Universiteit Utrecht

[Faculty of Science]
Information and
Computing Sciences

Opening Up The Box

A Comprehensive Evaluation of Shoebox Virtual Reality and its Function in Open Environments

Marco van Laar
ICA-3652246

March 19th 2014
Wolfgang Hürst

Table of Contents

1. Introduction

In the past decades the usage of 3D graphics and interactive worlds has led to significant milestones such as photo realistic CGI, lifelike simulation and fluid animation. Aside from increasing entertainment value, high fidelity graphics and simulations have been proven to result in better immersion and sense of presence in a virtual world [1]. At the core of this fidelity lies how humans interpret visual information; the more it resembles what we see in the real world, the more it is qualified as realistic. The fidelity, or degree of realism, of three dimensional graphics can be attributed to several key properties of an image called visual cues. The presence or absence of these cues determine whether our brain will think the image looks realistic, even when displayed on a 2D screen with limited color space or a piece of paper.

Visual cues can be divided into categories that signify their function in, or relevance to our interpretation of the image. For example, Cutting and Vishton (1995) had attempted to isolate and identify the most essential cues and split them into three separate categories [1]. The primary cues categorize the cues that create true depth in an image including but not limited to binocular disparity and vergence. The second category, pictorial cues, contain cues such as occlusion, relative size and density, among others. Their final category describes motion cues including motion parallax as well as perspective. Interestingly, modern desktops typically cannot display graphics that incorporate the primary cues due to the 2D screens that are used. Although stereoscopy has aimed to overcome this limitation it suffers from being inconsistent in its application due to the many different implementations (e.g. active vs passive) and has only become somewhat available for consumers over the last few years. But even in a static 2D image pictorial cues relay us spatial information about the environment allowing us to estimate size, distance, shape and other properties using pictorial cues (Figure 1).

Figure 1: Pictorial cues allow us to extract spatial information from an otherwise flat image.

http://3dfiction.com

To make up for the lack of true depth on 2D displays several new systems and techniques were developed, the most popular being the Fish-Tank and CAVE systems [2][3]. The Fish-Tank system (Figure 2) uses a camera that tracks the user's head and uses this position to adapts the perspective on the display so that it looks correct. This creates an interesting effect that does not emulate true 3D depth, but enhances the sense of three-dimensionality as a result of being able to look “into” the 3D virtual world [2][3]. The CAVE system (Figure 3) used a different approach in that it projected its graphics on the walls and floor of a room. In a specific study on task efficiency, users reported the CAVE system being favorable on certain accounts such as having a wider field of view and being able to walk around, but in general preferred the Fish-Tank approach because of having sharper images and generally being more efficient in its use [2]. Regardless, both systems introduced a relatively new method of interacting with a virtual 3D environment that increased the need for user participation, thus enhancing the experience.

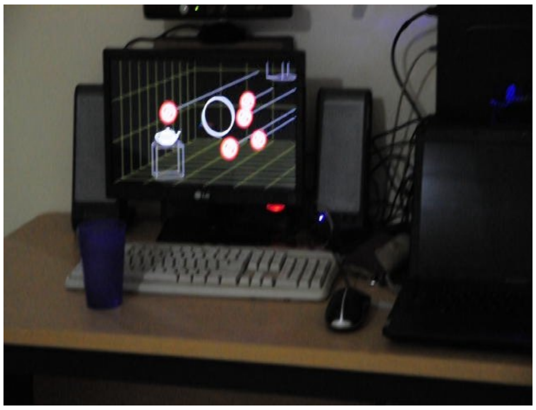


Figure 2: Fish-Tank for desktop uses a camera to track the user's head allowing the user to explore the virtual environment to a degree by changing perspective.

http://vimeo.com/51371465

Figure 3: The CAVE virtual environment system uses a projector to display images on the walls of a room. The world can be explored by looking around.

http://en.wikipedia.org/wiki/Cave_automatic_virtual_environment

Both Fish-Tank and CAVE were developed more than a decade ago, but more recent implementations prove visualization techniques are relevant still. For instance, the pCubee is a cuboid device with displays attached to each of its six sides that uses head-coupled perspective to create the illusion of it being a real box [5] (Figure 4). Another example is Tangible Windows that uses a large display in the form of a table in combination with head-coupled perspective to allow for the intuitive viewing and manipulation of 3D data [6] (Figure 5).

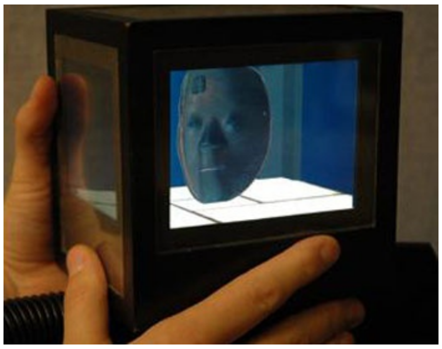


Figure 4: The pCubee by Staveness et al. (2009).

http://hct.ece.ubc.ca/cubee/

Figure 5: Tangible Windows by Spindler et al. (2012).

http://vimeo.com/51679775

In spite of such successes, implementations that require uncommon hardware are generally not accessible enough for the general public. Furthermore, each system has different costs and overhead that may form an obstacle for development [4]. However, even by just using a tablet or smartphone, the concept of Fish-Tank on mobiles shows potential as demonstrated by for example Li et al. (2012) [4] and Francone et al. (2011)[7]. In spite of this potential, their implementations are not without flaws. The main issue can be attributed to use of the device's camera. In contrast to desktop Fish-Tank implementations, the distance between the user and the camera is much smaller, creating a narrow field of view. This severely limits the angle from which the device can be viewed whilst still applying the correction. As even today not many mobile devices utilize cameras supporting a wide view angle, an alternative is desirable.

One advantage of mobile devices over desktops and other static approaches is their ability to rotate, or tilt. Whereas Fish-Tank traditionally fixes the screen in place and allows the user to move his or her head to look into the 3D world, a different approach would be to “fix” the user and move the device the instead. The result is an implementation referred to by Hürst et al. (2011) as the Shoebox VR model (Figure 6) [8]. Similar to the “Fish-Tank”, the term “Shoebox” is derived from a comparison to the real world. In case of the Fish-Tank, the user looks into the virtual world similar to how they would look into a fish tank e.g. by moving around while the fish tank stays fixed. In case of the Shoebox VR, the user looks into the virtual world similar to how they would into a shoebox diorama (Figure 7) e.g. by moving the box. In other words, due to the differences in size and weight compared to a large object like a fish tank, it makes more sense to rotate the box to change perspective because it's much easier and intuitive. The same applies to mobile devices in that tilting and moving it makes more sense than looking at it from different angles while holding it still.

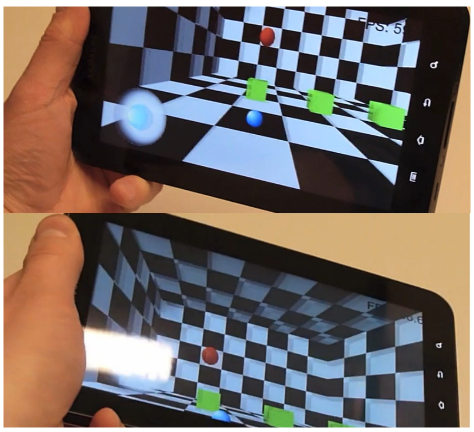


Figure 6: Shoebox VR: tilting the device adapts the perspective, revealing more of the virtual world while maintaining correct proportions.

http://vimeo.com/album/55295/video/19596206

Figure 7: The shoebox diorama analogy. The display represents the (fixed) “peep-hole” into the virtual world, allowing the world to be explored from different perspectives by moving the box.

http://www.californiapapergoods.com/shoe_box.shtml

There are many implementations that incorporate a perspective correction mechanic as is evident by the various systems mentioned in this chapter [2][3][4][5][6][7][8]. However, in spite of this, relatively little is known about the actual benefits and restrictions the head-coupled perspective mechanic imposes on practical use, especially on mobile devices. As such, we are primarily interested in determining the factors and variables that affect the experience of using the Shoebox VR in combination with a 3D application. We have conducted research of our own on topics such as interaction, depth perception and immersion in combination with the Shoebox VR for this purpose, but have not yet been able to contribute definitive results. We believe this may be attributed due to our conceptualization of the Shoebox VR model in that many of our studies have modeled the virtual environments as if they were inside an actual box or a similarly restricted space. Because such controlled work spaces make it easy to demonstrate the effects and the benefits of head-coupled perspective implementations, it raises the question if these environments are suitable for such experiments and studies. Therefore, with this research we intentionally "open the box" and step away from small spaces and movement restrictions. No longer do we consider the virtual environment to be "boxed off" by visible barriers. Instead, we allow the user to move about freely in the space along a 2D plane that acts as a floor to allow for active exploration.

With this freedom of movement comes user choice and responsibility in the form of navigation. Navigation encompasses many aspects of being present in, and interacting with a virtual environment including exploration and wayfinding. In order to navigate properly an individual must establish an adequate spatial comprehension i.e. understanding the physical structure of the environment [9][11][12][15]. With visualization techniques such as the Fish-Tank and Shoebox VR this spatial comprehension could possibly be improved as the studies of Arthur et al. (1993), Francone et al. (2011) and Li et al. (2012) indicate that users subjectively experienced an increased sense of 3D when using head-coupled perspective [3][4][7]. Whether or not their understanding of the 3D space is actually improved by the head-coupled perspective remains to be seen, but there is little reason to believe that head-coupled perspective cannot provide practical benefits to the exploration and interaction of and in 3D spaces. To confirm this and to better understand how head-coupled perspective can benefit exploration and interaction with a virtual 3D environment this study also includes a theoretical analysis on the mechanics and effects that make up head-coupled perspective visualization. This analysis will be featured in chapter 3 and can be considered a significant contribution towards fully understanding head-coupled perspective implementations and considerations.

Reviewing the scenarios in this chapter, head-coupled perspective shows both promising and limiting factors in its application. We will explore and evaluate these factors theoretically and by means of a user study to determine if head-coupled perspective implementations such as Shoebox VR are relevant for navigation in open 3D virtual environments. By doing so, we will gain a more complete understanding of how use of Shoebox VR affects a user's performance and experience in 3D applications.

Definition of Terms

Term	Definition
Navigation	The act of actively traversing an environment for the purpose of arriving at a particular place or finding a particular item with minimal delays.
Exploration	The act of actively traversing an environment for the purpose of learning its layout and structure as well as its key features.
Head-coupled Perspective	Tracking or approximating the position of the user so that displays, when viewed from an angle, show a natural and proportionally correct perspective.
Shoebox VR (Virtual Reality)	Head-coupled perspective implemented on mobile devices using orientation information acquired from the orientation sensors of the device such as the accelerometer, compass and gyroscope.
Mobile Device	A touch-enabled smartphone or tablet no larger than 10.1 inch in diameter outfitted with suitable orientation sensors and a display of at least 800 by 480 pixels in resolution capable of rendering and displaying 3D graphics at a minimum rate of 30hz.

2. Related Work

The topic of navigation and motion has been and is still being researched extensively due its relevance in many of our daily tasks. This naturally also includes research on exploration and navigation within 3D virtual worlds. However, there is little work that relates it to the use of head-coupled perspective, indicating it is still largely unknown what the effects of head-coupled perspective have on the general behavior and efficiency of movement and path planning (wayfinding) in virtual environments.

There are studies that explore the head-coupled perspective in combination with a certain system or model such as the pCubee [5], Tangible Windows [6] and CAVE systems [2]. In this regard, Fish-tank [3] is arguably the most important, as it the first head-coupled perspective implementation designed for computer desktops and is essentially what Shoebox VR and other implementations are derived or inspired from [3][4][5][6][7][8]. However, in spite of head-coupled perspective being available on more accessible platforms for a while now, there isn't much related work that goes into understanding the effects and benefits of its use in typical 3D applications, as most of the works only describe and discuss the technical realization and often disregard practical usage or useability issues. There are exceptions of course, for example Teather et al. (2008) performed a study on how exaggerating the head-coupled perspective affected users in performing several 3D positioning tasks [16]. Work by Li et al. (2012) used head-coupled perspective in combination with stereoscopy and showed that such advanced visualization techniques can improve depth perception in 3D applications [4]. Arguably the most relevant study is a work by Kulshreshth et al. (2013), who did a more elaborate research on the advantages of head-tracking in video games. Interestingly, the results showed that expert users (experienced gamers) did benefit from the head-tracking in two of the four commercial games (Arma II and Dirt 2, a sophisticated shooter and a racing game respectively), but it was never determined how or why the addition of head-coupled could contribute to a better performance or experience.

Works by Francone et al. (2011) and Hürst et al. (2011) have successfully demonstrated that head-coupled perspective can be applied to mobile platforms as well [7][8]. The gap between desktop and mobile platforms is not that large in terms of usability, but the difference in interaction and experience does suggest that results are not transferable directly. This argument is further supported by the differences in apparel and control between mobile apps and desktop programs. Although studying the differences in these platforms is not the focus of our study, we must understand how they correlate in order to meaningfully relate their results and findings to our own research. Because of this, a section of the theoretical analysis in chapter 3 will be dedicated to identifying key differences between platforms and what effects they may have on the usage and experience of head-coupled perspective implementations.

Finally, there are many studies on navigation and navigational aids in general [9][11][13][14][15], but none have combined it with head-coupled perspective. They instead aim to provide visual or audible cues to guide the user in the desired direction. Other works focus on factors such as perspective and camera control on navigation, including the works by Burelli (2011) and Ropinski (2005) that explore the possibilities of adaptive camera control and on-rails navigation respectively [20][21]. Studies by Salamin et al. (2006 & 2008), Bateman et al. (2011) and Claypool et al. (2009) tackle point-of-view related problems and solutions [22][23][24][25] which are relevant to our study because head-coupled perspective is not a replacement for traditional control schemes and instead should be considered more of an extension to perspective projection techniques. As such, understanding the effects of view and perspective on user performance and behavior are crucial for determining general cases where head-coupled perspective can be considered beneficial as well as identifying important factors and variables for the user study.

Our own studies on the Shoebox VR model have thus far tackled specific problems such as the issue of using the touch display while tilting the device, designing 3D user interfaces that can take advantage of Shoebox VR, the effect of high-speed object movement on use of the Shoebox VR by users and studying the effects of Shoebox VR on depth perception. All of these topics relate to our problem in some form or way as they include factors such as general usage of Shoebox VR in 3D spaces and the effects of motion, but navigation or free movement in open spaces has not been the topic of a study yet. Due to the lack of a solid theoretical foundation, we begin by describing and specifying known factors and variables that are relevant to the setup and implementation of a head-coupled perspective oriented study in chapter 3. We include the potential benefits and issues related to our topic i.e. navigation and open environments, and use the analysis as the foundation to initiate the first step towards exploring our problem by phrasing our research questions and hypotheses in chapter 4.

3. Theoretical Analysis

This chapter analyzes the head-coupled perspective model to contribute a complete overview over its technical details and important considerations to make when conducting a head-coupled perspective oriented study. We lack a solid and complete theoretical foundation for understanding head-coupled perspective and thus use this opportunity to explore and document all that we have learned from previous studies and experiences. By doing so, we contribute to the completeness of the theoretical and technical comprehension of the head-coupled perspective model and its implementations, so that future research can benefit from referencing and extending the theory as more detailed information becomes available.

We begin by detailing the technical implementation of Shoebox VR, followed by the practical uses and limitations that arise from use of head-coupled perspective and an analysis on the suitability of a variety of environments based on key attributes. Finally, we explore the possibilities and limitations of head-coupled perspective when combined with free movement in open spaces, to elaborate on our choice of topic for this experiment. Before we begin, we present a glossary of several keywords and concepts that will be used throughout this chapter:

Term	Definition
Perspective Projection	In 3D math, projection refers to the mapping of 3D points to a 2D plane. i.e. compressing a 3D space into a 2D one by eliminating one of its dimensions. Perspective projection is a specific approach to projection that emulates the human eye so that objects in the distance appear smaller and closer to the horizon than nearby objects. Projection is required for graphics because displays cannot (yet?) present true 3D space, only flat 2D images.
Off-axis Projection	Normally, projection uses a symmetrical view frustum so that the proportions are projected correctly onto the 2D image, preventing the image from becoming distorted. Off-axis projection moves the view frustum relative to its origin, creating a sheared frustum. As a result, the image appears stretched out, making it unsuitable for most typical cases where projection is used or required, but there are exceptions where its use can make up for technical or practical issues and limitations, such as when viewing a projected image from an angle on a 2D surface.
Perspective View	A point of view from where the virtual environment or world is observed. In some cases this view is directly associated with a virtual character. For example, a first person perspective view is used to look through the eyes of the virtual character and a third person perspective view keeps the camera focused on the virtual character from a distance. An overhead or isometric perspective view may also focus on a virtual character, but is typically used to observe large areas e.g. for strategic purposes.
Open and Closed Spaces	A closed space is a space that is visibly bounded by walls or other obstructions. An open space removes some or all of these boundaries or is of such a size that they are no longer noticeable and as such remove the sense of being "trapped". Colors are also known to have an effect on the perception of space, where darker colors are generally considered to make a space look and feel less expansive, whereas lighter colors may help create the illusion of a more spacious room. A visible sky or horizon are also known to establish a sense of freedom and space.

3.1 Technical Details

Although implementations of head-coupled perspective tend to vary to some degree, they generally implement it by modifying the perspective projection. Perspective projection is a method used to render images on the screen by simulating how the human eye works in the real world and uses a virtual camera to "see" or capture an area in the 3D space in order to convert it into a 2D image that can be displayed on a monitor or screen. This area is referred to as the view frustum and is created by specifying the six planes that bound it e.g. four planes are used to describe the vertical and horizontal field of view and the latter two limit how near and far objects are allowed to be rendered from (Figure 8).

Typically, the visible area, dubbed the view frustum, is symmetrical both horizontally and vertically. The actual width and height are determined based on the dimensions of the native resolution the display. The ratio of width to height is referred to as the aspect ratio. For example, an aspect ratio of 1.5 indicates that the width of the resolution is 1.5 times larger than its height. The view frustum assumes the same ratio of width to height as the resolution in order create an image that is proportionally correct and not distorted. Unless changing resolutions, the aspect ratio is fixed once it's been determined. The field of view is generally also determined beforehand, although it is more susceptible to change due to several factors such as the distance between the user and the display as well as preference. The virtual camera is moved around and rotated in order to explore the 3D space from different perspectives.

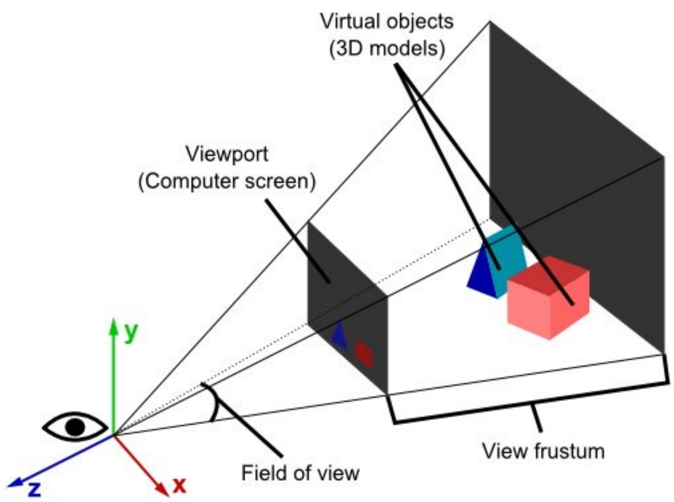


Figure 8: Perspective projection in 3D graphics.

http://www.real3dtutorials.com

Because the field of view and aspect ratio determine the general shape of the view frustum it is uncommon for it to change drastically throughout the course of an application. The primary cause for the lack of change is the assumption that the user and display are expected to remain in the same position with relation to each other. However, this assumption may be wrong e.g. when the user is not sitting or standing right in front of the display, but instead is viewing it from an angle. In order to create a correct perspective, the view angle to the display must be taking into account. Head-coupled perspective is an extension of perspective projection that acknowledges that this assumption may be wrong. By using varying methods of determining or approximating the position of the user relative to the display, head-coupled perspective implementations transform the view frustum and position of the virtual camera to best simulate the proper conditions.

Different methods have been used to acquire the user's relative position required to correct the perspective. Fish-Tank and other desktop based implementations often use camera tracking or variations thereof [2][4][8][29]. By fixing the position of this camera and using a face tracking algorithm to determine the position of the user's head in the captured image, it is possible to derive the view angle at which the user is viewing the screen. Several researchers have successfully made the transition to mobile devices by utilizing the camera on mobile devices. However, due to the distance between the user and the display being much smaller than a typical desktop scenario, the camera would need to support a much wider angle lens. Additionally, front cameras may vary in their position on the device which can further negatively affect the field of view if it is positioned on the side rather than the top. The Shoebox VR model uses a different method that uses hardware components such as the accelerometer, compass or gyroscope to get a reading of the device's orientation angles. If we assume the user holds the device in place whilst rotating it, the angles from the orientation sensors directly describe the viewing angle and thus require no additional computation. Furthermore, the assumption that the user will hold the device in place is a fairly safe one in that the typical usage of a mobile device does not include moving it in any significant way unless instructed to do so. However, it is not as safe as assuming the display will not move or rotate in a desktop environment and as such can be considered a disadvantage over desktop based head-coupled perspective models.

Regardless of the method used to determine the view angle, the technical implementation can be considered universal. Although theories on how the modifications to the perspective are to be made may vary, the general approach will employ the use of off-axis projection. Off-axis projection is an asymmetric variation of perspective projection that offsets the view frustum resulting in the frustum becoming skewed (Figure 9). This is used to correct the projection of the image when it is viewed from an angle and can often be seen on digital image projectors. The general concept is that the projector or camera must look at the image from the user's

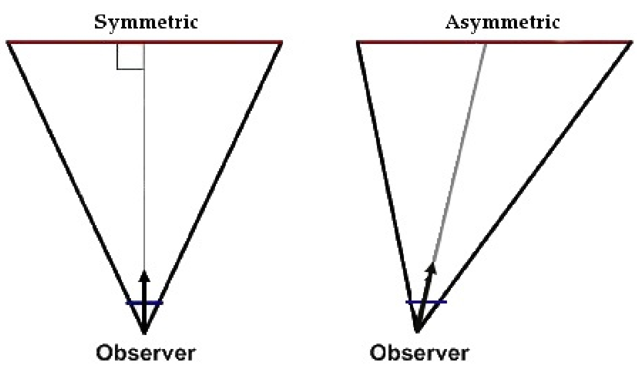


Figure 9: A top-view of the view frustums of symmetrical and asymmetrical projection methods. Off-axis projection is a form of asymmetrical projection that shifts the view frustum, causing the perspective to change in a way that transforms how the image is projected to a 2D surface.

Modified from http://www.vis-sim.com/vegaprime/faq_vp2.htm perspective and display that image rather than projecting it from its own position. Due to the misalignment of the user's point of view to the target 2D plane e.g. a wall or screen, the image is stretched to correct the proportions so that they look proper from the user's perspective. Because the screen surface of a display device cannot be extended, the horizontal field of view shrinks as a result of increasing the view angle. This implies a balance or limit to the extent to which the view angle may be increased before the benefits of the visual effect of off-axis projection is overshadowed by the increasingly smaller visible screen surface area. Although it is likely that this balance or threshold varies between scenarios, angles beyond 40 or 50 degrees should be considered a soft limit on practical use for general cases. Considering that this limitation cuts the effective use of Shoebox VR and other similar head-coupled perspective models in half, additional research in this area is required in order to fully understand the implications on technical and practical use.

In order to implement off-axis behavior in graphics libraries such as OpenGL, one must translate the frustum by redefining the left, right, bottom and top planes. However, in order to properly implement head-coupled perspective, one must also match the user's point of view by translating the camera orthogonally to the camera's eye vector (direction it is looking in). The reason why translation is used and not rotation is due to the fact that translating the camera does not affect the relative z-axis of the camera. To elaborate, if we were to rotate the camera around some point instead of translating it, it would move the camera forward in the world space due to the rotation arc (Figure 10). This could result in the camera passing through an object or ending up behind it, something that is not logically sound as the user could never do so without entering the virtual world or moving the camera by other means. Instead, moving closer and further away from the screen should modify the field of view as a result of the change in angles from the user to the edges of the viewport.

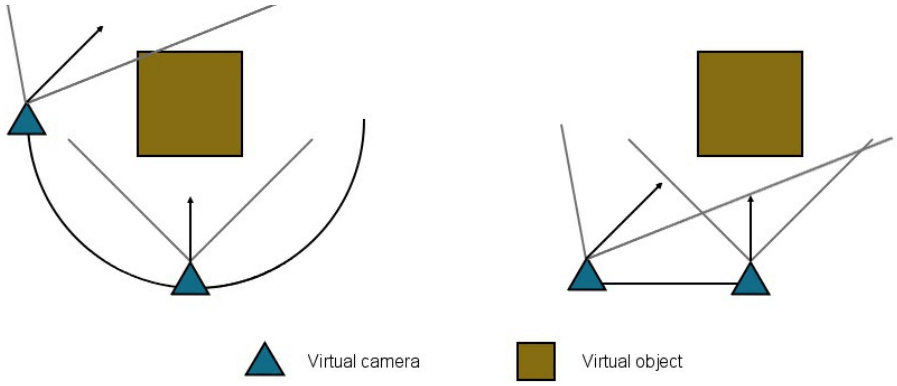


Figure 10: The difference between a rotation and translation of the virtual camera. The rotation may cause undesired effects as the camera is moved forward resulting from the arc of the circle. Translating the camera, on the other hand, results in the correct effect.

It is possible to scale the translation of the camera or view frustum to exaggerate the visual effect, but in order to recreate the proper perspective, the translation should directly correlate to the user's view angle towards the display. The translation of the virtual camera in the world space should also match the scale of the environment, because if the translation is too small, the change in perspective have less practical benefits and be more difficult to perceive. Conversely, if the translation is too large, the change in perspective will be too great, causing it to look unnatural and unrealistic. A proper method for determining a suitable scale for the camera's translation in world space is not yet known, but a useable range can easily be identified by means of subjective impressions.

In case of the Shoebox VR model, we convert the value of the accelerometer's relevant axis to degrees and then divide this value by 360 to convert the rotation angle to a concrete value describing the translation of the view frustum. However, when the view angle to the screen surface increases by rotating the device or changing position, the edges of the screen no longer align with the previous angles, implying a spherical interpolation is required to fully emulate proper conditions (Figure 11). The stereoscopic vision of humans further causes misalignment with the view angle, creating a variation in what

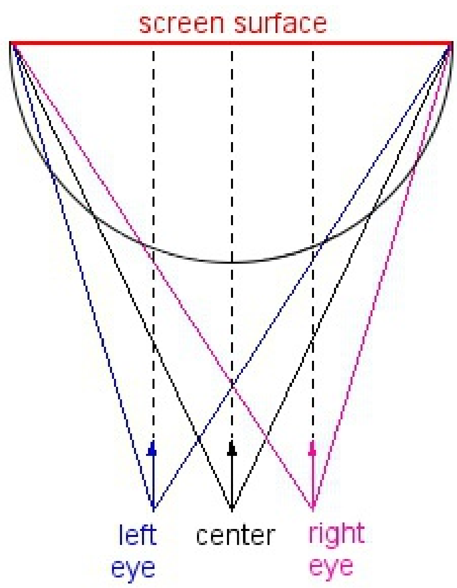


Figure 11: Rotating the device causes misalignment of the view angles. As a result of having stereoscopic vision, the misalignment further increases depending on the view angles of each eye.

Modified from http://paulbourke.net/exhibition/vpac/theory.html the left and right eyes perceive. One solution is to adapt to the eye with the lower view angle i.e. the eye with the best view of the screen. The easiest way to implement this is to express the difference between the center and the eyes as a factor < 1.0 where 1.0 is considered the center position, so that it scales the view frustum's translation to adapt to the position of the eye with the lowest view angle. This creates a more natural visual effect as the image rendered from the center position would result in a slightly stretched perception of the image. In order to determine a suitable factor, one can measure the distance between the center and the eyes relative to the width of the display. The distance to the display is not relevant in this regard as it would result in the view frustum's translation changing simply from moving the device closer or farther away. In typical cases, the factor will vary between values of 0.6 and 0.9 depending on the size of the screen. In case of doubt, a conservative value is generally preferable due to most users being more used to and comfortable with a “flat” and compressed image than an overly stretched one. Note, however, that a single fixed value is recommended for the purpose of a research study so that there can be no discrepancies between results even if multiple devices are used.

3.2 Potential Use and Ideal Situations

As a result of moving the camera during off-axis projection, the perspective changes in ways that increases the sense of depth and spatial comprehension by mimicking how perspective works in the real world i.e. changing the point of view. Perspective allows us to view an object from different angles so that we can combine this information into a 3D mental model. Video games have shown that by controlling the virtual camera directly, we can explore a virtual 3D space and build a mental model without the need to move our physical body to change perspective. However, these forms of interactions introduce a layer of abstraction that are easy to understand, but not necessarily intuitive because they far from resemble the methods of exploration in the real world e.g. walking around and turning the head to look into different directions. Visualization techniques such as head-coupled perspective create a more intuitive and realistic way to explore a 3D virtual space, thereby creating a more convincing virtual reality.

But there are also practical uses for head-coupled perspective. For one, it allows a limited form of exploration with minimal interaction. For desktop or monitor solutions that employ camera tracking, there is no need for any physical interaction at all. Furthermore, the exploration that off-axis allows for essentially consists of two parts. One part shifts the view so that we can see areas of the environment that would normally not be inside our field of view. The second part is that we can change our point of view, allowing us to look past objects that would normally obstruct our view, or more carefully examine an object. The latter is especially useful when true stereoscopy is absent. For example, it can be used to get a better impression of the relative sizes and distances of objects, thus enhancing depth perception on an otherwise flat and monographic display. In effect, this allows content displayed on a 2D device to portray characteristics of a 3D space, increasing the amount of information it can express.

The degrees to how much the understanding of the 3D space can be improved are however not yet clear, although logic suggests that the environment plays a large part in this. Our own experience and studies have shown that virtual environments with high degrees of motion parallax are ideal cases for use with head-coupled perspective because they make the visual effect extremely apparent as well as indisputably proving the sizes and positions of the objects in the 3D space. Conversely, environments with little to no occlusion and few objects that cannot be discerned from the background or are far apart make it very difficult to perceive and understand the visual effect. In this regard, adding visual cues such as pattern textures and shading may prove useful, but the primary focus should always be on how different an image a different perspective will give and how useful that perspective is in solving a task or objective. This consideration is the only valid way of determining the suitability of a virtual environment for optimal use with head-coupled perspective.

Another important aspect in how perceivable and beneficial head-coupled perspective is in a virtual environment is the placement of the virtual camera. Different perspective views are already used to suit certain situations best as they can effect the performance of the user in both positive and negative ways. For example, a third person perspective view is often considered to be beneficial to understanding how the player character relates to the environment as a result of being able to see the physical body and properties of the character. As such third person perspective views are often used in platforming and fighting games where the placement and state of the character are crucial to the user's performance. A first person perspective view is generally considered better at detecting detail and performing smaller precision movements making it more suitable for shooters that require a great degree of precision and attention to detail in order to quickly and effectively dispose of threats. Similarly, if a complete overview is considered important, the camera is better placed up high where it can oversee large parts of the environment [22][23][24][25]. The addition of head-coupled perspective to these perspective views has the effect of emphasizing the features of such views as a result of the virtual camera's extended functionality. For a first person perspective view, the camera's translation provides different points of view for better examining details and judging distances. Third person perspective views can benefit from a clearer understanding of the 3D space in which the character is present. Finally, an overhead or isometric perspective view can benefit greatly from the extended field of view by being able to view a substantially larger area without the need to relocate. However, as one increases the distance between the environment and the camera, one also likely decreases the visibility and usefulness of the “3D effect” as a result of reducing the effectiveness of the camera translation. This is likely the reason why many implementations are centered around the idea of a box because this puts the camera close to the environment, making it easy to create scenarios where the "3D effect" becomes easily visible [3][5][6][8][29]. Because of the different characteristics and uses of different perspective views, it is impossible to specify a universal setup. However, we do observe that the relation between the camera's translation as a result of head-coupled perspective and the distance to the objects forms a critical limiting factor on how visible the "3D effect" of head-coupled perspective can be.

To ensure proper use of ShoeboxVR, one must take care to select a suitable testing environment and appropriate hardware. For example, rotating the devices may cause glare or reflection in well lit rooms or outdoors. Devices should also be cleaned and maintained before each test session for hygienic and practical purposes, as users should be comfortable using the device. Finally, due to the reliance of Shoebox VR on orientation sensors such as the accelerometer, one must make sure the device is properly calibrated and that noise levels are within acceptable ranges.

3.3 Limitations and Issues

One crucial limitation of head-coupled perspective in its current state is that it only supports one user. The screen used can only render one image and that image will be visible to everyone. Hence, until we find a solution to this problem, it is best to find applications where one user is the norm. In this regard, using it for a mobile devices makes the limitation less apparent, because people rarely share one device. Even then, the Shoebox VR model is arguably more robust in that it can not get "confused" when there are multiple users. For desktop implementations that rely on camera tracking, it is possible and likely that more than one person will be captured on camera if there are multiple persons in the room. Depending on the intelligence of the tracking algorithm, it is possible that it will switch back and forth between persons. As a result, it may cause the image to change drastically as it repeatedly updates the perspective for several different positions. In contrast, Shoebox VR does not track the user in such a way and only relies on the rotation of the device to produce the effect. Although it will still look incorrect to one or both users due to the angle being misaligned, it will not cause the visualization to go awry.

A second argument that works more in favor of the use of mobile devices is that it's much easier to rotate a device in your hands than it is to move about in front of a camera. Aside from the issue of space, moving around is not easy to do when seated in a chair, thus limiting the scope of head-coupled perspective to a smaller area. Similarly, a large distance to the screen does not equal better, because it requires more movement to generate a noticeable effect on the image. Therefore, the ideal case would be close to screen while still being able to register plenty of movement. For mobiles this is the case due to the small distance between the user and the device and the ease to rotate the device to any angle. This allows mobiles to make the most of use of the perspective correction with the least amount of effort.

An argument that favors desktop environments is the fact that additional interactions do not use touch interaction. Although new desktop and laptops do sometimes come with a touch-enabled display, it is always possible to employ the use of peripherals i.e. input devices such as the mouse and keyboard. Where interaction with mobile devices is almost always limited to usage of the orientation and touch screen, desktop environments make use of more passive means of interaction that do not require direct interaction with the display itself or clutter the screen with virtual devices such as onscreen joysticks or keyboards. Therefore, taking one of them in order to implement head-coupled perspective further reduces the availability of options of being able to interact with the device. This in turn leads to limited complexity of applications or useability issues with the control scheme because of limited input sources. In this study we also tackle this problem to a degree due to the usage of touch controls in combination with Shoebox VR.

Several other limitations can be attributed to the size of or quality of the device. For example, smaller devices are easier to handle and hold due to their size, but may have issues with using touch controls because the touch interactions are likely to occlude the screen (Figure 12). Larger devices may bring benefits in terms of a larger display, but are generally not comfortable to hold for extended periods of time because of their weight. Use of Shoebox VR is also susceptible to these attributes and characteristics, as it is probable that tilting is easier on a small mobile phone than a large 10 inch tablet. However, experience or perception may be better on the larger device due to its bigger screen. As such, differences in device sizes do not only affect the visibility or clarity of the image, but also affect issues such as interaction, useability and portability.

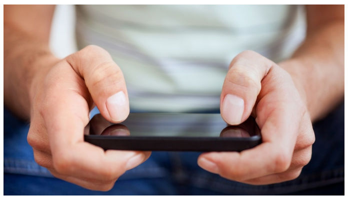


Figure 12: Device size affects interaction and perception.

http://mashable.com/2013/03/06/apps-for-downtime/

Finally, most cases are limited to the 2D plane of which the screen is part of meaning we cannot move past it. The only existing case that can be considered an exception to this is the pCubee by Stavness et al. (2010) that has screens on six sides that fully enclose the virtual environment [5]. This allows the object to truly be observed from almost any perspective, but limits the virtual space to that of the box. A typical flat screen display panel, on the other hand, is just exactly that: flat.

As for limitations specific to Shoebox VR, there known problems with the accelerometer being limited to certain angles that impose restrictions on how far the device can be tilted. As a result, the ideal default orientation is when it is held with the screen facing up as this enables rotations of 90 degrees in both horizontal and vertical directions. However, this is not necessarily a suitable or comfortable orientation to hold the device in, and may therefore form an obstacle for some scenarios. Additionally, orientation sensors are known to have certain levels of noise in their reported data and thus require a form of filtering to prevent jitter. It is best to avoid filtering using averages or similar approaches, as they tend to have the effect of reducing response times as a result of their filtering behavior. For our studies, we employ the use of a Kalman filter based on a recursive prediction algorithm that features both good response times and jitter removal. At this time, it is unclear if other types of orientation sensors, such as the gyroscope suffer, from similar restrictions.

3.4 Open versus Closed Environments

In our research, we want to evaluate user behavior and performance with relation to Shoebox VR when users are allowed to navigate freely. In order to support a reasonable freedom of navigation and exploration, the environment needs to be large enough to enable the user to make observations and choices with regard to their movements. Because of this, navigation and open environments are highly related to each other in that navigation and exploration are likely to be extremely limited in a small closed space such as the box-like environments in many of the head-coupled perspective related studies. However, due to the larger size of the environment, the relative scale of the player or its character decreases which in turn results in occlusion being more likely and a larger distance between obstacles and objects. Because of this, the visual effect of head-coupled perspective becomes less noticeable, as motion parallax is less likely to occur and the translation of the virtual camera becoming less useful. Furthermore, the walls and ceiling of the box-like environments act as visual cues and reference points that help estimate the relative distances and sizes of other objects. In open environments, these are either much farther away or not present at all, leaving only the floor in many cases as a reference plane. This imposes challenges and potential conflicts of interests that may deter sensible use of head-coupled perspective in such environments, but there are also cases where use of head-coupled perspective may be beneficial.

The most important advantage that may be gained from the addition of head-coupled perspective to free movement is the ability to explore by actually looking around during motion. Aside from being a relatively novel method of exploration for many people, it is also much more intuitive and enables the user to dynamically shift the perspective without influencing movement, something that is not possible with traditional turning controls as they directly affect the direction of motion in many cases. Therefore, in theory this will make it easier to look around and could allow for more efficient observation due to the wider field of view. Additionally, for small devices such as mobile phones, usage of touch screen interaction may cause the user's hand to obstruct the view, but use of Shoebox VR will allow the user to overcome this obstacle by tilting the device in the desired direction, revealing the area that previously occluded. As such, the negative effect is greatly reduced at little expense while allowing the user to continue his or her use of the touch screen.

However, with the added control also comes an additional load to the senses. Regardless of control scheme or platform, the user would need three forms of interaction to be able to make use of movement, turning and head-coupled perspective separately. With each added form of interaction, the cognitive load on the user increases as the complexity of the control scheme increases as well. Although experienced users such as gamers have shown to be very capable of dealing with great complexity, it remains to be seen how the average user will fare in these scenarios. Additionally, due to the overlap of functionality between head-coupled perspective and movement and turning, combined use may result in an unnatural or confusing motion. One such case we identified occurs when using manual turning whilst making considerable use of head-coupled perspective. The reason for this being confusing and unnatural is that the two different interactions stem from having virtual (turning the virtual camera) and real world (viewing the 3D space from an angle) actions occur simultaneously. Another potentially problematic case concerning this overlap is that the addition of head-coupled perspective introduces a new form of interaction that is rather limited compared to free movement and turning. Indeed, head-coupled perspective may be added as an extra tool and may prove very useful in this regard, but the question remains if users are capable and willing to make use of it when they can change the perspective just as easy, if not easier, by simply moving and turning the virtual camera. Especially for users who are trained in the use of traditional controls it is unsure if they are willing and capable of making use of head-coupled perspective.

3.5 Conclusion

We have reasoned that environments with high amounts of motion parallax are ideal cases for detecting and experiencing head-coupled perspective models and that objects close to the virtual camera can benefit most from the translation of the virtual camera that results from using off-axis projection. Because of this, we expect Shoebox VR to be most relevant for small environments with a high number of objects distributed equally over the space, rather than large open environments where distances between objects are generally larger. However, we have also discussed potential cases where a head-coupled perspective model such as Shoebox VR may prove useful for open environments, by allowing off-axis projection to serve as an exploration tool.

We have also determined that there are important considerations to make when evaluating head-coupled perspective models in certain setups due to variations in hardware and conditions. For Shoebox VR, differences between mobile devices are especially relevant due to the relatively high variation between devices compared to typical desktop environments. As such, one must take into account factors such as size, weight and image quality in order to be able to create a suitable and consistent setup for all devices.

Although we can employ logic to determine the general scenarios where head-coupled perspective can be considered beneficial or limited, we lack answers and data to make definitive claims. As such we will form our research questions and hypotheses in the next chapter to make a start in contributing to solving these open problems, followed by a detailed description of our experiment setup in which we further elaborate on the choices we've made regarding the theories discussed in this chapter.

4. Research Questions & Hypotheses

Our primary goal is to determine what effects Shoebox VR, an implementation of head-coupled perspective for mobiles, has on user behavior, performance and experience within a 3D virtual environment. Based on previous studies and research we have determined that there is not enough material exploring or covering its uses and benefits in open 3D virtual environments where the user can move and look freely. Thus we phrase our question as follows:

How does the addition of Shoebox VR affect movement behavior and navigation performance on mobile devices in open 3D virtual environments?

Before we can attempt to answer this, we must first specify what variables are relevant in our study and how Shoebox VR may affect it. We have specified several questions that divide these variables into objective and subjective categories. We do this because users may perceive their performance or behavior differently from what the statistics might suggest, or what we have observed. Finally, we quantify navigation performance as the efficiency of search and collect task, measured primarily by time and distance. Therefore, we state the following sub-questions:

Objective
Does the use of Shoebox VR encourage users to employ it for search tasks *in open virtual environments?*
Does the use of Shoebox VR reduce a user's search time *in open virtual environments?*
Does the use of Shoebox VR enable the user to choose more optimal paths *in open environments?*

Subjective
Do users experience search tasks with Shoebox VR to be more efficient or accurate *than without?*
Do users experience search tasks with Shoebox VR to be more enjoyable or immersive *than without?*

4.1 Hypotheses

To answer each question, we must prove a relation exists between the variables and use of Shoebox VR. For this purpose we phrase five null-hypotheses, so that we require high statistical significance to disprove these claims. For this purpose we specify the requirement of the statistical significance to be $p \leq 0.1$, a fairly liberal value that we motivate by arguing that the degree of randomization as part of the construction and population of the environments will likely have a negative effect on the consistency of the performance results. We phrase the null-hypotheses as follows:

Null-Hypothesis 1 (NH1):
Use of Shoebox VR does not *encourage users to employ it for search tasks in open virtual environments.*

Null-Hypothesis 2 (NH2):
Use of Shoebox VR does not *reduce users' search times in open virtual environments.*

Null-Hypothesis 3 (NH3):
Use of Shoebox VR does not *enable users to more optimally choose paths in search tasks.*

Null-Hypothesis 4 (NH4):
Users will not *experience navigation with Shoebox VR as more efficient or accurate than without.*

Null-Hypothesis 5 (NH5):
Users will not *experience navigation with Shoebox VR as more enjoyable or immersive than without.*

4.2 Measurement Specifications

For the first hypothesis, we measure use of Shoebox VR throughout the experiment and specify an average rating of 50% or more to be the threshold at which we consider the use of Shoebox VR to be encouraged. We argue that this value isn't expected to be higher due to the predicted difficulty of using Shoebox VR in combination with touch controls and the relative novelty of using Shoebox VR during motion. This value represents the relative time at which Shoebox VR is used by tilting the device at an angle of 10 degrees or higher.

For the second hypothesis, we measure the users' time in milliseconds throughout each round and consider an average reduction of 20% or more in round times to be an improvement resulting from the use of Shoebox VR.

For the third hypothesis, we measure the total distance the user has traveled in each round and consider an average reduction of 20% or more in distance values to be an improvement resulting from the use of Shoebox VR.

For the fourth hypothesis, we measure the subjective impressions of each user regarding the usefulness of Shoebox VR for completing the tasks and express this as a value ranging from -2 to +2, with a stepsize of 1. We consider an average value of +0.75 or higher to be a positive impression.

For the fifth and final hypothesis, we measure the subjective impressions of each user regarding the ease of use and fondness of the visual effect of Shoebox VR for completing the tasks and express this as a value ranging from -2 to +2, with a stepsize of 1. We consider an average value of +0.75 or higher to be a positive experience.

4.3 Alternative Hypotheses

Finally, we phrase five alternative hypotheses that we may accept if we are able to reject the null hypotheses based on these specifications:

Alternative-Hypothesis 1 (AH1):
Use of Shoebox VR encourages *users to employ it for search tasks in open virtual environments.*

Alternative-Hypothesis 2 (AH2):
Use of Shoebox VR reduces *users' search times in open virtual environments.*

Alternative-Hypothesis 3 (AH3):
Use of Shoebox VR enables *users to more optimally choose paths in search tasks.*

Alternative-Hypothesis 4 (AH4):
Users will *experience navigation with Shoebox VR as more efficient or accurate than without.*

Alternative-Hypothesis 5 (AH5):
Users will *experience navigation with Shoebox VR as more enjoyable or immersive than without.*

5. Experiment Setup and Implementation Details

This chapter covers the concrete details on the experiment setup and some of the implementation. We will begin by describing the basic setup of the experiment, the data we will collect and our choices regarding the open problems that were highlighted in the theoretical analysis.

5.1 Basic Setup

Our study is a user study, meaning we collect data from users participating in the experiment. In total, we conduct three user studies with the goal to evaluate user behavior and performance regarding the use of Shoebox VR in open environments. The first study is a pilot study used to test initial theories and find problems with the implementation and setup. The second study focuses on collecting quantitative data and implements the improvements from the preliminary study and is therefore considered the main study. The final user study is a smaller, but more extensive study that focuses on collecting detailed qualitative feedback from experts in order to allow for a more complete analysis of the results from the main study. We begin by describing the basic setup and detailing several implementation and design choices.

The software implementation is realized as an Android application that must be installed and run on the designated device. Quantitative data is recorded by the software implementation throughout the test session and is stored on the device so that we may collect and process it in a later stage. Qualitative data is collected by means of a small survey at the start and end of the session and is used to acquire information regarding the experience level of the user and subjective opinions on the usage of Shoebox VR during the test. Because the initial study is performed remotely, qualitative feedback is limited to the short survey embedded in the software application unless participants choose to provide additional comments and suggestions in an email. However, the main performance study is supervised and thus enables additional communication and discussion. The expert analysis study continues this trend and extends it by recording the observations and impressions of the user during the session and discussing them in more detail afterward. In this regard, this approach bears some resemblance to a heuristic evaluation in that it features close observation and an extensive discussion [30].

The application consist of several pseudo-randomized stages where the user is tasked to find and collect a number of pseudo-randomly placed items called packages by moving close to them using any combination of movement, turning and tilting controls. Each stage features a challenge or specific objective that the user must solve in the best way possible. Before the user is subjected to these stages he or she is first explained the concept of Shoebox VR and given the opportunity to experience it in a controlled environment. A training stage is also presented in order to allow the user to experiment and become accustomed to using the navigation and tilting controls. Each stage features a predetermined number of rounds to provide the user more than one opportunity to solve the challenge. To encourage and motivate users, their performance in each round is indicated by a rank score based on how well they completed the challenge based on our expectations.

5.2 User Target Demography

Our target user base will be focused on individuals approximately aged 15 to 30 years with varying degrees of experience with mobile devices. Due to availability, it is likely that these individuals will be almost exclusively male and have moderate to extensive experience with 3D applications on PCs or other multimedia entertainment systems. This makes our target users generally proficient and knowledgeable on usage of mobiles and navigation within 3D spaces, so that interaction is less likely to form an obstacle during the experiment.

5.3 Virtual Environment Specification and Design

All challenge stages feature the same basic environment setup which can be represented as a 2D overhead map (Figure 13). The area where the user can move in is a square section that is enclosed by four walls to limit the navigable space, because having too large of an environment increases the time required to complete the stage and introduces the risk of the user wandering too far off in search of a package or even becoming lost, thereby increasing the likelihood of errors and the magnitude in which they can affect the performance results. We employed the use of buildings as exterior decorations outside of the navigable space to combat the impression of a closed off environment as a result of having to limit the navigable space to a degree. Even if they cannot be reached by the user, their presence combined with the view of an open sky suggests there is a larger world outside the four walls, thus eliminating the sense of entrapment the user may have from being separated from the rest of the virtual world. The walls do not only serve to limit the navigable space in a practical and understandable way, but also hide the empty horizon from the user's view that may otherwise cause an unsettling sense of emptiness or simply be distracting. The user's starting position is always fixed to ensure equal starting conditions for all users and is placed in an outside cell facing inward to prevent the need to turn a full circle at the start of a round.

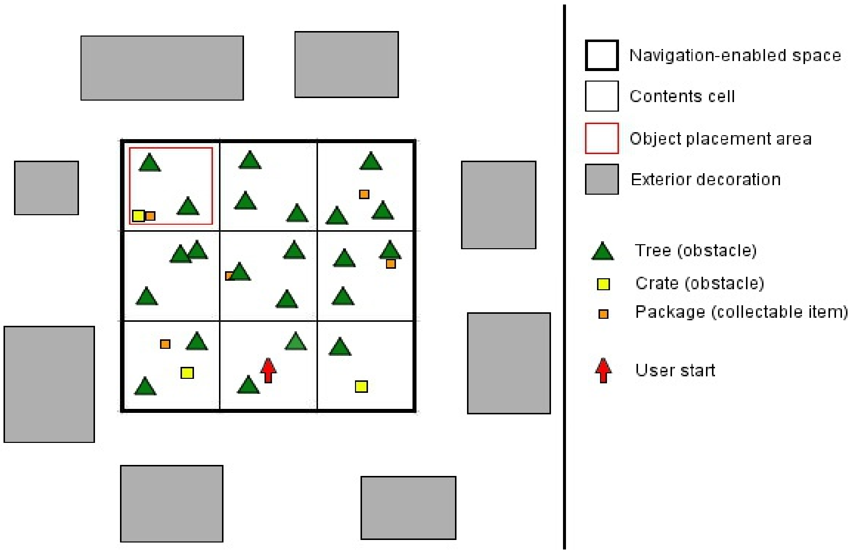


Figure 13: A simplified overhead projection of a sample environment shows how the navigable area is split into a smaller grid of content cells, allowing for a more uniform distribution of objects and packages.

The navigable area is split into a grid of nine smaller squares referred to as content cells. Each cell is able to generate up to three randomly placed solid obstacles in the form of a tree or crate object. Obstacles are not allowed to overlap and will be moved accordingly if this occurs. When placing these obstacles, a slight margin is taken into account to prevent overlap between two obstacles of neighboring cells or any of the outside walls. The grid is used to improve the uniformity of the distribution of objects in the environment and to employ several technical optimizations regarding rendering and collision detection. Each stage features the same base construction, so that only the pseudo-randomization of the obstacles and packages inside the cells may affect the environment. The cells are repopulated with new obstacles only when a new challenge stage begins. The collectable packages are generated at the start of each round.

The use of a primitive and pseudo-randomized environment prevents users from benefiting from a good memory or sense of direction. To elaborate, the short time the user is present in each stage combined with its contents being randomized prevent the users from being able to establish and make use of a comprehensive cognitive map. A cognitive map is a mental 3D model of the environment used to pinpoint important locations in the world space and to determine suitable paths to destinations. In other words, a cognitive map is used to navigate effectively an environment [1][9][10][11][12] and can therefore provide performance benefits for individuals that can establish such a mental model quicker or more detailed than others. Furthermore, studies show that there are noticeable differences between males and females that directly affect their navigation performance in new environments [27][28]. By creating a very simplistic setup where the aim is to find and collect items in a small pseudo-randomized environment, we eliminate any inherent navigation advantages or handicaps a person may have and thereby create a more leveled experience for all users.

Once the environment has been created, a second pass is used to select five cells in which the packages will be placed. The cell in which the user starts may not contain a package, leaving a total of eight eligible cells. Per generated package there is an equal probability of the package being attached to an arbitrary object in the cell or it being placed randomly on the ground within the area of the cell. In case of the latter, the row in which the cell is placed determines where the package is allowed to be placed. The top rows limit the area to the lower half of the cell so that the packages are more likely to be visible from the user's starting point. The middle rows allow the entire cell to be used and the bottom rows limits the area to the top half of the cell. The reason for this distribution is that packages placed on the ground are generally more visible than those attached to objects. By further limiting where they can be placed, we increase the probability of them being visible from the starting position by moving them in front of obstacles or out of the initial blind spot. This way, we ensure an initial target for the user to move to and to present them a choice: whether or not they choose to move toward this target or explore the environment for better candidates. Packages attached to an object are not affected by this distribution as they are already limited in their placement due to their attachment.

The reason why we make use of trees as the main source of obstacles in our environment is because earlier setups simulated urban environments with large buildings. Although we used decorations to fill up the space between these buildings, the relatively large scale of the buildings compared to the player character caused large amounts of the environment to be occluded from almost any position. As a result, it required the user to walk large distances in order to properly explore the environment and therefore imposed a severe limit on the practical use of Shoebox VR. As an alternative, we considered smaller, more densely populated indoor environments, but the risk of these environments ultimately resembling a closed off box and limiting the freedom of navigation were deemed too great. Thus we required outdoors environment with a higher density of smaller obstacles which is what led use to the premise of a forest or park themed environment. Although trees are not necessarily small, their lack of width means that it is possible to look past or around them using Shoebox VR, thereby creating a practical use and a noticeable motion parallax effect that would help users in understanding the effects and use of Shoebox VR. Other obstacles were considered, but would often look out of place due to the setting. Crates were eventually introduced in small numbers for the purpose of variation as the environment would otherwise be made up completely out of trees.

We employ the use of material textures and basic shading to reduce the level of abstraction of the visual representation of the environment, which allows the user to better identify the objects in the scene and to create a limited sense of immersion (Figure 14). Because we effectively prevent the user from establishing a cognitive map, it is crucial that we make the environment relateable and understandable to prevent confusion or disorientation [1][11]. A second benefit is that pattern textures and shading serve as visual cues that can enhance depth perception by creating a level of contrast in the image. One requires such cues because it remains extremely challenging to perceive any sense of size, shape or depth without them and would thus make the addition of head-coupled perspective largely pointless. Each challenge stage features slightly different textures to help distinguish the environments and affiliate a certain visual look with the challenges. The shading is limited to a simple diffuse shader with an ambient lighting level of 0.5 out of 1.0, thereby creating lighting conditions that were noticeable but never aided or interfered with the tasks.

Figure 14: Different stages make use of different textures to remind the user of the current challenge. The textures of important objects such as the collectable packages are not changed. Notice how use of material textures and diffuse shading provide additional contrast to the image and create a limited sense of immersion.

Due to Shoebox VR playing a prominent role in this experiment, it is important to detail the size and scale of the environment so that the setup can be recreated if so desired. Note that all values are in OpenGL units, which have no real-world comparison due to being completely virtual. The navigable area spans a total of 300 square units with trees varying in thickness between 2 to 4 units. The crate obstacles are cubes with sides of 4 units long and the packages are unit cubes. The player character model is represented as a red colored cuboid of 2 by 4 by 2, allowing the user to identify the character at all times and making it easier to understand what effects its physical properties would have during collisions opposed to more complex shape. The walls have a height of 10 units and surround the entire navigable area. Considering the sizes of this environment, the scale of the Shoebox VR camera's translation was set to a slightly conservative value of 5 in order to not risk exaggeration. The view frustum's translation was multiplied by a factor of 0.8 to account for the lack of stereoscopy.

5.4 Perspective View

The main objective in each stage is to find and collect certain relatively small packages by moving close to them. Because of this scenario, we chose a third person perspective due to being able to see the character's physical presence in the world, thereby increasing the user's environmental awareness. Furthermore, the distance between the camera and character creates an apparently wider field of view that can aid in exploring and finding packages whereas a first person perspective suffers from tunnel vision. Finally, the character model itself functions as a reference point for Shoebox VR, thus allowing the users to better understand the effect. Initial tests suggested that an overhead or isometric perspective view also had potential, provided we made several adaptations such as increasing the distances between the trees to prevent continuous occlusion from the treetops and hiding the packages directly beneath them.

However, such a perspective view was ultimately deemed to be unsuitable because it is more generally considered to be more situational than an on-the-ground perspective view and the adaptations would further remove us from a general case scenario. We opted to not to include multiple perspectives in our setup as it would introduce a new variable, thereby creating many possible new cases where the perspective could affect user behavior and performance.

The vertical field of view for the virtual camera is set at an angle of 30 degrees due to higher values creating noticeable distortions in the image in the corners of the screen on smaller devices. Limiting the field of view to lower values would have the opposite effect of narrowing the view too much, hampering the ability to find packages in the environment and making Shoebox VR possibly too powerful as a result. Sensitive users are also more likely to experience motion sickness with a low field of view. Therefore, a value of 30 degrees was considered to be most suitable value for our case and provides a well balanced view between the player character and the environment. Starting from the character's center, the virtual camera is placed 5 units behind the character and moved 2 units upward. The camera always points to the character's center, so that it rotates around the character and views it from a small distance with a slight downward angle.

5.5 Navigation Controls and User Interface

In order to navigate, the user requires a way to move and turn. Due to our target demography consisting mostly of experienced gamers, the concept of a virtual gamepad is a logical choice. Although there are other options, they often make use exclusive use of touch-dragging or point-and-click mechanics to simplify the control scheme [24][26]. As a result, precision is lost or the risk of Shoebox VR being more useful than the actual control scheme due to limitations becomes too great. Therefore, the virtual pad was chosen and implemented with several modifications to decrease the advantages of experienced users.

The primary modification limits the movement speed in order to reduce the cognitive load of processing motion in combination with turning or Shoebox VR. Because of the target demography's expected experience with 3D games, it is likely that at least some of the participating users will have a much better understanding of motion and movement controls. As such, these users would perform notably better by simply being able to move and turn more efficiently than others. By limiting the movement speed to a walking pace, we reduce the advantage of experienced users by giving all users ample time to observe and react.

The second modification limits the turning to a horizontal motion to reduce it to a one dimensional problem. Although packages are sometimes positioned on the ground, there is no direct need for the user to look up or down in order to spot or collect them. Initial tests suggested that the ability to look up and down using touch controls was able to cause severe disorientation and issues with controlling the movement of the player character which is why we chose to disable it. However, Shoebox VR may still be used to look up and down as it is not subject to this limitation. Note that this can be considered one of the primary reasons why vertical tilting is not tracked or collected when evaluating user behavior with regard to use of Shoebox VR.

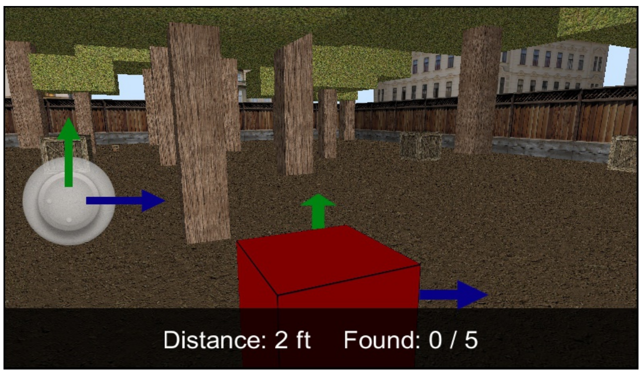


Figure 15: The virtual pad displays an image of an analog stick to serve as a reference point and visual aid for movement directions.

The virtual pad is distributed in two halves divided equally over the screen, with the left side controlling movement and the right side controlling the turning. Because the movement is completely unbound apart from being limited to a 2D plane i.e. the ground, a visual representation considered helpful in understanding the control scheme. As a result, whenever the user makes use of the left side of the screen, a semi-transparent image of an analog stick is used to represent the movement of the character. When the screen is first pressed, the virtual pad is centered on these coordinates, so that subsequent touch gestures control the movement direction (Figure 15). Releasing the left side of the screen stops the movement and removes the virtual pad from view. The direction in which the character moves is directly related to the direction the character is looking in, therefore requiring a way to turn. To keep the turning as simple as possible, the direction is changed by simply swiping the right side of the screen. Because of the relatively simplicity of this mechanic there no image used to visually support this action. Due to varying pixel densities across screens, the values for movement and turning were multiplied with the logical pixel density so that differences in interaction were limited to a minimum. Turning or using Shoebox VR does not move the player character and therefore does not affect distance measurements.

A semi-transparent visual information bar at the bottom of the screen is reserved for displaying relevant data and information regarding the stage challenge and completion status. Outside of the stages, the interface is mostly used to display text for informational and educational purposes, such as informing the user of the objective of the next challenge or indicating how well the user has performed in a round. At the beginning and end of the experiment, the interface is also used to gather information on the user's level of experience regarding 3D gaming and subjective impressions on Shoebox VR usage throughout the test session. Users can express their opinion or levels of experience by means of a quantitative rating system (Figure 16). For the experience ratings, values ranged from -1 to +1 for a total of 3 choices, with the survey ratings ranging from -2 to +2 for a total of 5 choices.

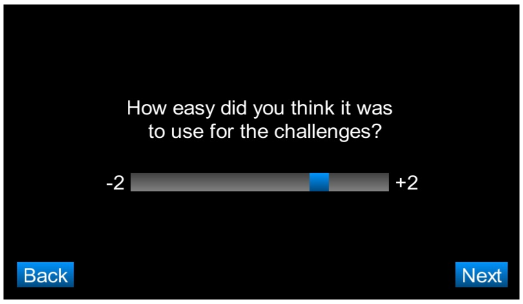


Figure 16: By moving the slider, users can express their thoughts by means of a rating system.

5.6 Sound Design

Sound playback was used to provide additional feedback to user actions or events. When users collect a package, the counter in the bottom right of the screen is incremented and a short clicking sound is played confirming the collection of the package. Additionally, a buzzer sound plays at the end of each challenge stage round when the objective is completed. No music or ambient sounds were used as they could be considered distracting by some users.

6. Pilot Study

The pilot study was aimed at exploring initial behavior patterns for determining interesting and promising cases for the performance study as well as detecting problems in the implementation and setup of the experiment. The participants group for this study consisted of 12 individuals that were asked to install the software on their devices and follow the instructions in order to complete the test. Because of the nature of remote testing, observations cannot be made and thus we are limited to the recorded data of the software. Furthermore, due to the relatively low size of the participants group and the inability to verify an optimal testing environment for the participants, we cannot make any definitive claims based solely on the results of this study and can therefore only discuss our findings in an informal manner.

6.1 Setup

The remote study consists of a total of four challenge stages in which navigation is required to complete it. Prior to these stages are a calibration setup and a Shoebox VR demo that allows the user to experience Shoebox VR in a controlled environment. Of these four stages, one is a training stage aimed at briefly explaining the goal of the experiment, the control scheme and the completion conditions to the user. During this training stage, the user is allowed to experiment with the movement, turning and tilting controls in order to get acquainted with the control scheme and to develop a certain degree of competence and efficiency before their performance becomes relevant. Usage of Shoebox VR during the training is briefly required in order to activate the navigation controls, thereby forcing the user to make use of it in order to continue. This was done so that the users would understand the use of Shoebox VR as an exploration tool before being presented with the ability to move and turn using the touch controls. Once the training stage is completed, the challenge stages follow after a clear indication that the performance of the user is now a primary concern.

6.2 Challenge Stages

The three stages each tackle a specific scenario that changes how the stage is solved best and are split between two rounds each for a total of six rounds to allow the user to improve his or her results the second try. Although we had considered using three rounds for each stage, we were very careful in increasing its overall duration, as the concept of a remote study could introduce the risk of users not completing the test if it took too long to complete. However, due wanting to verify some initial theories, we could not simply choose only one challenge for the stages. Similarly, we could not distribute different implementations among the participant group, as this would result in even smaller group sizes. Therefore, we willingly risk the effects of limiting our data in favor of testing more scenarios.

In spite of the differences in the challenges, all stages record the same data for each round of which we classify three as the main performance indicators: whether or not the first package collected was the closest to the start position, the total time taken and the distance traveled. This is done in order to be able to record a complete and comprehensive overview of the user's behavior during a round. However, their relevance depends mainly on the objective due to the conditions it creates. Additionally, we also record how much use the user makes use of Shoebox VR by recording the angles at which the device is tilted, although we limit the measuring to only horizontal tilting angles due to vertical tilting having no functional use and because vertical view angles on most devices generally being poorer than horizontal angles. We quantify Shoebox VR usage in a more meaningful way by creating four categories that measure how much time in milliseconds was spent in a certain angle range during each round. The first category limits the angle to 10 degrees or below, the second to 40 and the third to 60. The final category covers the angles higher than 60 degrees. These four categories effectively describe the usage as “Low”, “Average”, “High” and “Extreme” respectively. At the end of the round, these values are normalized by dividing them by the round time, thereby creating percentages of each category that portray the usage of Shoebox VR during that round much more accurately than a mean or median can.

The first challenge stage requires the user to locate and collect the package nearest to the starting location and is referred to as the “Closest” stage. There is no limit of any kind incurred on this stage, allowing the user to take as long as is needed before making his or her final decision. Therefore, time and distance are not highly relevant as performance indicators, although they may still provide insight into the user's behavior. The aim of this stage is to create a scenario where the user must judge relative distances at the best of their ability. Because the starting position is not indicated or highlighted in any way, it is best to remain in place and evaluate initial findings first. Here, Shoebox VR may be used to look around or change perspectives by a limited degree. Alternatively, the user may simply move to change perspective without the limitation, but risk the possibility of misjudging distances to the start position because of this. In this stage, the main performance indicator is whether or not the user collected the package closest to the start position. We expect users to perform well on this stage due to its simplicity and predict a usage of Shoebox VR of more than 50% of the round time.

The second “Timed” stage limits the round time to 30 seconds, forcing the user to think and act quickly. This stage simulates the need for a hasty resolution and induces a level of mental stress on the user that may cause him or her to no longer use Shoebox VR if they feel they are more efficient by restricting themselves to the navigation they are more accustomed to. The main performance indicator for stage is time, although distance may also be relevant as it could reveal poor path planning. If the user collected the package closest to the start position first, it may indicate that an educated choice was made in selecting the first target rather than rushing to whichever was spotted first. With the time limit set to 30 seconds, it may become highly challenging to collect all five packages, depending on how the packages are placed. It is therefore likely that many users will have great difficulty achieving this feat. We also expect low usage of Shoebox VR with an average of approximately 30% of the round time due to the pressure of the time limit and the limited experience with Shoebox VR in combination with movement.

The third and final stage challenges the user to minimize the distance traveled and is therefore dubbed the “Distance” stage. Because there is no time limit, the user is encouraged to observe the complete environment and plan their movements carefully prior to tasking themselves with the collection. Because not all packages are visible from the start position, the user is eventually required to move in order to change perspectives. Shoebox VR may be used as an interim solution due to its limited ability to look past or around some obstacles within the vicinity of the user. Alternatively, the user may travel short distances and look around during the pauses to get an overview of the package placement. This introduces the risk of engaging in pointless movements, however. Due to it's nature of analyzing the environment and judging distances it can be considered an extension of the “Closest” stage with a more difficult challenge and an added pressure component. The main performance indicator for this stage is distance traveled and is measured in OpenGL units, although the interface indicates as feet. A distance of less than 200 units is considered best, a distance of 250 is considered average and values above 300 are considered poor. Time may provide insight into how carefully the user chose to observe the environment before moving. If the user was patient and observant, it is also likely they collected the package closest to the start position first, therefore it is more relevant of a performance indicator than time for this stage. We expect users to vary in performance in this stage due to its dependence on path planning and keen observation, causing errors to have larger consequences. We also expect Shoebox VR usage to be highest in this stage with an approximate of 70% or higher, because of its benefits in completing the challenge and the need for careful observation.

6.3 Survey and Feedback

After the challenge stages are completed, the user is presented with three qualitative questions that they can answer by means of the rating system. The ratings range from -2 to +2 with intervals of 1, thereby creating five selections with 0 being the neutral and therefore default selection. By providing a 'back' button in the user interface, the users were free to revisit questions or review their ratings before completing the survey. This included a final screen that explicitly stated that it was the final screen. The questions are as follows:

- *How much would you say you liked the visual effect?*
  This question serves to answer how much the user enjoyed the visual effect of Shoebox VR, regardless of any performance issues or benefits they may have had using it.

- *How easy did you think it was to use for the challenges?*
  This question serves to answer how easy it was for the user to make use of Shoebox VR for completing the challenges, regardless of how useful they thought it was.

- *How useful do you think it was for the challenges?*
  This question serves to answer how useful the user thought Shoebox VR was for completing the challenges, regardless of how easy they thought it was.

Because we cannot rely purely on statistical data, these questions provide further insight into the user's behavior and, more importantly, opinion of Shoebox VR for the test. For example, users that made plentiful use of Shoebox VR but performed poorly may rate the *ease of use* of Shoebox VR negatively, indicating that they had difficulty in using it effectively. Similarly, users that made very little use of Shoebox and performed very well may rate the *usefulness* negatively, indicating that they saw little reason to use it. In some cases it's also reasonable to assume there is a correlation between the values, such as when somebody thinks the visual effect is useful and therefore rates both positively. All such cases contribute to determining if our hypotheses are valid as well as answering our research questions.

Although the test was performed remotely, several users were asked and able to give additional feedback by email or direct communication in case they found errors or had difficulty understanding some of the mechanics or explanations.

6.4 Results

Our first and foremost goal of this pilot study is to find behavior patterns and potential issues with the experiment setup. In this regard, we are first and foremost interested in how Shoebox VR was used throughout each stage. Therefore, we've tracked use of Shoebox VR angles during each round and categorized them in four categories. Each category represents the amount of time spent whilst tilting the device at a specific angle, relative to the time it took to complete the round. The "Low" category represents angles below 10 degrees, "Average" from 10 to 40, "High" from 40 to 60 and "Extreme" everything beyond that up to a 90 degree angle, which we consider the practical limit.

Based on the average values shown in Figure 17, Shoebox VR is used rather sparsely, averaging at around 31% of the round's time and rarely being used beyond a 40 degree angle. The latter is not that surprising considering that earlier studies have shown that tilting in combination with touch controls can make for a very uncomfortable experience as a result of the orientation of the device requiring increasingly unnatural hand and wrist movements the further the device is tilted. Furthermore, it is important to realize that turning using the touch controls is in essence more powerful than Shoebox VR because it is not limited to the tilting angles and thus allows for full circle rotation. In other words, the only way for Shoebox VR to be useful during motion is to survey the environment for (better) targets. However, does this not apply to the use of Shoebox VR when stationary, which is expected to occur most often. Therefore, the lower than expected values could be the result of the lack of inexperience with Shoebox VR in general or the perceived inability of Shoebox VR to provide any benefits for the search task.

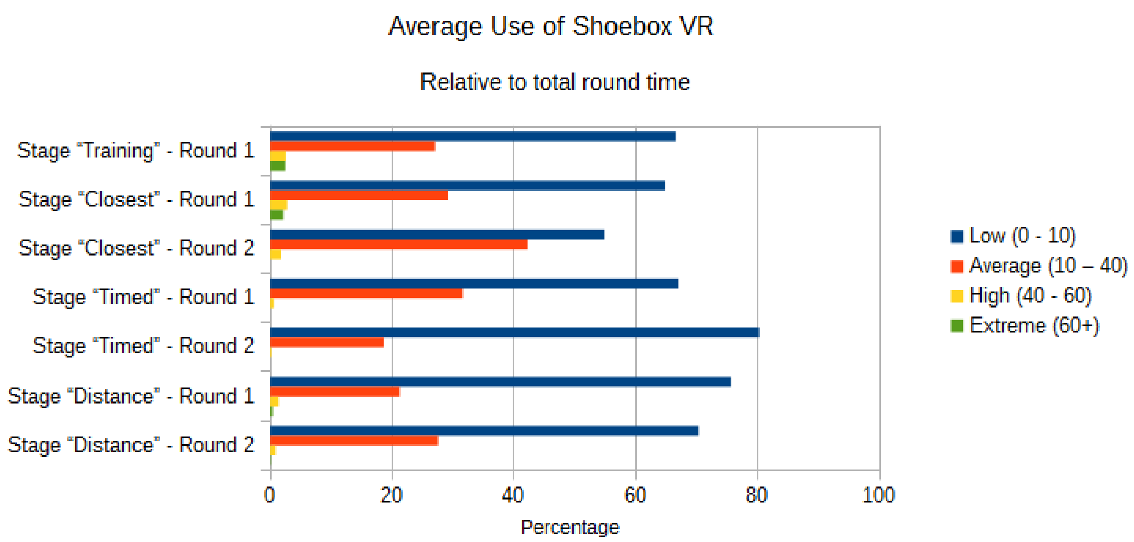


Figure 17: Average values seem to suggest a lower than expected usage of Shoebox VR in the stages, but it can be argued that the large amount of navigation required to collect the packages is the primary limiting factor due to the difficulty of using Shoebox VR with touch screen controls. Values represent the time spent in each angle category relative to the round time. The means of each category are 69% (Low), 28% (Average), 1% (High) and 1% (Extreme).

It is interesting to note that during training and the initial "Closest" stage, use of Shoebox VR is at its peak. The second round of the "Closest" stage in particular shows a slight spike in its use, although still largely limited to a 40 degree angle. The increased use likely results from users becoming accustomed to use of Shoebox VR and trying it out to see if it's useful for the challenge. In this regard, the "Closest" stage is an excellent opportunity to do so as it features a relatively simple challenge with no pressure component. However, the nature of the "Closest" stage also affects the results in that this stage requires considerably less movement than the other two and therefore does not suffer as much from the issue of using Shoebox VR in combination with the navigation controls. Combined with a lower average round time resulting from having to only collect one package, it is reasonable to assume that the higher than average use of Shoebox VR in this stage can be attributed to the lack of complexity of the stage and its objective rather than users seeing considerably more benefit from it compared to the other stages.

When the time pressure component is added, Shoebox VR usage approaches the average again with the second round having the lowest average use of all rounds. A possible explanation for this behavior could be that, because the time limit of 30 seconds makes this stage very challenging and stressful, users tend to fall back on their aptitude with traditional navigation controls resulting from their experience with 3D games. However, the movement speed was intentionally lowered to allow for plentiful use of Shoebox VR while moving in between target destinations, thus the decline in its use seems to suggest that the risk of using Shoebox VR during this stage was considered too great. Based on these arguments, it's reasonable to assume that having a time limit has a negative effect on use of Shoebox VR when the focus of the task is to find and collect items in the environment by means of movement.

Interestingly enough, the reduced use of Shoebox VR trend continues in the final stage, although it reverts back to average values for the second round. It is reasonable to assume that the experience with the "Timed" stage left users with a negative impression of Shoebox VR for stages with a pressure component, although this behavior seems to stabilize once the users appear to realize that Shoebox VR has more practical use in this stage resulting from the lack of a time limit. This grants them the ability to pause movement and make new observations before continuing, a far more suitable use for Shoebox VR that relies heavily on keen observation and exploration. Regrettably, due to only having two rounds, we cannot observe if this behavior would continue to lead to an increase of Shoebox VR or if it too would stabilize over time.

Because of the relatively small group size of 12 participants for this study, it is possible to more carefully examine individual behavior. This may help us identify if the low usage of Shoebox VR is due to the stages or because of personal choice. Figure 18 shows that users vary in their use of Shoebox VR indicating it is subject to choice. The possibility exists that the perceived usefulness or appeal of Shoebox VR heavily influences a user's choice of using it. Figure 19 shows a general negative impression and opinion of Shoebox VR that supports this argument It is especially worth noting that the *usefulness* of Shoebox VR is rated lowest, implying that the majority of the users had no interest in using Shoebox VR extensively for the purpose of locating the packages in the environment.

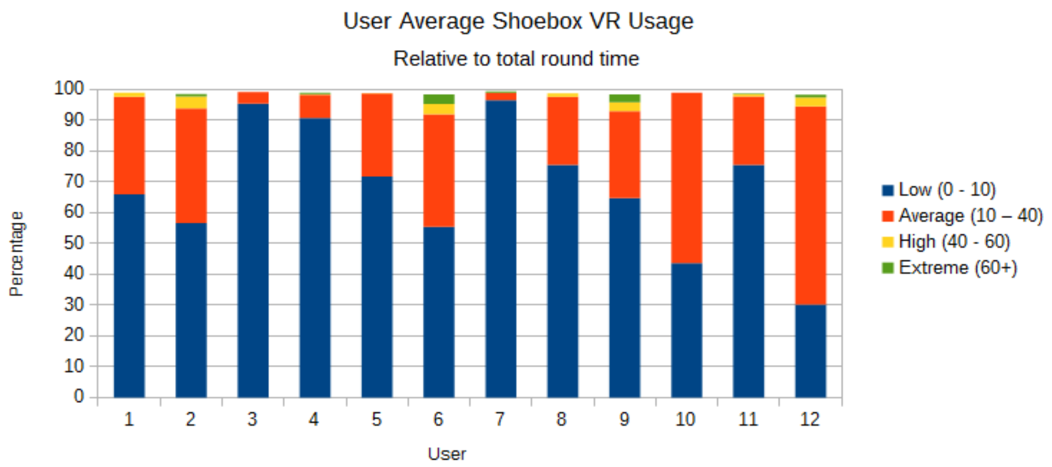


Figure 18: In general, use of Shoebox VR remains low with only two notable exceptions where Shoebox VR use exceeds an average of 50 percent. Note that not all totals reach an exact 100 percent due to rounding errors of the recorded data.

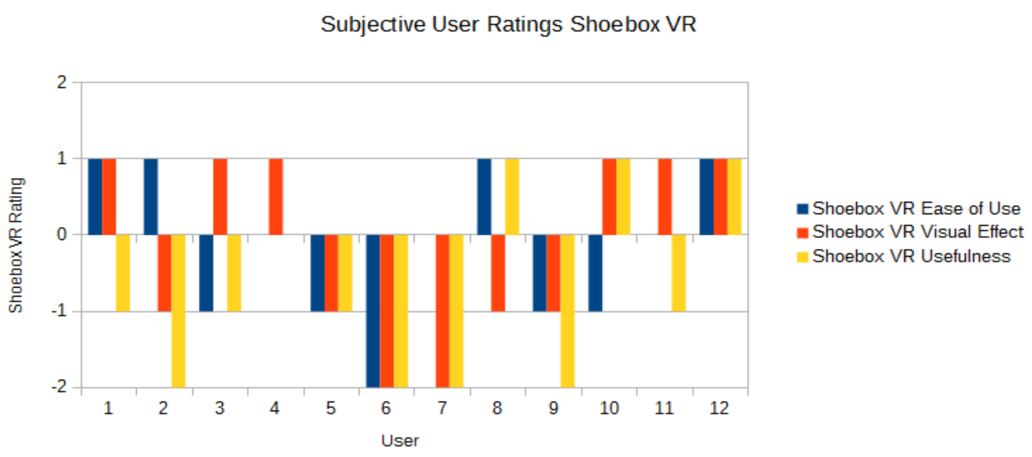


Figure 19: Users ratings on Shoebox VR vary, but overall portray a slightly negative image with means of -0.17, -0.17 and -0.75 for ease of use, the visual effect and usefulness respectively.

Figure 20 shows that all 12 participants have rated their 3D gaming experience on PCs and consoles to be medium or high with an average rating of 2.75 out of 3, as is expected of our target demography. However, their rated experience with 3D mobile games is considerately lower with an average rating of 1.83 out of 3. This further suggests that the users relied largely on their ability to use the touch screen navigation controls effectively for completing the challenges. Comparing the values from Figures 19 and 20 shows little correlation between experience levels and ratings, with no indication of any of the ratings being the result of inexperience with 3D gaming.

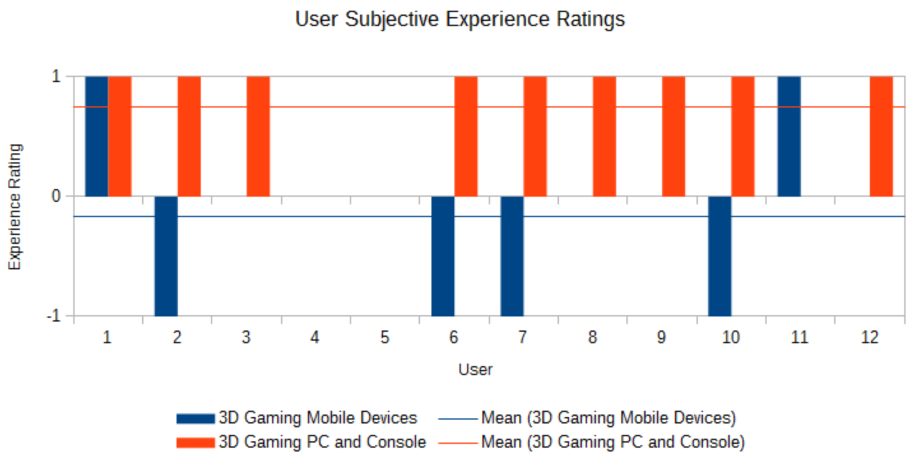


Figure 20: User rated their experience with 3D gaming on PCs and consoles considerably higher than their experience with 3D gaming on mobile devices with means of 0.75 and -0.17 respectively.

Up until now we have only evaluated use of Shoebox VR without looking at performance figures. The possibility remains that, despite low usage and generally negative ratings, use of Shoebox VR did contribute to better performance results. Figure 21 shows no significant improvement or degradation in performance, although use of Shoebox VR tends to vary to a degree between rounds. Note that relative performance is measured as the efficiency of the actual result of the main performance indicator compared to the ideal case. For the “Closest” stage, this is determined only by whether or not the correct package was collected. For the “Timed” stage, the performance is determined by how many packages out of five were collected within the time limit so that collecting two would yield a relative performance rating of 40%. For the “Distance” stage, the ideal case is a travel distance of 200 units or less with a maximum of 400, so that a value of 250 results in a relative performance rating of 75%.

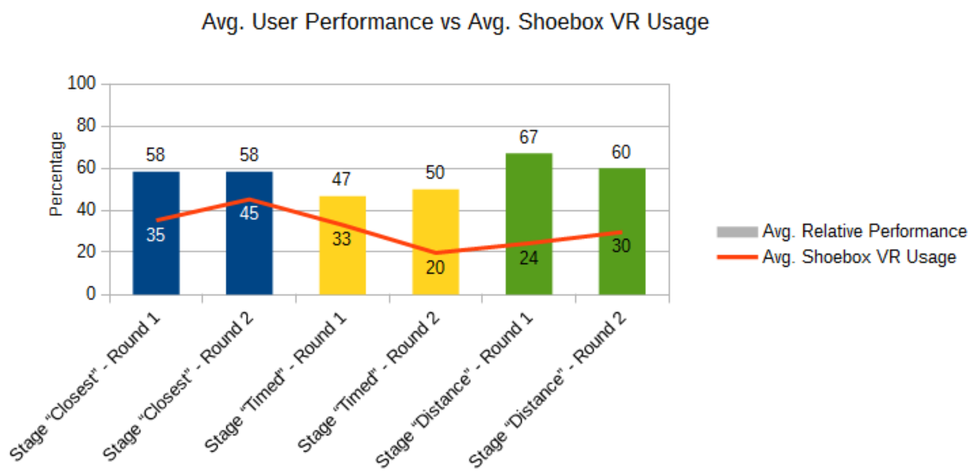


Figure 21: With the exception of the "Closest" stage, the curve shows a slight degrade in performance when Shoebox VR is used more. However, appearances deceive as the results between rounds of the same stage differ less than 10 percent in performance, which can be accounted to the pseudo-randomized generation of the stage and packages. Relative performance is measured as the efficiency of the actual result of the main performance indicator compared to the ideal case.

In spite of not being able to observe a clear trend on the effects of using Shoebox VR on performance, we cannot make any conclusive statements regarding their correlation due to once again being restricted to the limited number of rounds for each stage. As such we identify this is a crucial limitation and make notes on improving the setup for the performance study and expert analysis in this regard.

Based on the additional feedback given by some of the participants, minor issues were identified in the explanation of the navigation controls and user score ranking system, although none of these were considered an obstacle due to the addition of a training stage and the relative insignificance of the user score ranking system. Reevaluating the Shoebox VR usage statistics also revealed that there was not enough information recorded to identify if Shoebox VR was mostly used when stationary or if the users were using it in during motion. Finally, differences in devices were also determined to be relevant when evaluating performance and Shoebox VR use in that the size of the device and display may benefit or hamper the user in his or her ability to spot packages from a distance or use the touch controls whilst tilting the device.

6.5 Conclusion

Considering the results we can state that some of our expectations for the stages and use of Shoebox VR were not unfounded. To reiterate, we expected that for the “Closest” stage users would perform well and make fair use of Shoebox VR. With an average relative performance rating of 58% for both stages this is slightly lower than expected, but is still considered a positive result. We also observe that Shoebox VR usage is highest in this stage at 35% for the first round and 45% for the second, exceeding even the training stage where its use is required to continue. Conversely, we expected users to have difficulty completing the “Timed” stage and argued that, because of the prominent pressure component, Shoebox VR usage would be lowest for this stage. In terms of performance, this expectation came to be true, with relative performance ratings being the lowest of all stages and the only ones being at or below the 50% threshold. Shoebox VR usage also saw a significant drop in its usage during the second round which seemingly transferred to the final “Distance” stage. For this stage, we expected varying performance results and the highest use of Shoebox VR. While individual performance did indeed vary with a minimum distance of 182 units and a maximum of 346 units, Shoebox VR usage was lower than expected. We argue that this is due to the negative experience with the previous stage and the lack of accustomization resulting from only having two rounds to experiment with. As this is essentially true for every stage, we conclude that this setup is not suitable for determining if and how Shoebox VR affects user behavior and performance in open environments that require navigation, especially when considering the limited exposure of the users to Shoebox VR compared to traditional navigation controls.

Although the pilot study was never suitable for a comprehensive evaluation, it did highlight several issues in our setup. Thus we change our approach and limit the implementation to only one challenge stage with a higher number of rounds. For this purpose we select the “Distance” stage to be the best candidate, due to it effectively being an extension of the “Closest” stage and theoretically being the most interesting case because of the practical benefits Shoebox VR can have in this scenario and a relevant pressure component. As a result of there being no indication of Shoebox VR being in any way related to general performance we modify the setup of the performance study to include the evaluation of the presence and absence of Shoebox VR on user behavior and performance. We also consider the variations between handheld devices to be highly relevant for a behavior and performance oriented study and therefore include this component into our setup and evaluation for the main study. Additional modifications include improvements to the data recording in order to collect more detailed behavior data, use of images for explaining navigation controls over textual explanations and a boost to the Shoebox VR's scale to increase its usefulness when tilting by allowing the camera to translate larger distances. These modifications will be detailed further in next chapter covering the main study.

7. Main Performance and Behavior Evaluation Study

The main performance and evaluation study is a continuation of the pilot study, with the additions and improvements to the setup and implementation that will be detailed later in this chapter. However, this study is an offline supervised study and will feature three different mobile devices to be able to evaluate if the properties of these devices have an effect on the Shoebox VR usage or performance of the participants. Due to the focus on collecting quantitative data, discussions and feedback are not the main priority for this study and will thus be limited. Instead, the expert analysis is aimed at collecting detailed observations and discussions for arguing and motivating the results in this chapter and as a source of high quality information with regard to the technical and practical aspects of the use of Shoebox VR. Therefore, although we will discuss and interpret the results in this chapter, we will refrain from excessively speculating, as the expert analysis may provide additional insights into the problems.

7.1 Setup

For this study, we employ the use of three differently sized devices, namely a 4 inch device with a resolution of 854x480 pixels, a 7 inch device with a resolution of 1280x800 pixels and a 10 inch device with a resolution of 1024x600 pixels (Figure 22). Notice that, in spite of the 10 inch device being substantially larger in size compared to the 7 inch device, its resolution is lower. This results in a reduced pixel density, thereby creating a less sharp image compared to the 7 inch device. However, the larger screen may still be useful for detecting detail from larger distances due to the larger pixels. Rendering and computing performance is comparable despite the differences in internal hardware.

As a result of having three devices, it is possible to have up to three users participate at once. However, due to the sounds that are played by the application during the test, the users are separated by means of space and privacy screens in order to keep distractions to a minimum. Additionally, the test was conducted in a dim, but naturally lit room, with minimal sunlight to eliminate the negative effects of glare. Devices were checked and cleaned between uses to ensure proper functionality of both the device and the software.

The participant group consisted of a total of 24 individuals, of which the vast majority was male. None of these participants suffered from serious conditions that may have considerably impaired their vision, muscle movements or ability to understand the pragmatic aspects of the experiment. Out of these 24 individuals, 4 results are from the expert analysis, but this poses no problem or bias as there is no discussion or interruption during the first stage of the expert analysis in which the users are asked to complete the same test. Participants were free to ask questions during the session if they had trouble understanding the textual explanation or use of the controls for Shoebox VR and navigation and were asked to report unexpected behavior of the device or software such as crashes or accidentally minimizing the application.

Figure 22: A side-by-side comparison of the devices, showing differences in terms of screen size, color representation and image clarity.

7.2 Improvements and Modifications

Based on the pilot study, we have determined several flaws and issues with the setup and implementation of the experiment and have corrected these by making the necessary modifications. In the following section, we discuss the issues on an individual basis by motivating the reasons for the change and specifying how it is implemented.

Multiple challenges with each two rounds

Two rounds proved to be insufficient in order to effectively identify and evaluate trends and behavior over time. Although this was anticipated, the number of rounds was kept intentionally low to reduce the time of the total test and thus reduce the risk of users quitting before finishing due to boredom or other responsibilities. Therefore, we choose to have only one challenge stage and extend it by increasing the number of rounds from two to six. This is effectively the same number as before, but by reducing the problem to only one challenge stage, we are much more likely to be able to properly identify trends and behavior such as a learning effect or gradual familiarization. For the challenge objective, we select the “Distance” stage to the best choice, due to its relevance to the “Closest” stage and a more reasonable and relevant pressure component compared to the “Timed” stage with regard to the use of Shoebox VR.

Initial results suggest Shoebox VR has no effect on performance
Although we were unable to make any definitive statements, the initial results showed no indication of a correlation between general performance and use of Shoebox VR. In order to evaluate this in more detail, we disable Shoebox VR in three of the six rounds and compare their results, thus enabling us to do matched pair testing where users are both the sample and control group. This makes more sense than using two separate groups, because the initial results also showed a relatively large variation in use of Shoebox VR among participants. Therefore, it would be difficult to argue for the case of using two separate groups, as individual behavior is already likely to result in different performance and Shoebox VR usage statistics. By having each user act as his or her own control, we can evaluate and express the results more concretely as the differences between the two conditions with and without Shoebox VR. We represent each user's performance and behavior by taking the average of the three rounds in order to reduce the effects of outliers and other extreme cases that may arise from the randomization factor in the generation of the environment and placement of the packages, but do not discard the individual round samples as they are still relevant.

Instead of grouping the rounds with and without Shoebox VR together, we alternate them and randomly select the starting condition to make the users more aware of the change, as consecutive rounds with the same condition will likely result in habituation (decrease in responsiveness upon repeated exposure to a stimulus) that will in turn cause users who would start without Shoebox VR to be less motivated to change behavior as a result of the increased acclimatization to an environment without Shoebox VR. This would not be such an issue if the reverse were also true, but the removal of Shoebox VR is far less likely to have a negative impact on performance due to the navigation being required to complete the stage, whereas Shoebox VR is not. As we are already quite careful in reducing the positive effects and influences of Shoebox VR on navigation and search tasks, we should not further reduce the importance and efficacy of Shoebox VR and therefore eliminate this risk by alternating the rounds with and without Shoebox VR.

In order to accommodate the change in challenge stages, we extend the training stage by an additional round, as the “Closest” stage in the original setup served as an intermediate between the training and actual challenges. We also randomize and repopulate the entire environment with each round, instead of only randomizing the environment for each new challenge stage. Finally, the survey at the end of the test is extended to include the following questions:

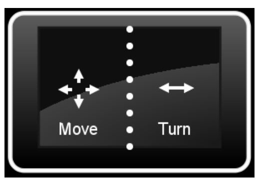
- *How much did you miss being able to use Shoebox VR during the rounds when you could not?*
  This question serves to answer if the user not only noticed, but also experienced the removal of Shoebox VR as a negative effect on their performance or enjoyment.

- *How did the size of your device affect your performance?*
  This question was aimed at determining if the user experienced the size of the device as a positive or negative influence on their ability to find and collect the packages. Instead of the minimum and maximum values being indicated as -2 and +2 as normal, they were changed to “Bad” and “Good” respectively, to eliminate the possibility to interpret the question as “To what degree did the size of your device affect your performance?”.

Shoebox VR use lower than expected limited to 40 degree angles
Although not necessarily a limitation, the overall average use of Shoebox VR at 30% of the round time was considered lower than expected with very tilting angles beyond 40 degrees. Therefore, we considered two options: to either reduce the camera translation of Shoebox VR to force users to tilt farther for the same result, or to increase it. Because the original translation scale at 5.0 was considered to be a fairly conservative value, making it unlikely that a higher value can be considered an exaggeration, we opted for the latter to encourage users to employ its use. The scaling value was changed from 5.0 to 7.5, increasing it by 50%.

Recorded data did not reflect if Shoebox VR was used during movement or only when stationary
We now record the percentage of distance covered while using Shoebox VR in order to be able see if users used Shoebox VR exclusively while stationary or if they were capable and willing to employ it during movement as well.

Textual explanations not always clear
One case that was mentioned more than once was the textual explanation regarding the navigation controls. The text explained how the screen was horizontally divided into two equal segments and that touch interaction with each segment resulted in different behavior i.e. the left side controlled movement and the right side controlled the turning. As this was not fully understood by some of the participants of the pilot study, the explicative text was replaced by the following image to help clarify the mechanics:



Additional Research Question
Due to the inclusion of different devices in our setup, we specify an additional research question specific to this study.

Do the differences in device sizes, specifically a 4 inch, 7 inch and 10 inch device, affect the performance of the search tasks and/or the usage of Shoebox VR in open environments?

However, as a result of having three sample groups with only 8 users each, it is highly improbable that we can prove any statistical significance and will therefore only discuss and elaborate on this problem informally. By doing so, we can still contribute the results, possibly providing motivation for a follow-up research or serving as a reference for future studies. Therefore, we do not phrase hypotheses, but instead make the following predictions:

Prediction 1 (P1):
There will *be a noteworthy difference between behavior with regard to use of Shoebox VR between the devices.*

Prediction 2 (P2):
There will not *be a noteworthy difference in performance with regard to time and distance between the devices.*

We define the level at which we consider the difference to be noteworthy to be at least 20%, as there is likely to be at least some variation in the results due to individual preference and prior experiences.

7.3 Results

Before we start evaluating results, let us first take a look at the experience levels of the participants in the main study (Figure 23). If we consider that the pilot study had average ratings of -0.17 for the mobile devices category and 0.75 for the PC and console category, it becomes clear that we are dealing with the same user demography. The only thing that stands out is the relatively high rating from the 10 inch group for the mobile devices category, but this is likely just a coincidence as devices were handed out randomly to the participants for each test.

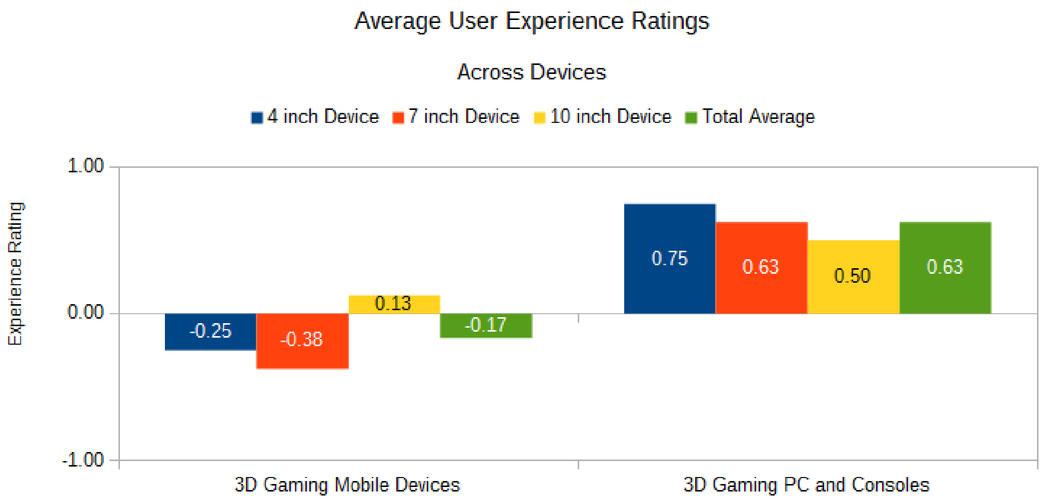


Figure 23: Average experience ratings are similar to the pilot study and represent our target demography fairly well. Further observation shows the almost linear decrease of experience ratings for the 3D gaming on PC and consoles category and a relatively high average rating for 10 inch device with regard to experience of 3D gaming on mobile devices.

Our research questions focus on general Shoebox VR behavior and performance indicators based on distance and time. Note, however, that time is considered a secondary factor, because there is no time limit and users are encouraged to plan carefully rather than finish early. Nevertheless, time can still be representational for certain behaviors e.g. using Shoebox VR only while stationary may increase the time, but in turn decrease distance. In the next sections, we will evaluate performance and behavior data in order to identify if such trends or patterns exist. We also perform significance analysis on the data in order to determine if any differences in the results between the rounds with and without Shoebox VR are sufficiently significant to draw conclusions from. Only then will we be able to reject a null-hypothesis and accept the corresponding alternate hypothesis.

Behavior Analysis

Our first research question focuses on Shoebox VR behavior and asks if users are encouraged to employ its use during the search tasks. As the results from the pilot study were lower than we had originally anticipated, we increased the translation strength of Shoebox VR by 50%, allowing it cover more ground. Therefore, we expected an increase in the use of Shoebox VR, but the data from Figure 24 shows nearly identical results. This may suggest that use of Shoebox VR is not appealing enough, either visually or functionally, for users to use it more than a third of the average round time. A possible explanation for this could be that navigation using the touch controls is simply considered more convenient or powerful by the users, even under the conditions where travel distance is to be minimized. Alternatively, Shoebox VR might have only been used during the moments where finding and locating the package is top priority, whereafter its use is subsequently discarded when a target for collection has been selected. If this is indeed the case, it is likely that Shoebox VR is used almost exclusively while stationary. However, Figure 25 shows that an approximate average of 27% of the total round distance is covered using Shoebox VR, suggesting that this theory may only be true for a certain number of users or that it simply is not the case.

In order to find out more, we've compiled a table that displays the average relative use of Shoebox VR per user and their ratings on Shoebox VR ease of use, visual effect and usefulness (Table 1). Note that results are color coded and represent an improvement (green) or declination (red) compared to the average values, with the exception of the user ratings where the colors represent a positive or negative rating respectively.

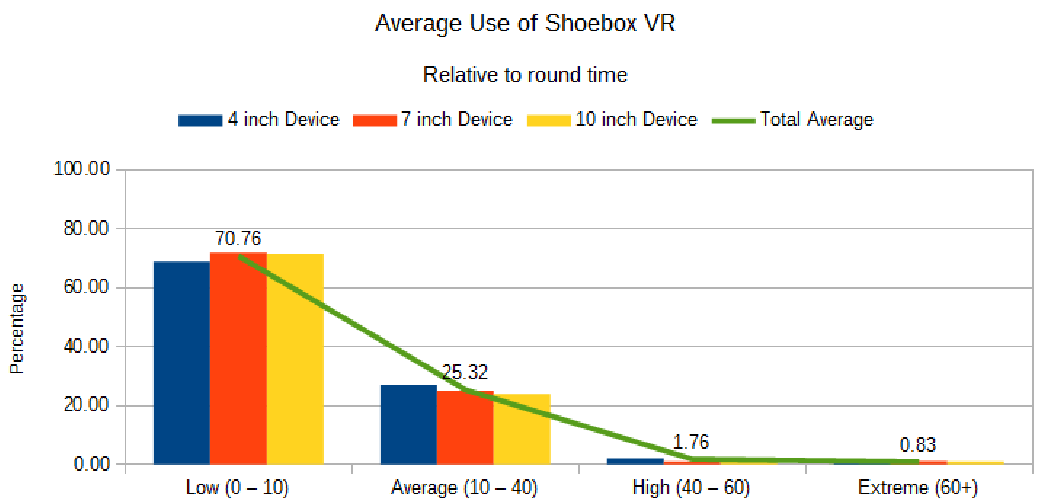


Figure 24: Shoebox VR use is roughly the same as during the pilot study, averaging at around 30% of the round time. Note that rounding errors of the recorded data result in a total sum of less than 100%.

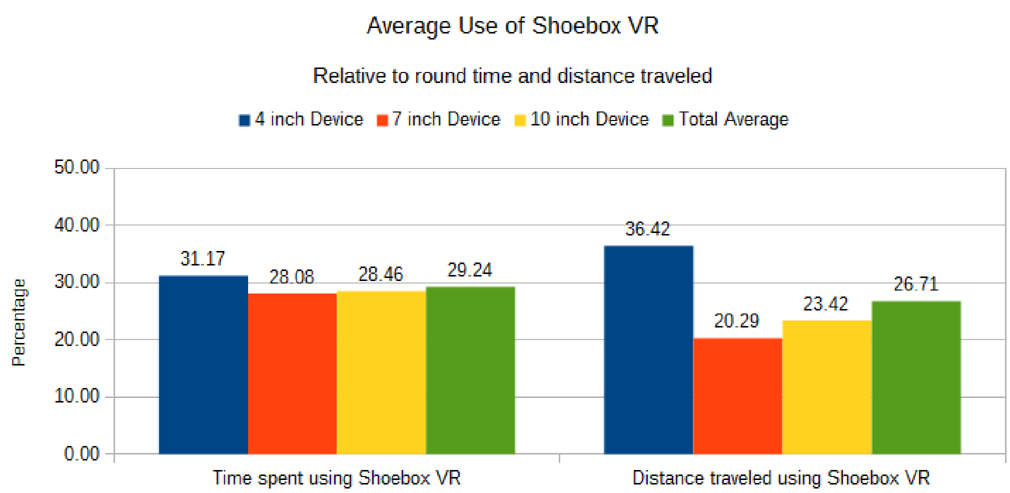


Figure 25: Relative use of Shoebox VR averages at around 30%, falling 20% short of our minimum threshold value of 50%. Values on Shoebox VR use during motion indicate that an average of approximately 27% of the total round distance is traveled while using Shoebox VR, with the 4 inch device category averaging more than 10% higher than the other categories.

4 inch									Averages
Average round time use of Shoebox VR	24%	3%	33%	51%	49%	31%	31%	27%	*31%*
Average distance covered with Shoebox VR	40%	51%	63%	52%	22%	23%	23%	19%	*36%*
Ease of use rating Shoebox VR	1	0	0	-1	2	2	-1	1	*0.50*
Visual effect rating Shoebox VR	-2	-2	-1	0	1	2	0	2	*0.00*
Usefulness rating Shoebox VR	-2	-2	0	0	2	1	-1	1	*-0.13*
7 inch									
Average round time use of Shoebox VR	26%	52%	44%	20%	7%	45%	26%	5%	*28%*
Average distance covered with Shoebox VR	22%	15%	56%	11%	2%	35%	21%	0%	*20%*
Ease of use rating Shoebox VR	0	1	0	1	1	-1	1	2	*0.63*
Visual effect rating Shoebox VR	-1	-1	-1	1	-1	-1	1	2	*-0.13*
Usefulness rating Shoebox VR	0	-1	-2	1	-1	-1	1	1	*-0.25*
10 inch									
Average round time use of Shoebox VR	2%	57%	54%	38%	38%	22%	10%	19%	*30%*
Average distance covered with Shoebox VR	35%	28%	43%	38%	25%	13%	0%	7%	*24%*
Ease of use rating Shoebox VR	1	-1	1	0	1	1	0	1	*0.50*
Visual effect rating Shoebox VR	1	-1	2	-1	1	1	1	-1	*0.38*
Usefulness rating Shoebox VR	0	-1	1	1	0	2	0	-1	*0.25*

Table 1: The average relative use of Shoebox VR per user and their ratings on the ease of use, visual effect and usefulness of Shoebox show no indication of patterns or trends, suggesting that there is no correlation between the users' use of Shoebox VR and their opinion of it. The only fairly consistent factor is the ease of use, with a positive average rating in all three device categories.

The data from Table 1 fails to show any notable trends or patterns that may indicate if the perceived usefulness or ease of use of Shoebox VR has had an effect on the average usage of Shoebox VR during the tasks. There is also no indication of the visual effect of Shoebox VR having any influence on the users' behavior with regard to its use. Therefore, we cannot say if the lower than expected usage of Shoebox VR is due to the complexity of the control scheme or lack of a practical use for solving the tasks. As a result, we look to the expert analysis in chapter 9 to provide more concrete details on what the cause may be.

Our first null-hypothesis (NH1) states that users are not encouraged to employ its use and we defined an average minimum relative usage value of 50% for this be considered false. With a total average of 30%, we are well outside the questionable range limit and therefore will not be able to reject null-hypothesis 1 (NH1). Furthermore, the distribution as shown in Figure 26 is not suitable for z- or t-testing, making it very difficult to determine the statistical significance of the data. Therefore, we cannot make definitive claims on whether or not the addition of Shoebox VR encourage users to employ its use for the challenges, although the results suggest that the null-hypothesis is likely to be true.

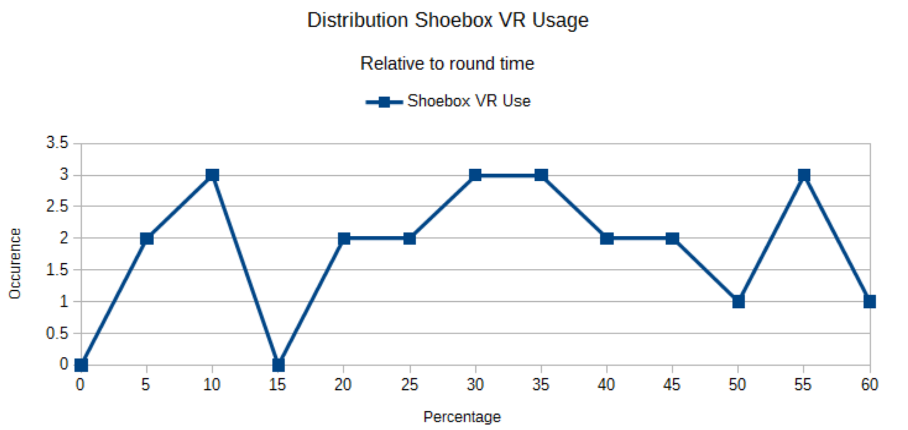


Figure 26: The distribution deviates quite firmly from a normal distribution, with multiple modes across the x-axis. The mean for this distribution is 29.24%, with a median of 29.17% and a standard deviation of 17.5%.

In terms of differences between devices, we see that subjective ratings on Shoebox VR are highest for the 10 inch device category, possibly due to a better experience or greater sense of immersion resulting from the larger screen, and that the average distance covered with Shoebox VR is highest for the 4 inch device category. The latter can be explained by reasoning that the smaller size of the 4 inch makes it easier to tilt larger angles compared to the larger devices. Based on these ratings so far, it would seem that the 4 inch device is best suited for using Shoebox VR on.

A final consideration to make is how the use of Shoebox VR transitioned over the three rounds in which it was enabled. Due to the starting condition being randomly selected, it is possible that users who started with Shoebox VR off for the first round were less motivated to continue using it throughout the other rounds. Conversely, users that started their first round with Shoebox VR enabled may feel more comfortable with its use and experience the loss of Shoebox VR as more of an issue than others.

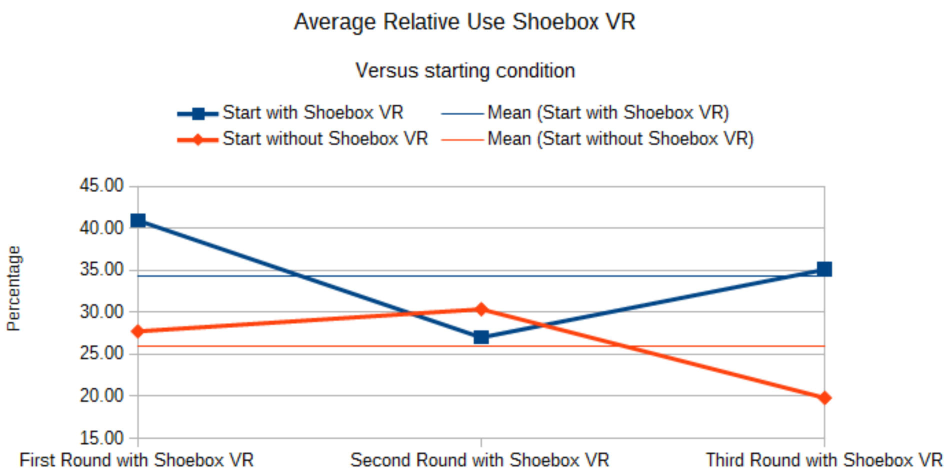


Figure 27: Starting condition does appear to have a slight influence on the continuous use of Shoebox VR with means of 34.08% for starting with Shoebox VR and 25.97% for starting without Shoebox VR.

Results from Figure 27 suggest that there may indeed be an effect caused by the starting condition, with the average being almost 10% higher for the group that started with Shoebox VR enabled. Also worth noting is how the third round for the starting condition without Shoebox VR is almost 15% lower than the second round whereas the transition between the first and second round resulted in an average increase of Shoebox VR usage. This may indicate that the users tried to use it more, but eventually felt like it was holding them back or wasting their time. The opposite appears to be true for the starting condition with Shoebox VR, where the use in the third round experiences a boost compared to the drop in the second round. This could suggest that users were initially unsure of the effects and benefits of Shoebox VR and tried to complete a round by relying on it less, but found its use to be more appealing or useful than expected. Ratings on whether or not the users missed being able to use Shoebox VR during the rounds in which it was disabled suggest that the appeal or perceived benefits of Shoebox VR were not experienced on a conscious level or are the results of chance (Table 2). Additionally, there is no visible consistent relation between the rating and the average use of Shoebox VR during the rounds.

Start with Shoebox VR													***Averages***
Average round time use of Shoebox VR	24%	51%	31%	27%	26%	52%	44%	20%	45%	5%	57%	38%	*34.08%*
Shoebox VR missed rating	-1	0	1	1	-1	-1	-1	1	-2	1	-2	-2	*-0.50*
Start without Shoebox VR													
Average round time use of Shoebox VR	3%	33%	49%	31%	7%	26%	2%	54%	38%	10%	19%	22%	*25.97%*
Shoebox VR missed rating	-2	-2	2	-1	-2	-2	-2	2	0	0	-2	1	*-0.67*

Table 2: Users that started with Shoebox VR enabled generally had higher usage values, although they generally did not miss Shoebox VR during the rounds where it was disabled. This suggests either an unconscious liking or perceived benefit from using Shoebox VR, or merely a coincidental development.

Performance analysis : Time

Our second research question focuses on the effects of Shoebox VR on time, so that we can evaluate if use of Shoebox VR can help the search for the packages and thereby reduce the time it takes to complete the round. In actuality, however, it is more likely that use of Shoebox VR under the circumstances of our experiment will lead to an increase of time, due to the importance of careful observation and path planning. Nevertheless, it is possible that use of Shoebox VR may help in spotting and locating packages and thus contribute to lower average round times.

One of the relations we are interested in is if Shoebox VR has a positive or negative effect on the round time and if use of Shoebox VR during motion may further influence this. Based on the results from Figure 28, there appears to be a slight indication that suggests that use of Shoebox VR increases the average round times. This is most noticeable for the 4 inch and 10 inch device categories and is likely to originate from their size differences. For example, the smaller screen on a 4 inch device may make it more difficult to spot and locate packages, thereby resulting in an increase in time due to requiring more effort to verify the existence and location of a package or simply not noticing them from large distances. Conversely, the large screen may aid in spotting and locating packages, so that less time is needed to find them all. This may explain why the results for the 7 inch device is relatively neutral, as it does not suffer nor benefit as much from its screen size.

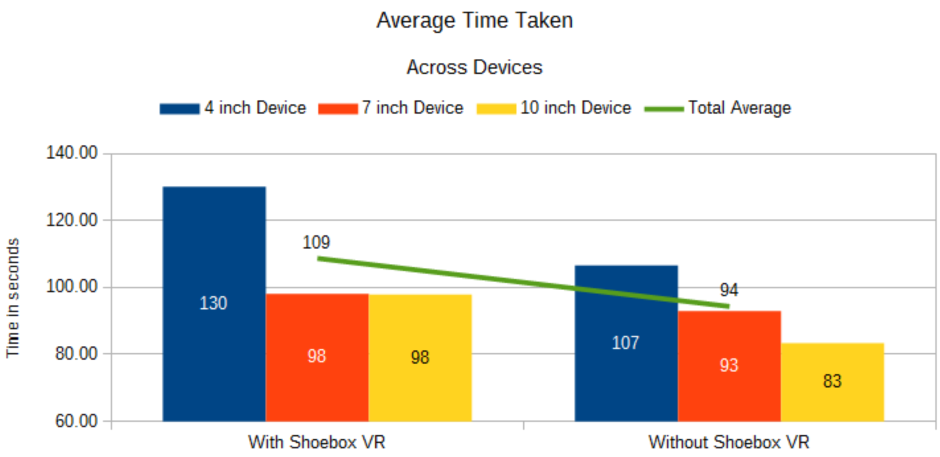


Figure 28: The 4 inch device category is most affected by the use of Shoebox VR. We also observe that averages are generally lower during the rounds where Shoebox VR is disabled, with the 10 inch category experiencing the second largest reduction in times and being the overall lowest.

It is understandable that use of Shoebox VR can increase the round time, as it takes time and effort to use Shoebox VR for exploration and observation. Logically speaking, this mostly applies to use of Shoebox VR when searching for a new target when stationary. Therefore, if the user is always one the move, time normally spent observing is no longer lost, thereby creating a case where a reduction in time is more plausible. However, this method of using Shoebox VR requires considerate skill and is very difficult to use effectively, making it unlikely that such cases exist in our tests. Based on the averages, Table 3 shows a slight trend where using Shoebox VR during motion appears to have a negative effect on the average time, which could also be seen to a limited extent in Figure 25. However, looking at individual cases, the effect is not consistent as there are cases in all category groups where the time both increases and decreases as a result of using Shoebox VR during motion at around or above average values. It is possible that the randomization in the environments may have caused this effect or that the difference in user skills is larger than expected. We will require the expert analysis to provide more information on this topic to help determine the cause for this variation in behavior.

4 inch									Averages
Average difference in time	-54	29	20	82	103	-18	26	1	23.58
Average round time use of Shoebox VR	24%	3%	33%	51%	49%	31%	31%	27%	31%
Average distance covered with Shoebox VR	40%	51%	63%	52%	22%	23%	23%	19%	36%
Shoebox VR Ease of Use	1	0	0	-1	2	2	-1	1	0.50
Shoebox VR Visual Effect	-2	-2	-1	0	1	2	0	2	0.00
Shoebox VR Usefulness	-2	-2	0	0	2	1	-1	1	-0.13
7 inch									
Average difference in time	-27	-11	0	13	-9	23	3	8	0.04
Average round time use of Shoebox VR	26%	52%	44%	20%	7%	45%	26%	5%	28%
Average distance covered with Shoebox VR	22%	15%	56%	11%	2%	35%	21%	0%	20%
Shoebox VR Ease of Use	0	1	0	1	1	-1	1	2	0.63
Shoebox VR Visual Effect	-1	-1	-1	1	-1	-1	1	2	-0.13
Shoebox VR Usefulness	0	-1	-2	1	-1	-1	1	1	-0.25
10 inch									
Average difference in time	55	-11	-3	51	27	8	-1	-12	14.46
Average round time use of Shoebox VR	2%	57%	54%	38%	38%	22%	10%	19%	30%
Average distance covered with Shoebox VR	35%	28%	43%	38%	25%	13%	0%	7%	24%
Shoebox VR Ease of Use	1	-1	1	0	1	1	0	1	0.50
Shoebox VR Visual Effect	1	-1	2	-1	1	1	1	-1	0.38
Shoebox VR Usefulness	0	-1	1	1	0	2	0	-1	0.25

Table 3: Time values show no correlation to use of Shoebox VR or subjective user ratings. The difference in time is measured between rounds without and with Shoebox VR, so that a negative value indicates an improvement (reduction) in the total average round time with relation to use of Shoebox VR.

Null-Hypothesis 2 (NH2) states that use of Shoebox VR will not reduce the average round times by 20% or more. Given that our results show an increase in time, we have no evidence that may suggest otherwise. Note that this does not mean we can prove the null-hypothesis, only that we cannot reject it. Compared to the statistics on use of Shoebox VR, the distribution of average time samples shows more resemblance to a normal or t-distribution (Figure 29). Although we only consider the averages as the representative values, we have also included the distribution of all time samples as a reference (Figure 30).

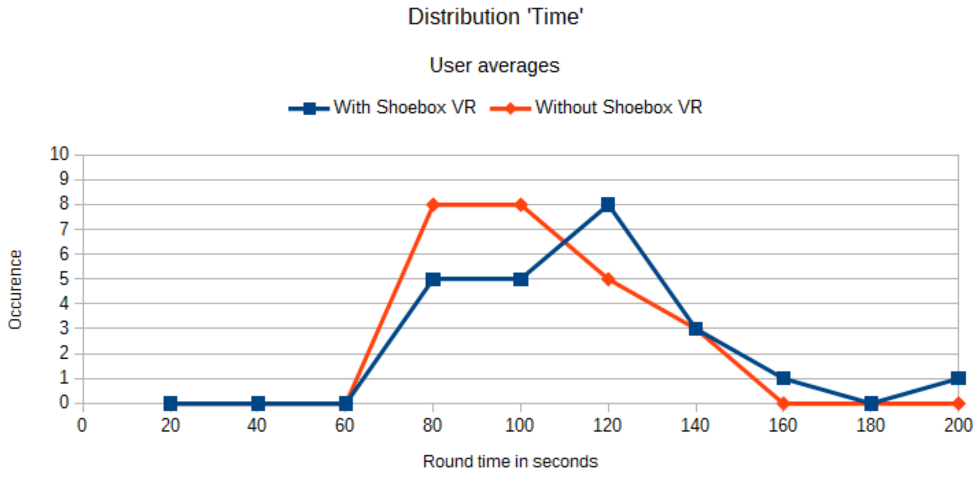


Figure 29: The distribution of average time values varies slightly between groups, illustrating their difference in means. The average time values with Shoebox VR have a mean of 105.85 seconds, a median of 106.33 seconds and a standard deviation of 39.03 seconds. The average time values without Shoebox VR have a mean of 93.24 seconds, a median of 92.67 seconds and a standard deviation of 21.18 seconds.

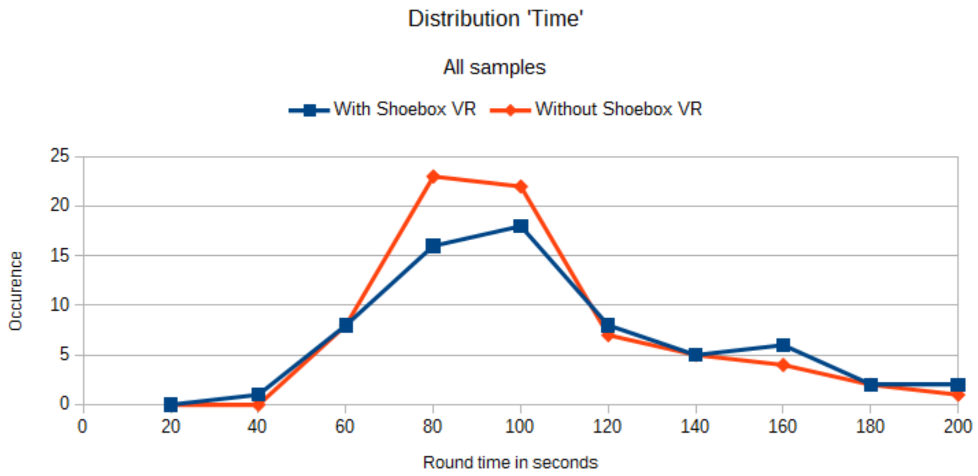


Figure 30: The distribution of all samples shows a smoother shape and further emphasizes the larger right tail. The time samples with Shoebox VR have a mean of 110.57 seconds, a median of 93.00 seconds and a standard deviation of 52.46 seconds. The time samples without Shoebox VR have a mean of 91.46 seconds, a median of 85.00 seconds and a standard deviation of 30.93 seconds.

Using a matched pair t-test on the average time values reveals a t-value of 0.03, which translates to a one-tailed p-value of 0.488 with a standard deviation of 30.53 and degrees of freedom at 23. Therefore, the probability of this result, assuming the null hypothesis, is 48.8% which does not allow us to reject the null-hypothesis.

Performance analysis : Distance

Our third research question aims to answer if Shoebox VR can improve performance with regard to path planning i.e. if use of Shoebox VR can help users pick better targets and thereby reduce their distance traveled. However, results in Figure 31 show no variation in distance whether Shoebox VR is enabled or not, which is somewhat surprising seeing as how there was a slight difference in average time values. What is also very interesting is the variation between distance with the 4 inch device category once again exceeding the total averages. In combination with the higher time averages for this device category, it would seem that the smaller screen causes the majority of the users to have difficulty with locating the packages. Perhaps they were aware of their overall poorer than average performance and were thus more motivated into using Shoebox VR in hopes it would increase the fruits of their efforts. Alternative, the relative ease at which they were able to use Shoebox VR may have lead them to be less careful or efficient in its use. Finally, there is also the possibility that use of Shoebox VR on smaller devices is simply ineffective for the purpose of finding small items in the environment despite best efforts. Given that use of Shoebox VR introduces noise in the form of motion even with the smallest of wrist or hand movements, it is not that unreasonable to assume that this noise makes it more difficult to spot the relatively tiny packages on the already small screen. Regardless of the cause, the lack of improvement in the other device categories suggests that spending more time does not achieve a reduction in distance traveled, even if logically speaking more careful observation should result in better path planning.

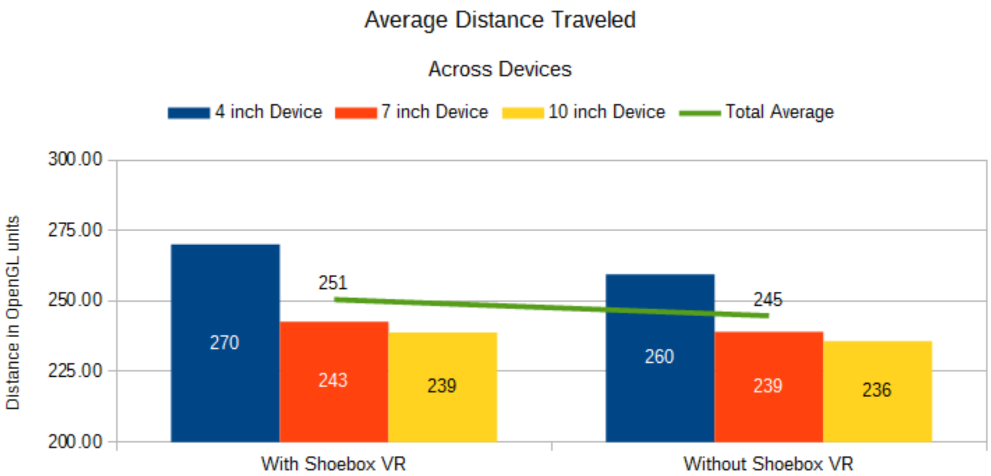


Figure 31: Distance performance between the device categories is most noticeable for the 4 inch device category as it performs relatively poorly compared to the other categories. In spite of this, it would appear that Shoebox VR has no direct effect on distance whatsoever, with averages with and without Shoebox VR being incredibly close.

Looking at the values of Table 4, we do not detect any consistent pattern or trend that suggests that the differences in distances are somehow affected or influenced by use of Shoebox VR during motion or the users' subjective ratings. In short, based on the result, Shoebox VR is not likely to affect path planning efficiency.

									Averages
4 inch									
Average difference in distance	-244	117	2	146	57	-45	28	23	10.58
Average distance covered with Shoebox VR	40%	51%	63%	52%	22%	23%	23%	19%	36%
Shoebox VR Ease of Use	1	0	0	-1	2	2	-1	1	0.50
Shoebox VR Visual Effect	-2	-2	-1	0	1	2	0	2	0.00
Shoebox VR Usefulness	-2	-2	0	0	2	1	-1	1	-0.13
7 inch									
Average difference in distance	-18	7	1	-28	-27	1	38	23	-0.37
Average distance covered with Shoebox VR	22%	15%	56%	11%	2%	35%	21%	0%	20%
Shoebox VR Ease of Use	0	1	0	1	1	-1	1	2	0.63
Shoebox VR Visual Effect	-1	-1	-1	1	-1	-1	1	2	-0.13
Shoebox VR Usefulness	0	-1	-2	1	-1	-1	1	1	-0.25
10 inch									
Average difference in distance	63	-39	-3	90	6	-44	-32	-15	3.29
Average distance covered with Shoebox VR	35%	28%	43%	38%	25%	13%	0%	7%	24%
Shoebox VR Ease of Use	1	-1	1	0	1	1	0	1	0.50
Shoebox VR Visual Effect	1	-1	2	-1	1	1	1	-1	0.38
Shoebox VR Usefulness	0	-1	1	1	0	2	0	-1	0.25

Table 4: The 4 inch device category has the highest average increase in distance, but the relative difference is so small compared to the mean of 250 (less than 2.5%) that it has virtually no significance. There also does not appear to be any correlation between the differences in distances and how much Shoebox VR was used during motion or user ratings on Shoebox VR.

We phrased null-hypothesis 2 (NH3) as the inability of Shoebox VR to reduce the distance traveled and specified a minimum reduction value of 20% as the threshold at which we consider the reduction significant. With no signs of a reduction (or an increase for that matter), we have no evidence to support our alternative hypotheses, which states that there is a reduction as a result of using Shoebox VR. Similar to the time evaluation, we provide distributions of the average distance values and the total set of samples (Figures 32 & 33).

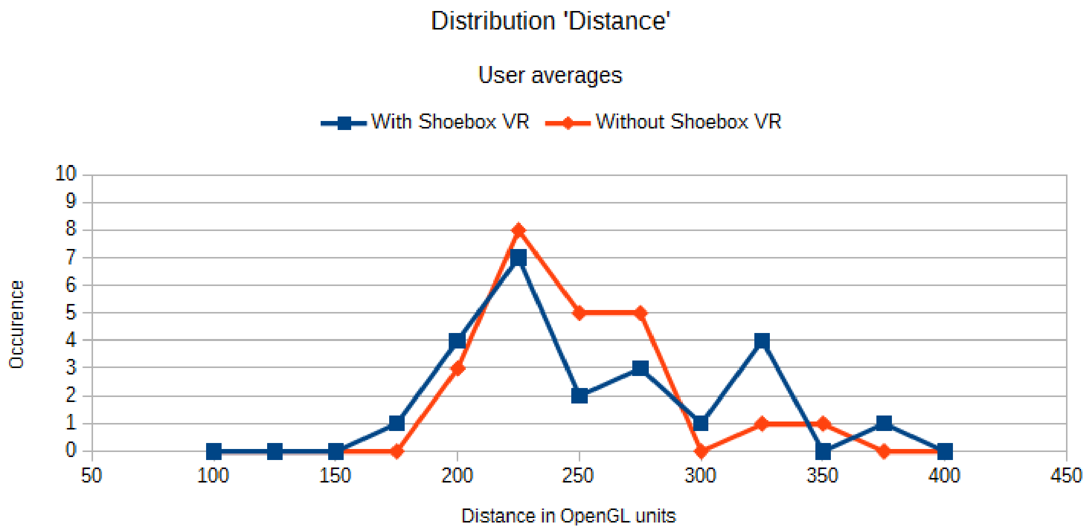


Figure 32: The distribution of the average distance values also shows a larger right tail. Additionally, the distribution of the averages with Shoebox VR are much less evenly distributed than those with Shoebox VR. The average distance values with Shoebox VR have a mean of 259.58 units, a median of 219.33 units and a standard deviation of 60.85 units. The average distance values without Shoebox VR have a mean of 245.15 units, a median of 238.00 units and a standard deviation of 56.32 units.

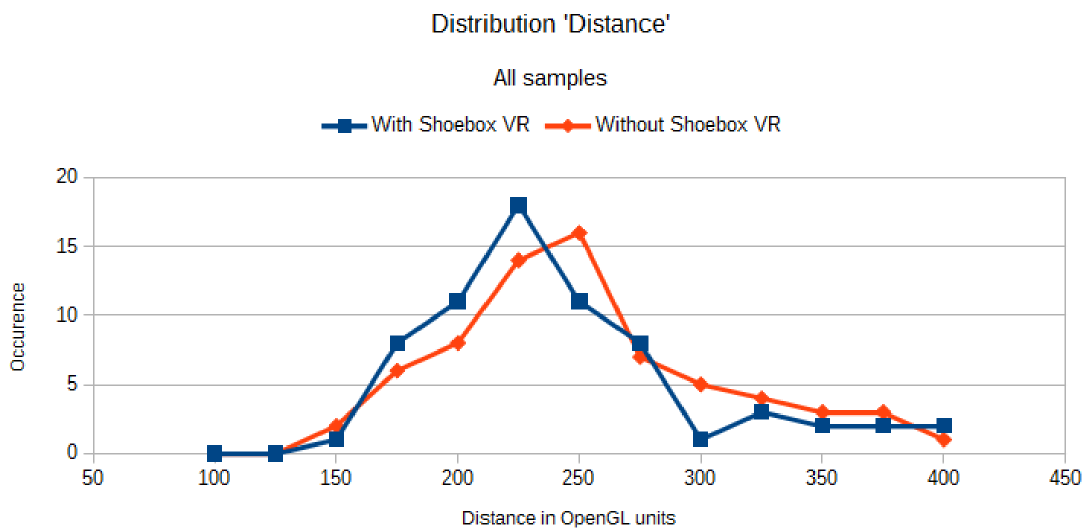


Figure 33: The distribution of all samples shows a smoother shape for distance as well. The distance samples with Shoebox VR have a mean of 246.99 units, a median of 220.50 units and a standard deviation of 88.21 units. The time samples without Shoebox VR have a mean of 252.08 seconds, a median of 233.50 units and a standard deviation of 89.11 seconds.

Using a matched pair t-test on the average distance values reveals a t-value of 0.35, which translates to a one-tailed p-value of 0.364 with a standard deviation of 58.07 and degrees of freedom at 23. Therefore, the probability of this result, assuming the null hypothesis, is 36.4% which does not allow us to reject the null-hypothesis.

User rating evaluation

Based on the developments thus far, it would seem as if Shoebox VR does not contribute to a better performance with regard to time or distance, but perhaps users may subjectively feel that it does, or that the overall experience is considered more enjoyable with the addition of Shoebox VR.

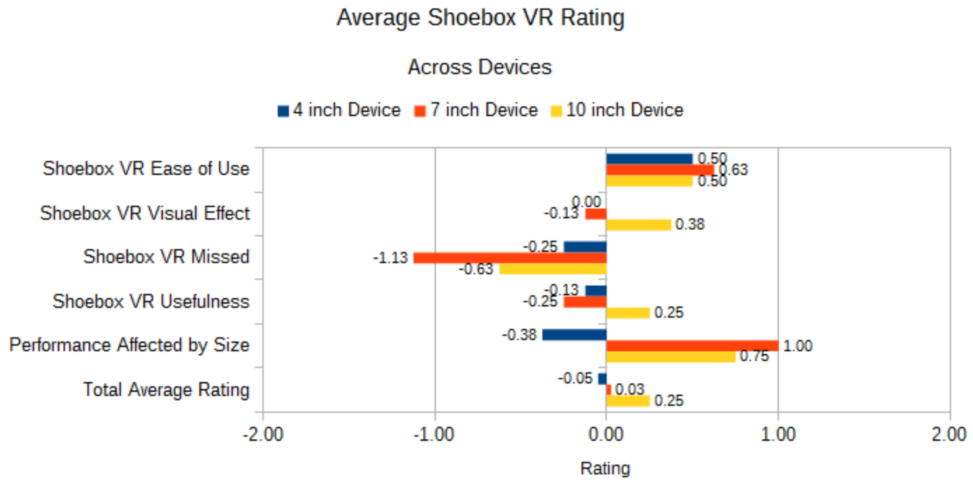


Figure 34: Ratings on Shoebox VR are substantially more positive than in the pilot study, with only the "Shoebox VR Missed" rating being universally poorly rated. For all other topics, the 10 inch device category results in exclusively positive ratings ≥ 0.25, indicating that the experience was generally best on the 10 inch device despite its lower pixel density compared to the other two devices.

The ratings in Figure 34 show fairly positive ratings, with the ease of use rated positive for all device categories. This may suggest that the difficulty and complexity of the control scheme was overestimated or that Shoebox VR is at least perceived as a simple mechanic. Also worth noting are the fairly high values for “performance affected by size” for the 7 inch and 10 inch devices, with the 4 inch device being extremely negative in comparison. This confirms that users did have a fair amount of difficulty using the 4 inch devices and explains the poorer results from this category. Based on the negative ratings of “Shoebox VR missed”, we also notice that users did not experience the removal of Shoebox VR as a negative influence on performance or appeal. Opinions on the Shoebox VR's visual effect vary, with the 7 inch and 10 inch device categories being opposed to one another and the 4 inch device resting at the neutral value of zero. We observe that the usefulness of Shoebox VR is rated fairly neutral, suggesting that the increase in translation scale of the camera did have a positive effect on the users' experience when compared to the same ratings of the pilot study. Finally, we recall that the average experience levels for mobiles were rated rather high for the 10 inch device category, which could subsequently have had an effect on the positive ratings.

Our fourth and fifth null-hypotheses (NH4 & NH5) state that users do not experience the addition of Shoebox VR as being beneficial to either their performance or personal enjoyment, with an average rating of +0.75 being required to be considered a positive impression. Not considering the total averages, the only topic that approaches or exceeds this threshold is that of “Performance affected by Size”, which is not necessarily directly related to Shoebox VR itself. Therefore, we argue that the overall impression is not a positive one, although we did see an improvement with regard to the ratings from the pilot study. As the rating values do not form a clear normal distribution (Figure 35), we cannot verify the statistical significance of the data. Regardless, it is safe to say that the overall ratings were lower than anticipated.

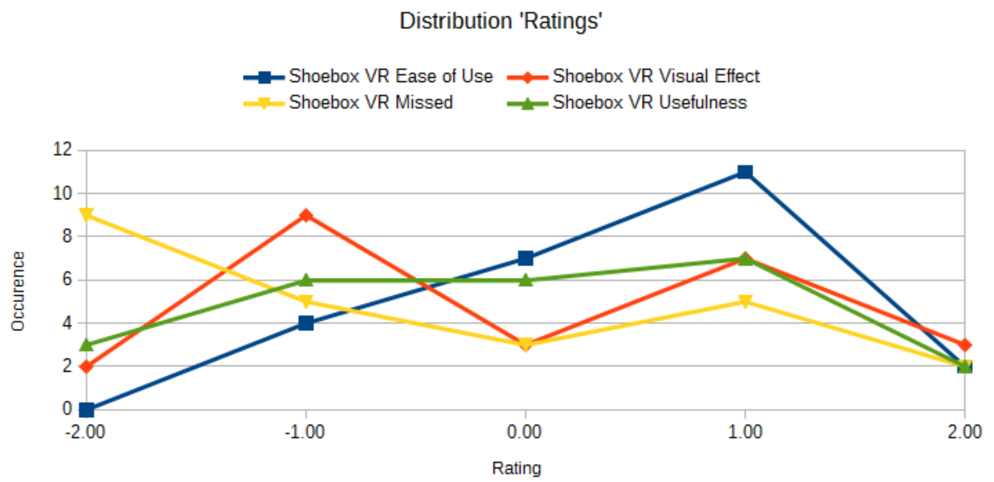


Figure 35: The distributions of relevant ratings are extremely skewed or do not resemble a normal or t-distribution.

Additional feedback and observations

When the test was completed, a short discussion often followed on how the users felt about Shoebox VR and its role in the experiment. A fairly commonly reported issue was that users expected the visual effect to be mirrored i.e. that the change in direction resulting from off-axis projection should have gone the opposite way. This effect has been reported in earlier studies on Shoebox VR as well and may be the result of the experience of the users with 3D games. To elaborate, some users prefer to to invert the x- or y-movements of their mouse or are used to a different scrolling system where dragging a scrollbar can cause the view to go up or down, depending on how you interpret the gesture. Therefore, it is possible that the users who reported the issue may have had other preferences or experiences with certain types of gesture interactions. Another possibility could be that these users were not fully able to relate the concept of Shoebox VR to that of an actual box and failed to understand the motions that result from tilting the device.

A second issue reported by some users concerned the difficulty of holding the device still enough to eliminate jitter and unintentional shifts in the view, especially with regard to the vertical axis. Although this problem is inherently related to the accuracy and responsiveness of Shoebox VR, the fact some the vertical motion was experienced as an annoyance can be largely attributed to its lack of function in the test. For the touch controls, we had disabled vertical turning to simplify the controls and to prevent disorientation or confusion that may have arisen by incorrect use. However, vertical motions were not disabled for Shoebox VR. Because of this, some users were annoyed by having to hold the device in place in order for Shoebox VR to not disrupt the view by looking too far up or down. Therefore, it may be worth considering disabling the vertical axis if it provides no functional use to the task, as otherwise it may be experienced as distracting or bothersome.

An interesting observation made during one of the tests revealed unexpected use of the 4 inch device where one hand was used to hold and support it while the other was used for movement and turning controls. Part of the reason why a virtual gamepad was implemented was so that users would be able to move and turn by using their thumbs, and tilt the device by means of wrist movements. Considering the demography's high experience with 3D gaming on PC or consoles, we had expected them to hold and use the device as they would a gamepad, so this observation came as somewhat of a surprise. Nevertheless, it is still possible to complete the task effectively, but use of Shoebox VR is made considerably more difficult by holding and using the device in this way.

When asked if users felt like it was difficult to use Shoebox VR in combination with the touch controls, responses varied. Some users explained that they had hardly used Shoebox VR to begin with and thus were not troubled as much, whereas others argued that it was simply not that difficult. When asked if they found Shoebox VR useful for finding the packages, most responses were fairly positive, although they were often conditional. For example, a small number of users stated that they only used Shoebox VR at the beginning of the round and when they could not find the last remaining package, whereas others reported using it exclusively during pauses in movement. There were also users who were generally not under the impression it helped them due to the ability to move and turn by simply using the touch controls. These varied responses show that not all users are equally fond, skilled or knowledgeable on the use of Shoebox VR for the search tasks, which in turn affects their behavior and performance data.

7.4 Conclusion

Due to the many factors that were involved in the setup and implementation of the experiment, we have set many boundaries and restrictions to limit our scope. Therefore, our experiment can only be considered an initial step towards evaluating head-coupled perspective behavior and performance in open environments and determining if this topic is interesting with relation to exploring the potential of Shoebox VR. However, results suggest very little effect of the use of Shoebox VR on search task performance with regard to time and distance in spite of careful considerations. This may be due to several reasons, such as the environment not being suitable enough for Shoebox VR or that the navigation controls for movement and turning are simply more useful and powerful compared to Shoebox VR and thus preferred for navigation in open environments. It is also possible that, due to being implemented on mobiles, Shoebox VR is generally less suitable for 3D exploration than head-coupled perspective models on other platforms because of the smaller screen size and difference in interaction methods. We will review these questions and uncertainties in the expert analysis and further inquire on the differences between the devices and the subjective impression of Shoebox VR on user performance and enjoyment to find more concrete answers to these problems.

Reviewing our predictions of P1 and P2, we determine that there is a difference in both behavior and performance in our results. We motivate that this is likely to be true for most cases by referring to the lower than average performance values of the 4 inch device category and the largely positive user ratings for the 10 inch device category. Therefore, although we cannot claim that this evidence is in any way conclusive, we answer the research question of this chapter by stating that there is a high probability of there being differences in behavior and performance with regard to Shoebox VR as a result of using different devices. We conclude by recommending future studies to carefully consider the importance and effects of selecting the hardware for their experiment, as initial results suggest that they may have an effect on how Shoebox VR is used and perceived by the users.

8. Expert Analysis

The expert analysis asked experts in the fields of 3D graphics and virtual environments to perform the test on all three devices in order to provide detailed and in-depth feedback on the experiment and the conditions to which the users are exposed. The aim of this study was to collect detailed qualitative feedback so that we may further explain the performance and behavior results from chapter 7.

8.1 Setup

The expert analysis test was taken by a total of four individuals, of which three were highly knowledgeable and experienced in the use and exploration of 3D graphics and environments. The fourth expert had great experience in use of a variety of mobile devices and thus was able to provide detailed feedback on general usage and considerations when using multiple devices. A session with three individuals during the main study was discussed to such a degree that it is comparable to the setup of our experiment analysis, so that we consider this session as a fifth expert. The setup of the analysis bears some resemblance to a heuristic evaluation as defined by Nielsen, et al. (1994) in that the expert users were instructed to complete the main experiment by themselves first, before engaging in a discussion. Where a heuristic evaluation is normally used to identify potential issues and oversights in user interface designs, we use it to get detailed qualitative feedback on use of Shoebox VR in open environments [30]. However, it should be noted that this is a purely informal test and that we cannot guarantee that the feedback and observations recorded during this analysis are representative of the general case. We merely wish to acquire more insight into the smaller details of the use of Shoebox VR that would normally not become noticeable during a more formal user study.

Each session begins by randomly selecting and presenting the expert user one of the three devices and having them complete the test independently. During this initial stage, the user is encouraged to make observations and notes on their findings, and to report them verbally so that the observer on site can record them. Note that this is not a dialog and that the user may not communicate directly with the observer. Only when the test has been completed may a discussion be started regarding the observations that were made by the user during the test. Once this has been completed, the user is given one of the remaining two devices and asked to complete a number of rounds while again making observations. By doing so, the user directly experiences the differences between the devices and is thus able to better form a more detailed opinion and impression on their usage. During this stage, discussion is allowed freely and the user is not required to complete the entire test a second time. Finally, the last device is tested in the same manner before discussing and commenting on conclusive statements. Depending on the length of the discussions and the speed at which the participant completes the tests, a session may take up to one or two hours.

The order in which the devices are selected is intentionally varied to reduce the risk of forming the same opinions. For example, if all users are given the devices in the same order, the probability of receiving the same feedback also becomes higher, as the experience between these transitions is more likely to be similar. Therefore, out of the four users, two started with the 4 inch device, one with the 7 inch device and one with the 10 inch device. The fifth expert study consisting of a total of three users did not follow this setup in that all users were assigned one of the three devices from the start and traded devices randomly thereafter. Although it was not clear beforehand how many participants would be available for this study, we were careful in choosing the 7 inch as the starting device due to its neutrality with regard to size and image quality (resulting from having the highest pixel density) compared to the other two devices. As such, switching from a potentially ideal case to cases that are theoretically inferior with regard to mobility or screen size would be very likely to result in mainly negative experiences.

However, our main goal is to evaluate user behavior and navigation performance with relation to the use of Shoebox VR. Although devices may play an important role in this regard, there are other factors that may affect its use such as the environment and addition of the navigation controls. Therefore, we categorize observations and discussion topics in the same way as we have categorized the results from chapter 7: behavior, performance and subjective rating. This way we can summarize the feedback and observations in a meaningful and way and provide structure to the statements and impressions.

8.2 Results

We evaluate the feedback of the expert users to find answers to our research questions and will refer to the data of chapters 6 and 7 where appropriate to provide a meaningful context or to support certain arguments and opinions with empirical evidence. We start by evaluating feedback and observations with relation to behavior i.e. how and when Shoebox VR is used, followed by a more in-depth analysis on if and how Shoebox VR may affect performance results of the test. Finally, we discuss how the expert users rated the properties and characteristics of each device and what they consider to be the preferred devices for the search tasks and use of Shoebox VR.

Behavior
For the behavior category, we aim to find concrete answers for the following questions:

- *Is Shoebox VR visually appealing or interesting enough to use it?*

In general, the expert users responded positively when first experiencing Shoebox VR in the demo. However, as was also reported in the main study, one user argued that the visual effect was inverted. When later asked as to why they were under that impression, the user was not able to provide an answer, stating that it “feels wrong”. During the challenge stages, some users reported that Shoebox VR was sometimes “annoying”. Upon inquiring during the discussion, the users found the change in view resulting from being unable to hold the device perfectly still to be distracting, although this was least visible on the 4 inch device. By the end of the session, one of the users felt that they had become somewhat accustomed to it and were less annoyed by it, implying that users vary in their ability to adapt (quicker) to this new method of interaction. This, combined with the previous reported issues, may explain why use of Shoebox VR can vary so strongly among users with a similar level of experience regarding 3D graphics and environments.

- *Is the use of Shoebox VR functionally appropriate and beneficial for the task or are navigation controls preferred?*

All expert users were of the opinion that, in terms of functionality, Shoebox VR could not be compared to the navigation controls due to its limitations. However, opinions on whether or not Shoebox VR was useful varied, with different cases being mentioned. One user stated that the use of Shoebox VR was particularly convenient for looking around trees, that would otherwise require movement. Two users praised the ability to look around while moving without it affecting their movement direction. One user found no real benefit in using Shoebox VR that could not also be accomplished by using the navigation controls at a slight distance penalty.

Upon combining the statements and opinions, we argue that the suitability of Shoebox VR in open environments is questionable. With the larger distances between objects and the reduced scale of Shoebox VR compared to a small box-like environment, it is more difficult to create scenarios where the use of Shoebox VR can truly provide benefits and advantages that navigation otherwise cannot. Although it is always possible to limit the freedom of movement in any space, and thereby subsequently making Shoebox VR more powerful, this is not always recommended or desired and may introduce the risk of over-accentuating the use of head-coupled perspective in such cases.

- *Is the use of Shoebox VR during motion more difficult than use while stationary?*

All users agreed that Shoebox VR is best used during the pauses in movement, as this removed the complexity of using Shoebox VR while simultaneously making use of the touch screen and allowed for more time for observation. It was argued that if movement is simplified or even automated, use of Shoebox VR during motion would be much less of an issue. It's also worth noting that some users experienced the use of Shoebox VR during motion to be more problematic or difficult than others, implying that skill and experience with navigating in 3D virtual environments may be crucial to effective use of Shoebox VR. Therefore, it is highly questionable if the average mobile user would be able to use Shoebox VR as effectively as the experts or the participants of the main study. As such, Shoebox VR is perhaps best used in scenarios with no movement or when movement is automated e.g. on-rails. This notion is supported by the fact that, in the main study, less than 30% of the average total distance traveled was covered using Shoebox VR. Generally speaking, using Shoebox VR during motion on the 4 inch device was considered to be the least difficult, although the small screen did cause some issues with the touch interactions.

Performance
The performance category is more focused on the use of Shoebox VR for finding the packages and if it is considered useful for this purpose.

- *How does use of Shoebox VR help in finding packages (view shift or translation)?*

The use of Shoebox VR for exploration can serve to purposes. The first is the change in view direction i.e. view shifting, for looking around. Because turning accomplishes the same effect, it is best used while moving as this does not affect the movement direction like turning does. The second is the camera translation that changes the perspective without moving the player character and thus is not penalized like movement is in terms of distance traveled. When presented with these options, the majority of users were of the opinion that the view shifting was generally more useful than the camera translation due to it wider applicability, especially for the 7 and 10 inch devices with the larger screen sizes. Where the translation was considered really only useful if there is something relatively close by blocking the view, the view shifting was useable everywhere. Conversely, the main benefit and purpose of head-coupled perspective is to provide different point of view in a realistic manner. Considering that the majority of users found the view shifting to be more useful, use of Shoebox VR in large spacious environments may not be ideal for demonstrating its main features.

- *Does use of Shoebox VR increase the time needed to find all the packages?*

Due to the use of Shoebox VR during motion being generally considered as difficult by both the expert users and the participants in the main study, the logical result is that Shoebox VR is mostly used when stationary. When asked if this takes more time than turning, several users responded by stating that they felt more comfortable using turning for the majority of the time and reasoned that the wrist and arm movements required to make effective use of Shoebox VR took more effort in comparison to the touch controls using only the thumb or finger. It was also argued that the addition of Shoebox VR adds another option for observation and exploration and that, seeing how it cannot replace movement or turning, it adds an additional few seconds to the clock each time it is used. Only one out of seven users did not perceive the use of Shoebox VR to be time consuming and motivated this opinion by stating that it should be used only for quick looks rather than extensive exploration.

- *Does the ability to use Shoebox VR during motion reduce the time needed to find all the packages?*

As most users did not often employ the use of Shoebox VR during motion for more than just a quick look, the general consensus was that it was too difficult to use to justify the advantages it may bring. Therefore, the majority responded by saying it would normally not reduce the time needed due to the difficulty of its use. In theory, however, most users agreed that being able to observe the environment would benefit the ability to find the packages faster.

- *Does use of Shoebox VR help in selecting shortest path routes?*

In response to this question, two out of the seven users stated firmly that it did not help them find or plan better paths that they would not have normally found by using the touch controls. The other users were largely vague and unsure about whether or not it helped them make better observations and planning decisions. They often repeated why they make use Shoebox VR e.g. the view shifting and/or camera translation, but were unable to determine if this actually benefited them in selecting better paths. Based on their distance averages, there did not seem to be much of a difference when comparing with and without Shoebox VR, which may explain why they were not able to answer the question, as there was no noteworthy improvement.

- *Does the ability to use Shoebox VR during motion reduce the effectiveness of finding shortest path routes?*

In spite of the general opinion being that use of Shoebox VR during motion is more difficult than using it while stationary, users were not quick to dismiss this as being true. Many scenarios were illustrated where use of Shoebox VR during motion did not result in a poorer route, such as when using it while moving to a distant target and finding a better candidate nearby or discovering a potential next best target. However, these are conditional cases that do not always occur. As such, the best answer to this question is that it's situational.

- *What is the default plan for each round?*

Each person repeatedly exposed to the problem of finding the packages in each round, develops a plan that includes how and when they are planning to move and when to make use of Shoebox VR. Because the environment is square and split up into smaller cells, it is possible that some users have become able to detect or even predict a pattern in package placement. If this is indeed true, it could explain why distances are so similar with and without Shoebox VR, because there would be little if any variation in the actual path for each round.

When asked what their plan was for each round, the majority of users responded by saying that they would first roam the outer edges of the environment while looking primarily forward and inward until they discovered a sequence of packages that could be traveled to fairly efficiently. In some cases this would be right from the start i.e. when a number of packages is visible right away, whereas in other cases this was more difficult, such as when the packages were hidden better. In general, the users only used Shoebox VR at the start of the round, occasionally when moving in between or arriving at packages and when trying to find the last few packages that were not part of their main route. However, this essentially describes the majority of a round and statistics only show an average of around 30% of the round time being used for Shoebox VR. It becomes more reasonable if we assume the use of the Shoebox VR to be quick and select, with the movement and turning controls being used the majority of the time. This is understandable, because a quick look, especially during movement, is more convenient and comfortable than keeping the device tilted for an extended period of time. As such, it doesn't necessarily mean that Shoebox VR is not used often, but it is instead used only in short bursts and during select moments. Understanding this, expecting Shoebox VR to be used at least for 50% of the round may have been unreasonable. In fact, if Shoebox VR usage is indeed primarily used during these moments, it could be argued that this is actually the best and most efficient way to use Shoebox VR for the purpose of exploration and finding packages. Behavior may change for larger environments or when movement is automated, but in retrospect the use of Shoebox VR by the expert users may have demonstrated its intended use quite accurately.

- *What are the primary differences between the rounds with and without Shoebox VR?*

Two users were under the impression that movement with Shoebox VR enabled was faster than without, likely the result of the extra motion of the image generated by the visual effect. One user also commented that this motion resembled "head bobbing in real life", due to the inability to move and keep the device perfectly still. As a result, this user experienced the removal of Shoebox VR as a limiting factor in terms of immersion, preferring to have it enabled at all times. In terms of performance, users felt that, in spite of some of the advantage and benefits that Shoebox VR could provide, they were able to complete the rounds without Shoebox VR just as efficiently. This is somewhat of a contradiction to earlier responses regarding if use of Shoebox VR enabled them to select more efficient paths, where the majority of the users had stated they were unsure any practical benefits. It is also possible that their confidence in their skill with the touch controls does not allow them to question their abilities, regardless if an external tool such as Shoebox VR is used or not. However, as the performance results showed little to no differences in terms of distance with and without Shoebox VR, it generally safer to assume that most users did not perceive the use of Shoebox VR to be sufficiently advantageous for planning optimal paths. Considering this, Shoebox VR appears to have served more as a generic augmentation for exploration and immersion than a practical tool for improving performance results.

Ratings

With the ratings category, we look at how the users perceived and experienced their use of Shoebox VR on all three devices and highlight the common views and opinions.

- *How are the 4, 7 and 10 inch devices experienced with relation to each other in terms of (screen) size?*

All users, except for one, found the 4 inch device to be the least practical. Hands and fingers would obscure the view of the screen and it was reported more than once that users accidentally crossed the dividing line between the left and right side of the screen, causing them to move or turn unintentionally. Conversely, many users found the 10 inch device to slightly too large and the 7 inch device to be of ideal size.

- *How are the 4, 7 and 10 inch devices experienced with relation to each other in terms of interaction?*

Directly related to the size of the device, the 4 inch device was considered the easiest to hold and tilt, but the small size of the screen made touch interactions occasionally problematic. The 10 inch device was difficult to hold and tilt due to its size and weight, although impressions and opinions were not as strong as expected, with many users not being too bothered by it. Again, the 7 inch device was considered to be ideal for use with both the tilting and touch controls.

- *How are the 4, 7 and 10 inch devices experienced with relation to each other in terms of image quality or clarity?*

Here, users all agreed that the image on the 4 inch device was too small, increasing the difficulty of finding the packages, especially when they were far away. As a result, one needed to be very close to the screen to the point where it was considered uncomfortable or straining. One user even referred to it as "pixel hunting". Users also commented on the relatively poorer image quality of the 10 inch device compared to the 7 inch device, resulting from the lower pixel density, but were not of the impression that it affected their ability to find or locate packages. As a result, the image quality of the 7 inch device was rated highest, but the 10 inch device was still preferred due to the size of the image.

- *What is the overall preferred device for performing the task?*

Many users voted the 10 inch device to be the best selection, in spite of claiming that the bigger image did not help them find or locate packages. A minority of 2 users voted for the 7 inch device, due to its versatility and lack of any disadvantages that were present on the 4 and 10 inch devices with regard to their size.

- *What is the overall preferred device for using Shoebox VR?*

One user preferred the 4 inch device, due to the ease of tilting, but would not consider it the best choice for the experiment as is. The other users all voted for the 7 inch device, due to the preferable screen size allowing for a clear and sufficiently large image while still being relatively easy to tilt. The 10 inch device was rated least favorable in this regard.

Final notes

In our experiment setup, we discovered that turning while using Shoebox VR resulted in a very unnatural motion and had expected some comments about this from either the participants of the main study or the expert users. However, nobody reported such an issue, either suggesting that nobody used it this way or that nobody was bothered enough by it to report it. This is a fairly interesting development, but should not dismiss the existence of the issue.

We were somewhat surprised that generally preferred the 10 inch device, in spite of being aware of the lower image quality compared to the 7 inch device. Although we had considered the possibility that the 10 inch would be preferred due to the image size, we did not predict the majority of users to be in favor of it.

In total, we conducted the expert analysis with 5 experts and considered this to be sufficient for two important reasons. The primary and most important reason is that the majority of the observations and feedback were in line with our expectations and that responses on major issues were generally the same. Second, Nielsen, J. (1995) states on his website on heuristic evaluations that he recommends "to use three to five evaluators since one does not gain that much additional information by using larger numbers." [31]. Therefore, combining the two results in a scenario where our number of 5 experts can be considered sufficient to highlight the most important of issues, observations and opinions.

Although the inverted motions of Shoebox VR was mentioned during the analysis, it never posed a problem for the expert users. Regardless, given how some users during main study felt like they were negatively affected by this, it may be worth to look into this in more detail in future studies to find out the cause of the problem and if there are ways to alleviate or eliminate the confusion.

8.3 Conclusion

By only looking at the results from chapter 7, we would have assumed that Shoebox VR was used in less than expected quantities because users were not aware of the benefits or did not consider them useful enough. Although the general opinion of the expert users is that Shoebox VR does not provide substantial improvements to their time or distance performance, they were able to describe scenarios where their use of Shoebox VR is in line with how we had intended it to be. As a result, we conclude that our conditions to reject the first null-hypothesis NH1 may have not have been realistic enough. However, if we disregard Shoebox VR usage statistics, both statistical data and the feedback from the expert users suggest that Shoebox VR ultimately does not make enough of a difference in terms of performance to be considered useful, which is further supported by the fact that neither the participants of the main study and the expert users felt particularly inconvenienced by the removal of Shoebox VR in between rounds. Instead, Shoebox VR is generally experienced as a nice visual effect that can enhance immersion and aid to some degree in exploration, which implies that open environments in combination with free moment may not be the ideal scenario for Shoebox VR to be used in.

In terms of devices, the 4 inch device is considered to be the least practical, as both the results from the main study and feedback from the expert users indicate that its small screen size causes issues with regard to using the touch controls and image size. In comparison, the 10 inch device is rated much more positively, in spite of being the most difficult to hold and tilt. The 7 inch device was rated almost equally poorly on average with the 4 inch device in the main study, although expert users generally considered the size and weight of the device to be ideal. Regardless, both the results from the main study and expert analysis show a clear preference and favor for the 10 inch device for the search task, suggesting that the focus on finding and locating the packages may have distracted from experimenting with Shoebox VR in favor of better performance. If this is indeed the case, this would further support the argument that Shoebox VR was not ideal for search tasks, and that it is perhaps more suitable for exploration and examination without the focus on performance.

9. Conclusion and Future Work

We have explored and evaluated the effects of the use of Shoebox VR on navigation performance and behavior in open environments. In spite of not being able to reject our null-hypotheses due to the significance of our results not being high enough, we have shown a high probability that user behavior is affected by the addition of Shoebox VR. In this regard, we have observed that use of Shoebox VR may cause a slight increase in time due to it being an extra form of interaction and method for observing. With limited experience and training in the use of Shoebox VR, users did not use Shoebox VR as much as we had anticipated, although feedback and observations from the expert analysis portrayed scenarios where its use conformed to our specifications and expectations. Therefore, we concluded that our minimum value of 50% of the round time for the use of Shoebox VR was not a realistic indication of encouragement, especially considering the users' high levels of experience and training with the traditional movement controls, and that their use required to complete the stages. As a result, we recommend a future study to perform more detailed research on how Shoebox VR is measured best and what realistic expectations are, based on the objective, environment and methods of interaction.

The extra time for which Shoebox VR was used did not appear to contribute to a better path planning, with no indication of distance results improving as a consequence of using Shoebox VR. This has lead to the realization that, despite our efforts of creating a noticeable visual effect and functional use for Shoebox VR, an open environment in combination with free navigation may not be the ideal scenario for Shoebox VR to be used in. The larger scale of the environment and increased distances between objects make it less likely for a noticeable motion parallax effect to occur and reduces the strength and benefits of the translation of the camera. However, considering that the “Closest” round of the pilot study has higher Shoebox VR usage than other stages suggests that other areas such as distance or size estimation experiments could still be interesting to research. Therefore, we cannot claim that the use of Shoebox VR in open environments combined with free movement has no potential, only that it is likely that Shoebox VR will not be used extensively if such movement is a fundamental part of the objective e.g. finding and collecting items. As such, we would not recommend a continuation of this study unless there is sufficient reason to believe that another environment or objective may fundamentally change user behavior with regard to use of Shoebox VR. Instead, we encourage research on how Shoebox VR can influence other forms of navigation, such as on-rails movement or a point-and-click design.

As part of our study, we have also informally evaluated the differences between a 4 inch, 7 inch and 10 inch device, and compared behavior, performance and user ratings. Based on these results, we have shown a high likelihood that there is a difference between such devices and their use, with performance and user ratings being relatively poor on the 4 inch device in comparison to the 7 inch and 10 inch device. We have also observed a high user rating for the 10 inch device, suggesting it is the most popular choice and argue that this is the result of its larger screen size, capable of aiding in the search tasks and improving the immersion and enjoyment of the use of Shoebox VR. Thus we recommend future studies on Shoebox VR to generally prefer the use of tablets over mobile phones, with a 7 inch device potentially being the ideal choice, unless there is indication of other devices being more suitable for the tasks.

Finally, we conclude by answering our research question “*How does the addition of Shoebox VR affect movement behavior and navigation performance on mobile devices in open 3D virtual environments?”* by stating the following: With the ability of Shoebox VR to transform and translate the view comes an alternative take on exploration that allows for a more intuitive method of 3D visualization and interaction. However, in spite of the benefits and advantages it can provide, users vary in opinion on practical use and general appeal, resulting is large variances in behavior with regard to the use of Shoebox VR. Furthermore, there is no indication thus far that use of Shoebox VR can increase performance of time efficiency or path planning when the objective is to find and collect a number of small, hidden items in an open environment. Therefore, the addition of Shoebox VR to traditional navigation controls appears to have a limited impact on user behavior or performance, in that its addition did not generally encourage users to employ its use or increase their performance ratings in our experiments. Although we cannot guarantee the consistency or probability of these results in other cases due to use of a single environment and objective, we have shown that there is reason to question the potential of Shoebox VR in open environments where unrestricted navigation is a key feature of the application.

References

1. Dalgarno B., Hedberg J. & Harper, H. (2002) The Contribution Of 3D Environments To Conceptual Understanding. In *Winds of Change in the Sea of Learning: Proceedings of the 19th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education Vol 1*, pp 149-158. Auckland, New Zealand, 2002. UNITEC, Institute of Technology.
2. Demiralp, C., Jackson, C. D. Karelitz, D. B., Zhang, S. & Laidlaw, D. H. (2006) CAVE and Fishtank Virtual-Reality Displays: A Qualitative and Quantitative Comparison. In *IEEE Transactions on Visualization and Computer Graphics Volume 12 – Issue 3*, pp 323-330. New Jersey, USA, 2006. ACM.
3. Ware, C., Arthur, K. & Booth, K. S. (1993) Fish Tank Virtual Reality. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, pp 37-42. New York, USA, 1993. ACM.
4. Li, I. K. Y., Peek, E. M., Wünsche, B. C. & Lutteroth, C. (2012) Enhancing 3D Applications With Stereoscopic 3D And Motion Parallax. In *Proceedings of the Thirteenth Australasian User Interface Conference (AUIC2012) Volume 126*, pp 59-68. Melbourne, Australia, 2012.
5. Stavness, I., Lam, B. & Fels, S. (2010) pCubee: A Perspective-Corrected Handheld Cubic Display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp 1381-1390. New York, USA, 2010. ACM.
6. Spindler, M., Büschel, W. & Dachselt, R. (2012) Use Your Head: Tangible Windows for 3D Information Spaces in a Tabletop Environment. In *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces (ITS '12)*, pp 245-254. New York, USA, 2012. ACM.
7. Francone, J. & Nigay, L. (2011) Using The User's Point of View For Interaction On Mobile Devices. In *Proceedings of the IHM '11 23rd French Speaking Conference on Human-Computer Interaction*. New York, USA, 2011, ACM.
8. Hürst, W. & Helder, M. (2011) Mobile 3D Graphics and Virtual Reality Interaction. In *Proceedings of the 8th Internation Conference on Advances in Computer Entertainment Technology*, Art. no 28. New York, USA, 2011. ACM.
9. Elmqvist, N. & Tsigas, P. (2006) On Navigation Guidance for Exploration of 3D Environments. Göteborg, Sweden, 2006. Chalmers University of Technology.
10. Kim, H.-K., Song, T.-S., Choy, Y.-C. & Lim, S-B. (2005) Guided Navigation Techniques for 3D Virtual Environment Based on Topic Map. In *Proceedings of the 2005 international conference on Computational Science and its Applications - Volume Part I (ICCSA'05)*, pp 847-856. Heidelberg, Berlin, 2005. Springer-Verslag Berlin Heidelberg.
11. Darken, R. P. & Peterson, B. (2001) Spatial Orientation, Wayfinding and Representation. In *Handbook of Virtual Environment*, pp 493-518. Monterey, California, 2001. Erblaum.
12. Darken, R. P. & Sibert, J. L. (1996) Wayfinding Strategies and Behaviors in Large Virtual Worlds. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp 142-149. New York, USA, 1996. ACM.
13. Chittaro, L. & Burigat, S. (2004) 3D Location-pointing as a Navigation Aid in Virtual Environments. In *Proceedings of the working conference of Advanced visual interfaces,* pp 267-274. New York, USA, 2004. ACM.
14. Giudice, N. A., Bakdash, J. Z., Legge, G. E. & Roy, R. (2010) Spatial Learning and Navigation Using a Virtual Verbal Display. In *ACM Transactions on Applied Perception Volume 7 – Issue 1*, Art. no. 3. New York, USA, 2010. ACM.
15. Nurminen, A., & Oulasvirta, A. (2008) Designing interactions for navigation in 3D mobile maps. In *Map-based Mobile Services: Design, Interaction and Usability,* pp 198-224. London, England, 2008. Springer.
16. Teather, J. & Stuerzlinger, W. (2008) Exaggerated Head Motions For Game Viewpoint Control. In *Future Play '08 Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, pp 240-243. New York, USA, 2008. ACM.
17. Kulshreshth, A. & LaViola, J. J. Jr. (2013) Evaluating Performance Benefits of Head Tracking in Modern Video Games. In *Proceedings of the 1st symposium on Spatial user interaction (SUI '13)*, pp 53-60. New York, USA, 2013. ACM.
18. LaViola jr., J. J., Keefe, D. F. (2011) 3D Spatial Interaction: Applications for Art, Design and Science. In *ACM SIGGRAPH 2011 Courses*, Art. no. 1. New York, USA, 2011. ACM.
19. Oulasvirta, A., Estlander, S. & Nurminen, A. (2008) Embodied interactions with a 3D versus 2D mobile map. In *Personal and Ubiquitous Computing Volume 13 – Issue 4*, pp 303-320. London, UK, 2009. Springer-Verslag London Limited.
20. Burelli, P. & Yannakakis, G. N. (2011) Towards Adaptive Virtual Camera Control In Computer Games. In *Smart Graphics*, pp 25-36. Berlin, Germany, 2011. Springer Berlin Heidelberg.
21. Ropinski, T., Steinicke, F. & Hinrichs, K. (2005) A constrainted road-based VR navigation technique for traveling in 3D city models. In *Proceedings of the 2005 international conference on Augmented tele-existence*, pp 228-235. New York, USA, 2005. ACM.
22. Salamin, P., Thalmann, D. & Vexo, F. (2006) The benefits of third-person perspective in virtual and augmented reality? In *Proceedings of the ACM symposium on Virtual reality software and technology (VRST '06)*, pp 27-30. New York, USA, 2006. ACM.
23. Salamin, P., Thalmann, D. & Vexo, F. (2008) Improved third-person perspective: a solution reducing occlusion of the 3PP? In *Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry (VRCAI '08)* – Article 30. New York, USA, 2008. ACM.

24. Bateman, S., Doucette, A., Xiao, R., Gutwin, C., Mandryk, R. L. & Cockburn, A. (2011) Effects of view, input device, and track width on video game driving. In *Proceedings of Graphics Interface 2011 (GI '11)*, pp 207-214. Waterloo, Canada, 2011. Canadian Human-Computer Communications Society School of Computer Science.
25. Claypool, M. & Claypool K. (2009) Perspective, Frame Rates and Resolutions: It's all in the Game. In *Proceedings of the 4th International Conference on Foundations of Digital Games*, pp 42-49. New York, USA, 2009. ACM.
26. Moerman, C., Marchal, D. & Grisoni, L. (2012) Drag'n Go: Simple and fast navigation in virtual environment. In *3DUI*, pp 15-18. IEEE.
27. Maguire, E.A., Burgess, N. & O'Keefe, J. (1999) Human spatial navigation: cognitive maps, sexual dimorphism and neutral substrates. In *Current Opinion in Neurobiology*, pp 171-177. London, 1999. Eslevier Science Ltd.
28. Sandstrom, N.J., Kaufman, J. & Huettel, S.A. (1998) Males and females use different distal cues in a virtual environment navigation task. In *Cognitive Brain Research Volume 6 – Issue 4*, pp 351-360. USA, 1998. Elsevier.
29. Lee, J. C. (2007) Head Tracking for Desktop VR Displays using the Wii Remote. http://johnnylee.net/projects/wii/
30. Nielsen, J. & Mack, R. L. (1994) Usability inspection methods. New York, USA, 1994. John Wiley & Sons, Inc.
31. Nielsen J. (1995) How to Conduct a Heuristic Evaluation. http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation.