Considerations for Targets in 3D Pointing Experiments

Wolfgang Stuerzlinger
Simon Fraser University
Vancouver, Canada
http://ws.iat.sfu.ca

Robert J. Teather
McMaster University
Hamilton, Canada
teather@mcmaster.ca

ABSTRACT
We identify various tradeoffs in the design of 3D pointing experiments based on Fitts' law and the ISO9241-9 methodology. The advantages and disadvantages of several approaches for such experiments are analyzed and compared against each other. Results of an experiment that investigates various visual aids are presented as evidence. We conclude with recommendations for 3D pointing experiments and avenues of future work.

Author Keywords
Fitts' law; 3D pointing; ISO 9241-9

ACM Classification Keywords
H.5.2. User Interfaces: Evaluation/methodology.

INTRODUCTION
Pointing at three-dimensional objects to select them is a fundamental task in 3D user interfaces and is analogous to 2D pointing in graphical user interfaces. However, 3D selection is complicated by a number of issues not found in 2D systems. First, 3D graphics systems use perspective; much like reality, far objects appear smaller, which may influence pointing task difficulty. Second, 3D systems often use stereo display to enhance depth perception. Third, there is no universally accepted 3D pointing technique or device. Moreover, there are many different selection methods for 3D targets. Some of the most popular ones require that the users hand or finger intersects the target in space, "laser pointer" techniques, where the user's hand/finger or a device shoots a virtual ray into the scene and the system then selects the first object along that ray, or touch-through, where the user touches the 2D projection of the 3D object on a touch screen.

PREVIOUS WORK
Point selection (or pointing) at targets in 2D has been thoroughly investigated using Fitts' law [16] and the ISO 9241-9 standard [20]. The ISO standard contributes the measure of throughput, which permits comparisons across user strategies by normalizing experimental error rates via a post-experimental accuracy adjustment. This is a primary advantage of the metric – it helps ensure consistency between studies and hence comparability.

In contrast, 3D point selection (sometimes simply "selection") is comparably less well-studied. Numerous 3D pointing studies have been conducted [4, 12, 22, 28, 36], yet 3D pointing is still not as well understood as its 2D equivalent [32]. Most direct 3D selection techniques fall roughly into two broad paradigms: ray-based techniques (including occlusion) and virtual hand techniques [1, 7, 15, 29]. Virtual hands use intersection of the hand/cursor with the target and thus require depth precision. Our current work focuses primarily on techniques requiring depth motions, and hence we mostly focus on virtual hand-based techniques here. Note that these not necessarily use the actual hand (or necessitate finger tracking). For example, our recent work used a tracked stylus to approximate the actual hand position.

Virtual hand (and ray-based) selection techniques are largely equivalent to pointing tasks, i.e., they specify a unique position (of an object) in the environment. While many 3D pointing studies have been performed in the past, we focus here mostly on those based on ISO 9241-9. A study of 3D pointing in a stereo display system [7] revealed that a 2D cursor based method was fastest, followed by a pen-based method and then a ray-based method. A recent investigation of cursor-based techniques for 3D pointing also found that methods that 2D cursor-based techniques perform best [38]. Such techniques include the mouse as well as ray-casting with a cursor at the position where the ray hits the screen or the scene. One of the noteworthy outcomes of this work is a 2D model that accounts for perspective distortion, which describes cursor-based pointing at 3D targets quite well. Moreover, displaying the cursor only to the dominant eye was found to significantly improve performance in a stereo display system, even for classic ray casting. A recent comparison of touch-based methods on stereo surfaces revealed that 2D touch, i.e. touch-through, works well for objects within approximately 10cm of the surface, but that 3D touch is better for objects that further from the display [11].

Research comparing pointing in the real world vs. virtual reality indicates that VR performance is substantially worse

[22, 27, 33]. Several factors contribute to this difference, notably including input latency and noise [36], tracker registration [33] and tactile feedback [12]. However, visual cues (e.g., depth cues and "artificial" feedback mechanisms) are critical as the largest differences occur during the correction phase of motion, where visual feedback is used in a tight feedback loop [22, 27].

Due to the direct correspondence between input and display spaces, target selection is affected by several visual cues and feedback mechanisms. Early work [4, 5] focused on stereo and head-tracking and found that target position significantly affected task completion time and accuracy. Depth movements were slower and less accurate than screen-parallel movements. Participants were better able to judge depth with stereo enabled. These results were later confirmed in a docking task [5], where stereo significantly reduced movement error in depth.

Other visual aids are important in 3D pointing. Partial target occlusion improves selection with volumetric cursors, especially when combined with stereo [42]. Visual feedback also improves object position memorization [13]. Color change is a commonly used visual feedback mechanisms. The recent Virtual Mitten technique [1], for example, uses color changes to indicate pressure applied to a handheld grip device. Other recent work focused on visual feedback for hand-based grasp techniques [30]. The authors report that participants preferred changing the selected object color, even though it did not necessarily offer the best performance. Similar approaches improved participant speed and accuracy of 2D pointing in sub-optimal viewing conditions [17]. However, highlighting selected targets in a 2D pursuit-tracking task did not improve performance [26].

ISSUES IN 3D POINTING EXPERIMENTS

While it may seem straightforward to extend the ISO9241-9 methodology to 3D pointing, there are several issues in the domain of touch-based systems, but also relevant in other input methodologies. Here we review the most important ones.

Stereo viewing

One of the main issues is the stereo conflict inherent to any selection interface presented in stereo 3D. There are two main problems. First, the human visual system unable to focus simultaneously at objects at different visual depths (e.g., a target presented behind a finger or stylus). Converging the eyes on one feature will "double" the image of the other, yielding diplopia. This especially impacts systems that use 2D touch (i.e., on a touchscreen) for interaction with objects presented in 3D at different depths from the screen [11]. Second, most stereo systems also suffer from the vergence-accommodation conflict. When focusing on a 3D target displayed on the screen viewers will see a blurred finger or when focusing on the finger, they will see a blurred target [9]. This impacts systems using direct interaction (e.g., 3D touch in space) with stereo-presented targets [37]. Note also that any issues in depth perception impact not only the initial ballistic phase of pointing motions (as the motor program may target the wrong location in space), but also the final correction phase (where visual cues are very important).

In pilot testing, we confirmed that stereo together with head-tracking was important for selecting the 3D position of an object. Participants were unable to reliably detect and touch the 3D position of an object using a tracked stylus without these cues. Using only one or the other was also insufficient, even in the presence of other cues (e.g., "support" cylinders, selection feedback, texturing, see below for discussion of each).

Cursor-based selection methods avoid the full 3D pointing problem, as they select objects visible from the viewer or along a ray. With this, a 2D manifold effectively describes everything that can be selected. Evidence confirms that a 2D model describes the performance of such techniques quite well [38]. Interestingly, displaying the cursor only to one eye cancels any negative effects of stereo conflicts. Yet, offset-based, i.e. cursor-based, methods do not work as well as direct touch methods for 3D pointing [10].

Recent work [38] indicates that using such cursor visualizations certainly help when targets are presented at a greater depth than that of the cursor. Note that such situations arise when attempting to use an OS system cursor (typically *not* displayed in stereo, but may be used by default, e.g., in stereo games), which is effectively presented to both eyes with zero-disparity. This yields a "stereo" cursor presented in the plane of the screen. However, there is some debate about potential negative effects of the one-eyed cursor, for example, eye fatigue. Schemali and Eisemann [31] report that the one-eyed cursor offered significantly worse performance than specially designed stereo cursors. They reason that this is due to eye fatigue. Our recent work attempted to replicate these results [40]. While they found small negative effects for using the one-eyed cursor in isolation from its benefits (e.g., in a binocular scene with zero-disparity) these effects were not significant.

Non-spherical hit distribution

In some ways more worrisome is that the fact that the distribution of 3D "hit" points in a 3D mid-air pointing experiment is not spherical. The most likely cause for this is depth perception inaccuracies. This is best illustrated by a analysis of 3D touch on a tabletop [9], see Figure 1. The main issue associated with such hit point distributions is that the notion of throughput in ISO9241-9 *relies* on a (at least approximately) spherical hit distribution for the effective measures [32]. Strong deviations from that distribution may invalidate the underlying assumption(s) that enable the combination of speed and accuracy into a single measure.

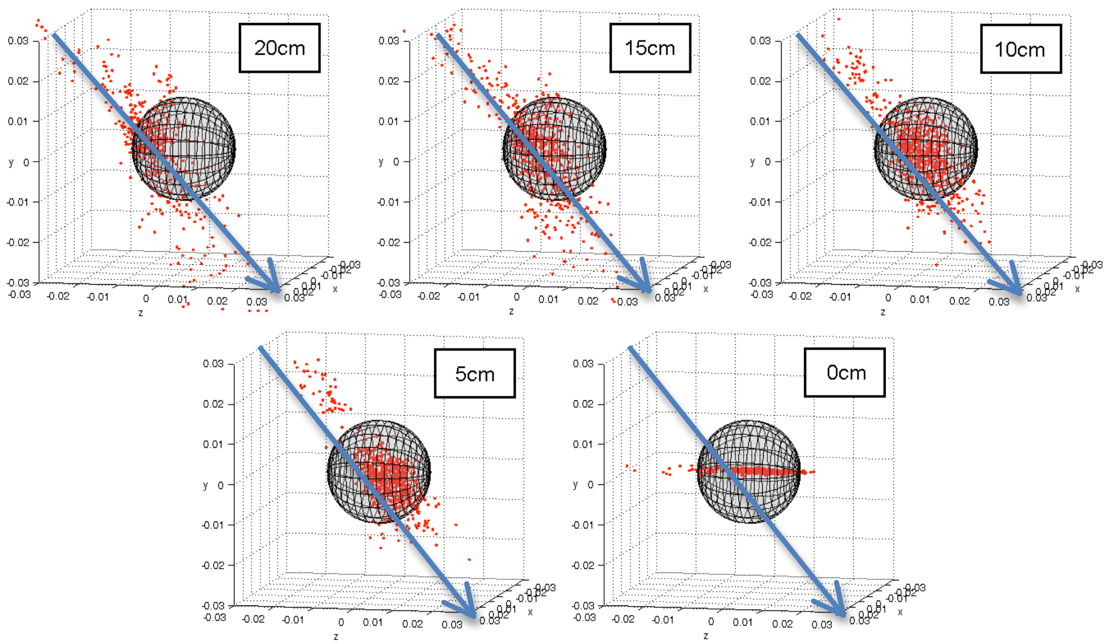


Figure 1. Distribution of selection points for a 3D touch technique at various heights above a display. Black wireframe spheres indicate the targets. The diagonal arrow illustrates the normalized view angle. Figure reproduced from [9], with permission.

Figure 2, for example, depicts selection coordinate distributions from a recent experiment on visual aids in 3D selection [39]. This scatter plot depicts the impact of selection feedback (discussed further below) on the distribution of selection coordinates. These selection coordinates are *roughly* spherical closer to the 0 cm depth (i.e., at the screen). Clearly the impact of the screen surface "flattens" the selection distribution somewhat in the 0 cm condition (the right-most column). However, for targets farther from the display, these distributions tend to become less round and more scattered. Other recent work [11] indicates that selection distributions can be far more oblong than spherical, with selection coordinates scattered along the movement axis, see [11].

Floating targets

Volumetric 3D targets are the natural extension of 2D ISO9241-9 targets. Displaying such targets as solid objects is not advisable, as the user can then not tell if the cursor is inside or behind the volume. Thus, most studies use semi-transparent volumetric targets. Yet, such transparent objects floating in space have few, if any, equivalents in the real world, i.e., they do not correspond to any real pointing task. The closest is popping soap bubbles. This reduces the external validity of volumetric targets. Note that Fitts' law [16] describes rapid aimed movements in the real world, so this is a concern. Based on this reflection, some research groups use other objects, such as cylinders, as bases or "pedestals" for the targets, to visually "anchor" the targets in 3D space.

Recent work [39] investigated the impact of floating targets on selection performance. Results of this study indicate that the presence of "support cylinders" did not significantly affect targeting performance. However, participant feedback almost universally favored the use of cylinders. All but two participants felt that the presence of cylinders helped them perform the point selection task used in the study, indicating "cylinders made the task a little bit easier" or "a lot easier". Consequently, this issue may warrant further study.

Target shapes

Another important question concerns the shape of the target area or volume for a 3D pointing motion. Here are the most relevant options, see also Figure 3:

- disc
- sphere
- hemi-sphere
- cylinder
- oriented cylinder
- oriented truncated cone

The differences between these targets become apparent when one compares the 3D target volume with their visual appearance from the user's view.

The advantage of the disc is that it is equivalent to a 2D target, which enables direct comparisons between 2D and 3D pointing. Any of the other target types suffer from the problem that one is comparing a target area against a target volume. A disadvantage of disc targets is that they are view dependent, i.e., their visual profile depends on the viewing angle.

Spherical targets are the natural 3D equivalent of 2D disc ISO9241-9 targets. A disadvantage of spheres is that one cannot simply put a sphere on top of the display itself nor on “pedestals”. In this situation, the user will then try to hit the sphere by touching the screen/surface (which is efficient), but fail to select the target as the sphere touches the screen/surface only at a single, infinitesimal small, point. One option is to use a hemi-sphere instead. Yet, this primitive has half the volume, which may distort the computation of various measures, including effective target widths.

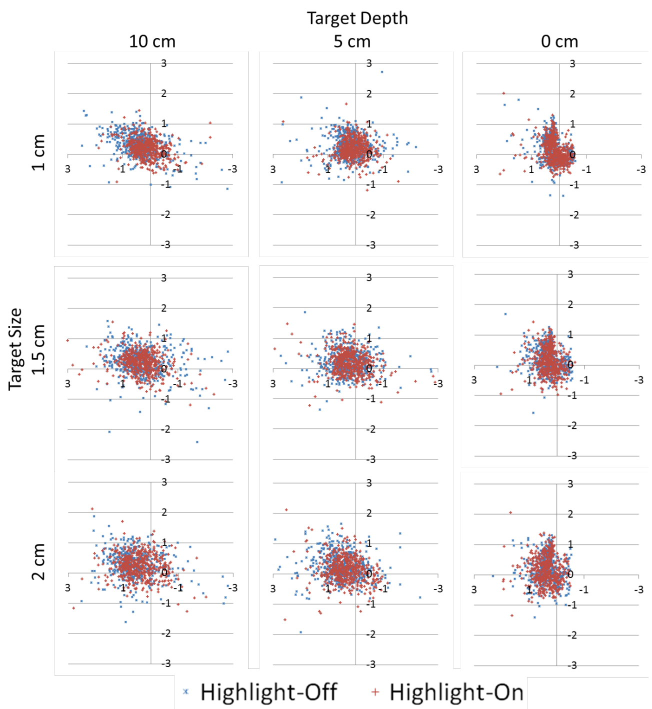


Figure 2. Selection distributions for separated by target depth and target size. Selections with feedback (highlighting) are shown in red, and without feedback (no-highlighting) are shown in blue. The horizontal axis is the *z*-axis of selection coordinates, while the vertical axis shows the *y*-axis (in cm).

Cylinders are the extrusion of a disc. A disadvantage is that the visible cross-section relative to the ideal is larger. To address this, one can rotate the cylinder towards the viewer to make the visual profile equal to a disc/sphere. The disadvantage of this approach is that an oriented cylinder close to the viewer will have a significantly smaller base relative to the top, which is closer to the viewer. To address this, one can use an oriented truncated cone, which appears as a disc when oriented towards the viewer.

Note that the natural selection distribution, as discussed earlier, may influence the decision of target shape. For example, given the inaccuracies of most 3D input devices, disc-shaped targets are unlikely to work reliably - selecting such a target in the absence of tactile feedback is likely impossible and may necessitate using a crossing paradigm experiment instead [2]. Similarly, while spheres appear a natural choice, selection coordinates tend to vary along the depth axis [39] or along the 3D movement vector [11]. Consequently, ellipsoids may also be a reasonable choice.

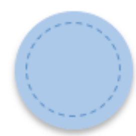
Shape	3D sketch	2D view
Disc		
Sphere		
Hemi-sphere		
Cylinder		
Oriented cylinder		
Oriented truncated cone		

Figure 3. Illustration of target shapes. 3D sketches shown in a side view for a viewer position from top right. 2D views shows what the viewer sees.

Selection Feedback

Several cues indicate when we have touched a target in reality, including tactile feedback and stereo viewing with correct vergence and accommodation. However, most VR systems do not present these correctly, if at all. Consider, for example, selecting a 3D target using a tracked finger. Due to the absence of tactile feedback, the finger will pass *through* the target. Stereo cues now indicate that the target is in front of the finger, while occlusion cues indicate the opposite – the finger *always* occludes the screen. Consequently, another means of selection feedback is required.

Recent work [11, 37] used target highlighting for this very reason. When the target is touched, it changes colour. This provides feedback that selection (e.g., via a button) will be successful, and helps the user choose between multiple targets. This has been shown to have a notable influence on 3D point selection performance [39]. In particular, highlighting was found to *increase* movement time significantly. However, it simultaneously cut error rates by about half. The authors report that it had a substantially stronger effect than cylinders for support, or texturing - neither of which significantly influenced the study results.

RECOMMENDATIONS

Based on the reflections above, we recommend addressing the above-mentioned issues as follows:

As the accommodation-vergence conflict is inherent to current stereo displays, it cannot be directly addressed. Yet, stereo is not the strongest depth cue [14]. Thus, one strategy is to use head-tracking (i.e. motion cues), textures, pedestals, a surrounding environment, and other methods to improve depth perception. Then users do not have to rely on stereo alone.

Placing targets on (textured) pedestals aids users in their depth perception and makes targets easier to hit. Consequently, we recommend the usage of cylinders or other objects to visually anchor targets in space.

Beyond this, we also recommend including a 2D pointing technique with any 3D study to increase external validity. Ideally, this should be a "best practice" comparison for 2D *and* 3D techniques.

Future Work

To address the issue of the non-spherical hit distributions observed in 3D pointing experiments, appropriate 3D generalizations of Fitts' law are needed, which take the depth dimension correctly into account. We plan on working on this in the future.

Any generalization of Fitts' law to 3D should use semi-transparent spheres or oriented truncated cones as targets. Yet, comparisons with 2D techniques require disc (or maybe hemi-spherical) targets. This calls for the development of methodologies that can compare 2D and 3D targets in one framework.

REFERENCES

1. Achibet, M., Marchal, M., Argelaguet, F., and Lécuyer, A., The Virtual Mitten: A novel interaction paradigm for visuo-haptic manipulation of objects using grip force, IEEE Symposium on 3D User Interfaces - 3DUI 2014, 59-66.
2. Apitz, G., Guimbretière, F., and Zhai, S., Foundations for designing and evaluating user interfaces based on the crossing paradigm, ACM Transactions on Computer-Human Interaction, 17, 2010, 9.
3. Argelaguet, F. and Andujar, C., A survey of 3D object selection techniques for virtual environments, Computers & Graphics, 37, 2013, 121-136.
4. Boritz, J. and Booth, K. S., A study of interactive 3D point location in a computer simulated virtual environment, ACM Symposium on Virtual Reality Software and Technology - VRST '97, 181-187.
5. Boritz, J. and Booth, K. S., A Study of Interactive 6 DOF Docking in a Computerized Virtual Environment, IEEE Virtual Reality Annual International Symposium, 139-146.
6. Bowman, D. A, Johnson, D. B., and Hodges, L. F., Testbed evaluation of virtual environment interaction techniques, ACM Symposium on Virtual Reality Software and Technology - VRST '99, 26-33.
7. Bowman, D. A. , Kruijff, E., LaViola, J. J. , and Poupyrev, I., 3D User Interfaces: Theory and Practice. Addison Wesley Longman Publishing Co., Inc., 2004.
8. Bowman, D. A., Wingrave, C. A., Campbell, J. M., Ly, V. Q., and Rhoton, C. J., Novel uses of pinch gloves for virtual environment interaction techniques, Virtual Reality, 6, 2002, 122 - 129.
9. Bruder, G., Steinicke, F., and Stuerzlinger, W., Touching the void revisited: Analyses of touch behavior on and above tabletop surfaces, Human-Computer Interaction–INTERACT 2013. 278-296.
10. Bruder, G., Steinicke, F., and Sturzlinger, W., Effects of visual conflicts on 3D selection task performance in stereoscopic display environments, IEEE Symposium on 3D User Interfaces - 3DUI 2013, 115-118.
11. Bruder, G., Steinicke, F., and Sturzlinger, W., To touch or not to touch? Comparing 2D touch and 3D mid-air interaction on stereoscopic tabletop surfaces, ACM Symposium on Spatial User Interaction - SUI 2013, 9-16.
12. Chun, K., Verplank, B., Barbagli, F., and Salisbury, K., Evaluating haptics and 3D stereo displays using Fitts' law, IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications, 53-58.
13. Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., and Looser, J., Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback, International Journal of Human-Computer Studies, 69, 2011, 401-414.
14. J.E. Cutting, P.M. Vishton, Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth, in Perception of space and motion, Academic Press, vol. 5, 69-117, 1995.

15. Dang, N.-T., A survey and classification of 3D pointing techniques, IEEE International Conference on Research, Innovation and Vision for the Future, 71 - 80.

16. Fitts, P. M., The information capacity of the human motor system in controlling the amplitude of movement, Journal of Experimental Psychology, 47, 1954, 381-391.

17. Fraser, J. and Gutwin, C., The effects of feedback on targeting performance in visually stressed conditions, Graphics Interface, 2000, 25-32.

18. Hoffman, D. M., Girshick, A. R., Akeley, K., and Banks, M. S., Vergence-accommodation conflicts hinder visual performance and cause visual fatigue, Journal of Vision, 8(3), 2008, 1-30.

19. Hubona, G. S., Wheeler, P. N., Shirah, G. W., and Brandt, M., The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization, ACM Transactions on Computer-Human Interaction, 6, 1999, 214-242.

20. ISO 9241-9 Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices: International Standard, International Organization for Standardization, 2000.

21. Kulshreshth, A. and LaViola Jr, J. J., Evaluating performance benefits of head tracking in modern video games, ACM Symposium on Spatial User Interaction - SUI 2013, 53-60.

22. Liu, L., Liere, R. v., Nieuwenhuizen, C., and Martens, J.-B., Comparing aimed movements in the real world and in virtual reality, IEEE Virtual Reality Conference - VR 2009, 219-222.

23. Lubos, P., Bruder, G., and Steinicke, F., Analysis of direct selection in head-mounted display environments, IEEE Symposium on 3D User Interfaces - 3DUI 2014, 11-18.

24. MacKenzie, I. S., Fitts' law as a research and design tool in human-computer interaction, Human-Computer Interaction, 7, 1992, 91-139.

25. MacKenzie, I. S. and Isokoski, P., Fitts' throughput and the speed-accuracy tradeoff, ACM CHI 2008, 1633-1636.

26. Mould, D. and Gutwin, C., The effects of feedback on targeting with multiple moving targets, Graphics Interface 2004, 25-32.

27. Nieuwenhuizen, K., Liu, L., Liere, R. v., and Martens, J.-B., Insights from Dividing 3D Goal-Directed Movements into Meaningful Phases, IEEE Computer Graphics and Applications, 29, 2009, 44 - 53.

28. Pawar, V. M. and Steed, A., Profiling the behaviour of 3D selection tasks on movement time when using natural haptic pointing gestures, ACM Symposium on Virtual Reality Software and Technology - VRST 2009, 79-82.

29. Poupyrev, I., Ichikawa T., Weghorst, S., and Billinghurst, M., Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques, Eurographics '98, 41-52.

30. Prachyabrued, M. and Borst, C. W., Visual feedback for virtual grasping, IEEE Symposium on 3D User Interfaces - 3DUI 2014, 19-26.

31. Schemali, L. and Eisemann, E., Design and evaluation of mouse cursors in a stereoscopic desktop environment, IEEE Symposium on 3D User Interfaces - 3DUI 2014, 67-70.

32. Soukoreff, R. W. and MacKenzie, I. S., Towards a standard for pointing device evaluation: Perspectives on 27 years of Fitts' law research in HCI, International Journal of Human-Computer Studies, 61, 2004, 751-789.

33. Sprague, D. W., Po, B. A., and Booth, K. S., The importance of accurate VR head registration on skilled motor performance, Graphics Interface 2006, 131-137.

34. Stuerzlinger, W., Considerations for targets in 3D pointing experiments, Workshop on Interactive Surfaces for Interaction with Stereoscopic 3D - ISIS3D 2013.

35. Surdick, R. T., Davis, E. T., King, R. A., Corso, G. M., Shapiro, A., Hodges, L., and Elliot, K., Relevant cues for the visual perception of depth: is where you see it where it is?, Human Factors and Ergonomics Society Annual Meeting, 1305-1309.

36. Teather, R. J., Pavlovych, A., Stuerzlinger, W., and MacKenzie, I. S., Effects of tracking technology, latency, and spatial jitter on object movement, IEEE Symposium on 3D User Interfaces - 3DUI 2009, 43-50.

37. Teather, R. J. and Stuerzlinger, W., Pointing at 3D targets in a stereo head-tracked virtual environment, IEEE Symposium on 3D User Interfaces - 3DUI 2011, 87-94.

38. Teather, R. J. and Stuerzlinger, W., Pointing at 3D target projections using one-eyed and stereo cursors, ACM CHI 2013, 159 - 168.

39. Teather, R.J. and Stuerzlinger, W., Visual aids in 3D point selection experiments, ACM Symposium on Spatial User Interaction - SUI 2014, 127-136.

40. Teather, R.J., and Stuerzlinger, W., Depth cues and mouse-based 3D target selection, Poster at the ACM Symposium on Spatial User Interaction - SUI 2014, 156-156.

41. Valkov, D., Steinicke, F., Bruder, G., and Hinrichs, K., 2D touching of 3D stereoscopic objects, ACM CHI 2011, 1353-1362.

42. Zhai, S., Buxton, W., and Milgram, P., The “Silk Cursor”: Investigating transparency for 3D target acquisition, ACM CHI '94, 459 - 464.

43. Zhai, S., Milgram, P., and Rastogi, A., Anisotropic human performance in six degree-of-freedom tracking: An evaluation of three-dimensional display and control interfaces, IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 27, 1997, 518-528.