)	Spatial Orientation, Spatial Visualization	Investigation of creativity, spatial ability, and visual cognitive with students' performance in design	Positive correlation exists between spatial abilities and visual cognitive styles
[30]	Sketching	Undergraduate Students	Quantitative: One-way ANOVA	None	Spatial Orientation	Identification of object drawings from scrambled order	In Spatial Relation, the way objects are parsed affects identification with improvised view
[31]	Sketching	Graduate Students	Qualitative: Protocol and Structural analysis	None	None	Identification of designers intent in sketching and non-sketching actions	Actions in sketching, along with verbalization, capture intent, process, and rationale
[24]	Sketching	Art Students	Quantitative: One-way ANOVA	Vividness of Visual Imagery Questionnaire (VVIQ), Questionnaire Upon Mental Imagery (QMI), Mental Rotation Test (MRT), Bricks Task (S(MA), Directed Mental Synthesis Task (DMS)	Mental Image	Mental image ability of experts vs novice	High imagery performance in experts
[36]	Sketching	Professionals	Qualitative: Protocol Analysis	None	None	Identification of techniques to capture cognitive behavior	Architects has longer dependencies, deeper thoughts than students in focus of attention
[114]	Sketching	Architecture Students	Qualitative: Thematic analysis	None	Visual Thinking, Mental Image	Impact of Immersive sketching on user's visual transformation, visual-spatial reasoning and visual memory and comprehension	3D sketching supports visual thinking and communication behaviors associated with PB sketching, CAD modeling, physical model-making and gesturing, all within the same space
[12]	Block Puzzle	Children under 15 years	Qualitative: Inductive Coding	None	Spatial Orientation	Spatial Thinking strategies for children in VR hand tracking	One/two finger pinch-based gesture for VR tracking
[40]	Block puzzle	Professionals	Qualitative: Behavioral Coding	None	Spatial Building	Reveal cognitive processes, capture building expertise	Cognitive process affects spatial construction tasks
[92]	Block puzzle	Professionals	Qualitative: Think-aloud observation	None	Divergent Thinking	Blocks replace sketching in design	LEGO substitute sketching for non-designers that supports divergent thinking in HCI conceptual ideation phase
[93]	Carton box design	Undergraduate Students	Quantitative: Factor analysis	Spatial Ability Test (SAT)	Spatial Relation, Spatial Orientation, Spatial Visualization	Effect of spatial abilities	Spatial Visualization affects box design performance
This paper	Multiple Assembling puzzles	Undergraduate Students	Qualitative: Protocol and Structural Analysis, Thematic analysis	Spatial Orientation Test (SOT) and Paper Folding (Vz-2)	Spatial Orientation, Spatial Visualization, Spatial Relation	Gestures that better express spatial thinking and enhance cognitive process	Spatial Thinking Gestures, Spatial Ability affects their Spatial Gestures

gestures, while Lee et al. [90] combined gesture elicitation with the Wizard-of-Oz technique, enabling collaborative development of gestures through performer-recognizer interaction. Past research has provided a comprehensive understanding of the evolution and application of gesture elicitation methodologies, offering key insights for advancing gesture design [143, 144].

Moreover, past work has attempted to quantify an individual's spatial thinking using hand actions. Examples include Verdine et al. [142], who analyzed children's behavior during free block play, evaluating their structural creations and correlating individual scores with spatial skill success. Carlson et al. [27] demonstrated that individuals with higher spatial abilities exhibit superior performance in assembling tasks, such as solving wooden puzzles. Other work explored the challenges of deciphering the connection between hand movements and cognitive processes during spatial tasks [7, 31, 39, 91]. For example, Chandrasegaran et al. [31] aimed to connect designers' actions to design intent and found gaps in gesture representations due to the time difference between gestures and verbal descriptions. Baykal et al. [12] explore the connection between hand gestures and spatial thinking. Yet, they focus on non-gravity spatial manipulations in VR. They emphasized potential benefits for children's spatial thinking but used tangram objects, which differ from other spatial manipulations like blocks regarding surface characteristics. In opposition to most of this work that focused on sketching or solving a single-block puzzle (Table 1), our work focuses on different 3D block-assembling puzzles and identifying the relationship between reflective hand actions and spatial abilities.

3 User Study

In this user study, we focus on how people use hand actions when assembling physical 3D objects, as that will allow us to uncover the reflective hand actions people adopt to help them in their spatial thinking.

3.1 Research Questions

Our research questions are the following:

- **RQ1:** *Are there hand actions that aid a person's spatial thinking when solving a 3D assembling task?* RQ1 aims to identify the Spatial Hand Actions, i.e., hand movements, that help people translate their spatial thinking thoughts into actions.
- **RQ2:** *How do Spatial Hand Actions relate to a person's spatial thinking abilities?* RQ2 aims to comprehend the relationship between a specific Spatial Hand Actions and a specific spatial thinking ability, e.g., spatial visualization, spatial orientation, and spatial relations. We call this relationship a Spatial Themes.
- **RQ3:** *Is there a difference in the Spatial Hand Actions used between people with high and low spatial abilities?* RQ3 aims to identify the differences in Spatial Hand Actions that exist based on the spatial ability of the individual.

We based RQ1 and RQ2 on previous work that found that hand movements and design intent are pivotal in enhancing spatial thinking during design tasks [7, 39, 59]. These movements aid in perceiving, manipulating, and reasoning about spatial relationships, which are crucial aspects of 3D design [10, 114]. RQ3, on the other hand, zeroes in on the influence of the user's spatial abilities on their performance in 3D assembling tasks. Previous research has illuminated the substantial impact of an individual's spatial thinking on their capacity to excel in such tasks [30, 59, 93, 107, 120]. Thus, by delving into these research questions, we aim to extend the knowledge in 3D design and spatial thinking, shedding light on the intricate connection between Spatial Hand Actions, spatial abilities, and design performance. Ultimately, our goal is to inform the creation of an adaptable UI for VR/AR capable of accommodating users with diverse spatial skills. The knowledge we acquire can also help other areas like spatial abilities education, as it will be possible to develop VR/AR applications that use Spatial Hand Actions.

3.2 Methodology

3.2.1 Participants. We recruited 18 participants (10 females, 8 males) aged 18 to 40 (M=24.83, STD=6.24). Recruitment was conducted through email advertisements sent via university listservs and flyers posted on university notice boards across various campus locations accessible to students and the general public. Eight participants had no undergraduate degree, six were master's students, and three were PhD students. Their backgrounds included computer science, health, environmental management, psychology, and business management. Among the participants, 16 were right-handed, while two were left-handed. We inquired about their familiarity with block construction and assembling wooden objects. Ten participants identified themselves as beginners, four as intermediate, and another four as experts in these tasks. Regarding their prior experience with 3D design, eight participants had no previous exposure, seven reported one to three years of experience, two claimed five years, and one nine years of 3D design experience. Lastly, these participants got their spatial ability measured using the Paper Folding test (VZ-2) [1, 49], which evaluates their aptitude for converting between 2D and 3D representations. Participants scored between 9 and 20 (M=14.05, STD=3.11). Participants also filled the Spatial Orientation test (SOT) [52], which assesses participants' ability to mirror and reverse mental images. Participants scored between 2 and 20 (M=13.88, STD=4.96). We chose those tests as they evaluate the skills we were interested in, e.g., VZ-2 for spatial visualization and relation and SOT for spatial orientation. Moreover, past work has used them in studies about the impact of spatial abilities on design tasks [10, 11].

3.2.2 3D Assembling Tasks. Participants solved four 3D assembling tasks while seated on a table, manipulating shapes of different textures and colors to form 3D objects. All four tasks are shown in Figure 2. *Tasks 1 and 3* require participants to arrange irregularly shaped components correctly to form an object. Each component has a distinct 3D volume and varying textures and shapes. Participants must use their visual synthesis and transformation abilities for these tasks. *Task 2 and 4* consists of random-colored pieces with irregular shapes. Participants must join these pieces together to construct a specific shape. Participants had to use their visual-spatial reasoning and comprehension abilities for these tasks. The tasks were based on Shah et al. [130]'s framework for visual thinking.

We decided to use real-world tasks instead of a VR or AR application to prevent issues with the technology. For example, previous work about hand movements in VR suggested using real-world objects to avoid technical problems [12]. Also, Jetter et al. [72] demonstrated that individuals lacking prior VR experience encountered difficulties while interacting with VR. Finally, as previous 3D sketching experiments [31, 92] found that real-life Lego blocks support the same divergent thinking as 3D sketching in VR due to their shared ability to externalize abstract spatial ideas and facilitate iterative exploration of spatial configurations, we expect our results to extend to immersive environments using naturalistic 3DUI.

3.2.3 Procedure. The experiment was conducted within a noise-free, empty room, ensuring an environment conducive to focused participation. Only one participant and a researcher were in the room during the study. Figure 1A shows the setup. We captured the participant's actions with two GoPro Hero 10 Black cameras [60]. One camera was affixed to a head mount with a strap, capturing the participant's egocentric view while they engaged with the design tasks (See Figure 1B). The second GoPro camera is mounted on a tripod, providing a top-down view of the participant's work (See Figure 1C).

Upon arrival, the participants reviewed and signed the consent form — the Institutional Ethics Review Board of *[REMOVED FOR REVIEW]* approved the study design. Then, participants listened to

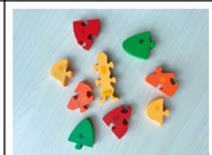
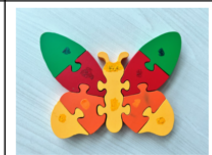
S. No.	3D Design	Spatial Ability	Scrambled Image	Build Image
Task 1	Wooden Block (Train)	Visual synthesis, Visual transformation		
Task 2	Lego Construction (Butterfly)	Visual Spatial reasoning, Visual comprehension		
Task 3	Wooden Block (Airplane)	Visual synthesis, Visual transformation		
Task 4	Lego Construction (Dinosaur)	Visual Spatial reasoning, Visual comprehension		

Figure 2: The four 3D assembling tasks used in the user study

a brief explanation of the experiment procedure and completed a pre-experiment questionnaire about their demographics and prior design experience. Next, participants answered the spatial abilities tests (SOT and VZ-2). Afterward, each participant was required to complete the four 3D assembling tasks described within five minutes while being recorded with the cameras. As participants engaged in the design tasks, they followed the "think-out-loud" method [113] so we could understand their thought processes. Past work has found that the "think-out-loud" methodology does not significantly disturb the cognitive process [138], and when managed appropriately, it can provide valuable insights without majorly disrupting task performance [32]. Therefore, we instructed participants to convey their explanations in a manner that a child or a visually impaired person could understand. In cases where participants did not talk, researchers asked questions to elucidate the reasoning behind each hand action used and to detail how specific properties of the object impacted their selection of a particular action. These questions were aligned with specific spatial factors identified through a literature review. See Table 4 for the questions.

After completing a task, participants took a 2-minute break before proceeding to the next task. After completing all tasks, participants completed a post-study questionnaire that took approximately 10 minutes. We also did a post-task interview with each participant to delve deeper into their thought processes during the design tasks. The interview questions were inspired by existing literature and categorized according to spatial thinking factors. See Supplementary Materials 8 for the questions. The interviews lasted around 20 minutes. As a result, the entire study lasted for a total of 60 minutes.

3.3 Data Analysis

This paper follows a multi-step process that involved various techniques previously used to study spatial thinking. Figure 3 shows the whole process. The participant's videos were automatically transcribed using the Microsoft Word Web App's transcription feature [105]. One author reviewed the transcripts to fix any transcription mistakes. For coding the videos, we used NVivo 12, released 2017 [48], which synchronizes the analysis of video footage and its corresponding transcripts for this step. We used Microsoft Excel [106] for the thematic analysis.

Four coders conducted the data analysis of the videos and interviews. Coder A ran the user study and defined the three types of spatial thinking abilities. Coder B had previous qualitative research experience, is a spatial thinking expert and provided guidance and insights after each step of the data analysis. Coder A and B are authors of the paper, and their involvement in the data analysis was to leverage familiarity with the data as key to analysis [17, 21]. Finally, we hired coders C and D for the qualitative analysis as they had experience with it. Still, they had no relationship to the paper or the research protocol, which increased the validity of our results.

3.3.1 Factor Analysis. The National Research Council [41] introduced three spatial contexts of learning to think spatially: life space (understanding the nature of space), physical space (representing spatial information), and intellectual space (reasoning about spatial

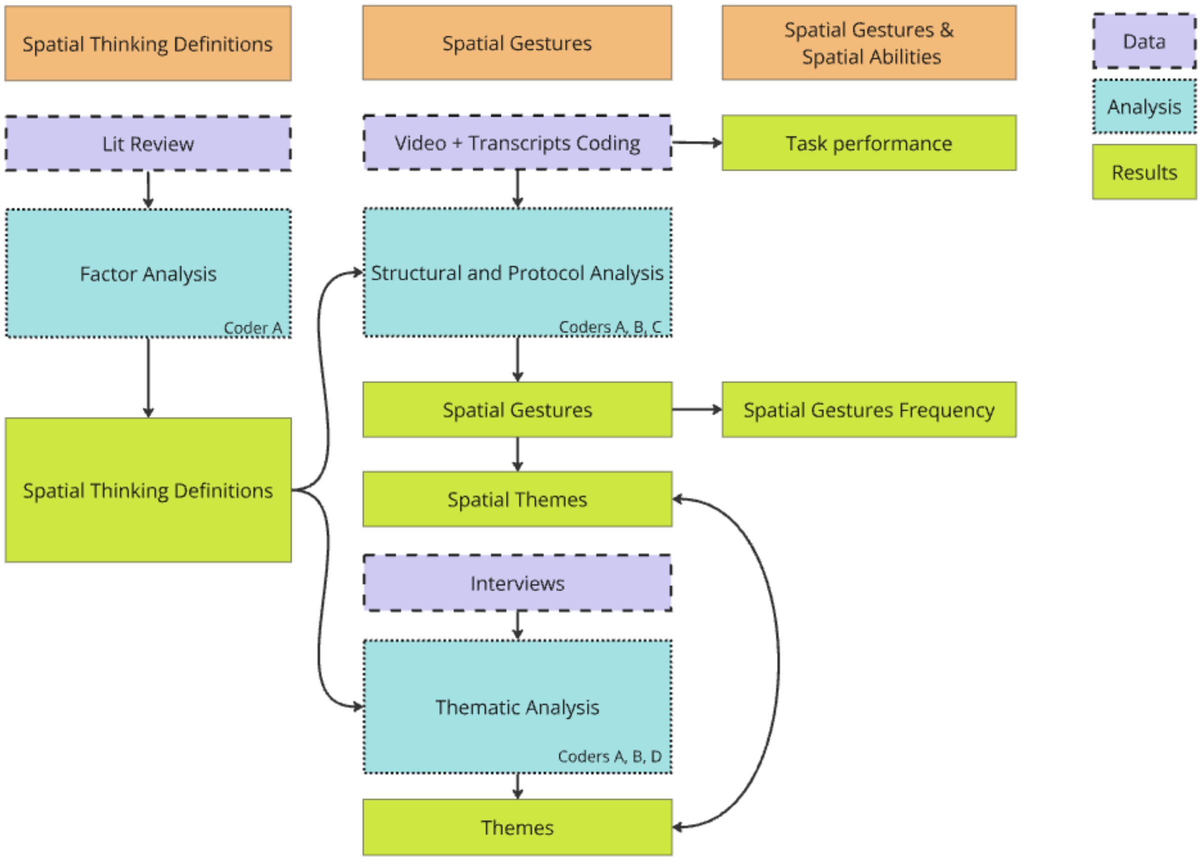


Figure 3: The Data Analysis Process

relationships). The authors had also provided interdisciplinary definitions of spatial thinking that is criticized for its lack of conceptual framework [55]. Further, spatial thinking was used as common umbrella term for the human abilities to direct optical stimuli received by the eye into the brain such as mentally visualizing and manipulating objects [112], adopting different perspectives [89], perceiving motion [119], and executing spatial actions [13]. Therefore, by utilization the term spatial thinking in various discipline, a variety of definition exist and substantial disagreement occurred about its factors, scale and connection with spatial abilities (see section 2) [23, 33, 50, 56, 63, 66, 68, 82].

With significant agreement exists around two foundational dimensions, spatial visualization (ability to mentally manipulate visual stimuli) and spatial orientation (ability to mentally align or rotate objects in space) and following similar to the previous work [28, 120], we break down the overall spatial thinking process into its contributing factors. Our goal was to define the theoretical framework we are going to guide the rest of the data analysis. Factor analysis has been a pivotal tool in identifying these dimensions, offering a systematic approach to grouping related test items and revealing distinct components of spatial ability. During the factor-analytical approach to refine and validate our understanding of spatial thinking, a numerous models emerged, ranging from two to thirteen factors. However, we focus on three key factors due to their centrality and alignment with our research objectives and with the consensus of spatial thinking expert. Table 2 shows the past work used in the *factor analysis* and which spatial thinking factors these past works majorly define. The identified factors are:

(1) **Spatial Visualization:** the cognitive ability to mentally manipulate and visualize movements, orientations, and spatial forms of objects.
(2) **Spatial Relations:** the cognitive ability to perceive and understand the spatial relationships and arrangements between objects or components within a scene.
(3) **Spatial Orientation:** the cognitive ability to mentally manipulate and perceive objects or stimuli from different perspectives or orientations, particularly from their point of view.

In this paper, we refer to these three factors as spatial thinking to avoid confusing definitions of spatial abilities that encompass more elements. For the full definition of these spatial factors used in the data analysis, see Supplementary Materials.

3.3.2 Structural and Protocol Analysis. Following an deductive methodology, we identify participants' hand actions, relate them to their thought processes, and divide them into the categories

Table 2: Meaning to the spatial ability factors extracted from the past literature

Paper Id	Spatial Relation	Spatial Orientation	Spatial Visualization	Visual Spatial Reasoning	Memory and Comprehension
[114]	reference frame of environment	mental-image manipulation	representing 3D shapes	externalizing, experimenting	recalling shapes
[30]	relationship among representation	imagine different orientations	Looking from different viewpoints	Communicate ideas	viewpoint-dependent features
[65]	intrinsic reference frames, egocentric reference frame	depend on egocentric	ability to imagine objects' movement	bresponse processes, time associated	latency identification
[33]	ability to find spatial relationships w.r.t subject	rotate 2D/3D figure spatiotemporal seq.	NA	generate solutions, evaluative aspect	match largest portion
[77]	NA	imagining movement relative to object-based frame	complex sequence imagination	apply logic, conventional rules	long term memory
[64]	picture plane rotation, mirror image reversal	direction of rotation, axis of rotation	folding and unfolding, complex movement parts	encoding, search, comparison, decision	NA
[139]	NA	egocentric spatial transformations	strategy differences in problem solution	viewpoint invariance	NA
[7]	speeded rotation, single-step mental rotation	imagine oneself reoriented within the array	NA	representing objects as contour	NA
[137]	reaction time, edges, region, color, texture	enhance the information recovery	solution strategies position estimation	not directly to spatial ability	NA
[63]	encoded w.r.t reference frames	imagine stimulus from another perspective	coordination, monitoring	NA	ability to recall the position of objects
[24]	angular discrepancy between two objects	views as dynamic information	NA	NA	NA
[59]	NA	enhance observer's ability, image features change	ability to recognize	NA	objects information
[20]	point of view, speedy	perspective to look 3D from vantage points	quantify the orientation changes	NA	spatial relations between objects
[40]	NA	engage in mental transformation	mental registration, understand orthogonal	NA	location, orientation of irregular shapes
[31]	longer duration	certain appearance moment in time	imagine final result, assembling parts	NA	perceptual Speed ability
[76]	natural-scrambled	estimate change in orientation	modifying projection visualization	visual memory, Perceptual Speed	ability to remember configuration
[42]	understand arrangements	mentally manipulate from ones prospective	NA	NA	NA
[153]	multi-step manipulation	egocentric spatial transformation	manipulate complex shapes	NA	NA
[87]	internal parts of complex shapes	view from multiple viewpoints	monitor change in orientation	NA	NA
[99]	spatial arrangements	help in match perspectives	estimate position without direct manipulate	related to visualization test	NA
[74]	coordination of multiple reference frames	2D/3D shapes	imagine final result of imagine part	NA	NA
[75]	object-centric volumetric primitives	dynamic information reduce difficulty	visualize movement	NA	NA

identified by the factor analysis. Based on previous work by Chandrasegaran et al. [31], this process was iterative. First, we used *structural analysis* [9, 64, 85, 87], which involved transcribing videos into written reports and segmenting them into two categories: participants' hand movements and design intent. Once we had these categories, we used *protocol analysis* [131] to understand their association. We focus on two key categories: participants' intentions, encompassing behavior or utterances reflecting specific design-oriented intents, and participants' hand actions, including nonverbal manual movements executed during the visual design task. Next, we will describe the four phases we followed:

(1) *Phase 1:* ***Defining hand actions categories via structural analysis***. Using *structural analysis*, coder A captured the hand actions and design behaviors participants did during the task. It's important to note that while all actions involve hand movements, not all hand movements constitute gestures. We distinguished actions from other hand movements based on their distinct stroke movement that conveys meaning, as defined by Vuletic et al. [145]. Following this description, coder A coded participants' hand actions based on their initiation and conclusion phases, excluding the intermediate rest or preparation positions.

(2) *Phase 2:* ***Establishing hand kinematics categories and linking them with the participants' design intent***. This phase uncovered hand movements associated with design intent during the participants' interactions with objects. Coder A and B used *protocol analysis* [131] to facilitate an in-depth exploration of participant's cognitive processes. They analyzed the sub-categories of hand actions, their design intents/behaviors, and their associations with spatial thinking

factors. Subsequently, they analyzed this data to uncover patterns of Spatial Hand Actions related to spatial thinking factors. This examination delved into how these hand actions influenced problem-solving approaches and the quality of 3D designs.

(3) *Phase 3: **Analysis of hand actions input and category refinement through discussion***. In this phase, coder A coded three participants using the taxonomy developed in *phase 2.* After this initial coding, coder B reviewed the codes. Any disagreements or inconsistencies were addressed and resolved through discussions and revisions of the coding guidelines within the coders, promoting the reliability of our coding system. This iteration was conducted multiple times to ensure that all categories were explored. Due to the intricate nature of hand action coding, our initial focus was on attaining semantic validity with the spatial thinking factor identified in the factor analysis step.

(4) *Phase 4: **Data Review and Reliability Assessment***. After formulating comprehensive descriptions for each subcategory, Coders A coded the remaining 15 participants, and coder C all 18 participants. The goal of this phase was to increase the validity of the results and the coding methodology, as it allowed both the coders to identify the existence of Spatial Hand Actions in other participants.

Finally, during each phase, coder B reviewed the coding process to detect and rectify any potential drift or bias that might arise over time. This proactive approach ensured that the coding scheme remained consistent and aligned with the research objectives throughout the analysis.

3.3.3 Thematic Analysis. The thematic analysis follows an inductive/deductive hybrid approach [123]. As a first step, inspired by Braun and Clarke [21, 22], we explore the collected data using an inductive approach to gain data-driven insights from analyzing audio transcripts. Then, utilizing spatial thinking theory as a framework, we did a deductive analysis of the insights to relate the identified themes to the spatial thinking factors identified in subsubsection 3.3.1.

Coders A and D coded the post-study interviews for the thematic analysis using a template with columns for transcript excerpts, selected phrases, codes and comments to code the data. Individually and inductively, they coded their transcripts to create a system to encode the data while keeping a list of this encoded data and their descriptions to track their process. Then, they shared the coded data and discussed the construction of themes. The themes were refined in conversations among coders A, D and B. Next, coders A and D returned to the definitions of spatial abilities from the factor analysis (subsubsection 3.3.1) to identify if there was a relationship between the identified themes and each factor following a deductive approach. This step aimed to determine if the identified themes aligned with the Spatial Themes discovered during the video coding analysis.

3.4 Structural and Protocol Analysis Results

After defining the hand actions categories, we observed 3141 actions that fit our definition. Figure 4 shows the resulting categories of each process step and how they relate to each other. Figure 5 shows the number and percentage distribution of various hand actions observed across all 18 participants.

3.4.1 Hand Actions and User Behaviours Categories. We identified the complete spectrum of hand actions participants employed while interacting with 3D assembling tasks. Our analysis shows that participants consistently engaged in standard hand movements like grabbing, rotating, shifting, dropping, aligning, inspecting, steadying, separating, and arranging objects. Additionally, we determined the frequencies at which participants employed these actions. See Figure 5. For example, the two most prevalent hand actions were grabbing 445 (14.16%) and rotation 504 (16.04%). The *Structural Analysis* yielded sub-categories of hand actions (such as grabbing, comparing, aligning, placing, moving, dropping, stability checking, and rotating) and participants' design intents/behaviors (such as building, relating, visualizing, forming imaginary elements, and 3D transformations). These sub-categories shed light on the connections between hand movements and design intent:

- **Grabbing**: Participants initiated this action by reaching for and grasping the target object. This action marks the onset of their task-related actions and is timed from the commencement of grabbing until the object is secured.
- **Comparison**: Participants often brought one or more objects closer to themselves. Although not always indicative of capturing an object, this action primed objects for comparison, aiding the overall task.
- **Aligning/Connecting**: Participants engaged in this action when aligning the edge of an object with a neighboring one. This action required the presence of at least one neighboring object and was executed accordingly.
- **Shifting**: This action denoted an intentional change in the strategy for manipulating an object. It could involve adjusting hand positions, reconsidering an object's location, or switching focus to a different object entirely.
- **Placing/Locating**: Participants performed this action when intentionally placing an object within or outside the designated working area. It reflected spatial reasoning for precise object placement within an estimated region of interest.
- **Moving**: This action entailed lifting and transporting an object, often to a different location rather than the region of interest for the task. The action ceased when the object was set down, marked as a distinct action.
- **Dropping**: Participants concluded various actions by releasing the object, denoted as the endpoint of the task-related action.
- **Stability Check**: This action was observed when participants used both hands to hold the fully assembled object, pausing briefly to ensure its stability.
- **Rotation**: An essential action in mental rotation, participants employed this to orient objects appropriately. It involved aligning the edges and corners of an object with other objects in the partially or secondary built structure.

3.4.2 Spatial Hand Actions. We used the hand kinematic categories, the users' verbalized design intent, and the spatial thinking factors defined in Section 3.3.1 to identify the actions associated with spatial thinking. We called these actions Spatial Hand Actions.

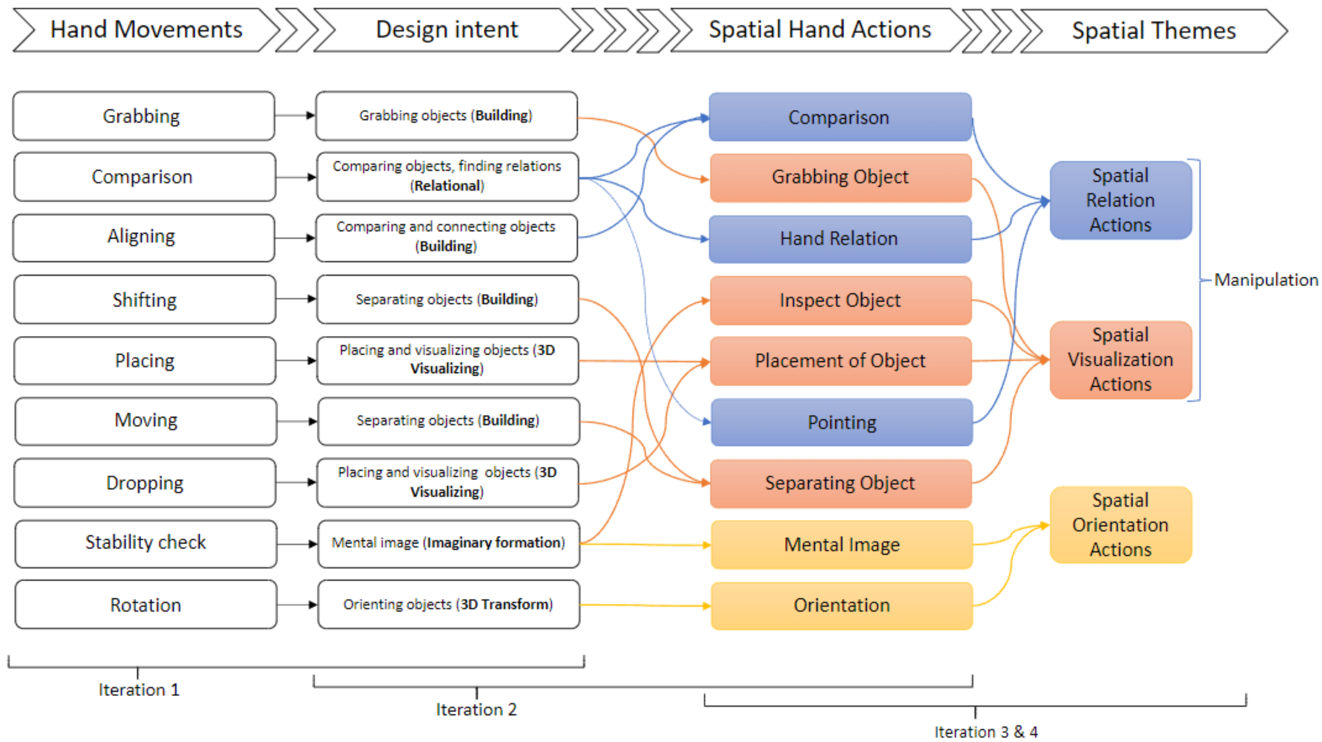


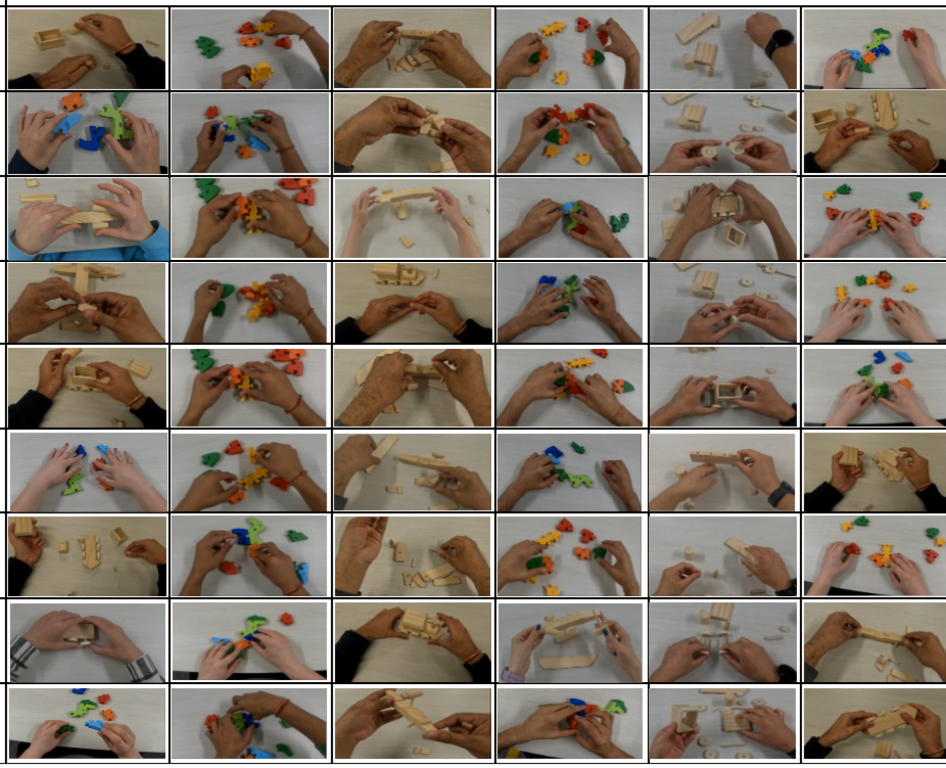
Figure 4: The results of each step of the Structural and Protocol Analysis. We identified three Spatial Themes encompassing nine Spatial Hand Actions, which result from the relationship between hand movements and design intent.

We identified that linking hand kinematics with design intent produces two distinct types of actions, each connected to a spatial thinking factor. The first category involved actions tied to object manipulation based on visual information, encompassing actions like grabbing, shifting, aligning, placing, and separating. The second category encompassed actions tied to object orientation and mental rotation based on dynamic information, such as rotating and mental imagination. These findings led us to identify nine Spatial Hand Actions associated with hand actions and design intent. The third column of Figure 4 depicts all Spatial Hand Actions and their relationship with the initially observed actions and user intent categories.

In this iteration, we had two different coders. We found a 20% difference in the number of codes identified by each coder. The two Spatial Hand Actions with a higher difference were Spatial Relation-Comparison with 39.85% and Spatial Visualization-Grabbing Object with 38.15%. The two Spatial Hand Actions with the lowest difference were Spatial Orientation-Mental Image with 15.73% and Spatial Relation-Placement Object with 16.24%. Additionally, Intra-class Correlation Coefficient (ICC) analysis was conducted to assess inter-coder reliability between Coder A, Coder B, and Coder C for the Structural and Protocol Analysis. The ICC coefficient was found to be 0.932 (95% CI [0.730, 0.984]), indicating excellent agreement between these coders in their ratings. These results demonstrate strong inter-coder reliability and support the validity of the coding process.

3.4.3 Spatial Themes. We categorize the Spatial Hand Actions into the three spatial thinking factors defined in subsubsection 3.3.1. These Spatial Themes are depicted in the fourth column of Figure 4. Next, we show the relationship between Spatial Hand Actions and Spatial Themes:

- **Spatial Relations Actions**: These actions were observed during 3D assembling tasks and encompass hand movements like comparison and aligning of hands. Users frequently attempted to bring both hands together, each holding different objects, and compared and identified the relationships between them. These actions are crucial in helping users envision relationships from multiple perspectives within an egocentric frame of reference.
- **Spatial Visualization Actions**: These actions were identified in conjunction with hand movements such as grabbing objects, stability checks, placing objects, shifting, and moving objects. As depicted in Figure 4, users employed these actions to manipulate 3D design components, aiding in their visualization and estimation of how these objects could be interconnected during the building process. These actions are predominantly guided by visual information within the 3D assembling task.
- **Spatial Orientation Actions**: These actions are primarily influenced by the need to orient objects in 3D space and mentally plan the subsequent steps required to complete the building task. Users often rotate components and estimate



Observed Actions	Examples of Hand Actions of 18 Participants in Iteration 1	Act. Count	Behavior found
Grabbing		445 (14.16%)	Grabbing objects (**Building**)
Comparison		470 (14.96%)	Comparing objects, finding relations (**Relational**)
Aligning		283 (9%)	Comparing and connecting objects (**Building**)
Shifting		248 (7.89%)	Separating objects (**Building**)
Placing		401 (12.76%)	Placing and visualizing objects (**3D Visualizing**)
Moving		235 (7.48%)	Separating objects (**Building**)
Dropping		314 (9.99%)	Placing and visualizing objects (**3D Visualizing**)
Stability check		241 (7.67%)	Mental image (**Imaginary formation**)
Rotation		504 (16.04%)	Orienting objects (**3D Transform**)

Figure 5: The result of the structural analysis [9], nine different hand actions with their corresponding user behaviors

their next course of action in the design task, emphasizing the importance of spatial orientation in their design process.

3.5 Thematic Analysis Results

The analysis of the themes found in the thematic analysis of the post-study interviews led us to better understand the relationship between Spatial Hand Actions and Spatial Themes. The findings also show the the relationship with the spatial ability factors identified in the factor analysis (subsubsection 3.3.3). The themes are as follows:

3.5.1 Theme 1: Spatial Design Synthesis. This theme concerns the participant's ability to identify the relationship between objects and visualize how these components fit together. The theme also encapsulates the idea of synthesizing visual information within a spatial context, where the participants analyze the components of a 3D design task, considering their spatial relationships, and generating models or solutions based on this synthesis of spatial elements. For instance, P2 said about their thought process during the task, *"First, I noticed the wheels because they were round and had holes. I separated them, and all the parts seemed very specific. I understood that they would come together to form a single object in the end, so I kept that in mind."* This statement underscores their adeptness at synthesizing visual cues, singling out visually distinctive components (e.g., wheels), and logically assembling them for a coherent final product. P15 mentioned, *"I was relating bigger pieces with being the main pieces and just thinking that the other ones should be connected or related some way."* This statement illustrates their engagement in visual interpretation by mentally correlating the large pieces as focal points and envisioning association with other elements. P6 also talked about visual synthesis, *"I spread the objects all over to gather ideas about what the object is and where it can fit. My thinking process aimed to find something that would attach most of the objects together."* Such a statement demonstrates their ability to thoroughly visualize the components and search for the next best-fit component to construct the design task.

3.5.2 Theme 2: Visual-Spatial Transformation and Reasoning. This theme concerns the participant's ability to imagine an element's appearance in another position, relate various interconnected 3D shapes, and comprehend the 3D spatial layout within a real environment. This theme help us understand whether participants mentioned analogical reasoning, pattern recognition, and the identification of irregularities while assembling the design components. This theme examines the participants' capacity to select secondary components that best complement partially constructed objects. Participants employed their hands and spatial thinking skills to assess object features, align components, and create coherent structures. This theme underscores the integration of visual and spatial cognition in the process of 3D assembling task. For instance, P2 explained, *"I noticed the wheels because they were round and had holes. I rotated the largest piece and discovered that there were holes*

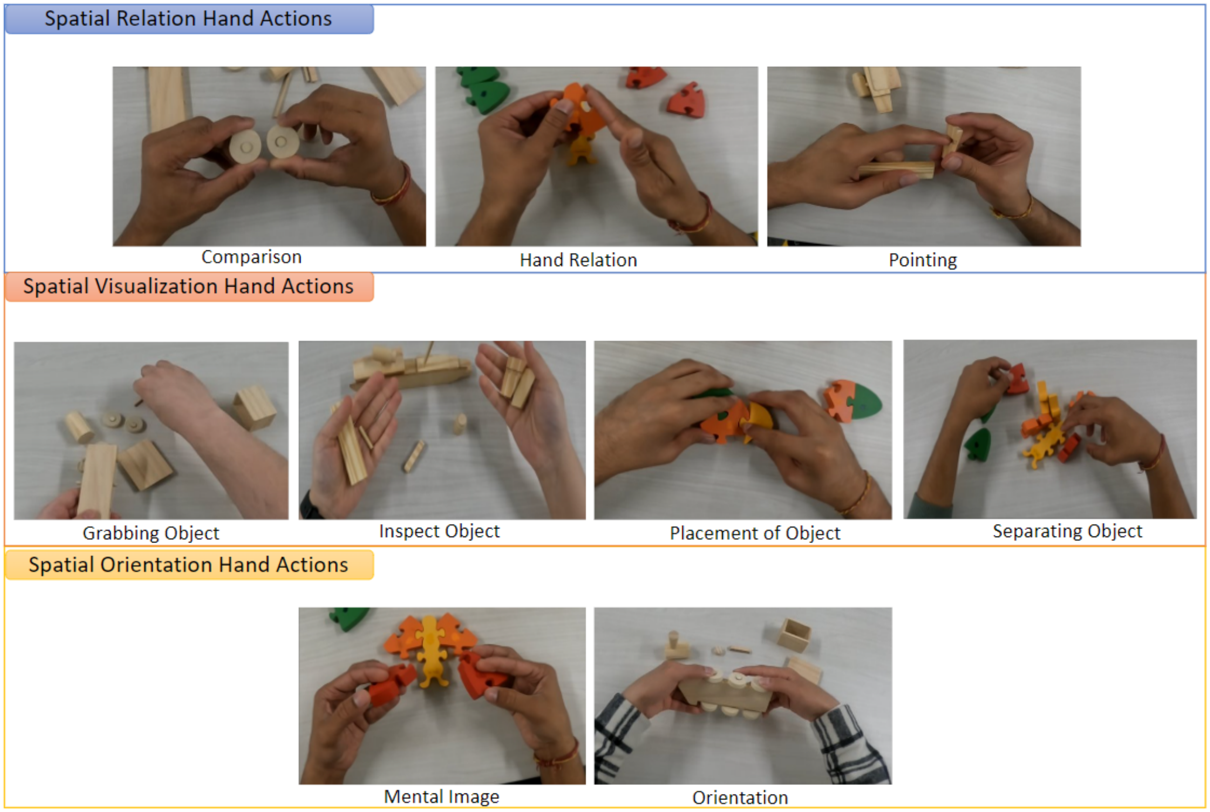


Figure 6: The nine Spatial Hand Actions identified in the user study, grouped by Spatial Themes. Refer to the video; the movement is impossible to capture in a static image.

all the way through, along with pegs that fit through easily, resulting in something resembling a vehicle chassis." This suggests that P2 carefully examined the components, selecting familiar and feature-rich ones, such as those with holes and indentations, and then rotated them along arbitrary axes, underscoring the rotation gesture's significance in the construction task [146]. P13 said about the orientation, *"I tried putting in every position, there was misalignment in the initial stage, but I had a feeling it should be similar so I just rotate it and saw if it's aligns with each other."* such statements exemplifies their spatial reasoning to rotate the pieces to find the correct alignment, demonstrating problem-solving approach. Some participants preferred to examine the components from different perspectives to explore their features for constructing the object. For example, P6 stated, *"It requires mental and physical effort; I had to rotate the objects and try to envision where they could fit."*

3.5.3 Theme 3: Spatial Perception and Mental Comprehension. This theme groups the participant's capacity to recall shapes from their experiences, understand the semantic relationships, and categorize various shapes within the context of the design task. This theme explores participants' ability to recall and mentally visualize an object while examining components from different viewpoints, particularly focusing on the most familiar shape within the design task. For instance, P2 stated, *"It can't be a car because a car doesn't really have six wheels. My guess is it's either like an 18-wheeler or a train."* This statement demonstrates the ability to memorize and connect components based on prior experiences and comprehend their relevance within the design task. Also, P4 mention mental imagery in her response, although it wasn't entirely accurate, as she herself stated, *"For the airplane wooden task, I assumed it was a boat while looking at the long piece of the object."* Similarly, P17 emphasized recalling objects based on specific shapes and colors of the components, stating, *"what I was trying to relate to previous things that I've seen in the past, like the shape look like a butterfly, or like the wheels. I knew that it was some kind of transportation."* However, despite forming mental images, he struggled to assemble the components for the design task. P3 mentioned, *"I can see a few meaningful patterns while connecting to similar objects that I know in daily life might help me figure out the other parts"*, but he couldn't establish a mental image from any partially built object. Additionally, P13 talked about the mental spatial perception, *"I looked for something that's similar, I saw round objects which I felt like wheels, so I thought that might be a moving object, the second was it had a shape, when I tried putting it together I felt it's a butterfly"* such statements shows their ability to identifies shapes to conceptualize objects with mental comprehension such as butterfly or wheels, during assembly.

3.5.4 Theme 4: Experimental-Based Learning. This theme concerns the participants with experimental-based strategies as placing or detaching secondary components form a primary or partially built object, engaging in trial-and-error, and facing challenges in creating a coherent mental image to estimate the final object. We found that participants are often applying experimental-based strategies as they are encountering difficulties in creating a mental picture of the final object. P18 mentioned, *"When I put two pieces together, they were fitting perfectly, but I had no idea what these shapes exactly is going to be. So I tried with other pieces."* We've discovered that participants who lack a natural inclination for visual synthesis, spatial transformation, and mental comprehension may find object manipulation challenging. Therefore, there's a requirement for spatial hand actions to aid individuals in understanding their cognitive processes while engaging in 3D assembly tasks.

4 Spatial Hand Actions Relationship to Spatial Abilities

Using a subset of participants from Study 1, we analyzed the relationship between Spatial Hand Actions and spatial abilities to gain a better understanding of how the individual's abilities of a person affect their use of Spatial Hand Actions. Study 1 primarily established the existence of Spatial Hand Actions as essential tools for spatial thinking during 3D assembling tasks. In this follow-up analysis, we aimed to investigate whether these actions were associated with an individual's varied spatial abilities, which are critical for them to perform complex design tasks. Based on past work [10, 11] and the relationship of the spatial thinking definitions used during the structural and protocol analysis (subsubsection 3.3.2), we expect that (High Spatial Ability) HSA participants were faster and used more Spatial Hand Actions than Low Spatial Ability (LSA) participants. Our approach was guided by Dolins et al. [45], highlighted using a subset of participants to explore nuanced spatial abilities while controlling variability. The authors emphasized carefully selecting participants to maintain balanced groups for comparative analysis. For this analysis, an additional coder (HCI researcher) and spatial thinking expert verified the association findings, enhancing the robustness of our results. Our findings allowed us to explore the variability in spatial abilities and their impact on the participants' frequency and type of Spatial Hand Actions.

4.1 Participants

This analysis was done with 12 participants from the 18 participants of the user study. To select this subset of participants, we first separated our participants into two groups: HSA and LSA. Similar to previous work [10, 11], HSA participants had both SOT test and VZ-2 test score higher than 15 out of 20. The rest of the participants were considered LSA. Then, we randomly selected six participants for each group to have a balanced population for our analysis.

4.2 Evaluation Metrics

We included the *task completion time* and *Spatial Hand Actions count* as metrics. For *task completion time*, we recorded the time participants took to complete each task, with a maximum allocated time of five minutes for each. For *Spatial Hand Actions count*, we used the number of Spatial Hand Actions each participant did.

4.3 Results

Data was pre-processed and plotted through JMP, and analyzed using an independent t-test in SPSS 28. We used Skewness (S) and Kurtosis (K) to analyze the normality of the data, i.e., when S and K values were within ±1 [62]. All data was normal except of hand relation, inspect object, placement of object, mental image and task completion time, thus we log-transformed the data. If data was still not normal, we utilized the Wilcoxon Signed-Rank Test to analyze it. Based on past work [10, 11] and the relationship of the spatial thinking definitions used during the protocol analysis, we were expecting that the HSA participants had a faster completion time and higher number of Spatial Hand Actions. Therefore, we used one-sided p value to test for significance. Statistical results are shown in Table 3.

4.3.1 Task Completion Time. There was a significant main effect on *Spatial Ability* groups (Figure 7Left). The Wilcoxon signed-rank test was conducted to compare task performance between the LSA and HSA groups. Our results found a significant difference in task performance between two groups ($Z = -3.05$, $p = 0.002$), with the LSA group exhibiting a lower median task completion percentage ($m = 22.50\%$) compared to the HSA group ($m = 82.50\%$). These findings suggest that individuals in the HSA group performed better and completed the task more efficiently compared to those in the LSA group.

4.3.2 Spatial Hand Actions Count. We found a significant main effect on *Spatial Ability* groups for four Spatial Hand Actions: 1) Comparison, Grabbing Object, Hand Relation and Mental Image. Specifically, participants in the HSA group did more of these actions than participants in the LSA group. The other Spatial Hand Actions did not show a significant difference between groups. See Figure Figure 7Right for the results.

5 Discussion

Through an in-depth multi-step analysis, we aimed to answer three research questions related to the existence of Spatial Hand Actions and their relationship with spatial thinking and people's spatial abilities in four different 3D assembling tasks such as wooden block (train, airplane) and Lego construction (butterfly, dinosaur) as depicted in Figure 2. Here, we discuss our results:

5.1 Spatial Hand Actions

This paper aimed to identify the existence of Spatial Hand Actions in physical environment, e.g., the hand actions people utilize to help their spatial thinking when doing 3D assembling tasks. Following a factor analysis from past studies (see subsubsection 3.3.1), we identified three spatial ability factors: spatial visualization [95], spatial orientation [65], and spatial relations [37, 86]. These factors are similar to Lohman et al. [96]'s model. While acknowledging the existence of other factors, Colom et al. [37] considered them less central to the core concept of *spatial abilities*.

Furthermore, we did a structural and protocol analysis of people solving four different 3D assembly tasks. Our results show the existence of nine Spatial Hand Actions (See Figure 6), which are naturalistic actions people make while trying to understand the nature of space, spatial information and relationships between the

Table 3: The Table Indicates the Descriptive Statistics and Results of the Inferential Test of the Participants' Spatial Abilities (Bolded Shows Significance)

S.No.	Parameters	HSA Mean (SD)	LSA Mean (SD)	Statistical Test	P Value	Effect Size
1	**Task Completion Time**	**278.04 (29.15)**	**294.53 (11.01)**	**Z=-3.06**	**p = 0.002**	**0.88**
2	**Comparison**	**13.17 (4.99)**	**6.17 (3.86)**	**t (10) = -2.71**	**p = <0.05**	**0.82**
3	Pointing	8.33 (1.75)	6.17 (4.40)	t (10) = -1.12	p = 0.14	0.04
4	**Grabbing Object**	**40.33 (5.28)**	**28.50 (5.16)**	**t (10) = -3.92**	**p = 0.001**	**0.63**
5	Separating Object	5.67 (1.96)	6.67 (3.07)	t (10) = 0.67	p = 0.26	0.07
6	Orientation	28.00 (15.19)	29.83 (20.06)	t (10) = 0.18	p = 0.43	0.82
7	**Hand Relation**	**6.67 (2.25)**	**1.83 (1.94)**	**t (10) = -3.9**	**p = 0.001**	**0.54**
8	Inspect Object	5.83 (3.25)	3.67 (3.33)	t (10) = -1.32	p = 0.11	0.02
9	Placement of Object	35.83 (17.29)	30.66 (9.73)	t (10) = -0.48	p = 0.33	0.48
10	**Mental Image**	**2.83 (1.60)**	**1.33 (1.05)**	**t (10) = -2.04**	**p = <0.05**	**0.52**

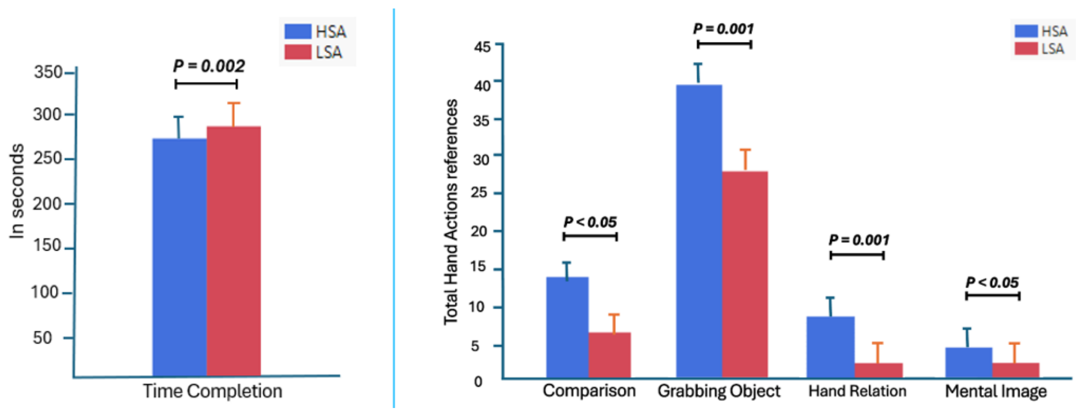


Figure 7: Left: Task Completion Time. Right: Significant Effect on Four Spatial Hand Actions by Spatial Ability Groups

different objects of varied volume that make a shape. Doing a more in-depth analysis, we found that although all these hand actions relate to spatial thinking, they have different uses. For example, the *Orientation Hand Actions* facilitates the synthesis of complex design elements by perceiving objects from various perspectives. Yet, the *Comparison Hand Actions* helps in relating 3D elements. Similarly, *Grabbing Hand Actions* helps in comprehending the spatial information of objects with respect to the nature of space. These results extend past work on the benefits of using hand actions when solving a spatial problem [34] by showing the existence of specific action related to different spatial thinking aspects. Ellis et al. [51] highlighted that spatial information about objects constrains thinking by shaping the relationship between visual perception and physical actions. Similarly, Gibbs et al. [57] found object perception is influenced by imagined physical manipulation, with participants interacting more quickly with objects suited to their natural grip patterns (e.g., a hammer) over precision grip (e.g., a screw). This supports the notion that *Grabbing Hand Actions* are inherently spatial, as they rely on visual perception to guide physical interaction, constraining thinking in a way that aids in understanding the spatial context of objects such as volume relative to the surrounding environment in 3D assembly tasks. Marsh et al. [97] also found a preference for two-handed gestures over single-handed ones when expressing 3D shapes and emphasized the effectiveness of user-defined gesture sets over pre-designed or arbitrary ones [111], further reinforcing the utility of naturalistic hand actions.

Our results also extend past work that studies naturalistic hand gestures used in 3D design tasks [7, 12, 31, 39, 59, 150, 152]. From this work, only Baykal [12] did a similar in-depth analysis of gestures that aid spatial thinking with tangram objects in VR. Yet, our work complements and extends their work by focusing on adult participants of different backgrounds and education levels and utilizing naturalistic hand actions in real 3D assembly tasks. Past work has found that spatial thinking and spatial abilities change as children grow into adults [15] and that a person's background affects their spatial abilities [73]. Thus, it is important to study the Spatial Hand Actions used by adults of different backgrounds in real 3D assembling task, mainly as this paper aims to utilize the insights of these actions for VR design applications, whose target audience is adult designers. They also utilize building blocks [12, 40] and Lego construction [92] as the design activity. Yet, we extend their work by focusing on design tasks that involve different spatial abilities besides mental rotation—see Figure 2 for the various spatial abilities our task requires.

In conclusion, by identifying nine unique Spatial Hand Actions, we answer **RQ1**.

5.2 Spatial Hand Actions Relationship with Spatial Thinking

The second goal of the paper is to identify the relationship between Spatial Hand Actions and spatial thinking factors. For this, we analyzed the participant's explanations of their hand movements during the design tasks using structural and protocol analysis. McNeill et al. [103]'s theory emphasized that gestures are inherently spatial and often produced during spatial tasks because they externalize the gesturer's cognitive process, reflecting how individuals mentally visualize and strategize around spatial problems [4]. Such existing evidence supports our findings, establishing a connection between the hand actions used in spatial tasks and spatial thinking, as they activate mental simulations of physical movements and spatial positions [58]. We also used thematic analysis on their post-study interviews to better understand their intentions towards each action with the consideration of spatial thinking factors extracted from factor analysis. Our results answer **RQ2** by grouping the nine Spatial Hand Actions into three Spatial Themes, directly mapped into the three primary spatial thinking factors. Here is a summary for each action type:

- **Spatial Relations Hand Actions:** allow people to create new reference frames by interacting with and repositioning elements in the scene. These Spatial Hand Actions help people understand the arrangements, positions, and connections among objects without changing their viewpoint of the 3D space. It is important to note that manipulation can happen as a change in position or a comparison of two elements or an element and the user's body. Chu et al. [34] also found such action is utilized to rapidly recognize the identity of a simple item form the same reference and others have utilized assembling actions such as fine-motion and screw to understand the spatial relation between the fingers of the hand and the objects [116]. This supports the existence of actions that helps in visualizing the relationships between the object and the life space.
- **Spatial Visualization Hand Actions:** aim to identify the relationship between objects, which our participants identified as an essential element of spatial visualization. Spatial Hand Actions in this category help people visualize the relationship between objects by manipulating them and to synthesize the relationship between these elements mentally. Such hand actions are the most supported in the literature, as Hegarty el at. [66] highlighted such actions as "the ability to mentally manipulate, rotate, twist, or invert objects without reference to one's self". These actions requires subjects to coordinate information during spatial transformation such as mentally folding a paper [88].
- **Spatial Orientation Hand Actions:** help people arrange objects in new ways to help improve their mental image of that object and facilitate the planning of subsequent steps. Our participants identified these Spatial Hand Actions necessary for spatial perception, transformation of the elements, and mental comprehension of the final shape. These actions also help people recall shapes from past experiences, understand semantic relationships, and categorize the shapes within the context of the design task. Existing evidence from Hegarty et al. [65] supports such actions as to make egocentric spatial transformations in which one's egocentric reference frame changes with respect to the environment, but the relation relation between the object and nature of space remains the same.

We extend previous work on the importance of spatial relations, spatial orientation, and spatial visualization on user performance during 3D assembling tasks [93] by showing that people use Spatial Hand Actions to describe spatial conditions and enhance their spatial thinking. In return, Spatial Hand Actions helps them engage more in advanced critical thinking [33] and displayed efficient mental manipulation of objects [65].

5.3 Difference in Spatial Hand Actions between HSA and LSA

RQ3 is about better understanding the relationship between Spatial Hand Actions and an individual's spatial abilities. Using a sub-set of 12 participants divided into HSA and LSA groups, we ran a statistical analysis on the number of hand actions and the task completion time to answer this research question. We found that HSA individuals did some of the Spatial Hand Actions more than LSA (Figure 7Right). For the *Spatial Relation Hand Actions*, we have Comparison and Hand Relation that help people compare objects and find their relations; for the *Spatial Visualization Hand Actions*, we have Grabbing that helps in visual perception to guide the physical interaction; and for the *Spatial Orientation Hand Actions*, we have Mental Image which helps them relate the objects to past experiences.

These hand actions are about manipulating the objects in 3D assembling tasks to make sense of them by comparing them to another object (comparison), body parts (hand relation), previous experiences (mental image) and guiding manipulation with visual perception (grabbing). Thus, our results show that HSA individuals better use the objects in the 3D assembling tasks to understand space and that their hand actions allow them to do this. It also shows that when the 3D elements are not enough to make sense of the objects, they use their past knowledge to identify their relationships. Moreover, despite recruiting two left-handed participants in the study, we did not observe any notable differences in the hand actions they performed or their spatial thinking processes compared to right-handed participants. Interestingly, other hand actions like pointing, inspecting, placing, separating, and orienting objects showed no difference between groups. We hypothesize that these hand actions are more related to building the shape than making sense of the objects. However, more research is needed to understand the differences between Spatial Hand Actions better.

Finally, our results also show that an individual spatial ability affects their task performance, as demonstrated by the HSA group having a lower task completion time than the LSA group. See Figure 7Left. These results verify previous work that found a strong relationship between task design, task performance and spatial abilities [10, 11, 33, 114]

6 Recommendations for Designers of 3DUI

Our results show the existence of Spatial Hand Actions that help people with their spatial thinking abilities during a 3D assembly task performed on a table while seated in a physical environment. We also show that an individual's spatial ability affects their hand actions. While these insights were derived from physical settings and have not yet been directly tested in immersive environments, they offer a foundation of exploring innovative hand-based naturalistic 3DUI. Here are four actionable recommendations for designers of naturalistic 3DUI for 3D design:

(1) **Consider using Spatial Hand Actions as interaction techniques**: Spatial Hand Actions are naturalistic and intuitive hand movements that people use to aid their spatial thinking when designing 3D objects in assembling task. These actions were identified in a user study with 18 participants of different backgrounds, education levels and spatial abilities, making them generalizable to the general population. Prior work has emphasized the cognitive role of hand actions in enhancing spatial reasoning [34, 58]. Designers should explore ways to map these actions to interaction techniques in 3DUI for enabling seamless object manipulation, spatial alignment and intuitive assembly workflows to aid in spatial thinking within 3D assembling task, as we recommend integrating Spatial Hand Actions in novel 3DUI for VR/AR.

(2) **Consider the role of Spatial Themes for 3D design tasks**: Our results identified a relationship between the Spatial Hand Actions and three specific spatial thinking factors called Spatial Themes (Spatial Visualization, Spatial Relations, and Spatial Orientation). In a post-study interview, our participants expressed the importance of these Spatial Themes in finishing the design task. Thus, we recommend implementing ways to help people rely less on their spatial thinking abilities, e.g., by creating ways to visualize the relationship between objects like visual guides. Assisting people to think spatially can enhance their comprehension and application of spatial concepts, offering a path to overcome spatial hurdles when working on 3D assembling tasks within immersive environments. These tools could help users externalize spatial concepts and navigates complex spatial relationships more effectively, as suggested by Goldin-Meadow et al. [14, 58]

(3) **Optimize the Scene for Enhanced Interaction**: Our analysis of the role of an individual's spatial ability shows that HSA participants utilize Spatial Hand Actions that allow them to use the objects, their position in the scene, and their hands to better understand the relationships between objects. In return, this increased their task performance. Based on these results, we recommend that future 3DUI for 3D assembly consider the elements shown in the scene and the user body and create ways for users to utilize them during their design tasks. This could include features such as augmented hand presence, scene-aware object interactions, and contextual prompts that guide users through 3D assembling tasks. Such scene-centric designs ensure that all elements are intuitively accessible, enhancing task performance.

(4) **Design Adaptable UI**: The next wave of 3D design applications for immersive technologies should consider the individual user's experiences and spatial abilities to tailor the interaction. For example, we found that individuals use their past experiences and own bodies to understand space better, the elements in it, and how they relate to each other. Therefore, users with LSA might benefit from additional scaffolding, such as step-by-step assembly guidance or enhanced haptic feedback. Drawing from elicitation studies that have guided user interface design (e.g., Rodriguez et al. [124]), our findings offer a road-map for an adaptable 3DUI that will lead to more efficient and user-friendly applications that dynamically respond to user actions. This can enhance their comprehension and application of spatial concepts, offering a path to overcome spatial hurdles when working on 3D assembling tasks within virtual reality.

In summary, this research bridges the gap by providing empirical evidence of the relationship between spatial abilities and hand action use, offering actionable insights for UI designers to enhance user engagement and spatial comprehension. Following similar work that uses elicitation studies to help design better user interfaces [124], our work can help designers better understand how specific hand actions have a cognitive meaning about space.

7 Limitations & Future Work

Through a user study with eighteen participants, this paper identifies the relationship between hand actions and spatial thinking. Although we employed an in-depth multi-step analysis, all our results are based on qualitative data analysis, which might have been affected by the individual's past experiences. Future work should verify our results using quantitative data like fMRI of the brain while solving 3D design tasks, similar to the work by Tung and Chang [133].

Another limitation of our study is that the participants solved real-life 3D puzzles, including Lego assembly and wooden blocks, which might affect the applicability of our results to immersive technologies. Yet, we based our tasks on previous work [31, 92] that already identified the compatibility between mediums and explored a broader range of 3D design activities, such as pattern blocks, puzzles, tangrams, and block design tasks [91, 108, 142]. We acknowledge that focusing solely on assembly tasks may have restricted the range of hand actions observed. Moreover, Baykal et al. [12] has suggested using real-world objects to avoid technical problems. In the future, we plan to repeat our experiment with a broader range of participants and in VR to verify our results.

8 Conclusion

In this paper, we identified the existence of Spatial Hand Actions and their relationship with spatial thinking and people's spatial abilities. Our findings offer insights into how people use their hands and elements in the scene to think spatially. Insights include the difficulty of visualizing mental images of the process and the importance of using Spatial Hand Actions for task success. These findings can inform the creation of a more user-friendly 3DUI for VR/AR interfaces tailored to specific individuals.

We aim to bridge a gap in developing VR/AR tools that utilize natural hand movements to interact with 3D objects, which is a key focus for technological advancement in this field. Additionally, our results benefit education related to spatial abilities and tangible user interfaces, as they can utilize these actions to enhance people's understanding of space.

References

[1] James Ainooson and Maithilee Kunda. 2017. A Computational Model for Reasoning About the Paper Folding Task Using Visual Mental Images. *Cognitive Science* (2017). https://api.semanticscholar.org/CorpusID:37541759

[2] Leila Alem and Jane Li. 2011. A study of gestures in a video-mediated collaborative assembly task. *Advances in Human-Computer Interaction* 2011, 1 (2011), 987830.

[3] Martha W Alibali. 2005. Gesture in spatial cognition: Expressing, communicating, and thinking about spatial information. *Spatial cognition and computation* 5, 4 (2005), 307–331.

[4] Martha W Alibali, Miriam Bassok, Karen Olseth Solomon, Sharon E Syc, and Susan Goldin-Meadow. 1999. Illuminating mental representations through speech and gesture. *Psychological Science* 10, 4 (1999), 327–333.

[5] Martha W Alibali, Robert C Spencer, Lucy Knox, and Sotaro Kita. 2011. Spontaneous gestures influence strategy choices in problem solving. *Psychological science* 22, 9 (2011), 1138–1144.

[6] John Aliprantis, Markos Konstantakis, Rozalia Nikopoulou, Phivos Mylonas, and George Caridakis. 2019. Natural Interaction in Augmented Reality Context.. In *VIPERC@ IRCDL*. 50–61.

[7] Remi Alkemade, Fons J Verbeek, and Stephan G Lukosch. 2017. On the efficiency of a VR hand gesture-based interface for 3D object manipulations in conceptual design. *International Journal of Human–Computer Interaction* 33, 11 (2017), 882–901.

[8] Sang-Gyun An, Yongkwan Kim, Joon Hyub Lee, and Seok-Hyung Bae. 2017. Collaborative Experience Prototyping of Automotive Interior in VR with 3D Sketching and Haptic Helpers. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (Oldenburg, Germany) *(AutomotiveUI '17)*. Association for Computing Machinery, New York, NY, USA, 183–192. doi:10.1145/3122986.3123002

[9] Wayne A Bailey and Edwin J Kay. 1986. Structural analysis of verbal data. *ACM SIGCHI Bulletin* 17, SI (1986), 297–301.

[10] Mayra Donaji Barrera Machuca, Wolfgang Stuerzlinger, and Paul Asente. 2019. The Effect of Spatial Ability on Immersive 3D Drawing. In *Proceedings of the 2019 Conference on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19)*. Association for Computing Machinery, New York, NY, USA, 173–186. doi:10.1145/3325480.3325489

[11] Mayra D. Barrera Machuca, Wolfgang Stuerzlinger, and Paul Asente. 2019. Smart3DGuides: Making Unconstrained Immersive 3D Drawing More Accurate. In *25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 37, 13 pages. doi:10.1145/3359996.3364254

[12] Gökçe Elif Baykal, Ali Leylekoğlu, Sude Arslan, and Demet Ozer. 2023. Studying Children's Object Interaction in Virtual Reality: A Manipulative Gesture Taxonomy for VR Hand Tracking. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 264, 7 pages. doi:10.1145/3544549.3585865

[13] Robert S Bednarz and Jongwon Lee. 2011. The components of spatial thinking: empirical evidence. *Procedia-Social and Behavioral Sciences* 21 (2011), 103–107.

[14] Sian L Beilock and Susan Goldin-Meadow. 2010. Gesture changes thought by grounding it in action. *Psychological science* 21, 11 (2010), 1605–1610.

[15] Vittorio Belmonti, Giovanni Cioni, and Alain Berthoz. 2015. Switching from reaching to navigation: differential cognitive strategies for spatial memory in children and adults. *Developmental Science* 18, 4 (2015), 569–586. doi:10.1111/desc.12240 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/desc.12240

[16] Zafer Bilda and Halime Demirkan. 2003. An insight on designers' sketching activities in traditional versus digital media. *Design studies* 24, 1 (2003), 27–50.

[17] Cindy M Bird. 2005. How I stopped dreading and learned to love transcription. *Qualitative inquiry* 11, 2 (2005), 226–248.

[18] Doug A Bowman, Ryan P McMahan, and Eric D Ragan. 2012. Questioning naturalism in 3D user interfaces. *Commun. ACM* 55, 9 (2012), 78–88.

[19] Doug A. Bowman, Ryan P. McMahan, and Eric D. Ragan. 2012. Questioning naturalism in 3D user interfaces. *Commun. ACM* 55, 9 (sep 2012), 78–88. doi:10.1145/2330667.2330687

[20] TJ Branoff and M Dobelis. 2012. The relationship between spatial visualization ability and students' ability to model 3D objects from engineering assembly drawings. *The Engineering Design Graphics Journal* 76, 3 (2012), 37–43.

[21] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[22] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis*. American Psychological Association, Washington DC, USA.

[23] Lorelle J Burton and Gerard J Fogarty. 2003. The factor structure of visual imagery and spatial abilities. *Intelligence* 31, 3 (2003), 289–318.

[24] Lucia Calabrese and Francesco S Marucci. 2006. The influence of expertise level on the visuo-spatial ability: differences between experts and novices in imagery and drawing abilities. *Cognitive Processing* 7 (2006), 118–120.

[25] Yvonne M Caldera, AnneDonald Mc Culp, Marion O'Brien, Rosemarie T Truglio, Mildred Alvarez, and Aletha C Huston. 1999. Children's play preferences, construction play with blocks, and visual-spatial skills: Are they related? *International Journal of Behavioral Development* 23, 4 (1999), 855–872.

[26] Carlos Carbonell-Carrera, Peri Gunalp, Jose Luis Saorin, and Stephany Hess-Medler. 2020. Think spatially with game engine. *ISPRS International Journal of Geo-Information* 9, 03 (2020), 159.

[27] Patrick Carlson, Anicia Peters, Stephen B Gilbert, Judy M Vance, and Andy Luse. 2015. Virtual training: Learning transfer of assembly tasks. *IEEE transactions on visualization and computer graphics* 21, 6 (2015), 770–782.

[28] John B Carroll. 1993. *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press, Cambridge, UK.

[29] Erica A Cartmill, Sian Beilock, and Susan Goldin-Meadow. 2012. A word in the hand: action, gesture and mental representation in humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, 1585 (2012), 129–143.

[30] Carolyn Backer Cave and Stephen M Kosslyn. 1993. The role of parts and spatial relations in object identification. *Perception* 22, 2 (1993), 229–248.

[31] Senthil Chandrasegaran, Devarajan Ramanujan, and Niklas Elmqvist. 2018. How Do Sketching and Non-Sketching Actions Convey Design Intent?. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 373–385. doi:10.1145/3196709.3196723

[32] Elizabeth Charters. 2003. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal* 12, 2 (2003).

[33] Ji Young Cho. 2017. An investigation of design studio performance in relation to creativity, spatial ability, and visual cognitive style. *Thinking Skills and Creativity* 23 (2017), 67–78.

[34] Mingyuan Chu and Sotaro Kita. 2011. The nature of gestures' beneficial role in spatial problem solving. *Journal of experimental psychology: General* 140, 1 (2011), 102.

[35] Cheryl A Cohen and Mary Hegarty. 2012. Inferring cross sections of 3D objects: A new spatial thinking test. *Learning and Individual Differences* 22, 6 (2012), 868–874.

[36] Fiona Coley, Oliver Houseman, and Rajkumar Roy. 2007. An introduction to capturing and understanding the cognitive behaviour of design engineers. *Journal of Engineering Design* 18, 4 (2007), 311–325.

[37] Roberto Colom, Mª José Contreras, Juan Botella, and José Santacreu. 2002. Vehicles of spatial ability. *Personality and Individual Differences* 32, 5 (2002), 903–912.

[38] Julián Conesa, Francisco José Mula, Kristin A Bartlett, Ferran Naya, and Manuel Contero. 2023. The influence of immersive and collaborative virtual environments in improving spatial skills. *Applied Sciences* 13, 14 (2023), 8426.

[39] Douglas Cooper. 2018. Imagination's hand: The role of gesture in design drawing. *Design Studies* 54 (2018), 120–139. doi:10.1016/j.destud.2017.11.001

[40] Cathryn S Cortesa, Jonathan D Jones, Gregory D Hager, Sanjeev Khudanpur, Amy Lynne Shelton, and Barbara Landau. 2017. Characterizing spatial construction processes: Toward computational tools to understand cognition. In *CogSci*. The Cognitive Science Society, London, United Kingdom, 246–251.

[41] National Research Council, Division on Earth, Life Studies, Board on Earth Sciences, Geographical Sciences Committee, Committee on Support for Thinking Spatially, and The Incorporation of Geographic Information Science Across the K-12 Curriculum. 2005. *Learning to think spatially*. National Academies Press.

[42] Nigel Cross. 2023. *Design thinking: Understanding how designers think and work*. Bloomsbury Publishing, NY, USA.

[43] Adriano da Silva Marinho, Uwe Terton, and Christian M Jones. 2022. Cybersickness and postural stability of first time VR users playing VR videogames. *Applied Ergonomics* 101 (2022), 103698.

[44] Eder de Oliveira, Esteban Walter Gonzalez, Daniela Gorski Trevisan, and Luciana Cardoso de Castro Salgado. 2020. Investigating users' natural engagement with a 3d design approach in an egocentric vision scenario. In *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*. IEEE, 74–82.

[45] Francine L Dolins and Robert W Mitchell. 2010. *Spatial cognition, spatial perception: mapping the self and space*. Cambridge University Press.

[46] Tobias Drey, Jan Gugenheimer, Julian Karlbauer, Maximilian Milo, and Enrico Rukzio. 2020. VRSketchIn: Exploring the Design Space of Pen and Tablet Interaction for 3D Sketching in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3313831.3376628

[47] Tobias Drey, Michael Montag, Andrea Vogt, Nico Rixen, Tina Seufert, Steffi Zander, Michael Rietzler, and Enrico Rukzio. 2023. Investigating the Effects of Individual Spatial Abilities on Virtual Reality Object Manipulation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*

(Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 398, 24 pages. doi:10.1145/3544548.3581004
[48] Andrew Edwards-Jones. 2014. Qualitative data analysis with NVIVO. *Journal of Education for Teaching* 40, 2 (2014), 193–195. doi:10.1080/02607476.2013.866724
[49] Ruth B Ekstrom and Harry Horace Harman. 1976. *Manual for kit of factor-referenced cognitive tests: 1976.* Educational testing service, NJ, USA.
[50] John Eliot and Mark Y Czarnolewski. 2007. Development of an everyday spatial behavioral questionnaire. *The Journal of general psychology* 134, 3 (2007), 361–381.
[51] Rob Ellis and Mike Tucker. 2000. Micro-affordance: The potentiation of components of action by seen objects. *British journal of psychology* 91, 4 (2000), 451–471.
[52] Alinda Friedman, Bernd Kohler, Peri Gunalp, Alexander P Boone, and Mary Hegarty. 2020. A computerized spatial orientation test. *Behavior research methods* 52 (2020), 799–812.
[53] Wen Gao and Shuanglin Wang. 1995. Independent hand gesture recognition in HandTalker. In *Image Analysis Applications and Computer Graphics*, Roland T. Chin, Horace H. S. Ip, Avi C. Naiman, and Ting-Chuen Pong (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 443–450.
[54] Nirit Gavish, Teresa Gutiérrez, Sabine Webel, Jorge Rodríguez, Matteo Peveri, Uli Bockholt, and Franco Tecchia. 2015. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. *Interactive Learning Environments* 23, 6 (2015), 778–798.
[55] Phil J Gersmehl and Carol A Gersmehl. 2006. Wanted: A concise list of neurologically defensible and assessable spatial-thinking skills. (2006).
[56] Philip J Gersmehl and Carol A Gersmehl. 2007. Spatial thinking by young children: Neurologic evidence for early development and "educability". *Journal of Geography* 106, 5 (2007), 181–191.
[57] Raymond W Gibbs Jr. 2005. *Embodiment and cognitive science.* Cambridge University Press.
[58] Susan Goldin-Meadow and Sian L Beilock. 2010. Action's influence on thought: The case of gesture. *Perspectives on psychological science* 5, 6 (2010), 664–674.
[59] Weilun Gong, Stephanie Santosa, Tovi Grossman, Michael Glueck, Daniel Clarke, and Frances Lai. 2023. Affordance-Based and User-Defined Gestures for Spatial Tangible Interaction. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1500–1514. doi:10.1145/3563657.3596032
[60] GoPro, Inc. . 2021. HERO10 Black camera. https://gopro.com/en/ca/
[61] Shouchao Guo, Xiao Wang, Wenbo Deng, Jialing Hong, Jiawen Wang, and Yonghe Wu. 2022. Whose spatial ability benefits from learning with 3D design? From the perspective of learning analysis. *Educational Technology & Society* 25, 1 (2022), 179–192.
[62] Joseph F. Hair Jr, William C. Black, Barry J. Babin, and Rolph E. Anderson. 2014. Multivariate Data Analysis.
[63] Devamardeep Hayatpur, Seongkook Heo, Haijun Xia, Wolfgang Stuerzlinger, and Daniel Wigdor. 2019. Plane, Ray, and Point: Enabling Precise Spatial Manipulations with Shape Constraints. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1185–1195. doi:10.1145/3332165.3347916
[64] Mary Hegarty and Maria Kozhevnikov. 1999. Types of visual–spatial representations and mathematical problem solving. *Journal of educational psychology* 91, 4 (1999), 684.
[65] Mary Hegarty and David Waller. 2004. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence* 32, 2 (2004), 175–191.
[66] Mary Hegarty and David A. Waller. 2005. *Individual Differences in Spatial Abilities.* Cambridge University Press, Cambridge, UK, 121–169.
[67] Ken Hinckley, Randy Pausch, John C Goble, and Neal F Kassell. 1994. A survey of design issues in spatial input. In *Proceedings of the 7th annual ACM symposium on User interface software and technology.* 213–222.
[68] Chun-Heng Ho, Charles Eastman, and Richard Catrambone. 2006. An investigation of 2D and 3D spatial and mathematical abilities. *Design Studies* 27, 4 (2006), 505–524.
[69] Ke Huo, Vinayak, and Karthik Ramani. 2017. Window-shaping: 3d design ideation by creating on, borrowing from, and looking at the physical world. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction.* 37–45.
[70] Hessam Jahani and Manolya Kavakli. 2018. Exploring a user-defined gesture vocabulary for descriptive mid-air interactions. *Cognition, Technology & Work* 20 (2018), 11–22.
[71] Hannu Järvinen, Ulysses Bernardet, and Paul F.M.J. Verschure. 2010. Interaction mapping affects spatial memory and the sense of presence when navigating in a virtual environment. In *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction* (Funchal, Portugal) *(TEI '11)*. Association for Computing Machinery, New York, NY, USA, 321–324. doi:10.1145/1935701.1935776
[72] Hans-Christian Jetter, Roman Rädle, Tiare Feuchtner, Christoph Anthes, Judith Friedl, and Clemens Nylandsted Klokmose. 2020. "In VR, everything is possible!": Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3313831.3376652
[73] Jamie J. Jirout and Nora S. Newcombe. 2015. Building Blocks for Developing Spatial Skills: Evidence From a Large, Representative U.S. Sample. *Psychological Science* 26, 3 (2015), 302–310. doi:10.1177/0956797614563338 PMID: 25626442.
[74] Keith Jones. 1998. Theoretical frameworks for the learning of geometrical reasoning. *Proceedings of the british society for research into learning mathematics* 18, 1-2 (1998), 29–34.
[75] Jinsheng Kang, Kang Zhong, Shengfeng Qin, Hongan Wang, and David Wright. 2013. Instant 3D design concept generation and visualization by real-time hand gesture recognition. *Computers in Industry* 64, 7 (2013), 785–797.
[76] Aisha Urooj Khan and Ali Borji. 2018. Analysis of Hand Segmentation in the Wild. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, UT, USA, 4710–4719. doi:10.1109/CVPR.2018.00495
[77] Sumbul Khan and Bige Tunçer. 2019. Gesture and speech elicitation for 3D CAD modeling in conceptual design. *Automation in Construction* 106 (2019), 102847.
[78] Hind Kharoub, Mohammed Lataifeh, and Naveed Ahmed. 2019. 3D user interface design and usability for immersive VR. *Applied sciences* 9, 22 (2019), 4861.
[79] Jong-Oh Kim, Mihye Kim, and Kwan-Hee Yoo. 2013. Real-time hand gesture-based interaction with objects in 3D virtual environments. *International Journal of Multimedia and Ubiquitous Engineering* 8, 6 (2013), 339–348.
[80] Mi Jeong Kim and Mary Lou Maher. 2008. The impact of tangible user interfaces on designers' spatial cognition. *Human–Computer Interaction* 23, 2 (2008), 101–137.
[81] Yongkwan Kim and Seok-Hyung Bae. 2016. SketchingWithHands: 3D Sketching Handheld Products with First-Person Hand Posture. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 797–808. doi:10.1145/2984511.2984567
[82] Doreen Kimura. 2000. *Sex and cognition.* MIT press, USA.
[83] David Kirk, Tom Rodden, and Danaë Stanton Fraser. 2007. Turn it this way: grounding collaborative action with remote gestures. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 1039–1048.
[84] Sotaro Kita, Martha W Alibali, and Mingyuan Chu. 2017. How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological review* 124, 3 (2017), 245.
[85] Maria Kozhevnikov, Mary Hegarty, and Richard E Mayer. 2002. Revising the visualizer-verbalizer dimension: Evidence for two types of visualizers. *Cognition and instruction* 20, 1 (2002), 47–77.
[86] Maria Kozhevnikov, Michael Kozhevnikov, Chen Jiao Yu, and Olesya Blazhenkova. 2013. Creativity, visualization abilities, and visual cognitive style. *British journal of educational psychology* 83, 2 (2013), 196–209.
[87] Maria Kozhevnikov, Michael A Motes, and Mary Hegarty. 2007. Spatial visualization in physics problem solving. *Cognitive science* 31, 4 (2007), 549–579.
[88] Maria Kozhevnikov and Ronald Thornton. 2006. Real-time data display, spatial visualization ability, and learning force and motion concepts. *Journal of Science Education and Technology* 15, 1 (2006), 111–132.
[89] Jongwon Lee and Robert Bednarz. 2012. Components of spatial thinking: Evidence from a spatial thinking ability test. *Journal of geography* 111, 1 (2012), 15–26.
[90] Sang-Su Lee, Jeonghun Chae, Hyunjeong Kim, Youn-kyung Lim, and Kun-pyo Lee. 2013. Towards more natural digital content manipulation via user freehand gestural interaction in a living room. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. 617–626.
[91] TaeYang Lee, Woori Noh, Yeji Kim, and Jun Dong Cho. 2016. Hide and Seek: A LEGO Village Game for Developing Spatial Ability of Children. In *Proceedings of the The 15th International Conference on Interaction Design and Children* (Manchester, United Kingdom) *(IDC '16)*. Association for Computing Machinery, New York, NY, USA, 408–413. doi:10.1145/2930674.2930715
[92] Annemarie Lesage, Hubert-David Au-Yeung, Simon Bourdeau, Béatrice C. Caron, and Pierre-Majorique Léger. 2019. Sketch or Play? LEGO® Stimulates Divergent Thinking for Non-sketchers in HCI Conceptual Ideation. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3290607.3313023
[93] Kun-Hsi Liao. 2017. The abilities of understanding spatial relations, spatial orientation, and spatial visualization affect 3D product design performance: using carton box design as an example. *International Journal of Technology and Design Education* 27 (2017), 131–147.
[94] Wanhong Lin, Lear Du, Carisa Harris-Adamson, Alan Barr, and David Rempel. 2017. Design of hand gestures for manipulating objects in virtual reality. In *Human-Computer Interaction. User Interface Design, Development and Multimodality: 19th International Conference, HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings, Part I 19.* Springer, 584–592.
[95] Marcia C Linn and Anne C Petersen. 1985. Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child development* 56, 6

(1985), 1479–1498.
[96] David F Lohman. 2014. Spatial abilities as traits, processes, and knowledge. In *Advances in the psychology of human intelligence*. Psychology Press, USA, 181–248.
[97] Timothy Marsh and Alan Watt. 1998. Shape your imagination: iconic gestural-based interaction. In *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No. 98CB36180)*. IEEE, 122–125.
[98] Gordana Marunic and Vladimir Glazar. 2013. Spatial ability through engineering graphics education. *International Journal of Technology and Design Education* 23 (2013), 703–715.
[99] Alexander Masurovsky, Paul Chojecki, Detlef Runde, Mustafa Lafci, David Przewozny, and Michael Gaebler. 2020. Controller-free hand tracking for grab-and-place tasks in immersive virtual reality: Design elements and their empirical study. *Multimodal Technologies and Interaction* 4, 4 (2020), 91.
[100] Erin McAweeney, Haihua Zhang, and Michael Nebeling. 2018. User-driven design principles for gesture representations. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–13.
[101] Karla K McGregor, Katharina J Rohlfing, Allison Bean, and Ellen Marschner. 2009. Gesture as a support for word learning: The case of under. *Journal of child language* 36, 4 (2009), 807–828.
[102] Robert H McKim. 1972. *Experiences in visual thinking*. Brooks/Cole Publishing Company, UK.
[103] David McNeill. 1992. Hand and mind1. *Advances in Visual Semiotics* 351 (1992).
[104] Daniel Mendes, Fabio Marco Caputo, Andrea Giachetti, Alfredo Ferreira, and Joaquim Jorge. 2019. A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In *Computer graphics forum*, Vol. 38. Wiley Online Library, 21–45.
[105] Microsoft. 2023. Transcribe Your Recordings: OneNote for Microsoft 365, Word Web App. https://support.microsoft.com/en-us/office/transcribe-your-recordings-7fc2efec-245e-45f0-b053-2a97531ecf57. https://support.microsoft.com/en-us/office/transcribe-your-recordings-7fc2efec-245e-45f0-b053-2a97531ecf57
[106] Microsoft Corporation. 2018. Microsoft Excel. https://office.microsoft.com/excel
[107] George B. Mo, John J Dudley, and Per Ola Kristensson. 2021. Gesture Knitter: A Hand Gesture Design Tool for Head-Mounted Mixed Reality Applications. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 291, 13 pages. doi:10.1145/3411764.3445766
[108] Anders I Mørch and Ingvill Thomassen. 2016. From Wooden Blocks and Lego to Minecraft: Designing and Playing with Blocks to Learn in a 3D Virtual World. In *CoPDA@ NordiCHI*. 61–67.
[109] Sundar Murugappan, HaiRong Liu, Karthik Ramani, et al. 2013. Shape-It-Up: Hand gesture based creative expression of 3D shapes using intelligent generalized cylinders. *Computer-Aided Design* 45, 2 (2013), 277–287.
[110] Miguel Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of pre-designed and user-defined gesture sets. In *Proceedings of the SIGCHI Conference on Human factors in computing systems*. ACM, 1099–1108.
[111] Miguel A Nacenta, Yemliha Kamber, Yizhou Qiang, and Per Ola Kristensson. 2013. Memorability of pre-designed and user-defined gesture sets. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1099–1108.
[112] Nora S Newcombe and Thomas F Shipley. 2014. Thinking about spatial thinking: New typology, new assessments. In *Studying visual and spatial reasoning for design creativity*. Springer, 179–192.
[113] Gary M Olson, Susan A Duffy, and Robert L Mack. 2018. Thinking-out-loud as a method for studying real-time comprehension processes. In *New methods in reading comprehension research*. Routledge, UK, 253–286.
[114] Alfred Oti and Nathan Crilly. 2021. Immersive 3D sketching tools: Implications for visual thinking and communication. *Computers & Graphics* 94 (2021), 111–123.
[115] Rivka Oxman. 2002. The thinking eye: visual re-cognition in design emergence. *Design studies* 23, 2 (2002), 135–164.
[116] George V Paul, Yunde Jiar, Mark D Wheeler, and K Ikeuchik. 1996. Modelling human assembly actions from observation. In *1996 IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems (Cat. No. 96TH8242)*. IEEE, 357–364.
[117] Siyou Pei, Alexander Chen, Jaewook Lee, and Yang Zhang. 2022. Hand interfaces: Using hands to imitate objects in ar/vr for expressive interactions. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–16.
[118] James W Pellegrino, David L Alderton, and Valerie J Shute. 1984. Understanding spatial ability. *Educational psychologist* 19, 4 (1984), 239–253.
[119] John W Philbeck and Jesse Sargent. 2013. Perception of spatial relations during self-motion. (2013).
[120] Marios Pittalis and Constantinos Christou. 2010. Types of reasoning in 3D geometry thinking and their relation with spatial ability. *Educational Studies in mathematics* 75 (2010), 191–212.
[121] Thammathip Piumsomboon, David Altimira, Hyungon Kim, Adrian Clark, Gun Lee, and Mark Billinghurst. 2014. Grasp-Shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 73–82. doi:10.1109/ISMAR.2014.6948411
[122] Charles Potter and Errol van der Merwe. 2001. Spatial ability, visual imagery and academic performance in engineering graphics. In *International Conference on Engineering Education, Oslo, Norway*. USA, IEEE, 239–269.
[123] Kevin Proudfoot. 2023. Inductive/Deductive Hybrid Thematic Analysis in Mixed Methods Research. *Journal of Mixed Methods Research* 17, 3 (2023), 308–326. doi:10.1177/15586898221126816
[124] Richard Rodriguez, Brian T Sullivan, Mayra Donaji Barrera Machuca, Anil Ufuk Batmaz, Cyane Tornatzky, and Francisco R Ortega. 2024. An Artists' Perspectives on Natural Interactions for Virtual Reality 3D Sketching. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–20.
[125] Jose Luis Rubio-Tamayo, Manuel Gertrudix Barrio, and Francisco García García. 2017. Immersive environments and virtual reality: Systematic review and advances in communication, interaction and simulation. *Multimodal technologies and interaction* 1, 4 (2017), 21.
[126] J de Ruiter. 1998. *Gesture and speech production*. [Sl: sn].
[127] Hannah Sampson, Dan Kelly, Burkhard C Wünsche, and Robert Amor. 2018. A hand gesture set for navigating and interacting with 3d virtual environments. In *2018 International Conference on Image and Vision Computing New Zealand (IVCNZ)*. IEEE, 1–6.
[128] Marc A Schnabel and Thomas Kvan. 2003. Spatial understanding in immersive virtual environments. *International Journal of Architectural Computing* 1, 4 (2003), 435–448.
[129] Stephanie Schwarz, Georg Regal, Marina Kempf, and Raimund Schatz. 2020. Learning success in immersive virtual reality training environments: practical evidence from automotive assembly. In *Proceedings of the 11th nordic conference on human-computer interaction: shaping experiences, shaping society*. 1–11.
[130] Jami J Shah, Jay Woodward, and Steven M Smith. 2013. Applied tests of design skills—part II: visual thinking. *Journal of Mechanical Design* 135, 7 (2013), 071004.
[131] Masaki Suwa, Terry Purcell, and John Gero. 1998. Macroscopic analysis of design processes based on a scheme for coding designers' cognitive actions. *Design studies* 19, 4 (1998), 455–483.
[132] Shun'ichi Tano, Shinya Yamamoto, Muhd Dzulkhiflee, Junko Ichino, Tomonori Hashiyama, and Mitsuru Iwata. 2012. Three design principles learned through developing a series of 3D sketch systems: "Memory Capacity", "Cognitive Mode", and "Life-size and Operability". In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, Seoul, South Korea, 880–887. doi:10.1109/ICSMC.2012.6377839
[133] Yu-Hsin Tung and Chun-Yen Chang. 2024. How three-dimensional sketching environments affect spatial thinking: A functional magnetic resonance imaging study of virtual reality. *PLOS ONE* 19, 3 (03 2024), 1–19. doi:10.1371/journal.pone.0294451
[134] Barbara Tversky. 2011. Visualizing Thought. *Topics in Cognitive Science* 3, 3 (2011), 499–535. doi:10.1111/j.1756-8765.2010.01113.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1756-8765.2010.01113.x
[135] Barbara Tversky. 2019. *Mind in motion: How action shapes thought*. Hachette UK.
[136] Barbara Tversky and Angela Kessell. 2014. Thinking in action. *Pragmatics & Cognition* 22, 2 (2014), 206–223.
[137] Remko Van der Lugt. 2005. How sketching can affect the idea generation process in design group meetings. *Design studies* 26, 2 (2005), 101–122.
[138] Maarten Van Someren, Yvonne F Barnard, and J Sandberg. 1994. The think aloud method: a practical approach to modelling cognitive. *London: AcademicPress* 11, 6 (1994).
[139] Steven G Vandenberg and Allan R Kuse. 1978. Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills* 47, 2 (1978), 599–604.
[140] Radu-Daniel Vatavu. 2013. A comparative study of user-defined handheld vs. freehand gestures for home entertainment environments. *Journal of Ambient Intelligence and Smart Environments* 5, 2 (2013), 187–211.
[141] M.C. Velez, D. Silver, and M. Tremaine. 2005. Understanding visualization through spatial ability differences. In *VIS 05. IEEE Visualization, 2005.* IEEE, MN, USA, 511–518. doi:10.1109/VISUAL.2005.1532836
[142] Brian N Verdine, Roberta M Golinkoff, Kathryn Hirsh-Pasek, Nora S Newcombe, Andrew T Filipowicz, and Alicia Chang. 2014. Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child development* 85, 3 (2014), 1062–1076.
[143] Santiago Villarreal-Narvaez, Arthur Sluÿters, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2024. Brave new GES world: A systematic literature review of gestures and referents in gesture elicitation studies. *Comput. Surveys* 56, 5 (2024), 1–55.
[144] Santiago Villarreal-Narvaez, Jean Vanderdonckt, Radu-Daniel Vatavu, and Jacob O Wobbrock. 2020. A systematic review of gesture elicitation studies: What can we learn from 216 studies?. In *Proceedings of the 2020 ACM designing interactive systems conference*. 855–872.

[145] Tijana Vuletic, Alex Duffy, Laura Hay, Chris McTeague, Gerard Campbell, and Madeleine Grealy. 2019. Systematic literature review of hand gestures used in human computer interaction interfaces. *International Journal of Human-Computer Studies* 129 (2019), 74–94.
[146] Quoc C Vuong and Michael J Tarr. 2004. Rotation direction affects object recognition. *Vision Research* 44, 14 (2004), 1717–1730.
[147] J. Weissmann and R. Salomon. 1999. Gesture recognition for virtual reality applications using data gloves and neural networks. In *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*, Vol. 3. 2043–2046 vol.3. doi:10.1109/IJCNN.1999.832699
[148] Aiden Wickey and Leila Alem. 2007. Analysis of hand gestures in remote collaboration: some design recommendations. In *Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces*. 87–93.
[149] Daniel Wigdor and Dennis Wixon. 2011. *Brave NUI world: designing natural user interfaces for touch and gesture*. Elsevier.
[150] Adam S Williams, Jason Garcia, and Francisco Ortega. 2020. Understanding multimodal user gesture and speech behavior for object manipulation in augmented reality using elicitation. *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (2020), 3479–3489.
[151] Jacob O Wobbrock, Meredith Ringel Morris, and Andrew D Wilson. 2009. User-defined gestures for surface computing. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1083–1092.
[152] P-H Won. 2001. The comparison between visual thinking using computer and conventional media in the concept generation stages of design. *Automation in construction* 10, 3 (2001), 319–325.
[153] Huiyue Wu, Weizhou Luo, Neng Pan, Shenghuan Nan, Yanyi Deng, Shengqian Fu, and Liuqingqing Yang. 2019. Understanding freehand gestures: a study of freehand gestural interaction for immersive VR shopping applications. *Human-centric Computing and Information Sciences* 9 (2019), 1–26.
[154] Yuan Yao, Po-Tsung Chiu, and Wai-Tat Fu. 2017. A Gestural Interface for Practicing Children's Spatial Skills. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion* (Limassol, Cyprus) *(IUI '17 Companion)*. Association for Computing Machinery, New York, NY, USA, 43–47. doi:10.1145/3030024.3038265
[155] Ying Yin and Randall Davis. 2010. Toward natural interaction in the real world: Real-time gesture recognition. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*. 1–8.

Supplemental Materials

All supplemental materials are available in PCS, including the definitions of spatial thinking from the factor analysis.

Spatial Thinking Abilities Definitions

Spatial Visualization This factor refers to an individual's cognitive ability to mentally manipulate and visualize movements, orientations, and spatial forms of objects [23, 75]. Spatial Visualization involves the capacity to imagine and mentally transform complex spatial information [28, 95]. It is helpful for tasks requiring complicated, multistep manipulation of spatial objects and analysis of relationships among different spatial representations, for example, manipulating the internal parts of a complex shape, folding and unfolding flat patterns, or visualizing a sequence of mental manipulations. One important aspect of Spatial Visualization is the ability to recognize and quantify changes in orientation within a scene [20]. This skill requires estimating one's position in relation to a static object or understanding the orientation changes without direct manipulation [33, 93]. Another important aspect is to imagine the final result after folding or assembling parts of an object. Finally, it is important to note that Spatial Visualization is distinct from mental rotation, although they share similarities [93, 96, 120]. Mental rotation primarily involves mentally rotating objects in space, while spatial Visualization focuses on manipulating and understanding spatial forms, including changes in orientation and the assembly or folding of objects.

Spatial Relations This factor refers to an individual's cognitive ability to perceive and understand the spatial relationships and arrangements between objects or components within a scene [37, 66]. Spatial Relations involves the capacity to recognize and interpret the spatial organization, positioning, and connections among objects, irrespective of their size or viewpoint. Spatial Relations is helpful for tasks requiring manipulating and analyzing object components' spatial arrangement, mental transformations, and the coordination of multiple reference frames [75, 119]. To solve these tasks, individuals need to update the relationships between different spatial frames of reference, including the intrinsic reference frames of objects, the egocentric reference frame centered on one's body, and the reference frame of the environment [93]. Examples of tasks that require Spatial Relations include arranging objects that were in disorder. One important aspect of Spatial Relations is the recognition of object-centered volumetric primitives, which allow people to identify objects from different viewpoints and sizes [66, 118]. This recognition is based on processing information from edges, color, texture, or non-accidental properties such as parallelism, collinearity, and cotermination [86, 96, 137]. However, the spatial relations among object features play a critical role in object identification, while part of identity may be less important.

Spatial Orientation This factor refers to an individual's cognitive ability to mentally manipulate and perceive objects or stimuli from different perspectives or orientations, particularly from their point of view (egocentric spatial transformations) [65, 82, 139]. It involves accurately imagining and visualizing how objects or stimuli appear when viewed from various angles or positions [33, 93]. Spatial Orientation also encompasses imagining how an object or scene would appear from another perspective. This involves mentally reorienting oneself within the object, allowing individuals to make accurate spatial transformations and gain a different viewpoint [30]. Spatial Orientation is helpful for tasks requiring identifying positions, matching perspectives, or making correct mental transformations of objects or stimuli in 2D or 3D space [12, 33]. Moreover, dynamic information decreases a task's difficulty by providing additional visual cues, allowing observers to recover shape information, find meaningful edges, segment scenes into discrete objects, anticipate views, and encode different views in visual memory [93]. One important aspect of Spatial Orientation is the ability to perform mental rotations, which entails rapidly and accurately rotating 2D or 3D figures [33]. This mental rotation process involves imagining movement relative to an object-based frame of reference, where the location of one object (or its parts) is specified with respect to other objects.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009

Table 4: Spatial questions

Spatial Ability	Spatial Questions	Previous work
General	What does you non-dominant hand do?	[93, 107]
Orientation	Why do you rotate it?	[10, 12, 33, 52]
Orientation	Why do you rotate it with that hand?	[33, 52, 93]
Visualization	Why did you place the object there?	[33, 36, 72, 130]
Visualization	Why are you ordering the objects like that?	[33, 65]
Relation	Why did you compare these two objects?	[12, 16, 72]
Relation	Do you see any relationship between these objects?	[16, 40, 114]

Table 5: Post-Interview Questions

S.No.	Question	Goal
1	For the task you like the most, could you walk me through your thinking process from the moment you start to the final completion?	Comprehension of the design task
2	How did you determine which block/piece should be placed next? What factors or criteria were you considering?	Comprehension of the design task
3	Did you have any strategies or techniques you use to mentally rotate or manipulate the block/pieces in your mind?	Spatial Orientation
4	How did you understand the relationships between elements?	Spatial Relations
5	What steps did you take to make sure the pieces fit together correctly and align properly?	Spatial Visualization
6	When encountering a mistake or misalignment in the assembly, how did you identify and correct it? Can you provide an example?	Comprehension of the design task
7	Do you have any other comments?	Feedback of study