Conveyor: A Dual-Task Paradigm for Studying VR Dialogue Interfaces

Patrick Dubois[1], Daniel J. Rea[1], Kevin Hoang[1], Meghan Chua[1], Danielle King[2], Corey King[2], James E. Young[1], Andrea Bunt[1]

[1]Department of Computer Science, University of Manitoba; [2]ZenFri

Abstract

VR applications can enhance players' sense of presence within virtual environments. A common scenario is to have a player working on a task, while simultaneously making dialogue selections using VR. We investigate a dual-task experiment design for this scenario, and an initial study using four interfaces for dialogue selection. We found that interface naturalness – one measure of immersion – seems to have a large role in player preference, regardless of selection speed.

Keywords: Virtual reality, dialogue selection.

1 Introduction

One compelling aspect of Virtual Reality (VR) is the immersion factor, where users can experience an increased sense of being present in the virtual world. Immersion has a number of potential user benefits including better emotional responses [7], skill retention [6], engagement [4], or enjoyment [6].

In some VR applications, like games [4], collaborative environments [5], or education tools [6], users can be performing a task in the virtual world while simultaneously conversing with a virtual agent. A challenge, however, is engaging the user in these dialogues without negatively impacting their immersion in the primary task. For example, traditional dialogue-selection interfaces may require diverting a user's attention from their main task in the virtual world. They might also involve input methods that are being used to complete main tasks, increasing cognitive demands and breaking immersion.

In this work, we investigate different selection mechanisms for dialogue-based interaction and their impact on tasks completed in a virtual environment. Techniques can leverage potentially intuitive 3D interactions, such as pointing, or gaze based interfaces, but can be challenging due to the imprecise nature of human movements [2,3]. Traditional menu and selection designs often result in equal or better performance to 3D designs, but they may break task immersion [1,2]. As our interest lies in both immersion, and performance for dual-task VR situations, we explore both natural, 3D interactions, as well as more traditional selection techniques. As a first step, we investigate a dual-task experiment design for VR with flexible difficulty, and conduct an initial exploration of conversation-picking interfaces in our dual-task VR environment.

2 Dual-Task Design for Conversation Interfaces in VR

Our dual-task experiment design had two main goals: 1) to provide a main task that is difficult and engaging enough such that the secondary, conversational task could impact both task performance and engagement, and 2) to have a conversation where user engagement is important, so users take care picking their dialogue.

{patrick.dubois, daniel.rea, young, bunt}@cs.umanitoba.ca,
{hoangk, umchua2}@myumanitoba.ca,
{danielle, corey}@zenfri.com

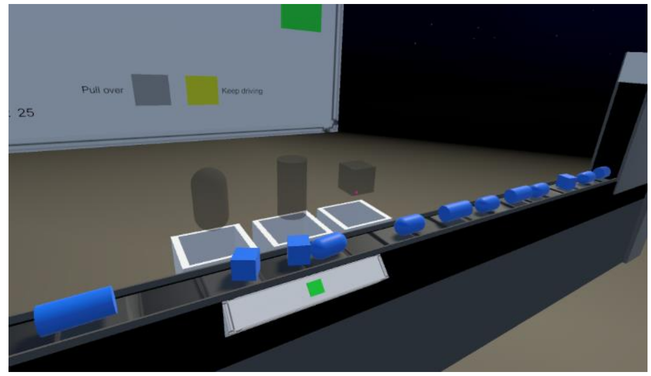

Figure 1. Our dual-task environment. Users sorted objects on a conveyor belt into bins and progressed through a story by selecting options displayed on the virtual wall.

For our main task, we explored many options and decided on a repeated, simple, and discrete task where we could scale difficulty: a sorting task. Users pick up objects from a conveyor belt, and place objects into different bins (Figure 1). Objects to be sorted are produced at a set rate and have a limited time to be sorted before they are removed at the other end of the conveyor belt. We note that the moving targets require hand-eye coordination, which may be difficult with the potentially detracting conversation task.

Our conversation task was a "choose-your-own adventure," where users' answers would lead them down different paths through a story. A narrator read part of the story, then the user then had to pick an option from a set of choices within a time limit.

Our dual-task design can be tuned in a number of ways to achieve a desired level of difficulty. We can change the number of categories to be sorted, the complexity of sort (e.g. red AND square objects in one bin), the production rate of objects, and time per object to be sorted. The moving speed and size of objects can be changed to increase the difficulty of picking objects up. The secondary task can also modulate difficulty, by having a limit on how long a user can take to make a choice, how many options are presented, and the complexity of the dialogue and choices.

3 Investigation of Dialogue-Choosing Interfaces

We ran a within-participant initial study with twelve participants (eight male) from a university community to investigate how four dialogue-choosing methods affected a user's immersion and performance within a virtual environment. Our users used each interface in a counter-balanced order.

3.1 Dialogue-selecting interfaces

Participants tested four dialogue-selection interfaces, selected through pilot testing. They were chosen to cover different levels of immersion using a range of common VR selection techniques:

Touchpad: The controller's circular trackpad selects a conversation option displayed in an analogous position (e.g. left on the pad chooses the left option).

Trigger: Virtual switches are placed in front of the user, and are activated by pressing a button.
Head Gaze: Gaze is used to point to a conversation option, with a controller button to confirm selection.
Voice Recognition: Options are read aloud. This was implemented with Wizard of Oz in this initial exploration.

3.2 Dual-task implementation

Participants completed four choose-your-own adventure stories, each with a different selection technique. Each story was approximate five minutes long, and was created by a professional video game story writer. Each story was narrated by a different person: two female, two male. Dialogue options were displayed to the user as text on a virtual board when the stories branched (Figure 1). The user would use an assigned technique to select an option.

Users were seated in front of the conveyor belt and could move both hands, which held the hand-tracking HTC Vive controllers, and their head, wearing the HTC Vive headset. Players had to pick up objects as they passed, by intersecting the virtual controller with an object and holding a button, moving the object to the correct bin, and dropping the object. Objects disappeared if they reached the end of the conveyor belt or fell to the floor.

3.3 Measurements

Our primary performance measures were: time to select conversation options, time to complete each story, number of sorting errors, and number of correct answers to post-task comprehension questions. Immersion was measured via a standard VR immersion questionnaire [4], which we modified to fit our specific VR application.

3.4 Procedure

Participants were introduced to the sorting task, the equipment, and the dialogue selections. This was followed by a practice sorting task, where participants could sort shapes for two minutes, without needing to select dialogue.

Each selection interface was explained before it was used, followed by another two-minute practice session with both tasks. This was followed by the real task, then the story's comprehension question, and post-condition questionnaire.

At the end, participants filled in both a demographics and overall experience questionnaire, where they ranked the interfaces in their order of preference and gave general feedback.

4 Results

A repeated-measures ANOVA (Figure 2) found a significant effects interface type on self-reported story engagement (F=3.6, $p<.05$, $\eta^2=.25$). Post-hoc tests, however, were not statistically significant. We also found a main effect of interface on how natural it felt to select a dialog option (F=4.5, $p<.01$, $\eta^2=.29$). Post-hoc tests found a trend for Voice to feel more natural than Touchpad ($p=.07$, mean difference of 1.4 ranks, 95% CI [-.09, 2.9] ranks) and Gaze ($p<.05$, mean difference of 2.1 ranks, 95% CI [.13, 4.0] ranks).

For performance metrics, we found a main effect of interface on dialogue selection time (F=30.7, $p<.001$, $\eta^2=.74$). Post-hoc tests found Voice slower than all interfaces, ($p<.001$, mean differences of 1.4s, 1s, and 2s. 95% CI of [-.09s, 2.9s], [-.42s, 2.4s], and [.13s, 4.0s] for Touchpad, Trigger, and Gaze respectively). All other tests were non-significant, and we observed large variances.

For preference data, Voice had seven favourite rankings, compared to three for trigger, and two for touchpad. No participant ranked Gaze as best, and no participant ranked Voice as worst.

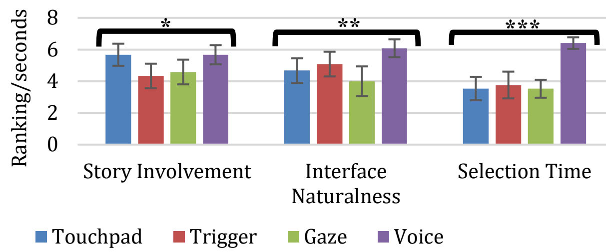

Figure 2. Significant results. Time to select is in seconds; others are rank. **$p<.01$, *$p<.05$, ***$p<.001$. Errors bars are 95% CI.

5 Discussion and Recommendations

Our results found higher ratings for involvement and naturalness for Voice, lining up with our interface-ranking data. This may not come as a surprise: users did not need to use any specialized controls to make their selections, and could keep sorting using both hands. However, it was the slowest to make selections, as Voice takes time to correctly read the option. This may imply that speed of entry was not considered important by participants in our task.

Gaze was also surprising: while users had the fastest selections, they rated it poorly in terms of preference, story involvement, and naturalness. It is possible that this is due to a "fat-finger problem" as our targets may have been small when compared to the average head jitter. Gaze may also be innapropriate in dual-task situations, as users' visual attention was needed for the main task.

We saw a lot of variability in performance data, implying that our sorting task might need refinement. We selected a sorting task because there are many parameters we could manipulate to control difficulty. However, despite extensive piloting, our final configuration seemed too easy for our users. This implies it is non-trivial to find a difficulty that requires some degree of the participant's attention, but not to the point they could not pay attention to the story, or feel discouraged. Methods to calibrate this difficulty is an open challenge.

6 Conclusion

We presented an initial testbed for a dual-task VR application, where a player needs to complete a primary task while making dialogue selections. We used this environment to investigate immersion using four different dialogue-selection methods, and found that study participants seem to prefer methods that allow greater immersion than efficiency. We plan to refine our scenarios to improve our understanding how selection interface impacts player immersion and task efficiency.

References

[1] D.A. Bowman et al. 2001. Design and evaluation of menu systems for immersive virtual environments. IEEE Virtual Reality: 149–156.
[2] J.C. Coelho et al. 2014. Pointing Task Evaluation of Leap Motion Controller in 3D Virtual Environment. Creating the Difference: 78–85.
[3] R. Dachselt et al. 2007. Three-dimensional menus: A survey and taxonomy. Computers & Graphics 31, 1: 53–65.
[4] C. Jennett et al. 2008. Measuring and defining the experience of immersion in games. Human Computer Studies 66, 9: 641–661.
[5] P. Koutsabasis, et al. 2012. On the value of Virtual Worlds for collaborative design. Design Studies 33, 4: 357–390.
[6] W. Schoor. 2006. VR Based Knowledge Transfer in Medical Technology and Techniques. Computational Science and Its Applications: 268–277.
[7] M.J. Schuemie, et al.. 2001. Research on Presence in Virtual Reality: A Survey. CyberPsychology & Behavior 4, 2: 183–201.