Challenges of Precueing Instructions for Compound Task Procedures in Mixed Reality

Ahmed Rageeb Ahsan
ahmedrageebahsan@ufl.edu
University of Florida

Andrew W. Tompkins
a.tompkins@ufl.edu
University of Florida

Eric D. Ragan
eragan@ufl.edu
University of Florida

Jaime Ruiz
jaime.ruiz@ufl.edu
University of Florida

Ryan P. McMahan
rymcmaha@vt.edu
Virginia Tech

Abstract

Augmented reality (AR) and virtual reality (VR) can enhance task guidance by overlaying visual information to improve efficiency and reduce errors. However, challenges remain in designing the appropriate presentation format and amount of information for real-time assistance. Prior research has shown benefits of visual cues in procedural tasks, but these findings are limited to simplified scenarios, highlighting a gap in understanding their effectiveness for complex, real-world applications. Therefore, we study visual design and cue effectiveness in the context of compound procedures encompassing subtasks and heterogeneous instructions. We present an experiment assessing different visual cues in VR to test a user's ability to harness distinct information streams for different tasks, separating cues for object search and object placement for multi-step procedures. The results show that even for compound tasks requiring processing of multiple types of information, the addition of simple interaction cues for individual subtasks did significantly improved task performance for both time and errors. However, in contrast to prior studies showing successful precueing of future steps in more simplistic tasks, the study did not find evidence of precueing with the more complex tasks.

CCS Concepts

• **Human-Centered Computing**; • **Mixed / Augmented Reality**; • **Computing Methodologies**; • **Perception**;

Keywords

Virtual Reality, Augmented Reality, Task Guidance, Precuing, Cueing

ACM Reference Format:

Ahmed Rageeb Ahsan, Andrew W. Tompkins, Eric D. Ragan, Jaime Ruiz, and Ryan P. McMahan. 2025. Challenges of Precueing Instructions for Compound Task Procedures in Mixed Reality. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Graphics Interface '25).* ACM, New York, NY, USA, 13 pages. https://doi.org/XXXXXXX.XXXXXXX

Graphics Interface '25, May 26–29, 2025, Kelowna, British Columbia

ACM ISBN 978-1-4503-XXXX-X/18/06
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

The ability to overlay additional information in a user's view within virtual reality (VR) and augmented reality (AR) offers substantial potential to enhance performance during procedural tasks. This is particularly relevant in task guidance, where users can receive real-time instructions and tips as they work through complex procedures (e.g., [37, 41, 57]). For example, VR training applications provide step-by-step guidance to help users learn new operations, such as vehicle handling [56] or safety procedures [58]. Similarly, AR can assist workers with mechanical repairs by displaying the correct sequence of actions and alerting them to errors [7, 57]. These mixed reality applications deliver real-time, step-specific guidance that aids users in executing tasks efficiently, even without prior experience.

However, communicating instructions is significantly more difficult for tasks that involve integrating multiple pieces of information and combining multiple actions. We refer to such tasks as *compound tasks*, which require users to perform multiple types of actions (e.g., searching for objects, manipulating items, and placing components) together or in near sequence. To communicate necessary details of multiple related actions, *compound instructions* provide step-by-step guidance that include different types of information—such as the combination of text and images to describe what items are required and how they are to be used in completing a compound task. Typically, understanding the multiple pieces of information together is necessary for meaningfully interpreting the task as a whole, but simultaneously processing multiple types of information is naturally more cognitively difficult than processing individual pieces of information.

A key design challenge is determining how much and what types of information to present, as the large amount of information needed to describe compound tasks open many options for presenting guidance. The use of *interaction cues*—visual or auditory signals added to guide user attention—can play a critical role in managing the challenges of distilling immediate individual actions within the overall procedure. For instance, highlighting objects, arrows indicating directions, and icons providing additional information can help users break down compound instructions into manageable parts, making it easier to understand and execute each step. Interaction cues are particularly beneficial in task scenarios where procedural accuracy and timely execution are crucial, such as in medical training, assembly lines, or emergency response simulations. By directing the user's attention to the next required action, these cues can reduce cognitive load, minimize errors, and enhance overall task performance. For instance, in a medical training simulation, an AR overlay could highlight the correct anatomical location for a surgical incision and indicate the sequence of steps. However, despite their potential, the specific impact of interaction cues in guiding users through compound instructions remains underexplored.

Prior studies have shown interaction cues to be effective for improving task performance for basic individual actions (e.g., search, navigation, target selection, object manipulation) [8, 21], but there is limited understanding of the effectiveness of different cues for cases involving compound tasks that require processing more information. A related design issue for task guidance instructions is how to provide hints about upcoming steps, which can provide context and potentially help the user plan ahead for even greater efficiency. This is referred to as *precueing* [13, 40, 51]. Support for precueing involves presenting cues or information about future steps in a sequence. Research has demonstrated that precueing is beneficial for enhancing task efficiency in simple tasks [25, 26, 52] that require one or very limited actions.

Despite the potential benefits of interaction cues and precueing, its effective implementation presents several challenges in mixed-reality applications for compound tasks involving greater amounts of information. Too much information may overwhelm or distract users, especially when using precueing to present information about multiple future steps. Conversely, insufficient information may hinder users' efficiency in navigating a sequence of steps quickly.

We present a VR experiment to study whether users can effectively take advantage of different types of interaction cues and information about future steps in cases when following compound task instructions. The experiment controls separate interaction cues for a specific subtasks (3D object search and placement) while requiring the interpretation of various forms of instructions (e.g., text, images, spatial relations). Further, this study assesses the effectiveness of precueing for future steps in the compound task. The findings provide novel evidence of the effectiveness of interaction cues for fundamental subtasks (search and placement) within compound tasks despite the existing overhead of processing compound instructions. In addition, the study did not find evidence that participants effectively utilized information about future steps to enhance their task performance. This finding raises concerns about the practical applicability of precueing techniques in complex tasks.

2 Related Work

The experimental design in the presented research was informed by relevant literature in interaction cues, task guidance for VR and AR applications, and task precueing.

2.1 Interaction Cues

To enhance task performance, previous researchers have frequently employed interaction cues within their studies while delivering task instructions. Interaction cue is a technique designed to direct users toward specific actions at particular points in the task workflow [21]. These cues play a pivotal role in indicating desired locations, facilitating the discovery of intended objects, and offering directional guidance [10]. Empirical investigations have consistently demonstrated a positive correlation between the incorporation of such interaction cues and elevated user performance [8, 21].

However, there are many types of interaction cues such as visual cues, auditory cues, and haptic cues. Among all of these cues, visual cues generally work better than others for virtual environments [27]. Therefore, use of visual interaction cues was deemed most suitable for our study. Different types visual cues are used for different type of purposes. The creation of effective interaction visual cues necessitates consideration of multiple design factors. Notably, attributes such as color, shape, and texture play a pivotal role in influencing the efficacy of interpretation [48]. In AR and VR research, diverse approaches have been adopted to convey specific information. Several researchers have opted for arrows to signify navigational directions [4, 22, 32] or points of interest [9, 19, 24, 37, 57]. Arrows have also been used to represent various actions like object insertion and movement [42]. Simple visuals are also common, such as circles [3, 53] or rectangles to indicate objects [7, 34]. Color coding and highlighting can accentuate specific objects [35, 44, 47]. Other research used box cues to indicate the focus plane in video see-through scenarios [11]. Animations of a virtual hand can serve as a cue (e.g., pointing to an object or selecting an object) [5, 36]. In scenarios necessitating a sequential order of objects for task completion, the use of numbers as cues is common to indicate a sequence [2, 29, 49].

The variety of methodologies employed by researchers in AR/VR underscores the significance of designing interaction visual cues that cater to diverse perceptual and functional requirements. Our research investigates the effectiveness of interaction cues for task guidance and the ability to precue future steps. Our study design employs commonly-used forms of interaction cues with the goal of testing a context where the task and instructions are easily understandable without extended explanation or practice.

2.2 Task Guidance Formats in VR and AR

Prior work has explored the use of AR and VR to provide users with guidance during tasks. For instance, Laviola et al. [23] simulated a LEGO assembly task and examining the effect of different combinations of visual AR assets on task performance. Based on their results, the authors recommend using minimal visual cues, such as circles and arrows, for locating objects, color coding for identifying objects, and drawings for identifying multiple objects. As another example, Webel et al. [53] developed a mobile-based AR system to provide maintenance task instructions. Their implementation employs a combination of textual explanations, 3D images, and cues to deliver the instructions effectively. While some implementations limit complexity by providing only text-based instructions (e.g., [35]), a considerable number of examples have combined both images and text to convey instructions (e.g., [3, 31, 39, 42, 47, 57]). A survey by Gattullo et al. [15] found that approximately 63.22% of AR for assembly and industrial applications employ visual aids—including drawing, auxiliary model pictures, and images of products to communicate instructions. In comparison, the review found 26.12% of work used text-based instructions. Thus, the most common approaches for instructions combine text, images, or a combination of the two.

Informed by the existing body of research, our study similarly assesses interaction cues in a task guidance context that combines both textual descriptions and visual illustrations to communicate instructions. As can be seen in Table 1 from the previously stated research, existing task-guidance applications commonly combine representations to present compound instructions.

2.3 Task Precueing

Precueing is a fundamental cognitive process of providing information about impending stimulus [25]. This information can be presented by variations including color, location, and orientation [59]. Precueing reduces reaction time and improves task time by: 1) enabling users to prepare for the pending stimulus, 2) serving as an alerting signal and 3) allowing users to anticipate the upcoming location and timing [17]. Additionally, research has demonstrated that precueing reactivates mental imagery resulting in an increased vividness and detail in mental imagery, translating into improved performance on cognitive tasks [13].

In the context of VR and AR, precueing is primarily used to improve task performance. For example, Liu et al. [25] examined the effects of utilizing a different number of precues for task-following tasks (e.g., simple actions at each of a series of locations). Their results indicated that multiple precues led to faster completion times for task-following activities. However, a follow-up study [26] of precueing for object manipulation tasks (involving object movement and rotation) found a single precue was more effective. This demonstrates the importance of attention to task complexity. In another example, Volmer et al. [51] examined precueing for a button pressing task and found drawing a line to the next step was highly efficient for enhancing task performance. Another study by Volmer et al. [52] added intermittent changes to different button stations as part of the task sequence, though the location changes were limited to discrete steps within the sequence.

In this paper, we examine if the benefits of precueing are still observed for tasks with multiple subtasks across a wider workspace for compound tasks. An important differentiation from prior work is the increased amount of task complexity and information processing in our study. Table 2 provides a summary of key differences among recent precueing studies. For example, prior studies [26, 51, 52] involve selection tasks with a limited amount of change to the user's viewpoint during the progression of subtasks. More specifically, Liu et al. [26] occurred completely within the seated user's viewpoint. Similarly, while Volmer et al. [52] required viewpoint changes, these were intermittent turns between a series of several tasks occurring within the user's viewpoint. In contrast, our study requires numerous viewpoint changes for the combination of referencing the provided instructions, searching for target objects from the surrounding environment, and placing the object at an instructed destination. We designed the procedure in our study to require processing multiples types of text and image information as part of the compound instructions. Further, performing the task required continual changes to the viewpoint for surrounding 3D search, and each step of the procedure required switching among multiple subtask types while processing the step instructions.

3 Experiment

We conducted a controlled experiment to evaluate whether participants could effectively use precueing hints for future steps to improve task completion. The study used a VR system to simulate AR and independently control object search cues, object placement cues, and the number of instruction steps shown at a time.

3.1 Research Goals and Hypothesis

This research investigates the implications of providing different types of visual interaction cues for user performance on compound procedures involving sequences of multiple steps. Common representations of instructions of steps within multi-step procedures consist of a combination of different types of information such as images and descriptive text. Table 1 shows examples from an analysis of instructional guidance applications from related work. Headworn VR and AR can offer convenience and aid task efficiency by presenting instructions and providing visual cues for the next steps in the procedure. However, due to the complexity of compound steps, there is a risk of providing more information than is helpful. Our experiment tests (a) whether visual cues for certain types of subtasks are more effective than others, and (b) whether users can effectively take advantage of information for future steps.

For our research goals, we address the following hypotheses:

- **H1:** *The inclusion of in-world visual cues will improve task performance (i.e., reduce errors and reduce completion time) on compound procedures with compound instructions.* We hypothesize that, despite the complexity of information processing for the compound task, the addition of visual annotations for objects search and placement subtasks will provide performance benefits over the presentation of instructions alone.
- **H2:** *The presence of in-world visual cues will reduce the need to refer back to full instructions.* We hypothesize that integrating visual annotations for relevant objects and locations will reduce the amount of details the user needs to remember when completing steps, which we expect will lead to differences in viewing and turning behavior.
- **H3:** *Visually presenting additional future steps in compound tasks will reduce task completion time.* While prior work has demonstrated the benefits of precueing upcoming steps in more simplistic selection [25] and object manipulation tasks [26], we test whether the benefits of precueing can be achieved in a compound task that requires parsing different types of information and while performing both search and object manipulation subtasks. If participants can process information for future steps in advance, they will be able to complete the task faster.
- **H4:** *The cognitive difficulty of performing a compound procedure will increase as the total amount of information increases.* We expect cognitive load to increase when more step instructions and visual in-world cues are provided. This hypothesis will test the risk of visual clutter or information overload associated with adding visual annotations and more step instructions to the user's view.

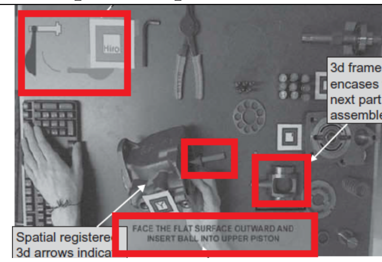
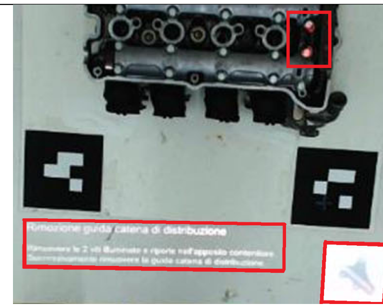
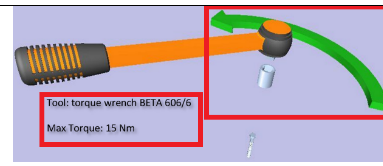
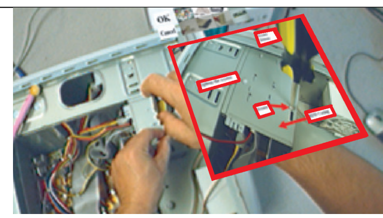
Paper	Description	Example of Step Instruction
Radkowski et al. [37]	Steps require manipulation of multiple objects described through textual instructions, visual diagram with annotation, and spatial indicators.	
Uva et al. [47]	Steps require manipulation of objects while a colored graphic sign is used to indicate the required object. There is also textual information that aids the participant to understand the action.	
Fiorentino et al. [14]	Instruction indicates the specific tools to use, the required motion, and amount of force.	
Yuan et al. [57]	Instructions include labels for multiple objects components while showing the required action with the recommended tool.	

Table 1: Examples from literature analysis shows previous mixed-reality task guidance applications that provide compound instructions. We designed the task in our study to require interpretation of multiple types of information for each step, reflecting the compound instructions used in guidance applications. Images are included under fair use for research with red annotations added to highlight variety of information design across applications.

3.2 User Task and Environment

To test our hypotheses, we designed an experiment with a task composed of a sequence of subtasks requiring object search, selection, and relational spatial positioning. Our experimental design closely resembles common procedures (e.g., assembly, repair, recipes) where users are guided through a sequence of tasks involving different objects. For example, if a mechanic tries to assemble a piston to an engine block, they need to locate the appropriate piston (which might be stored in a designated area), ensure that it matches the specifications for the engine being assembled, and finally ensure the selected piston must be positioned correctly within the engine block, aligning it with the cylinder bore. Another example is computer assembly requiring component identification and placement on the motherboard, which similarly involves locating and positioning specific objects in designated locations [37]. Notably, this multi-step sequence mirrors the fundamental structure of assembly tasks. The design of our study aims to capture a similar composition of compound actions as are commonly seen in real-world assembly procedures (see Table 1) and less commonly tested in precueing studies (see Table 2). Importantly, the study was designed to require textual instructions that provide details of how each step should be performed, as is typically provided by instruction manuals for procedures. To support the focus on assessing instruction-following and task performance for a generalized population, we designed an abstracted task with objects and movements that could be clearly described and easily understood by lay participants. Actions were simplified to avoid operational difficulty and to enable unambiguous measurement of success. Each trial consists of twenty tasks, where during each task participants were required to do three subtasks: (1) find an instructed target

	Study	Guidance Format			Subtask Types			Viewpoint Changes	Interaction Cue Format	Cue Result	Precue Result
		Interact cues	Text	Image	Selection	Search	Obj. Manipulation	Amount	Format	Effective	Effective
1	Liu et al. [25]	✓			✓			None	Geometric, Arrow	✓	✓
2	Liu et al. [26]	✓			✓		✓	None	Geometric	✓	✓
3	Volmer et al. [51]	✓			✓	✓		Limited	Geometric, Color. Arrow, Blink	✓	✓
4	Volmer et al. [50]	✓			✓	✓		Limited	Geometric, Color. Arrow, Blink	✓	✓
5	Volmer et al. [52]	✓			✓	✓		Occasional	Geometric, Color. Arrow, Blink	✓	✓
6	Our study	✓	✓	✓	✓	✓	✓	Frequent	Geometric, Number, Color	✓	

Table 2: Comparison of tasks in recent previous precueing studies. Note that for types of interaction cues in this table, *geometric cues* include lines, circles, and rectangles; and *arrow* type cues include standard arrows and alternative directional shapes.

object in the environment, (2) select and "pick up" the object, and (3) place the object in an instructed destination location.

The task was designed to require both physical turning and a limited amount of physical walking and reaching to find and move objects in a virtual work area. Figure 1 shows an overview of the virtual environment for the task. The virtual environment had the participant stand in front of a central (virtual) table to view instructions and place objects. Two shelves (one on each side of the participant's primary location) contained the necessary virtual objects for the task in addition to some that were not required.

For each trial, the setup had 26 unique objects distributed across two shelves. The starting object locations were randomly selected from 28 possible starting positions to provide minor variations or gaps in the object layout.

For the purpose of studying the presentation of task instructions, the object set needed to consist of visually distinct and easily recognizable objects that could be easily identified with textual names. Therefore, we selected object types based on various geometric solids (cube, cylinder, capsule, rectangle, pyramid, trapezoid, cone, octahedron, tetrahedron) and chess game pieces (king, queen, knight, rook) with a total of 13 possible unique shapes. Each shape was also present in two colors (red and dark blue) to create a total of 26 distinct objects. The colors were selected to increase accessibility for those with color impairment. To create the multi-step procedure, each trial consisted of 20 tasks that each involved one specific, unique object of a certain shape and color. To do this, for each trial, 20 objects were randomly selected from the set of objects, and each was assigned to a step in the procedure.

For each task, the study application presented visual instructions (see Figure 1) for which object was needed and where it should be placed on a 3x3 grid on the central table. Instructions were presented above the central table. Furthermore, we did not fix the instructional panel to participants' field of view (FOV) as there can maximum of four instruction panels at a time which may cause visual clutter, and participants can not see the environment properly. Moreover, previous research [55] suggests that world-anchored text is preferable over fixing text to participants' FOV as in a heads-up display (HUD).

Participants were required to successfully complete each task before the application updated the task instructions for future steps. The instructions included a simple text description indicating the name of the new object and where it should be placed relative to the object destination of the previous task (i.e., the number of squares to the left or right and up or down from the previous object on the grid). The exception was the first task; the first object was always to be placed at the center of the grid. Additionally, each instruction included an image of the object they would need to find.

To keep a similar level of difficulty and visual complexity for object search, each selected object taken from a shelf for each task was replaced with a virtual duplicate in its original location after completing the task. In addition, the object placed on the grid from the previous task would disappear after the successful placement of the next object, thus ensuring only one reference object would be on the grid at a time for each task. When a participant completes a step correctly, the corresponding object (e.g., cube, cylinder) adheres to the table, becoming immovable and non-interactive. Additionally, the object changes its color, and the instructional prompts are updated to guide the participant to the next task. These real-time feedback mechanisms ensure participants can easily identify their successful placement. Conversely, if a participant places the correct object in an incorrect location, an incorrect object in the correct location, or an incorrect object in an incorrect location, the object automatically returns to the shelves, and the instructional prompts remain unchanged. This feedback informs the participant that the step was executed incorrectly, prompting further attempts.

We structured the environment with the expectation that participants will utilize precueing, gather future elements together, and remember the future target. This will help the participants mitigate repeated instruction reading, and facilitate the efficient completion of tasks within minimal timeframes. The study task is designed as a procedural task requiring participants to complete steps sequentially.

However, participants had the option to exploit precueing future steps by keeping track of upcoming future object locations for faster selection. In addition, participants could optionally collect all objects indicated in the instructions and place them closer to the target table (i.e., still only moving one object at a time but completing

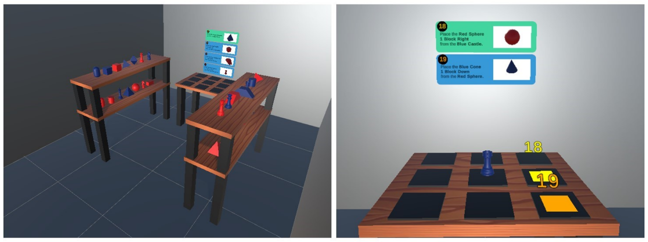


Figure 1: The virtual environment used for the study. The left image shows the layout, where participants were positioned between the two shelves with the placement grid straight ahead. The right image shows how the instructions were shown above the 3x3 placement grid. The number of instructions shown and the inclusion of placement cues (floating numbers and bright squares) depended on the experimental condition.

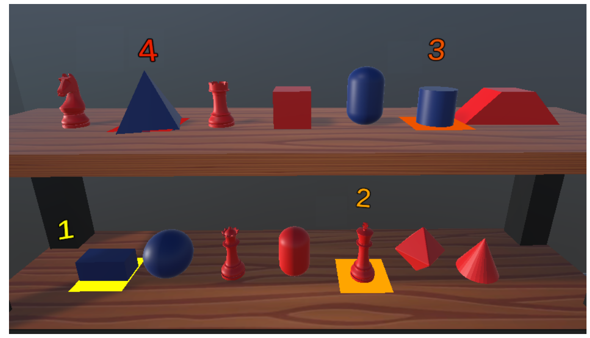


Figure 2: View of the shelves with target objects with search cues enabled. The numbers map to instructions, and the cue for the immediate step was always shown in yellow.

multiple search tasks as a batch) and then place them to their final destinations afterward. While participants were informed that they could use this strategy, it was not mandatory. Although this strategy had the potential to accelerate task completion, observations showed that only a few participants chose to utilize it.

3.3 Visual Cues for Task Guidance

To study the impact of interaction cues for 3D object search and placement, the experimental design adopts familiar visuals combining numbers and simple markers, as are commonly used in many mixed reality applications (e.g.,[2, 3, 7, 29, 34, 49, 53]). Though prior studies suggest leader lines to be preferred as type of cues [25, 51, 52] such findings were demonstrated for more simplified tasks completed within a constant viewpoint or with only occasional, distinct viewpoint switches to a new reference frame. In contrast, our research explores more dynamic search that requires continual viewpoint changes due to intermixed subtasks across locations. Leader lines would therefore not be suitable due to the amount of criss-crossing and visual clutter in compound tasks covering multiple sites. Moreover, our research aims to test user ability to utilize different types of subtask cues in the compound task and to independently control search cues and placement cues as separate independent variables. This would not be possible with line cues since a line would indicate both search and placement locations together (i.e., the two endpoints of a line), thus making it impossible to separately manipulate cues for the search and placement types. Nevertheless, based on the review presented by Gattullo et al. [15], visual location markers and numerical indicators are commonly used types of interaction cues, making them a suitable choice for our study.

In our task, we refer to *search cues* as visual annotations for the new object in the environment needed for the step. The search cue was comprised of two parts: 1) a colored square underneath the object to aid identification along with 2) a colored number (corresponding to the step number) floating above the object (see Figure 2). The color for each search cue was based on the step number, with the current step always shown in bright yellow, and colors for the upcoming steps followed a gradient from darker yellow, through orange, to red for later steps.

In addition, *placement cues* show the object's intended destination on the 3x3 grid. The visual design for the placement cues match the search cue; the same colored squared and numbers were positioned at the destination location (see Figure 1, right).

The design of our visual cues-color, number, and shape-was informed by prior studies and tailored to the task in our study through iterative pilot testing. The cues were selected to ensure clarity and support participants in completing the steps effectively. We use numbers serve as a critical component for step-tracking, allowing participants to quickly identify which step they are on. This aligns with previous research Mura et al. [29] demonstrating that numerical labels facilitate cognitive processing in sequential tasks. In our study, the numbers also correspond to step instructions displayed on the panel in front of participants, reducing cognitive load by directly linking instructions to actions. Moreover, the numbers provide an intuitive guide for object selection, as participants must match objects to the corresponding step numbers. Numbers also helped them to place the object on the right target location, which was labeled with the matching number for each object. The choice of colors for visual cues were chosen to provide high contrast and visibility in the test environment (we note that optimal color choices will depend on specifics of environment and display system). In our study, we use a gradient-based approach for step identification: the current step is always highlighted in bright yellow, while upcoming steps transition from darker yellow to orange and red. This gradient encoding not only enhances visibility but also provides a temporal cue, allowing participants to anticipate and prepare for future steps in the sequence. The color choice made it easy to distinguish the cue for the current step as a highly-salient yellow, which was not used for other objects or visual information in the environment. Finally, we use the rectangle beneath each object aids spatial recognition, helping participants identify the correct object and its location more easily. Previous studies [7, 37] used rectangles to highlight the current objects. In our case, the rectangle acts as both an object indicator (i.e., *which object?*) and a location indicator (i.e., *which target destination?*).

3.4 Experimental Design

The experiment controlled three independent variables following a 2x2x4 mixed design. The two types of visual cues were independently controlled between-subjects to separate cues for object identification for search and target location for placement, while the number of precued steps varied within-subjects. Below are descriptions of each variable and each level of the variables:

- **Search Cues:** Has two levels: the presence or absence of search cues, as illustrated in Figure 3.
- **Placement Cues:** Has two levels: the presence or absence of placement cues, as illustrated in Figure 3.
- **Number of Steps:** How many steps are shown at a time in the instructions list (as shown with green and blue rectangular panels in (Figure 3). This also controlled the number of tasks ahead to show integrated visual cues for search objects and placement location, if enabled per the experimental condition. This was either one steps, two steps, three steps, or four steps.

The *search cues* aid participants in locating the required objects, whereas *placement cues* guide them in correctly positioning these objects. Since these cues assist participants in distinct ways and were toggled independently according to the experimental conditions, ensuring that changes to each cue type did not affect the display of the other.

3.5 Equipment and Software

The virtual environment was developed using Unity (v2019.4.8f1) with the SteamVR Unity Plugin (v2.7.3). Our setup used the HTC Vive Pro Eye Head-mounted display (1440x1600 pixel resolution per eye with a 110-degree diagonal field of view) with two base stations, allowing for 3D room and head tracking. The study application was running at 120 frames per second for the experiment.

For object interaction during the task, the application used the virtual hand technique using a motion-tracked Vive Pro controller. The participant would receive vibration feedback when colliding with another object to indicate that selection was possible. Objects would snap to the center of the hand when selected.

3.6 Procedure

The study was conducted in person, with each study session starting with informed consent followed by a background questionnaire. The experimenter then helped familiarize the participant with the VR headset and eye calibration.

Next, the experimenter explained the study task and guided participants through a tutorial task consisting of a three-step procedure as a short example of the task. The tutorial task enabled or disabled the visual cues for object search and placement as appropriate for the participant's assigned experimental condition. Participants were free to practice using the tutorial and were encouraged to ask any questions. After the participant demonstrated the ability to successfully interpret the presented instructions, manipulate virtual objects, and place objects in the correct indicated positions, participants performed the primary task.

Each participant completed four procedure trials, each consisting of 20 tasks. Following the experimental design, the conditions for visual search and placement cues were consistent for all four trials, but the number of steps visible at a time (i.e., the number of instructions shown and the corresponding visual cues available for current and upcoming steps) varied following a balanced Latin square for trial ordering.

After each trial, participants completed the six NASA-TLX ratings and fifteen weighting questions using a desktop computer, which we used to calculate total workload as described in [18]. By conducting this questionnaire without the headset on, participants always had a brief break from the VR environment between trials. The study session took between 30–50 minutes, depending on task speed and choice of an optional break (a few minutes to stretch or use the restroom) after a trial.

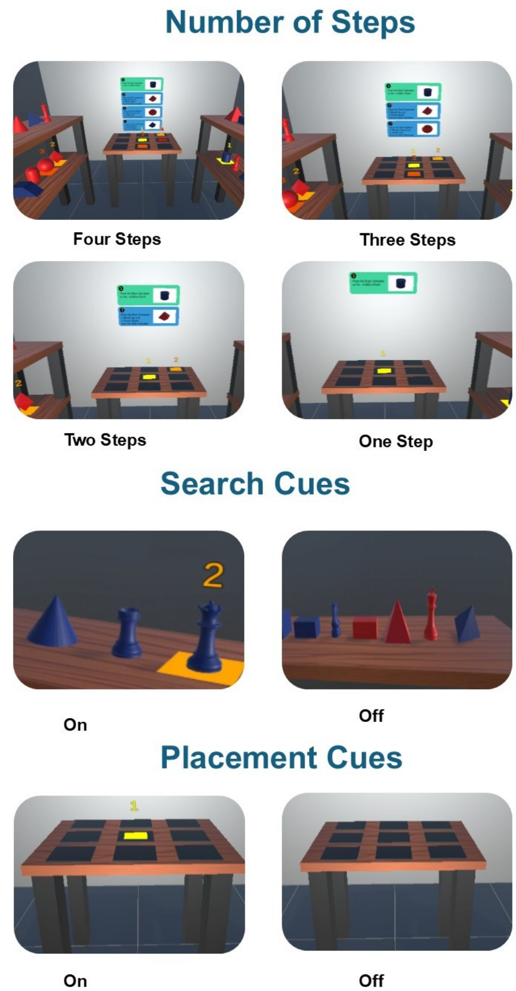


Figure 3: Figure showing various experimental conditions. The first two rows illustrate the "Number of Steps" conditions, ranging from 4 steps to 1 step. The third row presents the "Search Cue" conditions (Search Cue On and Search Cue Off) followed by the "Placement Cue" conditions (Placement Cue On and Placement Cue Off).

3.7 Participants

The study was approved by the ethics review board of our institution. The study sample included 48 participants with 34 participants self-reporting as male and 14 as female. We dismissed three participants for a total of 51 participants; two users could not complete the trial due to technical issues, and one participant reported headaches. But no one has reported any visual impairments. Participants' ages ranged 18–30 with a median age of 22. Participants were undergraduate or graduate university students at our institution. Among the 48 participants, 38 were studying in computing disciplines, 5 in engineering disciplines, and 5 in various other disciplines. Participants were compensated with in-course credit for approved courses.

4 Results

We investigated the impact of in-world cues, including both search and destination cues, as well as the number of steps, on user performance while conducting the tasks. Additionally, we assessed users' mental load through a post-task survey. We measure user performance through two dependent variables: task completion time and task errors.

4.1 Task Completion Time

Task time was analyzed based on the average time per step for the core steps of the procedure. Notably, the initial step of each trial was excluded, as the first instruction differed from subsequent tasks in that it could not describe a placement instruction relative to an object from the previous step. Subsequent steps involve interpreting spatial instructions relative to other objects (e.g., "Place the Red Cube 1 block right and 2 blocks down from the Red Cone"), while the first step involves a simpler instruction: "Place the Red Cube in the middle block." Consequently, the placement task is much easier in the first compared to the following steps. Therefore, during data analysis, the completion time for the first step was excluded from the total task time. We also omitted the final three steps of the procedure because the presentation of instructions for future steps varied when there were no more future steps. Therefore, the analysis focused exclusively on the task completion times for tasks numbered 2 through 17 of the 20 step sequence, ensuring that results for all conditions cover an equal number of steps.

Before analysis, we first checked for outliers, calculated as values falling outside the range of [Q1 − 3*IQR, Q3 + 3*IQR]. We removed outlier points associated with specific step completion times within trials. This approach resulted in a data loss of only 0.3% of steps.

Following outlier removal, time results are summarized in Figure 4. We conducted a three-way mixed ANOVA for the two between-subjects factors (both cues) and one within-subjects factor (number of steps shown). Prior to testing, we addressed test assumptions using Shapiro–Wilk test, Levene's test, and visual inspection of the data distribution; we applied a logarithmic transformation to the time data to address an identified normality issue.

The ANOVA results for task time are summarized in Table 3. The results indicate that the number of instruction steps shown did not significantly influence task time. Therefore, we do not find evidence to support hypothesis **H3**. Conversely, our analysis demonstrates that both destination and search cues exert a notable influence on time, supporting hypothesis **H1**.

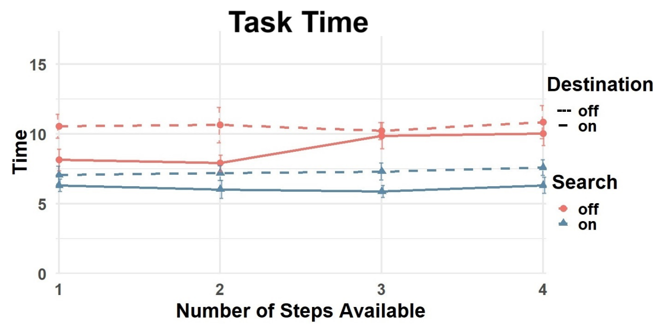


Figure 4: Time results as the average step completion time (in seconds) per condition. Search cues and destination cues each independently had significant effects for faster task time. Error bars show standard error.

For deeper investigation into the time results, we conducted an additional analysis to separate task times by distinguishing between subtasks during the study.

We also analyzed task times by separating subtasks into search and placement periods. *Search time* was defined as time when participants faced the shelves or when not holding virtual objects, while *placement time* was calculated as the time holding an object while facing the grid area before confirming placement. After following the same method previously described for the time results with outlier removal and assumption checks for normality and homogeneity of variance, a three-way mixed ANOVA showed significant reductions in search time (F(1, 44) = 33.28, $p < 0.001$) with search cues and in placement time with both search cues ($p < 0.001$) and placement cues ($p = 0.007$). In addition, as with the analysis of overall task times, no evidence of precueing was detected for the subtasks. No differences were detected for effects of the number of steps shown on either search time or placement time.

Time Results

Factor	Effect	F	p
Search Cue	**Cue on** < Cue off *	F(1, 44) = 33.28	< 0.001 *
	On: $M = 6.73$, $SD = 1.96$		
	Off: $M = 9.79$, $SD = 3.22$		
Destination Cue	**Cue on** < Cue off *	F(1, 44) = 6.68	0.007 *
	On: $M = 7.57$, $SD = 2.76$		
	Off: $M = 8.94$, $SD = 3.22$		
Steps Shown	Not significant	F(2.44, 107.52) = 1.49	0.227
	1 step: $M = 8.03$, $SD = 2.83$		
	2 steps: $M = 7.96$, $SD = 3.21$		
	3 steps: $M = 8.33$, $SD = 2.90$		
	4 steps: $M = 8.71$, $SD = 3.35$		

Table 3: ANOVA results for task time show the presence of search cues and destination cues each led to faster performance. Interaction effects were not significant and are not included in the table.

4.2 Task Error

The experiment counted task errors based on deviations from the prescribed task instructions. Errors were counted whenever either an incorrect object was used or when an object was placed at an incorrect grid location. Since participants were required to successfully complete each step before advancing, it was possible to make multiple errors for each step.

In preparation for analysis, we examined potential outliers following the same criteria as for the time results (see Section 4.1); no data points were considered as outliers. Due to wide variance and distribution skew, the raw data did not meet the assumptions for parametric testing. We therefore transformed the data using aligned rank transform (see [12, 54]) prior to factorial ANOVA testing.

Figure 5 shows the errors per cue conditions, and the outcomes of the ANOVA are detailed in Table 4. The test found the addition of *search cues* significantly reduced errors compared to *no search cues*. No effect on error was detected for *destination cues*. A significant interaction effect on error was detected between the *search cue* and *destination cue* factors, and a following post-hoc Tukey test showed significant results demonstrating the greatest reduction in errors when both search and destination cues were enabled together.

Thus, the results underscore the collective influence of both types of cues (search and destination) on the observed outcomes. Therefore, we accept our **H1** hypothesis in a qualified manner, acknowledging the significant roles played by source and destination cues while noting the lesser influence of the destination cue on participant performance.

The mixed ANOVA did not detect a significant effect on errors due to the number of steps shown, as the resulting 0.051 p-value was not statistically significant at the chosen 0.05 alpha threshold (see Figure 4). However, as the ANOVA analysis did not consider the ordinal nature of the different numbers of steps, we followed up with a correlation analysis. A Spearman correlation test between the number of steps shown and detected errors produced $rho = 0.155$ and $p = 0.032$, indicating a weak correlation of errors increasing with more steps shown. These findings provide some evidence that displaying too much information with interaction cues for future events may have detrimental effects. However, the effect is limited for the task of this study, as overall participant error was low.

In addition to total errors, we also considered the composition of the two types of errors, i.e., use of the wrong object for a step (51.43% of errors) or placing an object in an incorrect grid location (48.57%). The similar ratios of error types further reinforces the compound nature of the task required similar attention to the search and placement requirements. For completeness, we conducted an additional analysis with a three-way ANOVA to test for effects on error-type ratio due to the cues and number of steps. No differences were detected, as ratios of error types were similar across conditions.

Task Error Results

Factor	Effect	F	p
Search Cue	Cue on < Cue off *	F(1, 44) = 4.11	0.049 *
	On: $M = 0.74$, $SD = 1.89$		
	Off: $M = 1.22$, $SD = 1.44$		
Destination Cue	Not significant	F(1, 44) = 0.05	0.815
	On: $M = 0.81$, $SD = 1.34$		
	Off: $M = 1.15$, $SD = 1.98$		
Steps Shown	Not significant	F(3, 132) = 2.66	0.051*
	1 step: $M = 0.50$, $SD = 0.86$		
	2 steps: $M = 0.94$, $SD = 1.36$		
	3 steps: $M = 1.25$, $SD = 2.21$		
	4 steps: $M = 1.23$, $SD = 1.95$		
Search x Dest Cue	Interaction effect *	F(1,44) = 5.13	0.028 *
	(both cues) < (no cues) *		0.047 *
	(both cues) < (only dest. cues) *		0.002 *

Table 4: ANOVA results for task errors show the presence of search cues led to significantly less error. The interaction effect between search and destination was significant; other interactions were not significant and are not included in the table.

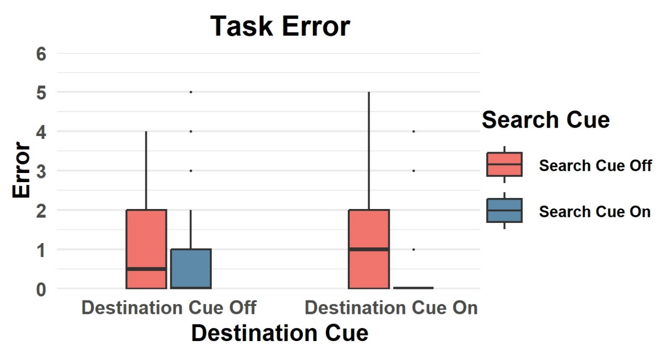


Figure 5: Average task error per condition for search and destination cues. The addition of search cues significantly reduced errors, while a significant interaction effect shows destination cues only significantly reduced errors when search cues were also used.

4.3 Mental Load

We analyzed participant scores from the NASA TLX questionnaire [18] completed on completion of each procedure. Analysis by three-way mixed ANOVA did not find significant differences due to the experimental factors. We also tested correlations between number of steps shown and TLX scores. No significant effects were detected, thus we did not find support for hypothesis **H4**. For brevity, we omit further statistical details of the non-significant effects in the paper.

4.4 Viewing and Turning Behavior

The study also tested how the visual cues affect users' viewing and head movement behaviors. Specifically, we address hypothesis **H2** by analyzing how often participants referred to the full instructions

Factor	Effect (Viewing Behavior)	F	p
Search Cue	Cue on < Cue off *	F(1,44) = 13.62	< 0.001 *
	On: M = 130, SD = 129		
	Off: M = 179, SD = 144		
Destination Cue	Cue on < Cue off *	F(1,44) = 33.70	< 0.001 *
	On: M = 107, SD = 124		
	Off: M = 202, SD = 136		
Steps Shown	Not significant	F(2.36, 103.67) = 0.18	0.868
	1 step: M = 142, SD = 123		
	2 steps: M = 160, SD = 128		
	3 steps: M = 175, SD = 185		
	4 steps: M = 141, SD = 106		
Search x Dest Cue	Interaction effect *	F(1,44) = 14.46	< 0.001 *
	*(both cues) < (no cues)**		< 0.001 *
	*(both cues) < (only dest. cues)**		< 0.001 *
	*(both cues) < (only sear. cues)**		< 0.001 *
	*(only dest. cues) < (no cues)**		0.009 *

Table 5: ANOVA results for viewing behavior for number of times participants looked at the step instructions. The only significant interaction effect was between source and destination cues, with additional p values included from a Bonferroni-corrected pairwise t-test. Other interactions were not significant and are not included in the table.

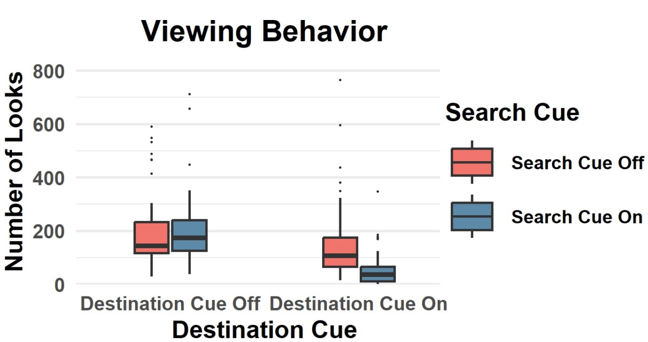


Figure 6: The results for viewing behavior show the number of times participants referenced the instructions during the entire procedure. Both types of cues significantly reduced how often participants looked at the full step instructions, and a significant interaction effect found the combination of both cues together was better than having either single cue alone.

(i.e., the central panel with textual descriptions and images) during the task, as the presence of interaction cues was expected to reduce the need to review the written step details. Analysis of viewing behaviors used application logs of participants' 3D physical head movements over time using the motion tracking system during the study. Processing the head rotation data, we calculated a count of the number of times a participant looked toward the instruction list above the target grid. This measure was calculated when the gaze direction moved within a 30-degree horizontal range and within a 25-degree vertical range around the central instruction location; this range was chosen to exclude viewing periods when participants were searching the shelves or looking down at the target grid. The instructions were required to be in view for a minimum duration of 0.1 seconds to filter tracking noise or accidental recounts.

Before analysis, we conducted outlier checks and applied outlier removal following the same approach described in Section 4.1. Figure 6 graphically shows the viewing behavior results. For parametric statistical testing, a logarithmic transformation was used to normalize the data to meet assumptions. The outcomes of the three-way mixed ANOVA are presented in Table 5.

Both the *search cue* and *destination cue* variables had a significant influence on participants' gaze direction towards the panel. A noteworthy finding emerges with respect to the interaction effect between the *search* and *destination* cues, signifying a substantial influence on participants' visual focus. A Bonferroni-corrected pairwise t-test confirmed a statistically significant reduction in visual engagement when both cues were activated, supporting **H2**.

We note the *number of steps visible* did not significantly affect viewing behavior.

4.5 Summary

In summary, our results provide partial support for our initial hypotheses. H1 is supported, as both search and destination cues significantly improved task performance by reducing task completion time and errors. H2 is also supported, with evidence showing that the presence of in-world visual cues reduced the need to refer back to the full instructions, as participants exhibited significantly fewer glances toward the instruction panel when both cues were present. However, H3 is not supported, as the number of future steps shown did not significantly impact task completion time, indicating that precueing did not provide measurable performance benefits in the context of our study. Finally, H4 is not supported, as the results did not indicate a significant increase in cognitive load with the addition of more visual cues and step instructions, suggesting that the provided information did not overwhelm participants. Overall, our findings highlight the benefits of incorporating visual cues in complex procedural tasks.

5 Discussion

In this section, we discuss the results holistically along with implications for mixed reality applications.

5.1 Summary of Results

Despite the task complexity for the compound task requiring continual viewpoint changes, multiple types of subtasks, and compound instructions, the results show that both search cues and destination cues significantly improved task completion time. The results demonstrate the value of including supplemental interaction cues for subtasks within a more compound task, as are common in guidance scenarios (see Table 1).

Participants were able to successfully leverage the additional visual cues without interference with their ability to process the text and image instructions. This suggests that the presence of interaction cues are recommended in instructional and guidance applications despite the already high level of information presentation.

Search cues expedited identification of object locations, reducing the need for repetitive glances at instructions, while destination cues enhanced placement efficiency by providing target locations without the need for parsing instructions. While both types of cues caused significant improvements in task speed, Figure 4 shows a greater benefit from search cues for the studied task due to the importance of locating the target item from the spread of possible objects across the two virtual shelves. Performance benefits were similar in terms of task errors, where a significant interaction between search and destination cues indicate that the fewest errors occur when both cues were used together (see Figure 5). Search cues also significantly reduced errors on their own, though having destination cues on their own did not.

This result aligns with findings from previous research that studied more simplistic tasks (see Table 2). Volmer et al. [51] demonstrated that visual cues, such as lines, colors, blinking, and arrows, reduce task completion time compared to conditions without cues. In a subsequent study, where task difficulty was slightly increased, they observed similar outcomes, reaffirming that visual cues consistently enhance task performance over no-cue conditions [52]. Other studies [25, 26] reported analogous findings, showing that the presence of visual cues improves task efficiency and accuracy. Our study further extends these results to compound tasks with greater similarity to those found in mixed-reality guidance applications (see Figure 1), demonstrating that a combination of fundamental subtask cues can improve task completion time and reduces error rates for more involved procedures.

Interestingly, the precueing results of our study differed from prior studies without compound tasks and instructions (see Table 2). Our study did not detect evidence of benefits of precueing future steps in addition to the current step. Though the lack of significant finding does not indicate an impossibility to achieve successful precueing for compound procedures, the contrast with prior work (i.e., [25, 26, 51, 52]) demonstrate practical challenges in supporting precueing for more complex procedures—that is, those that require processing more information, continual changes to the viewpoint, and multiple subtask types.

The comparison of studies in Table 2 can help inform possible reasons for unsuccessful precueing in this study. Participants needed to parse textual instructions, consult images, continually scan the 3D surroundings, and frequently switch among subtasks during the steps of the procedure. We expect this complexity of the procedure to increase the base difficulty and cognitive demand beyond the tasks of prior precueing studies. Precueing on its own can already increase cognitive load, meaning that the addition of precues may be a detriment rather than an aid in compound tasks [45]. Complex instructions often require significant cognitive effort to process, understand, and retain [43]. When processing new information, if people are presented with too much information upfront, they may struggle to prioritize and integrate it effectively [43]. Thus, while intended to aid comprehension, precueing can inadvertently overload cognitive resources.

We see evidence of similar undesirable effects from our study. While the time results show relatively consistent task times for the different number of steps (see Figure 4). The error results showed a trend with errors increasing with additional steps shown (though not significant at $p = 0.051$).

The results also had a significant correlation indicating more errors when more steps were shown. This trend agrees with the notion that as the amount of presented step information increases, it also increases visual clutter and the risk of misinterpreting the cues. However, the self-reported mental workload ratings from the NASA TLX did not show the same relationship, which suggests participants likely did not consciously consider the addition of extra information to be a greater burden.

5.2 Implications for Design and Future Research

The data presented in Table 1 illustrates the intricate nature of real-world mixed reality task guidance systems, where participants engage in complex interactions with multiple objects, potentially employing various tools. Such tasks necessitate continuous adjustments to the participants' viewpoint. Conversely, Table 2 showcases previous precueing studies, which primarily concentrated on simpler, distinct interactions within the same viewpoint like button pressing, path following, etc., used the cue for that type of interactions rather than for a more involved complex set of instructions.

The lack of effective precueing in our study suggests it may be challenging to adapt existing knowledge of precueing to more complex or realistic procedures. Previous studies [25, 26, 51, 52] focused on more discrete tasks that could be completed within the same viewpoint or limited viewpoint switching. Importantly, these examples also did not require parsing additional compound information from instructions, as is desirable for many guidance and assembly applications [15]. As many real-world applications that seek to benefit from instructional task guidance involve more advanced tasks [28, 33], our research explores interaction cues for compound tasks, pushing the boundary for task guidance in more complex procedures.

Despite the results of our study providing strong evidence that the interaction cues for search and object placement are effective for improving task performance, participants did not benefit from seeing future instructions or the visual cues for upcoming steps. This raises questions about the effectiveness of precueing in real-world applications that involve compound tasks. Even in our task, complexity was bounded by the consistency of instruction format and subtask components for each step. Many real-world contexts will only increase in complexity and variability. As seen in prior AR and VR guidance applications, designers often hope to present instructions that include more information (e.g., [1, 20]). A limitation of our study (and perhaps any study of interaction cues) is effectiveness will depend on the specifics of the task and environment, which leads to uncertainty for generalization to other contexts. From the results of our study, we can hypothesize that the utility of information will decrease as more pieces of information require simultaneous attention and integration.

The results of our presented experiment do not, however, necessarily mean that it is impossible to precue future steps for more complex procedures. One possibility is the need for alternative interaction cues that are tailored to assist future guidance for different types of tasks. In our study, we opted to employ basic, commonly used interaction cues (numbers and visual markers) that could be easily understood with minimal practice. A review of prior literature also indicates that designers and researchers tend to adopt basic designs for interaction cues [1, 15, 46]. Especially in compound tasks, we suspect that visually minimalistic cues will be preferable to reduce visual clutter, distraction, and cognitive load while providing straightforward interpretation.

Nevertheless, empirical research continues to advance the field's knowledge of the effectiveness of different types of interaction cues for 3D interaction [6, 38]. For instance, Gruenefeld et al. [16] developed the EyeSee360 visualization for cueing users where to look within their 360-degree field of regard. This guidance cue uses latitudinal and longitudinal coordinates to convey what direction to look. Despite its complexity, the EyeSee360 cue had significantly higher accuracy than simpler techniques (e.g., 3D arrow or 2D wedge). For further research of precueing for compound tasks, it would be valuable to consider experimentation with alternative cues, particularly those that explore the design space recently introduced by Muresan et al. [30]. The results of our study motivate further exploration of precueing with diverse interaction cues for an expanded range of tasks.

6 Conclusion

Our study of precueing for compound multi-step tasks provides empirical results to suggest both search and destination cues had positive effects on task completion time and error rate. However, showing future steps did not help task performance for the compound task of the experiment, and the results caution that showing future steps may even increase the chance of errors for such procedures. The results of the presented study should be interpreted in the context of other studies of precueing and interaction cues in mixed reality, as the lack of successful precueing combined with the negative result of errors demonstrates the importance of continuing empirical studies of task guidance designs to more complex task and application contexts.

References

[1] Medhat Alaker, Greg R. Wynn, and Tan Arulampalam. 2016. Virtual reality training in laparoscopic surgery: A systematic review & meta-analysis. *International Journal of Surgery* 29 (2016), 85–94. https://doi.org/10.1016/j.ijsu.2016.03.034

[2] Andrea Albarelli, Augusto Celentano, Luca Cosmo, and Renato Marchi. 2015. On the Interplay between Data Overlay and Real-World Context Using See-through Displays. In *Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter* (Rome, Italy) *(CHItaly 2015)*. Association for Computing Machinery, New York, NY, USA, 58–65. https://doi.org/10.1145/2808435.2808455

[3] Markus Aleksy, Elina Vartiainen, Veronika Domova, and Martin Naedele. 2014. Augmented Reality for Improved Service Delivery. *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, 382–389. https://doi.org/10.1109/AINA.2014.146

[4] Ahmed Alnabhan and Brian Tomaszewski. 2014. INSAR: Indoor Navigation System Using Augmented Reality. In *Proceedings of the Sixth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness* (Dallas/Fort Worth, Texas) *(ISA '14)*. Association for Computing Machinery, New York, NY, USA, 36–43. https://doi.org/10.1145/2676528.2676535

[5] Huidong Bai, Gun Lee, and Mark Billinghurst. 2014. Using 3D Hand Gestures and Touch Input for Wearable AR Interaction. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 1321–1326. https://doi.org/10.1145/2559206.2581371

[6] Felix Bork, Christian Schnelzer, Ulrich Eck, and Nassir Navab. 2018. Towards Efficient Visual Guidance in Limited Field-of-View Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 24, 11 (2018), 2983–2992. https://doi.org/10.1109/TVCG.2018.2868584

[7] C. Chen, J. Hong, and S. Wang. 2014. Automated positioning of 3D virtual scene in AR-based assembly and disassembly guiding system. *The International Journal of Advanced Manufacturing Technology* 76 (02 2014). https://doi.org/10.1007/s00170-014-6321-6

[8] N. Cooper, F. Millela, I. Cant, MD. White, and G. Meyer. 2021. Transfer of training-Virtual reality training with augmented multisensory cues improves user experience during training and task performance in the real world. *PLoS One* 16, 3 (24 3 2021), e0248225. https://doi.org/10.1371/journal.pone.0248225

[9] Saverio Debernardis, Michele Fiorentino, Michele Gattullo, Giuseppe Monno, and Antonio Uva. 2014. Text Readability in Head-Worn Displays: Color and Style Optimization in Video vs. Optical See-Through Devices. *IEEE transactions on visualization and computer graphics* 20, 01 (2014), 125–39. https://doi.org/10.1109/TVCG.2013.86

[10] Kody R. Dillman, Terrance Tin Hoi Mok, Anthony Tang, Lora Oehlberg, and Alex Mitchell. 2018. A Visual Interaction Cue Framework from Video Game Environments for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173714

[11] Christoph Ebner, Shohei Mori, Peter Mohr, Yifan Peng, Dieter Schmalstieg, Gordon Wetzstein, and Denis Kalkofen. 2022. Video See-Through Mixed Reality with Focus Cues. *IEEE Transactions on Visualization and Computer Graphics* 28, 5 (2022), 2256–2266. https://doi.org/10.1109/TVCG.2022.3150504

[12] Lisa A Elkin, Matthew Kay, James J Higgins, and Jacob O Wobbrock. 2021. An aligned rank transform procedure for multifactor contrast tests. In *The 34th annual ACM symposium on user interface software and technology*. 754–768.

[13] Peter Fazekas and Bence Nanay. 2017. Pre-Cueing Effects: Attention or Mental Imagery? *Frontiers in Psychology* 8 (03 2017). https://doi.org/10.3389/fpsyg.2017.00222

[14] Michele Fiorentino, Antonio E. Uva, Michele Gattullo, Saverio Debernardis, and Giuseppe Monno. 2014. Augmented reality on large screen for interactive maintenance instructions. *Computers in Industry* 65, 2 (2014), 270–278. https://doi.org/10.1016/j.compind.2013.11.004

[15] Marco Gattullo, Andrea Evangelista, Antonio Emmanuele Uva, Michele Fiorentino, and Joseph L Gabbard. 2022. What, How, and Why are Visual Assets Used in Industrial Augmented Reality? A Systematic Review and Classification in Maintenance, Assembly, and Training (From 1997 to 2019). *IEEE Transactions on Visualization and Computer Graphics* 28, 2 (Feb 2022), 1443–1456. https://doi.org/10.1109/TVCG.2020.3014614 Epub 2021 Dec 30.

[16] Uwe Gruenefeld, Dag Ennenga, Abdallah El Ali, Wilko Heuten, and Susanne Boll. 2017. EyeSee360: Designing a Visualization Technique for out-of-View Objects in Head-Mounted Augmented Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (Brighton, United Kingdom) *(SUI '17)*. Association for Computing Machinery, New York, NY, USA, 109–118. https://doi.org/10.1145/3131277.3132175

[17] Steven Hackley. 2008. The speeding of voluntary reaction by a warning signal: Presidential Address, 2006. *Psychophysiology* 46 (10 2008), 225–33. https://doi.org/10.1111/j.1469-8986.2008.00716.x

[18] Sandra G. Hart. 1986. NASA Task Load Index (TLX). NASA Ames Research Center Moffett Field, CA United States. https://ntrs.nasa.gov/citations/20000021488

[19] Steven Henderson and Steven Feiner. 2011. Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. *IEEE Transactions on Visualization and Computer Graphics* 17, 10 (2011), 1355–1368. https://doi.org/10.1109/TVCG.2010.245

[20] Petr Hořejší. 2015. Augmented Reality System for Virtual Training of Parts Assembly. *Procedia Engineering* 100 (2015), 699–706. https://doi.org/10.1016/j.proeng.2015.01.422 25th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2014.

[21] Xinyu Hu, Alec Moore, James Coleman Eubanks, Afham Aiyaz, and Ryan P. McMahan. 2020. Evaluating Interaction Cue Purpose and Timing for Learning and Retaining Virtual Reality Training. In *Proceedings of the 2020 ACM Symposium on Spatial User Interaction* (Virtual Event, Canada) *(SUI '20)*. Association for Computing Machinery, New York, NY, USA, Article 5, 9 pages. https://doi.org/10.1145/3385959.3418448

[22] Kit Lam, Lik-Hang Lee, and Pan Hui. 2021. A2W: Context-Aware Recommendation System for Mobile Augmented Reality Web Browser. https://doi.org/10.1145/3474085.3475413

[23] Enricoandrea Laviola, Michele Gattullo, Vito Modesto Manghisi, Michele Fiorentino, and Antonio Emmanuele Uva. 2022. Minimal AR: visual asset optimization for the authoring of augmented reality work instructions in manufacturing. *International Journal of Advanced Manufacturing Technology* 119, 3-4 (2022), 1769–1784. https://doi.org/10.1007/s00170-021-08449-6 arXiv:2111.14920

[24] Jae Yeol Lee, Dongwoo Seo, and Gue Won Rhee. 2008. Visualization and interaction of pervasive services using context-aware augmented reality. *Expert Syst. Appl.* 35 (2008), 1873–1882. https://api.semanticscholar.org/CorpusID:20334786

[25] Jen-Shuo Liu, Carmine Elvezio, Barbara Tversky, and Steven Feiner. 2021. Using multi-level precueing to improve performance in path-following tasks in virtual reality. *IEEE Transactions on Visualization and Computer Graphics* 27, 11 (2021),

4311–4320.

[26] Jen-Shuo Liu, Barbara Tversky, and Steven Feiner. 2022. Precueing Object Placement and Orientation for Manual Tasks in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (11 2022), 3799–3809. https://doi.org/10.1109/TVCG.2022.3203111 arXiv:36049002

[27] G. F. Meyer, S. M. Wuerger, F. Röhrbein, and C. Zetzsche. 2005. Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Exp Brain Res* 166, 3-4 (10 2005), 538–547. https://doi.org/10.1007/s00221-005-2394-7

[28] Alec G. Moore, Xinyu Hu, James Coleman Eubanks, Afham Ahmed Aiyaz, and Ryan P. McMahan. 2020. A Formative Evaluation Methodology for VR Training Simulations. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 125–132. https://doi.org/10.1109/VRW50115.2020.00027

[29] Michela Mura, G. Dini, and Franco Failli. 2016. An Integrated Environment Based on Augmented Reality and Sensing Device for Manual Assembly Workstations. *Procedia CIRP* 41 (12 2016), 340–345. https://doi.org/10.1016/j.procir.2015.12.128

[30] Andreea Muresan, Jess Mcintosh, and Kasper Hornbæk. 2023. Using Feedforward to Reveal Interaction Possibilities in Virtual Reality. *ACM Trans. Comput.-Hum. Interact.* 30, 6, Article 82 (sep 2023), 47 pages. https://doi.org/10.1145/3603623

[31] J.J. Ockerman and A.R. Pritchett. 1998. Preliminary investigation of wearable computers for task guidance in aircraft inspection. In *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No.98EX215)*. 33–40. https://doi.org/10.1109/ISWC.1998.729527

[32] Jason Orlosky, Kiyoshi Kiyokawa, and Haruo Takemura. 2014. Managing Mobile Text in Head Mounted Displays: Studies on Visual Preference and Text Placement. *ACM SIGMOBILE Mobile Computing and Communications Review* 18 (06 2014), 20–31. https://doi.org/10.1145/2636242.2636246

[33] Dhaval Parmar, Jeffrey Bertrand, Sabarish V Babu, Kapil Madathil, Melissa Zelaya, Tianwei Wang, John Wagner, Anand K Gramopadhye, and Kristin Frady. 2016. A comparative evaluation of viewing metaphors on psychophysical skills education in an interactive virtual environment. *Virtual Reality* 20 (2016), 141–157.

[34] Nattapol Pathomaree and Siam Charoenseang. 2005. Augmented reality for skill transfer in assembly task. *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.* (2005), 500–504. https://api.semanticscholar.org/CorpusID:31766269

[35] Roberto Pierdicca, Emanuele Frontoni, Rama Pollini, Matteo Trani, and Lorenzo Verdini. 2017. The Use of Augmented Reality Glasses for the Application in Industry 4.0. 389–401. https://doi.org/10.1007/978-3-319-60922-5_30

[36] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-Defined Gestures for Augmented Reality. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 955–960. https://doi.org/10.1145/2468356.2468527

[37] Rafael Radkowski, Jordan Herrema, and James Oliver. 2015. Augmented Reality-Based Manual Assembly Support With Visual Features for Different Degrees of Difficulty. *International Journal of Human-Computer Interaction* 31 (01 2015). https://doi.org/10.1080/10447318.2014.994194

[38] Patrick Renner and Thies Pfeiffer. 2017. Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 186–194. https://doi.org/10.1109/3DUI.2017.7893338

[39] Mark Rice, Hong Huei Tay, Jamie Ng, Calvin Lim, Senthil Kumar Selvaraj, and Ellick Wu. 2016. Comparing Three Task Guidance Interfaces for Wire Harness Assembly. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 2279–2284. https://doi.org/10.1145/2851581.2892347

[40] David A. Rosenbaum. 1983. The Movement Precuing Technique: Assumptions, Applications, and Extensions. In *Memory and Control of Action*, Richard A. Magill (Ed.). Advances in Psychology, Vol. 12. North-Holland, 231–274. https://doi.org/10.1016/S0166-4115(08)61994-9

[41] Sajay Sadasivan, Joel Greenstein, Anand Gramopadhye, and Andrew Duchowski. 2005. Use of eye movements as feedforward training for a synthetic aircraft inspection task. *CHI 2005: Technology, Safety, Community: Conference Proceedings - Conference on Human Factors in Computing Systems*, 141–149. https://doi.org/10.1145/1054972.1054993

[42] Giulia Scurati, Michele Gattullo, Michele Fiorentino, Francesco Ferrise, Monica Bordegoni, and Antonio Uva. 2018. Converting maintenance actions into standard symbols for Augmented Reality applications in Industry 4.0. *Computers in Industry* 98 (02 2018). https://doi.org/10.1016/j.compind.2018.02.001

[43] John Sweller, Jeroen J. G. Van Merrienboer, and Fred Paas. 1998. Cognitive Architecture and Instructional Design. *Educational Psychology Review* 10 (09 1998), 251–. https://doi.org/10.1023/a:1022193728205

[44] Christina Trepkowski, Alexander Marquardt, Tom Eibich, Yusuke Shikanai, Jens Maiero, Kiyoshi Kiyokawa, Ernst Kruijff, Johannes Schöning, and Peter König. 2021. Multisensory Proximity and Transition Cues for Improving Target Awareness in Narrow Field of View Augmented Reality Displays. *IEEE Transactions on Visualization and Computer Graphics* PP (09 2021), 1–1. https://doi.org/10.1109/TVCG.2021.3116673

[45] Junichi Tsurukawa, Mohammed Al-Sada, and Tatsuo Nakajima. 2015. Filtering visual information for reducing visual cognitive load *(UbiComp/ISWC'15 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 33–36. https://doi.org/10.1145/2800835.2800852

[46] Konstantinos Koumaditis Unnikrishnan Radhakrishnan and Francesco Chinello. 2021. A systematic review of immersive virtual reality for industrial skills training. *Behaviour & Information Technology* 40, 12 (2021), 1310–1339. https://doi.org/10.1080/0144929X.2021.1954693 arXiv:https://doi.org/10.1080/0144929X.2021.1954693

[47] Antonio Uva, Michele Gattullo, Vito Manghisi, Daniele Spagnulo, Giuseppe Leonardo Cascella, and Michele Fiorentino. 2018. Evaluating the effectiveness of spatial augmented reality in smart manufacturing: a solution for manual working stations. *The International Journal of Advanced Manufacturing Technology* 94 (01 2018). https://doi.org/10.1007/s00170-017-0846-4

[48] I. Vermeir and G. Roose. 2020. Visual Design Cues Impacting Food Choice: A Review and Future Research Agenda. *Foods* 9, 10 (19 10 2020), 1495. https://doi.org/10.3390/foods9101495

[49] Giuseppe Vignali, Massimo Bertolini, Eleonora Bottani, Luciano Donato, Alessandra Ferraro, and Francesco Longo. 2018. Design and Testing of an Augmented Reality Solution to Enhance Operator Safety in the Food Industry. *International Journal of Food Engineering* 14 (02 2018). https://doi.org/10.1515/ijfe-2017-0122

[50] Benjamin Volmer, James Baumeister, Raymond Matthews, Linda Grosser, Stewart Von Itzstein, Siobhan Banks, and Bruce H Thomas. 2022. A comparison of spatial augmented reality predictive cues and their effects on sleep deprived users. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 589–598.

[51] Benjamin Volmer, James Baumeister, Stewart Von Itzstein, Ina Bornkessel-Schlesewsky, Matthias Schlesewsky, Mark Billinghurst, and Bruce H. Thomas. 2018. A Comparison of Predictive Spatial Augmented Reality Cues for Procedural Tasks. *IEEE Transactions on Visualization and Computer Graphics* 24, 11 (2018), 2846–2856. https://doi.org/10.1109/TVCG.2018.2868587

[52] Benjamin Volmer, Jen-Shuo Liu, Brandon Matthews, Ina Bornkessel-Schlesewsky, Steven Feiner, and Bruce H Thomas. 2023. Multi-Level Precues for Guiding Tasks Within and Between Workspaces in Spatial Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 29, 11 (2023), 4449–4459.

[53] Sabine Webel, Timo Engelke, Nirit Gavish, Manuel Olbrich, and Carsten Preusche. 2013. Augmented Reality Training Platform for Assembly and Maintenance Skills. *Robotics and Autonomous Systems* 61 (04 2013), 398–403. https://doi.org/10.1016/j.robot.2012.09.013

[54] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[55] Julia Woodward, Jesse Smith, Isaac Wang, Sofia Cuenca, and Jaime Ruiz. 2020. Examining the Presentation of Information in Augmented Reality Headsets for Situational Awareness. In *Proceedings of the 2020 International Conference on Advanced Visual Interfaces* (Salerno, Italy) *(AVI '20)*. Association for Computing Machinery, New York, NY, USA, Article 57, 5 pages. https://doi.org/10.1145/3399715.3399846

[56] Jianwei Yang, Fu Liu, Jinhai Wang, Ziming Kou, Aihua Zhu, and Dechen Yao. 2020. Effect of virtual reality technology on the teaching of urban railway vehicle engineering. *Computer Applications in Engineering Education* 29 (12 2020). https://doi.org/10.1002/cae.22371

[57] M. L. Yuan, S. K. Ong, and A. Y. C. Nee. 2008. Augmented reality for assembly guidance using a virtual interactive tool. *International Journal of Production Research* 46, 7 (2008), 1745–1767. https://doi.org/10.1080/00207540600972935 arXiv:https://doi.org/10.1080/00207540600972935

[58] Dong Zhao and Jason Lucas. 2015. Virtual reality simulation for construction safety promotion. *International journal of injury control and safety promotion* 22 (01 2015), 57–67. https://doi.org/10.1080/17457300.2013.861853

[59] Xiaoyu Zhuang and Thomas V Papathomas. 2011. Cue relevance effects in conjunctive visual search: cueing for location, color, and orientation. *Journal of Vision* 11, 7 (2011), 6.