3D UIs are uniquely able to achieve superior interaction fidelity, and this naturalism can be a huge advantage.

BY DOUG A. BOWMAN, RYAN P. MCMAHAN, AND ERIC D. RAGAN

Questioning Naturalism in 3D User Interfaces

USER INTERFACES (UIs) based on buttons, keyboards, mice, and joysticks have dominated both the general-purpose computing world and the more specialized video game domain for decades. But recent years have seen a revolution, as designers have embraced spatial input for their UIs. These interfaces are described in many ways: gesture-based, motion-controlled, direct, controller-less, and natural. But they are all characterized by the use of spatial input in a physical three-dimensional (3D) context—what we call a 3D user interface[a] or 3D UI.[4]

The most widely known 3D UI example is the use of 3D input in console gaming systems. Nintendo's introduction of the Wii Remote—a wireless handheld device that uses accelerometers to sense 3D motion and an infrared camera to sense pointing direction—sparked a revolution in console gaming.[36] Game designers used the new motion capabilities to allow players to use realistic physical actions to swing virtual golf clubs, roll virtual bowling balls, punch out virtual enemies, and point directly to menu items on the screen. Instead of arbitrary mappings of buttons and joysticks to on-screen actions, games began using natural motions, allowing anyone to easily join in.

Following this trend, other major game-console makers have developed their own 3D input systems. Sony developed the Move controller[36]—a device similar in form factor to the Wii Remote, but with additional sensors and capabilities for true six-degrees-of-freedom (6-DOF) input. In other words, the Sony system can determine the absolute 3D position and orientation of the Move controller, which allows even more precise forms of 3D interaction. For example, the Move could be used to draw a 3D sketch directly in 3D space, where the position of the virtual pencil is determined directly by the position of the user's hand. Microsoft entered the motion gaming market with the Kinect,[36] which uses infrared lights and cameras to determine the 3D position and pose of the entire body without a

» key insights

■ Spatial input technologies allow unprecedented levels of interaction fidelity in UI design. However, natural 3D UIs must still be designed carefully, especially since no UI can replicate the real-world experience exactly.

■ Studies show that high levels of naturalism can enhance performance and the overall user experience, but moderately natural 3D UIs can be unfamiliar and detrimental to performance. Traditional, less natural, interaction styles can provide good performance, but result in lower levels of presence, engagement, and fun.

■ The hyper-natural design approach offers realistic interactions and enhanced abilities that avoid some of the unwanted constraints of the real world.

---

a The term "3D user interface" may also be used to refer to an interface using 2D input devices to interact with a 3D virtual world, but in this article, our discussion is limited to the definition given in the text.

controller. Knowing the body's pose affords interesting new forms of interaction. In particular, since video games often use an avatar to represent the player, the movements of the avatar's body can now be controlled directly by the player's physical movements.

It is not only in gaming that we see the emergence of 3D UIs. Mobile applications on smartphones and tablets are also joining the trend. While much attention has been focused on the multi-touch interfaces found on recent mobile devices, most of these devices also contain sensors—such as accelerometers, gyros, compasses, GPS receivers and cameras—that can be used for 3D interaction. For example, apps such as "Real Swing Golf: Putting" allow the player to use a real golf swing to determine the power of a shot, in much the same way one might use the Wii Remote. Even more compelling are mobile augmented reality applications. Such applications use sensor data to determine where you are and in which direction you are looking so that virtual information or objects can be combined with

a view of the real world. For example, amateur astronomers can point their smartphones up to see star names and constellation outlines overlaid on the night sky using apps like "Pocket Universe: Virtual Sky Astronomy."

In considering these high-profile 3D UIs, we find that almost all of them strive to use 3D interaction to make the interface more natural. We define the naturalism, or the interaction fidelity, of a UI as the objective degree with which the actions (characterized by movements, forces, or body parts in use) used for a task in the UI correspond to the actions used for that task in the real world.[20] Swinging the arms to hit a virtual golf ball is clearly more natural than pressing a sequence of buttons to accomplish the same task. "Looking through" a smartphone to see virtual information overlaid on the real world is more natural than browsing a panoramic image using buttons and menus on a touch screen, because the former uses head and body movements to look at a particular part of the environment, as one would in the real world. The level

of naturalism depends, of course, on both the interaction technique and the task context—a steering wheel metaphor is natural when the task is driving a virtual vehicle, but would not be natural for the task of shooting a virtual basketball.

Some actions (for example, teleporting to a new location in a video game) do not lend themselves easily to natural physical mappings. In these cases, UI designers usually resort to traditional interaction using buttons and the like, or some arbitrary 3D action is used (for example, many Wii games use the "shake" gesture when a natural mapping is not available). These techniques have low levels of interaction fidelity. For other tasks, current devices cannot easily sense the natural movements. For example, in one of the minigames in "Wii Sports Resort," the player must pedal a bicycle. The natural mapping using foot and leg motions is impractical, so the designers instead substituted hand and arm motions that mimic the real-world foot and leg motions. On the interaction fidelity scale, this

technique is more natural than buttons and joysticks, but is less natural than riding an exercise bike to power the virtual bike.

These examples raise some interesting questions about naturalism and 3D UIs:

▸ Are 3D UIs inherently more natural than traditional UIs?

▸ Should designers strive primarily for high levels of naturalism in their 3D UIs, or are there other interaction design criteria that are more important?

▸ Does a more natural interface result in better performance, greater user engagement, or increased ease of learning?

▸ In cases where the most natural mapping cannot be used, is it better to use a moderately natural technique, or are traditional techniques more appropriate?

In this article, we will explore these questions and the findings of 3D UI researchers who have studied naturalism. Although the high-profile 3D UIs in gaming and mobile applications have only emerged in the last six years (the Wii was introduced in 2006), researchers have actually been studying 3D UIs for much longer. An international community of 3D UI researchers have been studying various forms of 3D input, designing and evaluating different ways to map that input to actions within user tasks, and developing guidelines and principles for effective 3D UIs. Thus, we will begin with a historical look at the evolution of both natural and "magical" 3D interaction techniques.

History of Naturalism in 3D UIs

We can trace the roots of research on 3D UIs to the 1960s. In 1965, Ivan Sutherland described "The Ultimate Display."[33] In this visionary paper, he not only foresaw that computers could be used for more than number crunching, but also that users could use their whole bodies, tracked by the computer, to interact with highly realistic virtual worlds. Three years later, Sutherland published the first paper describing a tracked head-mounted display (HMD), with which people could use natural head movements to view virtual and augmented 3D environments.[34]

3D UI research continued sporadically over the next 30 years, but there was no focused research community or systematic exploration of the 3D UI design space until the 1990s. At that time, there was a growing interest in immersive virtual reality (VR), which uses 3D graphics, 3D tracking systems, and HMDs or stereoscopic projection displays to place the user inside a virtual world. In desktop computing contexts, 3D interaction was an optional novelty that could be explored. In immersive VR, it was an absolute necessity, since mice and keyboards do not work well when the user is standing up, walking around, and unable to see the physical world. In the beginning, VR research focused on technical issues. But as people started to use VR for more complicated applications, such as scientific visualization,[9] 3D modeling,[24] and interactive education,[12] researchers needed to focus on user interface design in order to make such applications usable. While the general principles of human-computer interaction were helpful in 3D UI design, more focused and specific techniques, guidelines, and theories were needed.

Out of the VR research community, then, grew a group focused on 3D UIs. Today, this community has an annual international conference (the IEEE Symposium on 3D User Interfaces) and a mailing list[b] comprising almost 500 researchers. Much of the research in this area was summarized in the 2005 book *3D User Interfaces: Theory and Practice*.[4]

From the early days of 3D UI research, there seemed to be two design approaches for 3D interaction techniques: one that tried to design techniques with as much interaction fidelity as possible, and another that tried to enhance usability and performance through the use of "magic" techniques.[31] Magic techniques might be intentionally less natural, or they might enhance natural interactions to make them more powerful. Below, we discuss both natural and magic techniques for three of the so-called "universal" 3D UI tasks: travel, selection, and manipulation.

Travel, the task of moving the viewpoint (or avatar) through the virtual 3D environment, is conceptually very simple, but the design space for travel inter-

_____
b   http://www.3dui.org/

action techniques is surprisingly large.

An obvious natural travel technique is to turn the head and physically walk in order to look around and move through the environment. In fact, head tracking was used for this purpose from the earliest days of 3D UIs.[34] Studies have found that physical turning and walking can enhance spatial orientation and movement understanding (for example, Chance et al.[10]) This technique only works, however, in very limited situations. First, the user's position and orientation must be tracked, requiring a complete 6-DOF tracking system. Second, the user must be able to see the display from any location and while looking in any direction, requiring an HMD or a fully surrounding set of display screens. Third, the virtual environment must be smaller than the tracked area in order to allow the user to physically walk to all virtual locations.

Since these three conditions are very rarely met, designers desiring a natural travel technique have come up with many approximations to real walking. For example, redirected walking techniques[29] allow physical walking at a one-to-one scale through a theoretically infinite virtual environment by rotating or otherwise modifying the environment to keep the user from walking out of the tracked area; however, these techniques are not yet able to achieve this goal for arbitrary environments with tracked areas of any size. Another, less natural, approach is to use a proxy for walking such as walking-in-place,[35] which retains the physical leg motions but does not provide the same proprioceptive cues to help users understand their movements. Researchers have also designed specialized locomotion devices, including treadmills,[11] giant "hamster balls,"[23] and even robots that circulate continuously to form an infinite floor.[16] None of these have made their way into common use.

Another natural approach to travel is vehicle simulation. High-end flight simulators, driving simulators, and tank simulators can have very high levels of interaction fidelity, since they use physical cockpits with exact replicas of the actual controls used in the vehicle.[8] These are obviously specialized systems, however, which cannot

be used for general-purpose 3D UIs. Simpler vehicle simulations, such as bicycles,[7] have been explored for general travel tasks.

Magic travel techniques exhibit great diversity. The simplest of these techniques, called "steering,"[4] allows continuous travel at a constant speed in one or more directions (forward, back, left, right). Most video games use some variant of a steering technique. An important consideration is which direction is considered to be forward: the direction the user is looking, the direction the user is pointing, or the direction toward the center of the display. In general, separating the viewing direction from the travel direction is considered beneficial and more flexible.[3] Another set of magic travel techniques called "target-based travel" only require the user to specify a point of interest, after which the viewpoint is moved smoothly to the new location (for example, Hachet et al.[14]). Target-based techniques reduce mental load for the user, but may not be as flexible or as natural for all applications.

A final category of travel techniques combines elements of both magic and naturalism. For example, one can scale up the movements of the user to allow physical walking through larger virtual worlds,[15] but this may reduce precision and the user's ability to determine walking distance. Manipulation-based travel is based on hand movements rather than body movements or button presses. For example, the user can "grab the air" with both hands to move through the virtual world as if pulling on an invisible rope.[19] Though fatiguing, these techniques allow physical movement to be mapped to virtual travel without the need for large tracking areas or complex locomotion devices.

Selection, the task of picking one or more objects from the environment,[4] is also a fundamental task in many 3D UIs. It has been studied extensively, and the number of different 3D selection techniques is very large.

We do not typically think of "selecting" things in the real world, but it is analogous to touching or pointing at something. It is not surprising, then, that natural selection techniques are usually based on one of these two metaphors. The most common technique

Magic techniques might be intentionally less natural, or they might enhance natural interactions to make them more powerful.

based on touching is called the "simple virtual hand,"[4] in which the user controls a virtual hand directly with real hand movements, and causes the virtual hand to touch a virtual object in order to select it. The primary issue with simple virtual hand is that it requires the user to be within arm's reach of the object in order to select it; otherwise, the user will have to travel until the object is nearby. We are used to this restriction in the real world (I cannot pick up a book without walking over to the shelf), but we also seek ways to circumvent it (for example, remote controls). In a virtual world, where travel can be cumbersome, having to move close to an object to select it is often unacceptable.

Techniques based on pointing can be used from a distance, while still being considered natural. The canonical pointing technique in 3D UIs is ray-casting,[4] in which the user controls the direction of a virtual light ray or laser beam with physical hand movements, and intersects an object with the ray to select it. This technique works well in many situations, but it can be difficult to select very small objects,[27] since small hand rotations can result in large movements of the end of the ray, and only objects that are at least partially visible can be selected.

Most magic selection techniques are also based on touching or pointing metaphors,[27] but are "hyper-natural" (that is, they use natural movements but make them more powerful by giving the user new abilities or intelligent guidance). For example, in the Go-Go technique[26] normal physical arm extension causes the virtual arm to extend far into the environment, allowing objects to be selected by touching even from a distance. Rays that snap to objects,[39] volumetric rays,[13] and rays that bend around obstacles[25] enhance the pointing metaphor, making it easier for users to point precisely even with small targets or occlusion.

Manipulation. Closely related to selection is the task of manipulation, in which the user modifies the position, orientation, scale, or shape of virtual objects.[c] Manipulation requires selec-

---

c   Other property changes can also be considered manipulation. The focus here is on basic geometric manipulation.

tion (just as throwing a ball requires picking it up first), and many selection techniques can easily be extended to the manipulation task.

Again, the most natural manipulation technique is based on the simple virtual hand. After touching the object to select it, the user can position and rotate the object directly with physical hand and arm movements. Just as before, the limited reach of the human arm is the primary limitation of this technique. Typically, the virtual hand is considered to be rigid, which makes certain rotations very difficult to perform; allowing fingertip manipulation is technically challenging,[18] but makes the technique even more natural.

Ray-casting can also be used for manipulation, by attaching the object to the end of the ray. Unlike ray-casting selection, however, ray-casting manipulation is not a natural technique.

Magic manipulation techniques generally take the hyper-natural approach described above. The Go-Go technique[26] can be used for both selection and manipulation. HOMER[2] combines pointing-based selection with virtual hand-based manipulation at a distance. Techniques like World-in-Miniature[32] use the simple virtual hand to manipulate proxy objects—miniature copies of the actual targets—in order to achieve precision and a natural feel no matter the distance to the target or its size.

Evaluations of selection and manipulation techniques (for example, Poupyrev et al.[27]) have generally shown that magic, hyper-natural techniques outperform their more natural counterparts. 3D UIs of this sort can make performing tasks in the virtual world easier than in the real world, which is a strong argument for the magic approach. Still, it may be that these magic techniques reduce the user's feeling of presence in the virtual world, their understanding of their actions, or their ability to transfer actions they have learned back to the real world.

In summary, 3D UI designers have explored a wide variety of natural, magical, and hyper-natural interaction techniques. Many successful techniques are hyper-natural, using natural-feeling actions to control magical interactions with the virtual world. But the existing literature does not answer all of our questions about the effects of interaction fidelity. Comparisons of hyper-natural techniques with natural ones are often more about the capabilities of the techniques than about the inherent effects of naturalism (for example, it is clearly faster to select several distant objects with Go-Go than with simple virtual hand, simply because Go-Go allows selection at a distance; this does not mean a natural design approach necessarily leads to poor performance). Studies focused on the level of interaction fidelity itself are needed to understand its effects more fully, as we discuss here.

Evaluating the Effects of Naturalism

In a series of experiments conducted by our group at Virginia Tech, we have studied the effects of the level of interaction fidelity on user performance, presence, engagement, and preference, for a wide range of tasks. Some of the studies also varied the level of *display fidelity* (also called immersion[5]) because we hypothesized that the type of display being used might have an influence on the effects of interaction fidelity. The results of these studies shed light on the questions we posed earlier regarding the nature and value of natural UIs. The accompanying table summarizes our high-level findings for various user tasks.

Benefits of natural travel. Proponents of natural travel techniques often claim these techniques will result in higher levels of spatial understanding because they use physical turning,

Benefits and limitations of natural 3D interaction for particular user tasks, taken from our prior research.

Task	Benefits of naturalism	Limitations of naturalism
Viewpoint rotation	Users prefer physical turning.[32]	Users prefer virtual turning to a combination of physical and virtual turning.[32]
	Natural turning techniques have better performance than virtual turning for visual search.[38]	
Viewpoint translation/travel	Head tracking can improve spatial understanding and detailed spatial judgments.[33]	The benefits of head tracking may depend on other factors, such as stereoscopic display.[33]
		Moderately natural techniques can have poorer performance than traditional techniques.[38]
Manipulation	Natural techniques improve performance of complex manipulation tasks.[34]	Highly natural techniques have limited range.[28]
	Hyper-natural techniques enhance users' abilities.[28]	Hyper-natural techniques often reduce precision.[36]
	It is possible to design hyper-natural techniques that feel natural and have high levels of precision.[36]	
Vehicle steering	Higher levels of interaction fidelity can be more fun for users.[3]	Moderately natural techniques can have poorer performance than traditional techniques.[3]
Aiming	Highly natural aiming techniques can have better performance than mouse-based techniques.[38]	
Multiple tasks	High levels of interaction fidelity, when paired with high display fidelity, can have very good performance.[38]	High levels of naturalism may not be beneficial if the overall interface is unfamiliar.[38]
	Users feel that highly natural techniques are more engaging and induce higher levels of presence.[38]	

leaning, crouching, and walking motions, which provide proprioceptive cues. A number of prior experiments have tested this hypothesis. For example, physical turning can allow users to make better direction estimations than virtual turning (for example, Chance et al.[10]), and head position tracking can provide better understanding of complex 3D spatial structures.[37]

One of our early studies[1] examined users' preferences for physical and virtual turning, which is an issue that is relevant for all sorts of displays. Fully surrounding displays such as HMDs and 360-degree surround-screen displays can make use of purely physical rotation, but designers may also opt to provide virtual rotation to reduce physical fatigue or simply for users' convenience. Most displays, however, do not provide a 360-degree surround, and so will require some virtual turning. Is the best strategy to turn physically as much as possible, only using virtual turning when the edge of the display is reached, or should all turning be virtual to avoid inconsistency?

In our experiment, participants traversed a virtual corridor with many turns, and we noted how often physical and virtual turning were used. Two displays were used in the study: an HMD (full surround) and a surround-screen display with three walls (270-degree surround). We found that users turned physically more often in the HMD, but often used virtual turning in the surround-screen display, not only when it was required (when the edge of the display was reached) but also when physical turning was possible. This suggests that users may prefer to use the less natural virtual turning technique consistently rather than using a combination of physical and virtual turning.

Providing positional head tracking can enable users to travel naturally by leaning, crouching, or walking. Although this does not allow long-distance navigation due to limitations on the size of the tracking volume, it can be an effective way to make small changes to the viewpoint. In an experiment examining the effects of display and interaction fidelity on small-scale spatial judgments,[28] we asked participants to examine visualizations of complex underground cave systems.
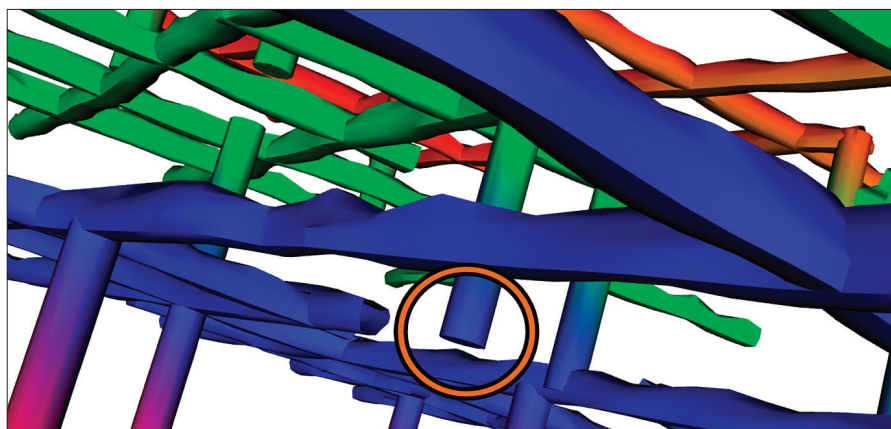

Figure 1. Small-scale spatial judgment task in the cave visualization study.[28] An example of a small gap between a vertical tube and a horizontal level of tubes is circled.

The task was to determine whether horizontal "tubes" connected two "layers" of the cave structure or not. Some tubes were connected, while others had slight gaps or breaks that were difficult to see except from certain angles (see Figure 1). We varied both interaction fidelity (head tracking vs. purely virtual travel) and display fidelity (stereo vs. mono and 270-degree surround vs. 90-degree surround).

We found the use of head tracking produced significantly fewer errors for the small-scale spatial judgment task, and that head tracking in combination with stereoscopy was significantly faster. Using head tracking allowed participants to quickly and precisely move their viewpoints, and to understand how far they were moving, resulting in easy comprehension of the changing visual imagery on the screen. These findings reinforce the benefits to spatial understanding of natural travel techniques.

Benefits of natural selection and manipulation. In another study,[22] we looked at the tasks of selection and manipulation, and evaluated interaction fidelity by comparing two hyper-natural techniques (Go-Go and HOMER) with a more traditional technique based on the mouse and keyboard. We also varied the level of display fidelity by comparing a single screen vs. four surrounding screens and stereo vs. mono graphics.

The more natural techniques used a handheld 6-DOF tracker and familiar hand and arm movements to select objects (by touching in the case of Go-Go and by pointing in the case of HOMER) and manipulate their position and orientation. The less natural technique used a mouse to control an on-screen cursor used to select objects via clicking, and the mouse and keyboard in combination to control the position and orientation of the objects. The mouse and keyboard mappings were based on common desktop techniques like those used in 3D modeling software.

Users performed the task of selecting 3D letter-shaped objects, then moving and rotating them until they were inside slightly larger versions of the same shapes (see Figure 2). This challenging task required users to manipulate all six degrees of freedom.

The results of this experiment were very clear: the more natural techniques significantly outperformed the less natural technique, regardless of the level of display fidelity, which had no measurable effect on performance. These findings indicate increased interaction fidelity can have a positive impact on efficiency for difficult 3D tasks. In particular, the manipulation task in this study required users to control both position and orientation at the same time in order to be successful, meaning interaction techniques that allow integrated control of all six degrees of freedom would perform better than techniques that separated them.[17] In other words, the more natural techniques were a better match for the task. These results also support the use of the hyper-natural design approach.

The enhanced abilities provided by hyper-natural techniques often come with a cost, however. In the case of HOMER, one of the costs is a loss of precision due to the scaling applied

to hand movements during manipulation. We designed and evaluated a technique called scaled HOMER[38] in an attempt to address this problem. Scaled HOMER uses a large scaling factor when the user's hand is moving quickly, which allows for long-distance object manipulation, but applies a small scaling factor when the user's hand is moving slowly, which allows for very precise manipulation. The nice thing about this mapping is that users naturally move faster when they are trying to move an object large distances, and naturally slow down when they are trying to be precise. Thus, the technique has the same natural feel as the original HOMER, while improving precision. This technique is an example of adding enhancements that can aid the user in task performance without a reduction in perceived naturalism.

Natural vs. traditional vehicle steering. While the prior experiments all investigated the effects of naturalism on "universal" 3D UI tasks, we have also examined its influence on application-specific tasks. One such study took place in the context of a racing game (see McMahan[21] for complete details). Motivated by the success of the Nintendo Wii, we looked for a game that provided multiple interaction techniques at different levels of interaction fidelity for the same task. Mario Kart Wii, a popular racing game, proved to be useful for this purpose.

In Mario Kart, the primary task is steering a vehicle around a track. Since the game franchise has had versions on several prior consoles, it has well-designed traditional controls using a thumb-controlled joystick for steering. The Wii version continued to provide these controls, but also added the option of using the Wii Remote (sometimes embedded in a plastic steering wheel) for more natural steering control.

Our experiment compared four different steering techniques: two traditional techniques (based on slightly different controllers) and two natural techniques (Wii Remote with and without the wheel prop). To remove the influence of game AI, we used the "Time Trial" mode, which simply measures lap times without any other racers on the course and without random power-ups.

The results showed the less natural, joystick-based techniques were faster and more accurate. Players were able to drive more precisely using a less natural interface. At a high level, there could be several explanations for this result. We could conclude that increased interaction fidelity is harmful to performance for steering tasks, but we think it is more likely the Wii Remote-based techniques were *not natural enough*. The Wii Wheel is not mounted to a fixed base—the user holds it in mid-air—and it also provides no force-feedback or re-centering as a real steering wheel would. In addition, the Wii Remote has some latency that may cause players to oversteer. Another interpretation is that the joystick-based techniques are actually more precise for the task of steering. This is an intriguing possibility, since it is known that small muscle groups such as those in the hand can be faster and more precise than the large muscle groups used to turn a steering wheel.[40] Per-
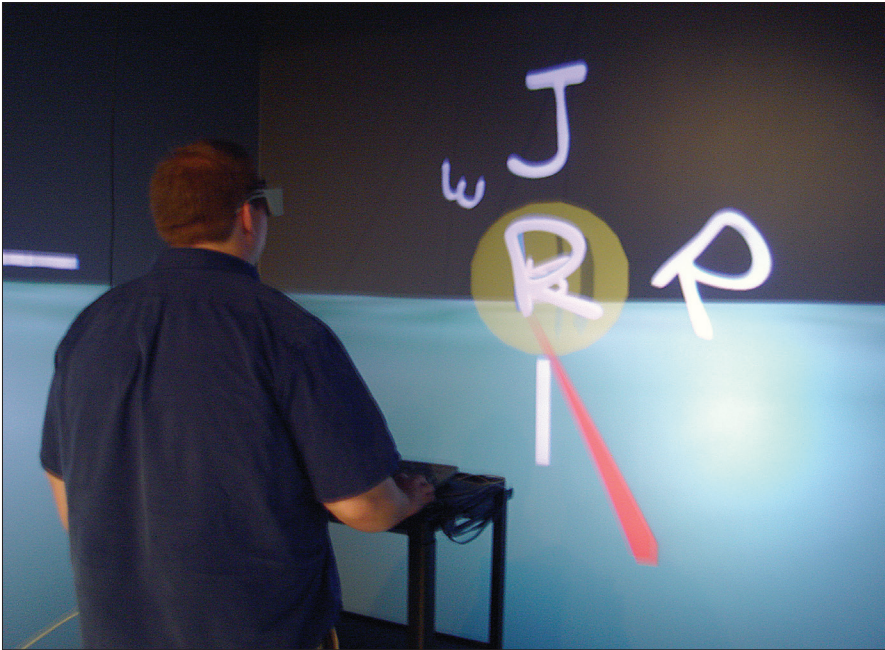

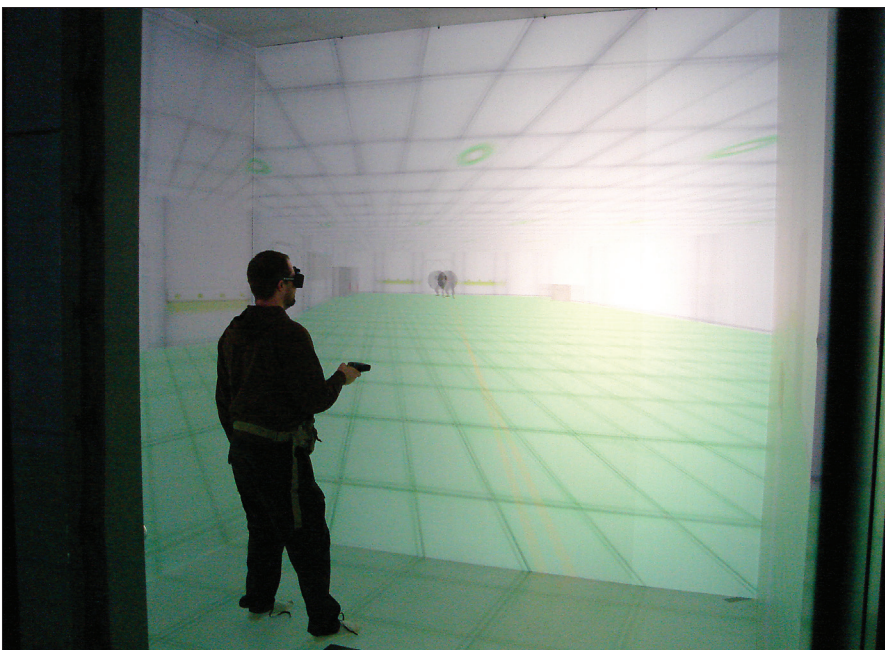**Figure 2. User manipulating letter shapes in the selection and manipulation experiment.**[22]

Figure 3. Playing a first-person shooter game in the DiVE.

haps cars of the future should be out-fitted with joysticks instead of steering wheels! Finally, it might be the case that the steering task in the game is not the same as steering a car in the real world, so that even a fully natural technique would not necessarily improve performance.

Regardless of the interpretation, this experiment showed that simply making interaction techniques more faithful to the real world does not ensure gains in performance (players, however, did rate the Wii Wheel as the most fun technique). In fact, increasing interaction fidelity without making the technique fully natural was actually harmful to performance in this study.

Influence of interaction fidelity in first-person shooters. Our most recent studies[20] have focused on the combined effects of interaction fidelity and display fidelity for the "first-person shooter" (FPS) style of games. We chose the FPS genre because of its demanding interaction requirements, variety of user tasks (including travel, visual search, aiming, and firing), and relevance to serious gaming applications such as military training.

To achieve the highest-possible level of display fidelity in these experiments, we used the DiVE system[d] at Duke University. The DiVE is a completely enclosed cube-shaped 3D projection environment that displays high-resolution stereoscopic 3D graphics on all six sides of the cube and provides 6-DOF wireless tracking (see Figure 3). With this system, we could achieve a very wide range of display and interaction fidelity levels, and therefore could simulate many different possible gaming setups.

In the first study, we wanted to explore the general effects of interaction fidelity and display fidelity, and find out whether one influenced the other. Thus, we designed two levels of each variable, representing "low" and "high" fidelity.

The low interaction fidelity condition used a typical mouse and keyboard interface for FPS games, with the mouse being used to turn, aim, and fire, and the keyboard to travel through the virtual world. The high

d http://vis.duke.edu/dive/

We found the use of head tracking produced significantly fewer errors for the small-scale spatial judgment task, and that head tracking in combination with stereoscopy was significantly faster.

interaction fidelity condition used a tracked handheld controller for direct aiming and firing, and a technique called the "human joystick" for travel. In the human joystick technique, the user would physically step in the desired travel direction, with movement starting once the user stepped outside a small circular area, and the speed of movement proportional to the distance from the center. Although this technique is not highly natural, it has higher interaction fidelity than the keyboard technique due to its use of physical leg movements with direction mapped directly to the environment. A more natural technique such as redirected walking was not practical in the DiVE.

The low display fidelity condition used a single screen of the DiVE without stereoscopic graphics. It therefore also required a method for rotating the view, so we provided a technique that turned the viewpoint when the cursor was near the edge of the screen. The high display fidelity condition used all six screens of the DiVE with stereoscopic graphics enabled, so users could turn physically to view the environment in different directions. This meant that for the mouse and keyboard conditions, users had to be able to turn the mouse and keyboard with them; we placed the devices on a turntable for this purpose.

Participants were placed in an FPS game that required them to navigate several rooms with varying shapes, sizes, and obstacles, destroying "bots" (enemies) along the way. We measured performance metrics such as completion time, shooting accuracy, and damage taken. We also used questionnaires to ask participants about their sense of presence,[30] engagement with the game,[6] and opinions of interface usability.

Performance results were strongly in favor of two conditions: the condition with low display fidelity and low interaction fidelity, and the condition with high display fidelity and high interaction fidelity. These conditions are similar to traditional gaming setups and high-end VR setups that simulate the real world as closely as possible. The other two combinations were unfamiliar to users (despite the fact that they were trained on each combina-

tion and practiced it before completing the trials for that condition); these mismatched conditions resulted in poor performance.

Thus, we took two primary lessons from the first FPS study. First, the effects of interaction fidelity could be dependent on other factors, such as the level of display fidelity. Second, familiarity, rather than interaction fidelity alone, may be the best predictor of performance and usability—the low/low condition was familiar from the PC gaming context, while the high/high condition felt familiar since it simulated the real world.

To explore these effects in a deeper way, we conducted follow-up studies that allowed us to assess individual aspects of interaction fidelity and their influence on the component tasks of an FPS game: long-distance travel, maneuvering (short movements to adjust the viewpoint or avoid an obstacle), searching for enemies, aiming, and firing. We separated the interaction techniques into three parts: travel technique (either keyboard-based or the human joystick), turning technique (either virtual turning without fully surrounding screens or physical turning with fully surrounding screens), and pointing technique (standard mouse, an enhanced mouse technique where the mouse cursor also moves based on physical body turning, and direct pointing at the display).

For travel, our results indicate the keyboard-based technique outperformed the human joystick technique for both long-distance travel and maneuvering. Precise and rapid changes in travel direction proved difficult with the more natural technique due to its reliance on large body movements for control. For searching and aiming, physical turning proved to be quicker than the less natural virtual turning technique. For aiming and firing, direct pointing was the best technique.

From these follow-up studies, we learned that highly natural interaction results in better user performance than traditional, lower-fidelity techniques for our FPS tasks. Similar to the steering study, however, we found traditional interaction techniques can outperform higher-fidelity techniques

Is there an inherent "goodness" to natural mappings, or can we make interaction in the virtual world "better" than interaction in the real world?

that are only partially natural, such as the human joystick technique.

Is Naturalism Worthwhile?
So what have we learned about natural UIs? Many designers' first instinct when developing a 3D UI is to make it as natural as possible—to increase the interaction fidelity so the mapping is as close to the real-world action as it can be. Is there an inherent "goodness" to natural mappings of this sort, or can we make interaction in the virtual world "better" than interaction in the real world? Building on the benefits and limitations of naturalism presented in the accompanying table, we conclude by discussing what we know and what we believe about natural interaction.

Traditional interaction techniques (that is, those not involving 3D interaction) are limited in their potential for naturalism. A travel technique using a game controller cannot approximate real-world walking in the same way that travel based on physical movements can. In our selection and manipulation experiment, there was no way to design a mouse and keyboard interface that was a direct match for the integrated 6-DOF manipulation task, and this resulted in poor performance. But we have also seen that traditional interaction techniques are not always inferior. In the Mario Kart study, the joystick-based techniques performed better than their more natural counterparts, and in the FPS studies, familiar mouse and keyboard interfaces were sometimes just as good as or better than the natural techniques. Traditional UIs have the additional advantage of minimal hardware and sensing requirements, and being well established and ubiquitous. On the other hand, natural techniques may be seen as more fun and engaging.

Natural 3D interaction can be quite beneficial, although this seems to depend on the context and the level of interaction fidelity. Travel based on head tracking was highly beneficial to spatial judgments in the cave visualization study. In the 6-DOF manipulation experiment, the more natural techniques were simply a better match for the task. On the other hand, the natural steering techniques in the

Mario Kart study performed worse than traditional techniques. Highly natural turning, aiming, and firing methods in the FPS studies were very successful, but the moderately natural human joystick approach was difficult to use and hard to understand even after instruction and practice. Simply increasing the level of interaction fidelity does not seem to be sufficient in all cases to ensure usability and performance; naturalism is most effective when very high levels of fidelity can be achieved, and when the resulting interface is familiar to users.

If they are well designed, techniques based on the hyper-natural, magic design approach can feel natural and familiar, while avoiding some of the unwanted side effects of replicating the real world exactly, and providing users with enhanced abilities that improve performance and usability. For example, the Go-Go and HOMER techniques performed well in the 6-DOF manipulation study, allowing the users to feel they were directly manipulating the virtual objects in their hands but without requiring them to travel within reach of the objects to do so. But these enhancements are not without a cost: in the case of Go-Go and HOMER, scaling of hand motions is used to allow flexibility of placement, which reduces precision. The human joystick technique allows long-distance travel without fatigue, but is harder to control than real walking would be. Techniques like scaled HOMER, however, show the potential to design hyper-natural techniques that feel familiar and provide enhancements without sacrificing precision.

Overall, the literature and our studies suggest designers' instinct to strive for natural interaction has merit. Natural UIs are, well, *natural*, and they come easily even for novice users. For certain tasks, like pointing, turning, and 6-DOF manipulation, humans have finely honed abilities that are hard to beat with any other interaction style, as long as the tracking system delivers high-quality data. For certain applications, such as training, using a natural UI helps ensure the training will transfer to the real world. Even when a task would be hard to perform in the real world, hyper-natural UIs can mitigate the difficult aspects of the task while maintaining a natural feel.

But this research also shows that natural UIs are not trivial to design. Making an interaction slightly more natural may actually reduce usability if the resulting UI is unfamiliar and does not achieve high levels of fidelity. And even techniques that seem to be a good approximation of the real world may have negative effects due to small differences, such as the lack of force feedback and the slight latency in the Wii Wheel for steering. Making the choice to use a natural UI does not reduce the importance of the designer; indeed, the designer must be very careful to make good choices when the UI must differ from reality. It is also important to remember that some tasks have no real-world counterpart that can be used as the basis for a natural UI, even though metaphors can often be used to design interactions that are understandable based on real-world experiences.

Another way to frame this question is as a choice between traditional and 3D UIs. When designers can select between traditional interaction styles (for example, mouse and keyboard, game controller), 3D interaction, and other forms of input, what should they choose? The research presented in this article shows 3D UIs are unique in their ability to achieve high levels of interaction fidelity, and that this naturalism can be a significant advantage. For this reason, we expect to continue to see growth in the use of 3D UIs, not just for gaming and mobile applications, but also in many other application domains.

Future Directions

There is still much to be learned about the influence of natural interaction in 3D UIs. Many additional tasks and application contexts must be explored before we can definitively answer questions about the merits of high interaction fidelity.

One important issue is that interaction fidelity is not binary. It would be erroneous to simply say one UI is natural and another is not. Rather, as we have done in this article, we should speak of a continuum of interaction fidelity, with parts of an interface (specific interaction techniques) falling at different locations on the interaction fidelity scale.

But the definition of naturalism is even more nuanced, because the overall level of interaction fidelity is composed of many different elements. Given any pair of interaction techniques for a given task, it may be difficult to say which one has higher interaction fidelity. One may replicate movement trajectories more exactly, while the other more closely approximates the force required of the user, for example. Thus, as we have done previously with display fidelity,[5] we are developing a framework[20] to describe interaction fidelity as a set of components. With this framework, we can design experiments to examine the influence of individual components of interaction fidelity, and gain a deeper understanding of how they relate to performance, usability, and other aspects of the user experience.

References
1. Bowman, D., Datey, A., Ryu, Y., Farooq, U. and Vasnaik, O. Empirical comparison of human behavior and performance with different display devices for virtual environments. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2002), 2134–2138.
2. Bowman, D. and Hodges, L. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the ACM Symposium on Interactive 3D Graphics* (1997), 35-38.
3. Bowman, D., Koller, D. and Hodges, L. Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Proceedings of the Virtual Reality Annual International Symposium* (1997), 45–52.
4. Bowman, D., Kruijff, E., LaViola, J. and Poupyrev, I. *3D User Interfaces: Theory and Practice.* Addison-Wesley, Boston, 2005.
5. Bowman, D. and McMahan, R. Virtual reality: How much immersion is enough? *Computer 40*, 7 (2007), 36–43.
6. Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K.M. and Pidruzny, J.N. The development of the Game Engagement Questionnaire: A measure of engagement in video game playing. *Journal of Experimental Social Psychology 45*, 4 (2009), 624–634.
7. Brogan, D., Metoyer, R. and Hodgins, J. Dynamically simulated characters in virtual environments. *IEEE Computer Graphics & Applications 15*, 5 (1998), 58–69.
8. Brooks, F. What's real about virtual reality? *IEEE Computer Graphics & Applications 19*, 6 (1999), 16–27.
9. Bryson, S. and Levit, C. The virtual wind tunnel. *IEEE Computer Graphics and Applications 12*, 4 1992), 25–34.
10. Chance, S., Gaunet, F., Beall, A. and Loomis, J. Locomotion mode affects the pdating of objects encountered during travel. *Presence: Teleoperators and Virtual Environments 7*, 2 (1998), 168–178.
11. Darken, R., Cockayne, W. and Carmein, D. The omnidirectional treadmill: A locomotion device for

virtual worlds. In *Proceedings of the ACM Symposium on User Interface Software and Technology* (1997), 213–221.

12. Dede, C., Salzman, M. and Loftin, R. ScienceSpace: Virtual realities for learning complex and abstract scientific concepts. In *Proceedings of the Virtual Reality Annual International Symposium* (1996), 246–252.

13. Forsberg, A., Herndon, K. and Zeleznik, R. Aperture-based selection for immersive virtual environments. In *Proceedings of the ACM Symposium on User Interface Software and Technology* (1996), 95–96.

14. Hachet, M., Decle, F., Knodel, S. and Guitton, P. Navidget for easy 3D camera positioning from 2D inputs. In *Proceedings of the IEEE Symposium on 3D User Interfaces* (2008), 83–89.

15. Interrante, V., Ries, B. and Anderson, L. Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. In *Proceedings of the IEEE Symposium on 3D User Interfaces* (2007), 167–170.

16. Iwata, H. Walking about virtual environments on an infinite floor. In *Proceedings of IEEE Virtual Reality* (1999), 286-293.

17. Jacob, R. and Sibert, L. The perceptual structure of multidimensional input device selection. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (1992), 211–218.

18. Kitamura, Y., Higashi, T., Masaki, T. and Kishino, F. Virtual chopsticks: Object manipulation using multiple exact interactions. In *Proceedings of IEEE Virtual Reality* (1999), 198–204.

19. Mapes, D. and Moshell, J. A two-handed interface for object manipulation in virtual environments. *Presence: Teleoperators and Virtual Environments 4*, 4 (1995), 403–416.

20. McMahan, R. Exploring the Effects of Higher-Fidelity Display and Interaction for Virtual Reality Games. Ph.D. dissertation, Virginia Tech, Blacksburg, VA, 2011.

21. McMahan, R., Alon, A., Lazem, S., Beaton, R., Machaj, D., Schaefer, M., Silva, M., Leal, A., Hagan, R. and Bowman, D. Evaluating natural interaction techniques in video games. In *Proceedings of the IEEE Symposium on 3D User Interfaces* (2010), 11–14.

22. McMahan, R., Gorton, D., Gresock, J., McConnell, W. and Bowman, D. Separating the effects of level of immersion and 3D interaction techniques. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (2006), 108–111.

23. Medina, E., Fruland, R. and Weghorst, S. VIRTUSPHERE: Walking in a human size VR hamster ball. *Human Factors and Ergonomics Society Annual Meeting Proceedings 52*, 27 (2008), 2102–2106.

24. Mine, M. ISAAC: A meta-CAD system for virtual environments. *Computer-Aided Design 29*, 8 (1997), 547-553.

25. Olwal, A. and Feiner, S. The Flexible Pointer: An interaction technique for augmented and virtual reality. In *Proceedings of the ACM Symposium on User Interface Software and Technology* (2003), 81–82.

26. Poupyrev, I., Billinghurst, M., Weghorst, S. and Ichikawa, T. The Go-Go Interaction Technique: Non-linear mapping for direct manipulation in VR. In *Proceedings of the ACM Symposium on User Interface Software and Technology* (1996), 79–80.

27. Poupyrev, I., Weghorst, S., Billinghurst, M. and Ichikawa, T. A framework and testbed for studying manipulation techniques for immersive VR. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (1997), 21–28.

28. Ragan, E., Kopper, R., Schuchardt, P. and Bowman, D. Studying the effects of stereo, head-tracking, and field of regard on a small-scale spatial judgment task. *IEEE Transactions on Visualization and Computer Graphics* (In press, 2012).

29. Razzaque, S. Redirected walking. Ph.D. dissertation, University of North Carolina, Chapel Hill, NC, 2005.

30. Slater, M., Usoh, M. and Steed, A. Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments 3*, 2 (1994), 130–144.

31. Smith, R.B. Experiences with the alternate reality kit: An example of the tension between literalism and magic. *SIGCHI Bull. 17* (1986), 61–67.

32. Stoakley, R., Conway, M. and Pausch, R. Virtual reality on a WIM: Interactive worlds in miniature. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (1995), 265–272.

33. Sutherland, I. The ultimate display. In *Proceedings of the IFIP Congress* (1965), 505–508.

34. Sutherland, I. A head-mounted three-dimensional display. In *Proceedings of the Fall Joint Computer Conference* (1968), 757–764.

35. Usoh, M., Arthur, K., Whitton, M., Bastos, R., Steed, A., Slater, M. and Brooks, F. Walking > Walking-in-place > Flying, in virtual environments. In *Proceedings of ACM SIGGRAPH* (1999), 359–364.

36. Vaughan-Nichols, S. Game-console makers battle over motion-sensitive controllers. *Computer 42*, 8 (2009), 13–15.

37. Ware, C. and Franck, G. Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphic 15*, 2 (1996), 121–140.

38. Wilkes, C. and Bowman, D. Advantages of velocity-based scaling for distant 3D manipulation. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (2008), 23–29.

39. Wingrave, C., Tintner, R., Walker, B., Bowman, D. and Hodges, L. Exploring individual differences in ray-based selection: Strategies and traits. In *Proceedings of IEEE Virtual Reality* (2005), 163–170.

40. Zhai, S., Milgram, P. and Buxton, W. The influence of muscle groups on performance of multiple degree-of-freedom input. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (1996), 308–315.

Doug A. Bowman (bowman@vt.edu) is the Director of the Center for Human-Computer Interaction and a professor of computer science at Virginia Tech, Blacksburg, VA.

Ryan P. McMahan (rymcmaha@utdallas.edu) is an assistant professor of computer science at the University of Texas at Dallas.

Eric D. Ragan (eragan@vt.edu) is a Ph.D. candidate in the Center for Human-Computer Interaction and the Department of Computer Science at Virginia Tech, Blackburg, VA.