

The Effects of Changing Projection Geometry on Perception of 3D Objects on and Around Tabletops

MIGUEL A. NACENTA, University of St Andrews
MARK HANCOCK, University of Waterloo
CARL GUTWIN, University of Saskatchewan
SHEELAGH CARPENDALE, University of Calgary

Displaying 3D objects on horizontal displays can cause problems in the way that the virtual scene is presented on the 2D surface; inappropriate choices in how 3D is represented can lead to distorted images and incorrect object interpretations. We present four experiments that test 3D perception. We varied projection geometry in three ways: type of projection (perspective/parallel), separation between the observer's point of view and the projection's center (discrepancy), and the presence of motion parallax (with/without parallax). Projection geometry had strong effects different for each task. Reducing discrepancy is desirable for orientation judgments, but not for object recognition or internal angle judgments. Using a fixed center of projection above the table reduces error and improves accuracy in most tasks. The results have far-reaching implications for the design of 3D views on tables, in particular, for multi-user applications where projections that appear correct for one person will not be perceived correctly by another.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; *Virtual reality*; *Touch screens*; *Collaborative and social computing*; Empirical studies in visualization; • **Computing methodologies** → **Perception**; *Virtual reality*

Additional Key Words and Phrases: Interactive tabletops, 3D perception, 3D representation, 3D visualization, projection, point-of-view, center-of-projection, fish-tank virtual reality

ACM Reference Format:

Miguel A. Nacenta, Mark Hancock, Carl Gutwin, and Sheelagh Carpendale. 2016. The effects of changing projection geometry on perception of 3D objects on and around tabletops. *ACM Trans. Comput.-Hum. Interact.* 23, 2, Article 11 (May 2016), 54 pages.
DOI: <http://dx.doi.org/10.1145/2845081>

1. INTRODUCTION

The range of applications that are being built for tabletop displays is rapidly increasing, and now includes systems that use 3D objects and environments [von Zadow et al.

This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Digital Surface Software Application Network (Surfnet), by the Canadian Networks of Centres of Excellence (NCE) through the Graphics, Animation and New Media (GRAND) Network, by Sheelagh Carpendale's Canada Research Chair (CRC) on Information Visualization, by Sheelagh Carpendale's NSERC/Alberta Innovates Technology Futures (AITF)/Smart Technologies Industrial Research Chair on Interactive Technologies, and by the Scottish Informatics and Computer Science Alliance (SICSA).

Authors' addresses: M. A. Nacenta, School of Computer Science, University of St Andrews, St Andrews KY16 9SX, United Kingdom; email: mans@st-andrews.ac.uk; M. Hancock, Department of Management Sciences, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada; email: mark.hancock@uwaterloo.ca; C. Gutwin, Department of Computer Science, University of Saskatchewan, 110 Science Place, Saskatoon, Saskatchewan S7N 5C9, Canada; email: gutwin@cs.usask.ca; S. Carpendale, Department of Computer Science, University of Calgary, 2500 University Dr. NW, Calgary, Alberta T2N 1N4, Canada; email: sheelagh@cpsc.ucalgary.ca.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2016 ACM 1073-0516/2016/05-ART11 \$15.00

DOI: <http://dx.doi.org/10.1145/2845081>



Fig. 1. Two perspectives on the same presentation of a 3D model on a table; from one side the image looks correct, but from the other the image is distorted.

2013; Hancock et al. 2007; Wilson et al. 2008; Agrawala et al. 1997; Hancock et al. 2010]. Some 3D application areas can benefit greatly from the combination of 3D data presentation and the natural collaboration and face-to-face communication afforded by tabletop systems—examples include planning of surgical procedures with 3D body models, urban planning discussions using models of real-world structures, and collaboration over 3D visualizations.

These applications rely on a person's ability to accurately perceive 3D structures in the flat, 2D representations displayed on a screen. Perceiving 3D structure from a 2D representation is not an uncommon task: we do it every day when we view photos (printed or displayed on a monitor), look at 3D diagrams of buildings or objects (e.g., the instructions of build-your-own furniture), or play video games. However, there are specific problems associated with visualizing 3D content on horizontal displays; in particular, images on tabletops can be looked at from any direction, unlike most vertical displays or paper documents where there is a clear *up* direction. This problem has been identified for 2D content [Scott et al. 2003], but it is much more evident when a 3D representation is used. Several difficulties can arise, including distortion, problems in recognizing objects or assessing angles, and communication errors when referring to objects (e.g., “the buildings that are tilted towards me” in Figure 1 left will not make sense for an observer looking from the other side—Figure 1, right).

Although advanced display technologies exist that could address this problem by providing different images for each participant (e.g., Matshushita et al. [2004] and Karnik et al. [2012]), or by generating volumetric images [Geng 2013], at this point in time these are expensive, impractical, and uncommon compared to ubiquitous flat single-image displays.

There is currently very little understanding of how different kinds of 3D projections affect people carrying out tasks in tabletop environments equipped with single-image flat displays. This work provides a comprehensive set of empirical data that can be used by designers of 3D tabletop systems. We carried out four studies that evaluated the effects of different perspectives and projections on several perceptual tasks that are fundamental to working with 3D content.

We tested two different *types of projection* (perspective and parallel), three different locations of the viewer relative to the point of projection, which we call *discrepancy* (no discrepancy, medium discrepancy, and large discrepancy), and two conditions of *motion parallax* (i.e., whether real-time head-motion parallax was available or not). Our three perceptual tasks were: *estimation of object orientation*, which is fundamental for comprehending spatial relationships in the model; *estimation of internal angles*, an integral part of understanding the structure of 3D objects; and *object recognition*, a high-level task required for effective navigation and manipulation of 3D content.

The findings from the four studies provide a wealth of new information about how people interpret and work with 3D content on tabletop displays. There are three main results from the studies:

First, as discrepancy increased (that is, as the projection point moved further away from the observer), people's error in estimating orientation significantly increased, but this effect was not present for estimation of internal angle or object recognition.

Second, when parallel projection geometry is used in combination with a projection point above the table (between the two viewers), participants more accurately estimated internal angle and orientation, and better recognized objects. This result contradicts two common beliefs: that designers must minimize discrepancy to improve interpretation, and that perspective projections provide a more accurate representation of 3D content.

Third, participants performed significantly better when interactivity was provided through direct interaction (i.e., touch-based rotation of objects) than when participants manipulated the view through their head motion (i.e., motion parallax) or when no interactivity was available. This suggests that problems of interpretation can also partly be mitigated by providing direct interaction.

The empirical evidence provides several new principles for the design of 3D projections on horizontal surfaces. Overall, our studies show that optimizing the perspective for one person at a shared table is likely to cause major problems for the others in the group, especially in tasks that involve estimation of object orientation, but that providing a neutral projection above the table and using parallel projection geometry can mitigate the problems. In addition, applications that are sensitive to errors must be designed with great care, since some projection types saw large estimation and recognition errors (e.g., up to 60° orientation error), and object recognition rates that were no better than chance. In addition, the performance differences we found must be balanced against several design trade-offs, including the cost and feasibility of introducing viewer tracking (necessary to provide discrepancy-free projections), and the problem that an optimal projection for one viewer may imply the worst possible view for another.

The rest of this article is organized as follows: first we provide additional information on the problem of 3D projection for tabletop surfaces, and introduce important concepts and terms that are used in the studies; second, we present a summary of fundamental concepts underlying 3D representation and perception and review related work from the areas of interactive displays, information visualization, and perception research; third, we report on the four studies; and fourth, we provide an overall discussion of our findings and implications for designers of 3D tabletop systems.

2. THE PROBLEM OF VIEWING 3D CONTENT ON TABLETOPS

When representing 3D content on horizontal surfaces, one must decide on a particular projection—that is, a method of mapping points from the 3D model to points on the 2D plane.¹ There are two main kinds of projection: perspective and parallel. A perspective projection maps points to the 2D plane by creating straight lines (rays) from every point in the 3D model to a single origin point; the intersection of the rays with the 2D tabletop determines the geometry of the model on the horizontal surface (see Figure 2, left). The single origin point is called the center of projection, or CoP. Perspective projection is used in most 3D applications of computer graphics and virtual reality (VR).

The second main type of projection is parallel projection [Carlbom and Paciorek 1978], which creates rays that are parallel to one another from each point of the 3D model (Figure 2, right); the rays project in a fixed direction instead of converging to a CoP, and thus, a parallel projection does not have a true CoP (it has a *direction of projection*—although for simplicity, in this article, we will refer to both as CoP). Parallel

¹http://en.wikipedia.org/wiki/3D_projection; note also that this term does not refer to projector-based displays (which are often used to implement tabletop systems).

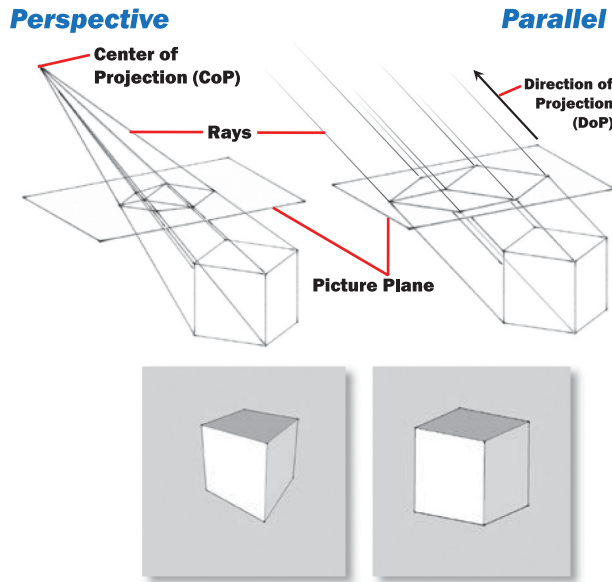


Fig. 2. Principles of perspective projection (top left), parallel projection (top right), a cube represented in perspective projection (bottom left), and a cube represented in parallel projection (bottom right).

projection is often used in construction and engineering drawings because the image maintains scale and parallelism from the 3D model.²

In a tabletop projection of 3D content, there are, therefore, two important reference points: the point or direction from which the projection emanates (the CoP), and the location of the viewer's eyes (i.e., the *point of view*, or PoV). If the center of projection and the point of view are different (i.e., there is a *discrepancy* between the CoP and the PoV), then the viewer must look at an image that does not correspond geometrically to what a real object would project in their retinas from their current perspective.³ As a consequence, the user's view of the projected image will appear distorted (see Figure 1).

Distortion caused by discrepancy can lead to misperception of objects. Accurate perception of 3D objects is critical in many applications such as medical image visualization [Kitamura et al. 2007; Bichlmeier et al. 2009], and industrial or architectural applications [Kruger et al. 1995; Agrawala et al. 1997]. Moreover, since the distortion effects will be different for different viewers, this can lead to miscommunication between collaborators, who might not be aware of how their point of view affects what they perceive. It is not uncommon to refer to and communicate about objects through attributes such as orientation and size, which can both be affected by perspective distortion.⁴

²Several types of parallel projections exist. For example, non-orthographic oblique projections such as cavalier “bends” one of the axes at a non-90° angle in its representation. However, most of the common projections used in current graphics are either perspective, or sub-types of parallel projections (e.g., isometric projection is a parallel projection with a CoP at 45° with respect to all the main references axes).

³This geometrical correspondence is only part of the match that we can achieve between the retinal images generated by real objects and their representations. Other cues such as focus blur cannot be matched by planar monocular projections.

⁴It is technically possible to provide a different image for each viewer of the horizontal display, (e.g., by using a light field display or a multi-image display [Grossman and Wigdor 2007; Kitamura et al. 2001; Holografika Kft 2011]). However, these technologies are uncommon in current collaborative settings due to price, limitations in angle of view, and because they can require tracking the positions of all participants or

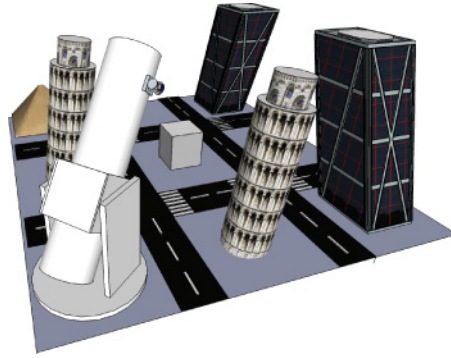


Fig. 3. 3D model used as example in the rest of the article.



Fig. 4. Left: CoP and PoV both above the table. Right: CoP above the table and PoV at the end of the table.

Displaying 3D objects on tables, therefore, presents new problems for designers. Choices made in creating the 3D image such as the location of the CoP, and whether a perspective or parallel projection is used, can change the appearance of the resulting scene. To illustrate the details of the problem, consider the 3D model in Figure 3, which contains a collection of buildings, many of which are not entirely perpendicular to the ground.

Typically, 3D graphics applications use a CoP that is in the center and directly in front of the screen (above the table for horizontal displays). Figure 4, left shows what one would see if hovering above the table; in this image, the CoP and PoV coincide. Figure 4, right shows a view of the same scene but from the end of the table. There is now a discrepancy between the PoV (at the end of the table) and the CoP (still above the table).

Ideally when standing beside a digital table one would like to get a good 3D impression of this model. For instance, one might expect to see the model as shown in Figure 5, left. This image again has a CoP and PoV that coincide, but this time both are located at the end of the table. However, if one were to walk to the opposite side of the table one would get the experience shown in Figure 5, right.

To further illustrate the problem, this same model is projected into the 2D images shown in Figure 6 using different projection geometries and CoPs. The letters A-F correspond to the same labels in Figure 7, which shows what the images would look

rendering images for all possible points of view. The overwhelming majority of tabletop systems currently installed and in production are single image, that is, monocular.



Fig. 5. Left: CoP and PoV both at the near end of the table. Right: CoP at far end of the table, PoV at near end.

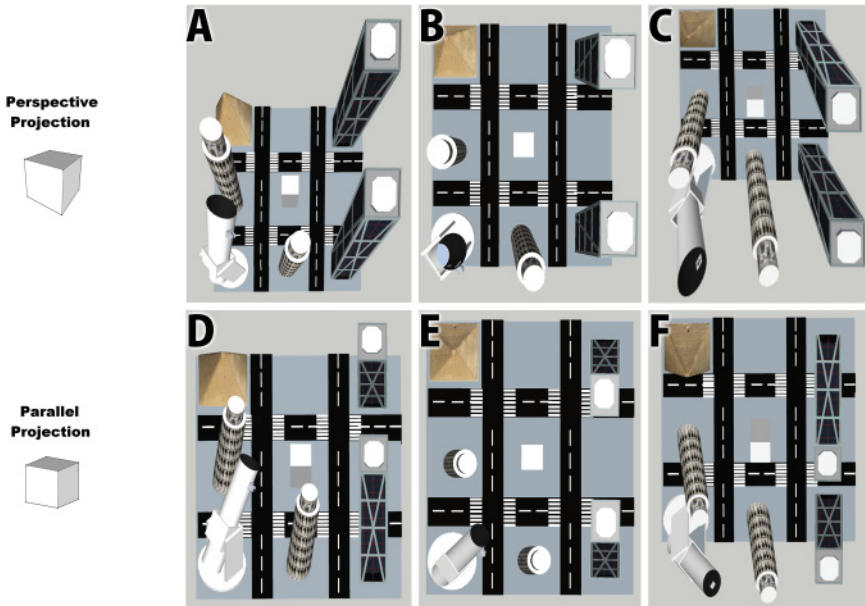


Fig. 6. Different CoPs and projection types, with PoV above the table. (A) CoP at South end of the table, (B) CoP on above the table, (C) CoP at North end, (D) CoP at South end, (E) CoP from top above, (F) CoP from at the North end. (A)–(C) use a perspective projection, while (D)–(F) use a parallel projection.

like when observed from different PoVs at a tabletop display. These representations coincide with the conditions of our empirical study.

The problems described above are the subject of our studies below; first, however, we provide additional background on representation and perception of pictorial space.

3. BACKGROUND AND RELATED WORK

Our work draws from three main areas of study: perception, graphics, and human-computer interaction. The following sections offer first a basic overview of the concepts that are necessary to describe our research and present a summary of the state-of-the-art.

3.1. Depth Cues and Perception

The visual system relies on several sources of information—visual cues—for volumetric perception; according to Cutting and Vishton [1995], there are nine basic and independent sources of information for depth perception: occlusion, relative size, relative

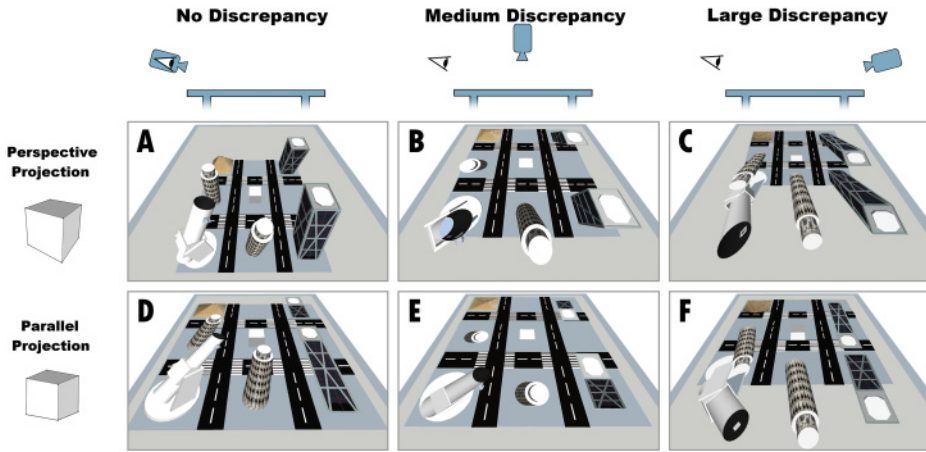


Fig. 7. Six combinations of discrepancy amount and projection type, seen from the PoV of a viewer sitting at the South end. (A)–(F) correspond to the same projections in (A)–(F) of Figure 6.

density, height in visual field, aerial perspective, motion perspective, convergence, accommodation, and binocular disparity, stereopsis, and diplopia. Not all of these sources need to be present simultaneously; in fact, the first four, which together form what is understood as linear perspective [Carlson and Paciorek 1978], are sufficient to provide a perception of depth, as is evident from our ability to perceive depth in photographs and 3D technical drawings. Furthermore, some of these cues are relevant only for certain ranges; for example, accommodation is only effective in the first few meters, while aerial perspective—the loss of air transparency shown by very distant objects due to atmospheric light scattering—only applies to objects very far away. Most common existing displays only portray linear perspective, with a few offering binocular disparity and motion perspective. A few other cues such as simulating accommodation via blur are being investigated but are still in early stages of development [Mauderer et al. 2014].

In this research, we focus on the most common display scenario—the widely installed, monoscopic tabletop that relies on non-stereoscopic display technologies. The depth cues that can be applied in this setting are those that arise from linear perspective (occlusion, relative size, relative density, and height in visual field), with the addition of motion perspective if the objects are in motion or if the viewer can move. In this last case, providing correct viewer-centric motion perspective (motion parallax) requires tracking the 3D location of the viewer’s eyes in real time.

3.2. The Perception of Pictorial Space

Scientists and artists have long investigated how pictures represented on a 2D plane are able to create the impression of depth—what is called by psychologists and artists the *pictorial space* [Jones and Hagen 1978]. Some researchers indicate that this is possible because pictures can accurately reproduce on the retina some of the necessary cues of 3D perception (e.g., the foreshortening of objects due to linear perspective and the visual change in receding textures [Cutting 1987, 1988; Sedgwick 1993; Ellis et al. 1991; Saunders and Backus 2006]).

However, as we stated in the previous section, flat projections of 3D space only create retinal images equivalent to the 3D scene if the PoV of the observer is located at, or close to, the CoP used to generate the image [Sedgwick 1993]. When PoV and CoP are at very different angles to the picture plane, or are at very different distances (e.g., when

we look at photographs on a table, a painting from a lateral point of view, or a movie from the side aisle), the differences between correct and distorted retinal projections can be very large. If pictorial perception is dependent only on the geometry of the projected retinal image, this should result in the perception of a space that is deformed compared to the depicted space [Goldstein 1987; Sedgwick 1993; de la Gournerie 1859].

Regardless of the distortion, observers are remarkably good at perceiving a relatively accurate pictorial space [Vishwanath et al. 2005]. However, there is still not a definitive explanation from the perception research community of the underlying processes that support correct space perception from geometrically incorrect retinal images, which is often referred to as space constancy. Some suggest that the visual system corrects distortions based on geometrical information from the represented scene, for example, the visual system might be assuming that certain angles are straight [Perkins 1973], that objects are rigid [Cutting 1987], or that certain converging lines on the picture are actually parallel in the real scene [Saunders and Backus 2006]. Others propose that information about the correct CoP can be recovered from perceptual information about the surface where the picture is projected, such as accommodation [Cutting 1997; Pirenne 1970] and the shape of the frame of the picture [Koenderink et al. 2004].

Although the perception of pictorial space is relatively stable regardless of the discrepancy between the locations of the CoP and the PoV, the relationship between the pictorial space and the physical space is not equally stable. In particular, the perceived orientations toward the physical space of elements within the picture plane can vary depending on the position of the observer [Goldstein 1991; Cutting 1988; Sedgwick 1993]. This effect is best exemplified by the famous U.S. recruiting poster of Uncle Sam, in which he points directly at the observer regardless of how far she is or how oblique she stands to the plane of the poster. For elements within the picture that point perpendicular to the picture plane (e.g., Uncle Sam's finger), the perceived orientation always follows the observer regardless of its position, and therefore, the perceived orientation of the object with respect to the plane of the image can vary almost 180°. For objects that do not point perpendicularly to the picture plane, the possible variation in the perceived angle is reduced; at the extreme (objects that are aligned with the picture plane) geometrical accounts of orientation perception predict that the pointing direction will not vary with changes in the PoV [de la Gournerie 1859; Cutting 1988]. This effect is referred to in the literature as the differential rotation effect (DRE) [Goldstein 1987] or the *de la Gournerie* effect [Cutting 1988]. The DRE is also subject to scientific controversy; experiments have shown that the geometrical predictions do not necessarily fit all data, especially for a very oblique PoV [Ellis et al. 1991]. The possible causes might be found among the cues that cause the pictorial space constancy discussed above (e.g., frame and perceptions of the picture surface through binocular cues [Vishwanath et al. 2005]).

Besides angle, the PoV can be different from the CoP in terms of distance; that is, if the projection is not observed at the right distance, it will not be geometrically correct either, although the distortion produced is more subtle than with angular discrepancy. Taken to the limit, a parallel projection is made from an infinitely far PoV with an appropriately powerful magnification. This is the reason why images from optical systems that are remotely located and have powerful zooms (e.g., satellite imagery) are almost equivalent to parallel projections. The perceptual differences in the perception of 3D objects in the continuum between parallel and close perspective projection were investigated by Hagen and Elliott [1976]. They found that parallel projections are perceived as more natural than their perspective counterparts. This finding contradicts the hypothesis that geometrically correct perceptions will generate more natural images, although it is not yet known if this holds when the graphics are interactive and motion cues are present. Steinicke et al. [2011] have also investigated perceptual and

manipulation effects of overlapping vs. distant (but mostly aligned) PoV and CoP for monitor-based situations, although they discuss these discrepancies in terms of differences between the *geometric field of view*, and the *display field of view*. Their results contradict Hagen and Elliott's results in that they found that overlapping PoV and CoP (or coincident GFOV and DFOV in their terms) are perceived as more realistic.

3.3. Depth From Motion and Action

Any single static view (e.g., a photograph) cannot provide a complete representation of the 3D structure of an object; parts of the object are bound to be hidden, ambiguously interpretable, and only partially visible from any single point of view. Recognizing an object that has been captured from a different point of view requires processing time, and is prone to errors [Wexler et al. 2001; Diwadkar and McNamara 1997; Biederman and Gerhardstein 1993]. Many experiments show that image sequences of the object that reveal its different surfaces are integrated by the visual system to contribute to the perception of 3D structures [Rogers and Graham 1979; Simons et al. 2002; Sun et al. 2004]. Moreover, some evidence suggests that motion cues are more dominant for depth perception than binocular cues [Arthur et al. 1993], although this evidence has also been contested for different tasks. See [Arsenault and Ware 2004] for a review and evidence.

The changes in the appearance of a 3D object that generate a sequence of images can result from changes in the position and orientation of the object itself, the physical point of view of the observer, the virtual point of view in the virtual space (i.e., virtual space navigation), and the center of projection of the scene. Additionally, in computer-based systems, these changes can be caused by interactive control by the viewer, automatic system change (e.g., in an animation or through a piece of experimental software), or both. To make things even more difficult, different display systems couple different combinations of the points of view and the center of projection to each other. For example, in fish-tank VR [Ware et al. 1993], the movement of a view (the physical PoV) is coupled to both the Virtual PoV of the Scene and the CoP.

Many display systems and experiments have explored the many combinations in the source of visual change and coupling between CoP and PoVs. Our investigation focuses on display setups that make sense in collaborative large display systems where viewers will be moving around the space. As detailed below in the descriptions of the experiments, we include conditions where self-motion of the viewer does not change the projection of the objects (a baseline, no-motion condition); where the object can be looked at from different points of view controlled by the head motion of the viewer (equivalent to fish-tank VR); where besides head-coupled motion, the viewer can manually manipulate the position and orientation of the object; and where there is depth motion, which is generated externally without the viewer's intervention.⁵

3.4. Collaborative and Volumetric Display Systems

An obvious way to address the problem of discrepancy for collaborative displays is the use of hardware systems that can either provide a true volumetric display where light is produced or reflected at the location in the space where the object is meant to be [Grossman and Wigdor 2007], or can simulate the different light patterns coming into the retinas of individual viewers. These light patterns can be simulated through a variety of technologies, including head-mounted projections (e.g., polarized glasses that filter the adequate image for each person/eye, or screens that provide a separate image to each eyeball [Kiyokawa et al. 2003; Höllerer et al. 1999]), clever filtering

⁵Appendix A, included in the ancillary materials, provides a summary of existing research in this area through a classification of existing systems and experiments. The classification enables systematic comparison between existing research and our results, which are also elaborated on in the Discussion section.

techniques on the source content (e.g., spatial occlusion [Kitamura et al. 2001], the use of polarization [Sakurai et al. 2008; Karnik et al. 2012], directionally heterogeneous diffusion [Matsushita et al. 2004]), or a combination of several of these techniques [Kulik et al. 2011]. A recent survey of stereoscopic and 3D display techniques is provided by Geng [2013].

Although the state of the art is rapidly advancing, the use of these techniques is still practically limited, expensive, or both. For example, volumetric displays based on rotating devices are expensive, noisy, and tend to have low resolution. Less mechanically based approaches, such as providing separate images for each person, are expensive and introduce technical challenges, like image cross-talk, especially if the system is meant to be binocular and displays can be oblique (which makes the use of polarization problematic). Transparent personal head mounted displays are becoming cheaper, but are still far from ubiquitous and many people will prefer not to have to wear this hardware. Autostereoscopic lenticular displays are also dropping in price, but tend to only provide multiple images in one axis, as they are meant to be mounted vertically for entertainment. Due to existing limitations in multi-user stereoscopic and volumetric display technology, and since the existing base of displays is still overwhelmingly flat and single image, the study of the effects of perception of 3D images in the latter kind of displays is of great relevance.

3.5. Perception and Projection in HCI Research

The topics of projection, perspective distortions, and perception have also received considerable attention in HCI and display research. The VR research has identified *single-screen multi-user 3D rendering* as an important challenge for shared-display VR systems [de Haan et al. 2007; Simon and Scholz 2005].

Although the distortion in the perception of 3D elements produced by PoV-CoP discrepancy is often acknowledged (e.g., de Haan et al. [2007], Simon and Scholz [2005], and Arsenault and Ware [2004]), proposed solutions have, so far, only addressed the issue of object manipulation. Previous solutions are based on providing correct perspectives only for the visual elements of the interaction techniques while keeping the CoP in a fixed position [Simon and Scholz 2005], or as an interpolation of the PoV of the different participants for the 3D scene [de Haan et al. 2007]. Hancock et al. [2007] have also studied effective 3D object manipulation through direct multi-touch input (which due to its direct nature eliminates the problem of discrepancy in the interaction technique visuals). More recently, several groups of researchers have started to explore techniques where perspective manipulations and object movement are used to facilitate selection and manipulation of 3D objects [Möllers et al. 2012; Valkov et al. 2012].

Our studies differ fundamentally from this work in that we focus on the perception of 3D objects, not on manipulation performance. Our approach is justified because perception itself is important for the use of 3D systems, but also because perception is a fundamental part of interaction that is bound to affect how manipulation takes place (e.g., through perspective distortion [Arsenault and Ware 2005]).

Another related strand of research studies the effects of projecting 2D information in oblique angles and rotations. Wigdor et al. [2007] studied how the slant of the surface affects perception of length, angle, and area for 2D data, and Nacenta et al., showed how perspective distortion affects 2D perception and interaction tasks, and how it can be improved through perspective-compensation techniques [Nacenta et al. 2007].

4. STUDY OF 3D PERCEPTION IN TABLETOP DISPLAYS

As discussed in our review of related work, different ways to project 3D information can have a strong impact on how 3D objects are represented, yet little is known about

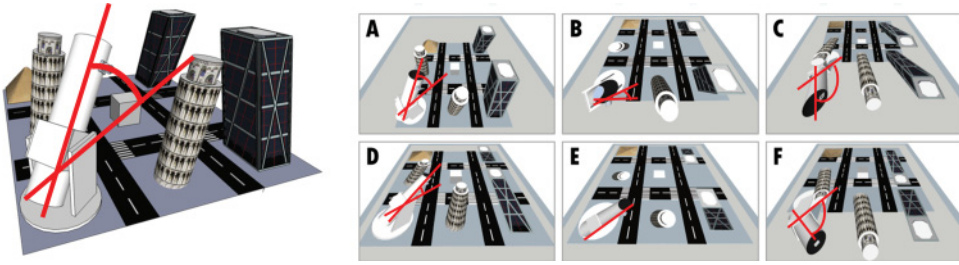


Fig. 8. Any angle in a 3D model is susceptible to change when represented with different projections. Incorrectly estimating the model's angles from the model's representation indicates a failure to interpret the representation of the model accurately. (A)–(F) correspond to the same projections in (A)–(F) of Figure 6.

how these factors affect collaborative tabletop scenarios or how the factors interact with each other. This article contributes a study involving four experiments designed to inform design choices for 3D tabletop applications. In the following subsections, we describe the high-level design of the four experiments and their common elements, including independent variables, measures, hypotheses, and apparatus. The details of each experiment and its results are discussed separately in Sections 5–8.

4.1. Tasks

We selected three perceptual tasks that are at the core of more complex perceptual activity on tabletops. The first and second tasks were tested in Experiments 1 and 2; the last task was tested in Experiments 3 and 4.

Orientation Estimation—estimating the orientation of 3D objects from the pictorial space into the physical environment; i.e., estimation of the *perceived external angle*. This task is fundamental in recognizing the spatial relationships between objects represented in the virtual space and in the real space. This task is particularly relevant for scenarios where the real world and the virtual world interact with each other. Examples where this task will be common are systems using tangibles, and ubiquitous computing scenarios where information on displays relates to external objects, devices, or people.

Internal Structure Estimation—estimating the internal spatial relationships of virtual elements within the pictorial space. Specifically, the relative positions and angles of the components of objects within the pictorial space (see Figure 8). This task is crucial for the correct perception of 3D objects and 3D scenes represented in the pictorial space; if the positions and internal angles are not correctly estimated, this indicates that the viewer does not perceive the same 3D structure that is intended by the representation.

Object Structure Recognition—being able to identify two objects that have the same 3D shape as equal to each other in structure. This is a higher level task because (a) it likely involves several ways to perceive the internal 3D structure of the object (e.g., angles and distances), and (b) it involves the comparison of two perceived objects. This task was selected to represent the generic ability to recognize objects through their 3D structure. Although the task of recognizing a 3D object from memory is somewhat different than comparing two objects that are visible, we believe this is a reasonable first approximation, since there are no evident reasons to believe that projection will affect the two tasks differently.



Fig. 9. The experimental space photographed from behind the shoulder of one of the participants. The experimenter sat perpendicular to the long side of the table.

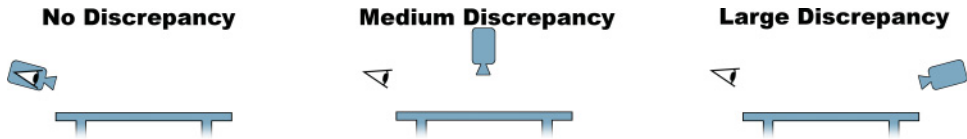


Fig. 10. The three levels of discrepancy studied across all experiments. The camera icon represents the position of the CoP, the eye icon represents the position of the PoV.

4.2. Common Apparatus

For all four experiments, participants stood or sat at an end of a 146cm \times 110cm bottom-projected table with a resolution of 2800 \times 2100 (19 pixels/cm; 2 \times 2 tiled projectors with 1400 \times 1050 pixels each). The PoV of each participant was tracked using a ViconTM motion tracking system with markers placed on hats that the participants wore throughout the experiment (Figure 9). Participants took part in the experiment in pairs, at opposite ends of the table. Although measurements were individual (per participant) and each participant's task was independent, the presence of another person was useful to simulate conditions of large-discrepancy and parallax imposed by others.

4.3. Main Independent Variables (Factors)

The four experiments share the manipulation of the main factors. The primary factor is *discrepancy*, defined as the difference between the CoP5F and the PoV. We explore three levels of discrepancy (Figure 10): none (the CoP coincides with a person's PoV); medium (CoP is directly above the table); and large (CoP is at someone else's PoV). Note that the medium-discrepancy condition is not just an interpolation of the two extremes; projections that are centered on and perpendicular to the display surface have been used often in 3D tabletop applications (e.g., Bichlmeier et al. [2009] and Hancock et al. [2010]).

Figure 11 shows the fixed parameters of the projection used throughout the experiment for medium discrepancy: the CoP was placed 41 cm above the center of the table; since the table was 110cm wide, this implied a view angle of 106°.

A secondary factor is motion parallax—that is, whether the CoP dynamically followed the participant's PoV, resulting in perspective changes as the participant moved their head. This corresponds to self-generated motion parallax as described in Section 3.3. The motion parallax condition was tested both when the CoP moved with the

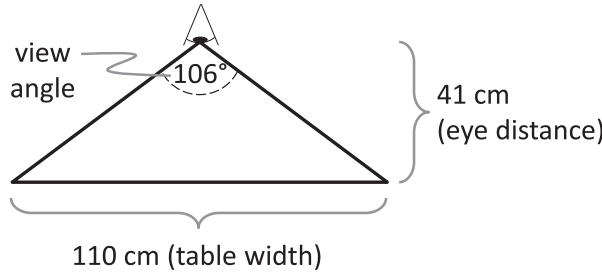


Fig. 11. Fixed projection parameters for the medium discrepancy condition. The remaining conditions used the same near plane, but the CoP was matched to either the participant (in the no-discrepancy condition) or their partner (in the large-discrepancy condition).

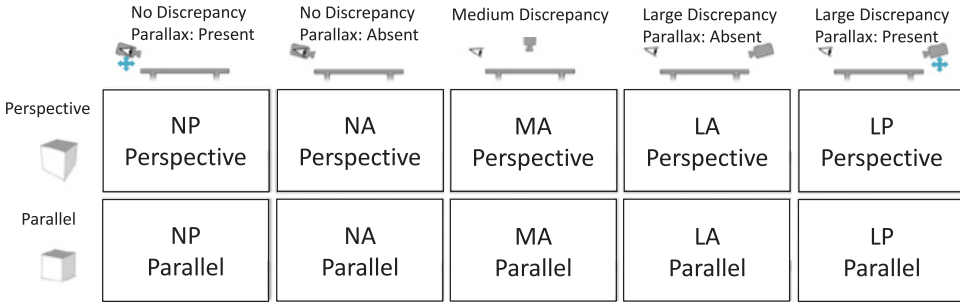


Fig. 12. The experiment has two main factors: parallax-discrepancy (horizontal), where discrepancy increases from left to right and parallax conditions are at the extremes, and projection type (vertical).

participant’s PoV and when the CoP moved with a different person’s PoV. Note that with a medium discrepancy parallax is not used, since the CoP is fixed above the table.

The first two factors combine into five discrepancy-parallax conditions: NoDiscrepancy/Parallax, NoDiscrepancy/NoParallax, MediumDiscrepancy/NoParallax, LargeDiscrepancy/NoParallax, and LargeDiscrepancy/Parallax. For a full depiction of all conditions and levels, see Figure 12.

An additional factor in our study was *projection geometry*, which had two levels: perspective and parallel. The combination of all factors, with discrepancy and parallax rolled into one (discrepancy-parallax) and fully crossed with projection geometry, results in a design with $5 \times 2 = 10$ cells, i.e., ten different combinations of levels (see Figure 12).

We decided to combine the parallax and discrepancy dimensions into one (as opposed to considering a $3 \times 2 \times 2$ model with discrepancy, parallax, and projection type as factors) because parallax can be considered as a continuous extension of discrepancy (i.e., parallax alters discrepancy), and because there is no parallax condition for medium discrepancy.

In Steinicke et al.’s [2011] terminology, the variation of our factors keep the geometric and display fields of view as similar as possible in the NoDiscrepancy conditions (with the parallax conditions adjusting this in real time), whereas the MediumDiscrepancy condition has different frustums (with the display frustum being an *asymmetric display frustum*) and the LargeDiscrepancy conditions have two asymmetric (but opposed) display frustums.

To create LargeDiscrepancy conditions, we used the PoV of the other participant. This means that for LargeDiscrepancy conditions, content was rendered for the other participant. Although the LargeDiscrepancy conditions, especially with parallax, might

seem somewhat artificial and designers will predictably try to avoid them, we decided to include them for three main reasons. First, they represent situations and system designs that are not far-fetched, such as a non-tracked observer entering the space or a more sophisticated perspective design where individual objects are projected differently depending on who is currently manipulating them (but perceivable by everyone). Second, they offer completeness from the experimental point of view and valuable empirical results for the 3D perception community. Third, it enables testing participants in pairs, which provides a slightly more ecologically valid simulation of multi-user environments.

4.4. Dependent Variables (Measures)

The main dependent variables recorded across all experiments were accuracy, defined specifically for each task, and task completion time (TCT). We also measured subjective assessment of each condition with two seven-point Likert-scale questions: one about the difficulty of the task, and one about overall preference.

4.5. Common Hypotheses

All four experiments tested four primary hypotheses: the first concerns the general effect of discrepancy, the second (subdivided in two) tests the combined effect of projection type with discrepancy, the third (subdivided in two) tests the effect of parallax, and the fourth examines the nature of the medium discrepancy conditions as a special case.

H1: As the discrepancy increases, the ability to make spatial estimates degrades.

H2: Perspective projection makes the effect of discrepancy more extreme.

H2.A: With no discrepancy, perspective projection improves spatial estimation.

H2.B: With discrepancy, perspective projection degrades spatial estimation.

H3: Parallax makes the effect of discrepancy more extreme.

H3.A: With no discrepancy, motion parallax improves spatial estimation.

H3.B: With discrepancy, motion parallax degrades spatial estimation.

H4: Medium discrepancy (CoP above) is a special case that improves spatial estimation.

The exact meaning of *improving or degrading spatial estimation* depends on the measure and the task: better spatial estimation implies better accuracy, less time, less perceived effort, and greater preference. Different measures can conflict (e.g., a condition might be faster but less accurate); therefore, we analyze the contributions to the hypotheses from all measures. All four experiments were also designed to test the secondary hypothesis.

H5: The use of motion parallax requires more time to complete the task.

4.6. Common Analysis

For each of the objective measures (errors and time), we first performed omnibus RM-ANOVA with factors parallax-discrepancy and projection (5 discrepancy-parallax \times 2 projection). Different experiments had additional main factors depending on their specific task. We tested for sphericity and applied the Huynh-Feldt correction to the degrees of freedom where necessary.

In addition to the omnibus tests, we performed a series of planned comparisons of selected conditions. These were specified before the start of the analysis and, if we found an interaction between the factors, we performed these tests separately on each type of projection geometry (perspective or parallel). If the interaction was not significant, we aggregated perspective and parallel conditions before the post-hoc analyses. Post-hocs used Holm-Bonferroni adjustments. Figure 13 summarizes the a priori selected post-hoc tests.

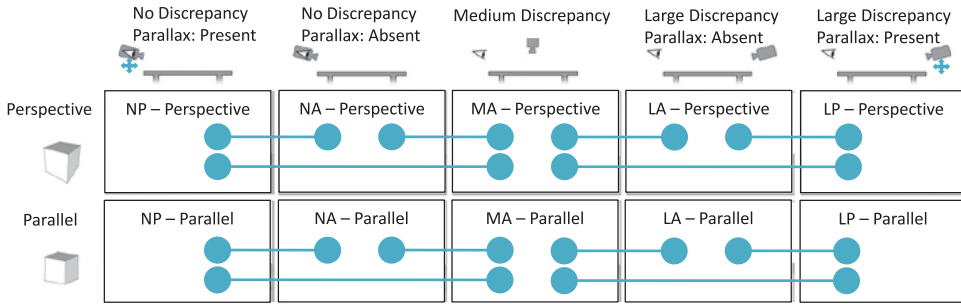


Fig. 13. Graphical illustration of the *a priori* selected post-hoc statistical tests.

Subjective measures were analyzed by first transforming the numerical data using the ART procedure [Wobbrock et al. 2011], which allowed us to test for main effects as well as interactions. If the omnibus test yielded a main effect of discrepancy-parallax, we performed the same comparisons between conditions as described above for the objective measures. As above, if an interaction between the two factors was found significant, the selected pairwise post-hoc tests were performed separately for the parallel and the perspective projection geometries. The post-hocs were Wilcoxon non-parametric tests performed on the non-transformed data. The analysis was designed to avoid the assumption of normality.

5. EXPERIMENT 1: ORIENTATION ESTIMATION

This experiment was designed to test people's perception of the orientation of 3D virtual objects with respect to their physical surroundings. That is, when an object is rendered so as to point "out" of the display, how well can people perceive its direction?

5.1. Participants

Twenty-four participants (11 female, 13 male) completed this experiment (ages 19 to 36, $Mdn = 28$, $SD = 4.5$). Participants performed a six-item Purdue spatial rotation test [Bodner and Guay 1997] previous to the experiment (correct answers, out of 6: $M = 4.56$, $Mdn = 5$). Participants were recruited in pairs and performed the experiment in pairs (one on each side of the table). The pairs were three all-female, four all-male, and five mixed gender. This study was conducted together with Experiment 4, and so upon completing this experiment, the same pair proceeded to Experiment 4.

5.2. Apparatus

In addition to the common apparatus, Vicon-tracked markers were placed at the end of a wand which was attached to a corner of the tabletop through a string (see Figures 14 and 15). This allowed us to accurately measure the 3D angle of the string, that is, the angle determined by the line from the tip of the wand to the corner of the table. Participants manipulated their own string with their own wand in order to record answers about the angle of the target.

5.3. Task

Participants were asked to determine the orientation of virtual objects on the table. For each trial, two 3D target objects were displayed on the screen (one per participant, always displayed on the half of the display at the participant's right). Targets were either in the "near" half or the "far" half of the table. Each object was a long thin cylinder inside a shorter thicker cylinder with the same axis. To provide their answer for each task, participants moved the tip of the wand, until the string joining it with

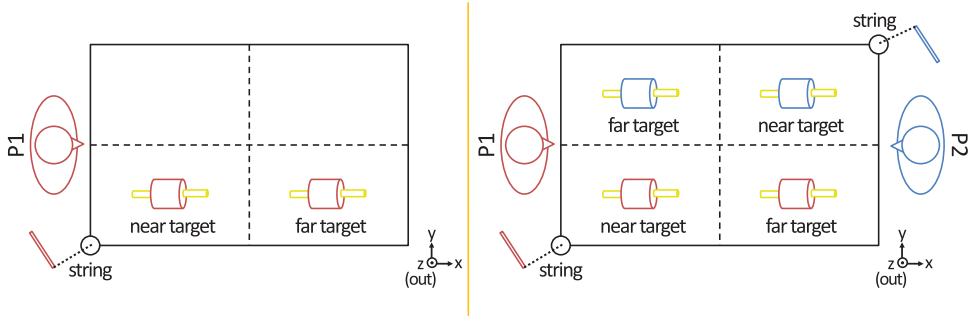


Fig. 14. Diagram of the experimental setup for Experiment 1. The left diagram displays elements that concern only the task of participant 1. The right diagram shows how the two participants were actually tested in pairs, with targets on the bottom corresponding to the participant on the left, and targets on the top corresponding to the participant on the right.



Fig. 15. The participant specified perceived angle by manipulating the angle of a string attached to the right corner of the table and the tip of a wand (string highlighted in yellow). In this simulation, the participant is judging the red and yellow object shown at the quadrant of the surface closest to him.

the corner of the table was oriented at the same angle as the main axis of the target (Figure 15).

Once both participants had indicated the angle and pressed a “done” button on the table surface with their free hand, the next trial would begin. Although the presentation of conditions and the orientation of the objects were randomized, the orders that the two simultaneous participants saw were complementary to each other’s (e.g., a no-discrepancy trial for one participant would coincide in time with a large discrepancy trial for the other).

To avoid interference or copying, in each trial each participant had to judge the orientation of a separate object. The two spaces not occupied by target objects were filled in with two reference objects represented using the same geometry as all other virtual objects on the table. Reference objects are four cube frames in a square configuration (see Figure 15). These were meant to provide additional visual context to the target objects.

5.4. Procedure and Experimental Design

In each of the ten *discrepancy-parallax* \times *projection* combinations, participants performed all combinations of the remaining factors in random order. Participants carried out six practice trials (one for each combination of location and angle), and twelve

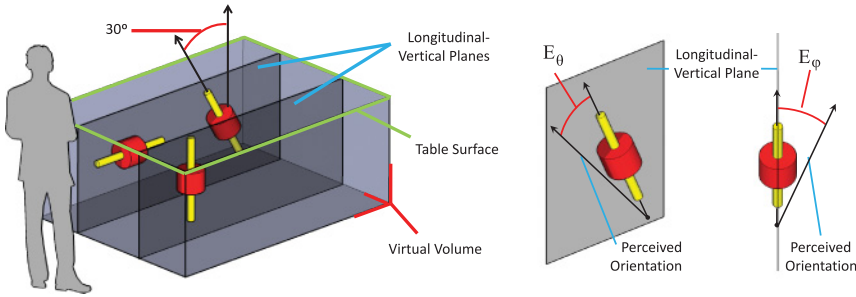


Fig. 16. (Left) The virtual volume in which the target objects were placed. (Right) The two measures of error taken in the experiment: longitudinal (deviation within the vertical longitudinal plane), and transversal (deviation with respect to the projection of the estimation onto the longitudinal vertical plane).

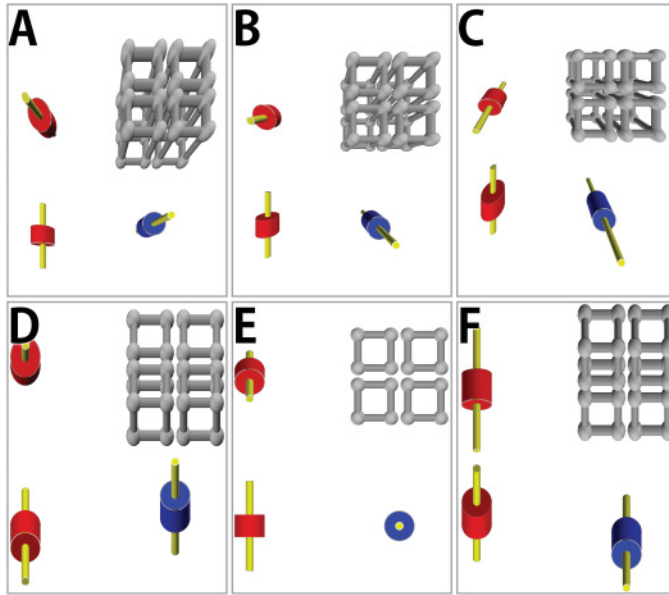


Fig. 17. Table renderings of target objects in each of the orientations. Notice that in a real trial, only two objects (one blue and one red) would appear simultaneously, and another set of gray cubes would take the place of the third object. Figures (A)–(F) follow the same arrangement of condition as in Figure 6.

testing trials (two for each location/angle combination) for each *discrepancy-parallax* \times *projection* combination, for a total of 180 trials.

As mentioned before, targets could appear either in the “near” or “far” halves of the table, as shown in Figure 14; this additional factor was controlled and randomized. Targets were also shown in three different angular orientations, and in two locations. As shown in Figures 16 and 17, targets could be at either 0° (lying flat on the table and pointing towards the end of the table where the participant was located), 60° (pointing upwards towards the end of the table), or 90° (pointing straight up from the table). Targets never leaned to the left or right. We chose to manipulate only the longitudinal direction of targets because the PoV and CoPs are mostly oblique in that direction, and therefore, this is the manipulation most likely to cause relevant perceptual differences.

To summarize, the study tested 5 discrepancy-parallax conditions \times 2 projection geometries \times 2 stimuli locations \times 3 angles in a fully crossed within-participants design.

The ten discrepancy-parallax and projection geometry pairs' order of presentation were counterbalanced between participants using a random Latin Square. The different target orientations appeared in randomized order to prevent participants from deriving angle from order.

5.5. Measures

We measured error in angle and TCT. The error in angle had two parts (Figure 16, right).

E_θ : the error within the longitudinal vertical plane (e.g., if the target pointed toward the end of the table with an angle of 60° , and the participant held the wand such that the string had an angle of 65° in the longitudinal vertical plane, the error was 5°).

E_φ : the error between the estimation and the projection of the estimation on the longitudinal plane (transversal error, or left-to-right error). Notice that the targets did not lean in this dimension, but some projections can make targets appear to lean.

The Euclidian combination of both angles sums up to the full error of the angle.

5.6. Experiment 1 Specific Hypotheses

This experiment was designed to isolate errors along the longitudinal vertical plane due to the change in discrepancy. In addition to H1–H5 (see Section 4.5), we derived two secondary hypotheses.

H6: The type of projection geometry will affect E_φ (i.e., errors outside the plane in which the angle is varied).

H7: E_θ will be least when the object's angle is horizontal (0°) and most when the object is vertical (90°). This hypothesis tests the DRE [Goldstein 1987] (Section 3.2).

5.7. Results

The omnibus RM-ANOVAs for this experiment had a 5 discrepancy-parallax condition \times 2 projection geometry \times 2 location \times 3 angle design. The results are discussed grouped by measure and then by hypothesis. For brevity's sake, some hypotheses that were not supported by the tests are not discussed explicitly for all measures.⁶

5.7.1. Error Within the Longitudinal Plane (E_θ). The omnibus ANOVA of the main error measure shows strong main effects of discrepancy-parallax ($F_{1.662,38.222} = 127.395$, $p < 0.001$, $\eta_p^2 = 0.847$) and projection type ($F_{1,23} = 12.621$, $p = 0.002$, $\eta_p^2 = 0.254$), as well as main effects of location and angle, as expected (location: $F_{1,23} = 40.569$, $p < 0.001$, $\eta_p^2 = 0.638$; angle: $F_{1.295,29.791} = 15.764$, $p < 0.001$, $\eta_p^2 = 0.407$). Many other interactions between factors are significant, but we will only discuss the most relevant.⁷

Projection type and discrepancy-parallax interact with each other ($F_{3.331,76.721} = 22.825$, $p < 0.001$, $\eta_p^2 = 0.498$), which we then analyzed through planned post-hoc tests displayed in Figure 18 and Table I.

H1 is generally supported by the longitudinal error data; large discrepancy conditions (both with and without parallax) resulted in much larger errors than no-discrepancy conditions. These are supported by most of the relevant post-hoc tests. With perspective projections, the medium discrepancy condition resulted in errors roughly between the no-discrepancy and the large discrepancy conditions, which suggests a linear relationship between discrepancy and error. However, as we will discuss below, with the parallel projection, the medium discrepancy is equivalent to the no-discrepancy conditions.

⁶A table summarizing the details of the statistical tests for all measures can be found in Appendix B, Experiment 1, which is included with the ancillary materials of this article.

⁷The rest of the test results are also available through the Appendix.

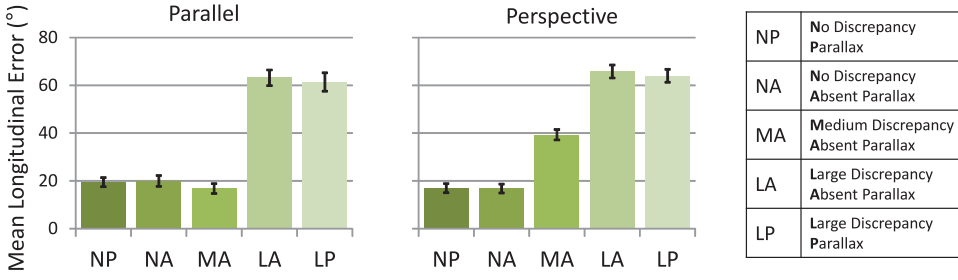


Fig. 18. Average errors within the longitudinal plane for the orientation estimation task. Error bars indicate standard error.

Table I. Planned Comparisons for Experiment 1 (Orientation Estimation Task)

	E_{θ}		E_{ϕ}		TCT	
	Parallel	Perspective	Parallel	Perspective	Parallel	Perspective
NA to MA	$p = 0.131$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.005$	$p = 0.814$
MA to LA	$p < 0.001$	$p < 0.001$	$p = 0.015$	$p = 0.001$	$p = 0.008$	$p = 0.830$
NP to NA	$p = 0.706$	$p = 0.803$	$p = 0.112$	$p = 0.084$	$p = 0.006$	$p = 0.022$
LA to LP	$p = .525$	$p = 0.387$	$p = 0.005$	$p = 0.231$	$p = 0.011$	$p = 0.008$
MA to NP	$p = 0.247$	$p < 0.001$	$p = 0.005$	$p < 0.001$	$p < 0.001$	$p = 0.059$
MA to LP	$p < 0.001$	$p < 0.001$	$p = 0.430$	$p = 0.001$	$p < 0.001$	$p = 0.003$

p -values highlighted with bold type indicate significant tests (α values are adjusted using the Holm-Bonferroni procedure).

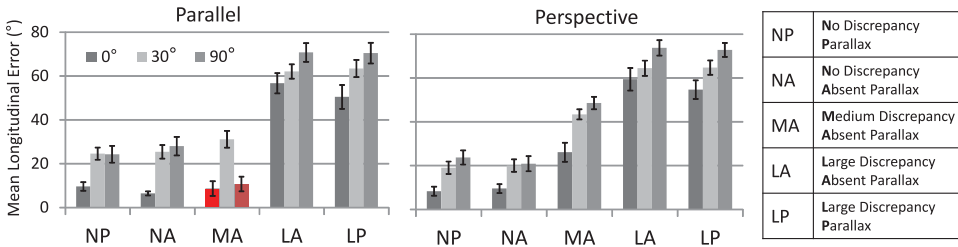


Fig. 19. Average errors within the longitudinal plane (per angle) for the orientation estimation task. Error bars indicate standard error.

H2.A and H2.B are not supported by the data, since the error levels of no-discrepancy conditions and large discrepancy conditions compared across projections are almost equivalent. H2 is, therefore, not supported. H3.A and H3.B are also not supported, since none of the parallax vs. no-parallax comparisons showed significant differences.

H4 (whether the medium-discrepancy case is special) is supported but only for the parallel projection geometry: medium discrepancy in the parallel condition has a lower average error than the no-discrepancy conditions and is statistically indistinguishable from them, and statistically different from the large discrepancy cases. Figure 19 shows the source of the main difference between the projection types: 0° and 90° objects were judged very accurately in the parallel projections, but not in the perspective projection.

H7 (a verification of the DRE) is generally confirmed by the results, and clearly visible in Figure 19, where the 0 angles (parallel to the surface of the table) showed almost always smaller errors than the 30° and 90° conditions.

A further analysis of the effect of location also shows a remarkable difference between the parallel and perspective conditions (Figure 20). With a perspective projection, the medium discrepancy condition is similar in its low errors to the no-discrepancy

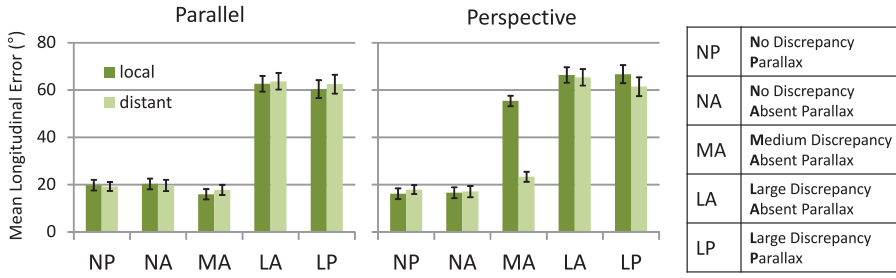


Fig. 20. Average errors within the longitudinal plane (per location) for the orientation estimation task. Error bars indicate standard error.

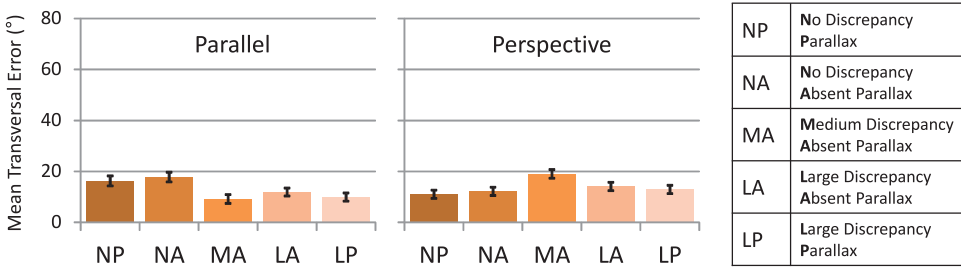


Fig. 21. Average transversal errors for orientation estimation. Error bars show standard error.

conditions when the objects are placed further away from the observer, whereas the average errors are equivalent to the large discrepancy conditions when the objects are further away. The parallel projection, however, appears immune to this effect.

5.7.2. Transversal Error (E_ψ). Analysis of the transversal error measures showed main effects of discrepancy parallax ($F_{2.804,64.502} = 3.604, p = 0.02, \eta_p^2 = 0.135$), location ($F_{1,23} = 10.519, p = 0.004, \eta_p^2 = 0.314$), angle ($F_{1.415,32.538} = 36.332, p < 0.001, \eta_p^2 = 0.612$), and an interaction between projection type and discrepancy parallax ($F_{1.886,43.275} = 19.64, p < 0.001, \eta_p^2 = 0.461$). The errors in the transversal error are (as expected) much lower than in the longitudinal dimension because the discrepancy between PoV and CoP is aligned in the longitudinal axis. However, as shown in Figure 21, different projections and levels of discrepancy produce distinct effects.

H7 is supported by the data because, although the overall comparison between parallel and perspective projections was not significant, the interaction was, and the separate analysis of the transversal error for the parallel and perspective conditions (Figure 21 and the post-hocs) shows a much different pattern. Although the size of errors is much smaller in the transversal direction (the scale of the bar chart is kept the same to avoid misleading impressions), the parallel non-discrepancy conditions show significantly larger errors, whereas for perspective projection it is only the medium discrepancy condition that is singled out with significantly more errors than all the other conditions.

5.7.3. TCT. We performed a similar analysis on TCTs. The omnibus ANOVA showed a strong main effect of discrepancy-parallax ($F_{4,92} = 12.414, p < 0.001, \eta_p^2 = 0.351$) and projection type ($F_{1,23} = 12.621, p = 0.002, \eta_p^2 = 0.354$), but no significant main effects of location and angle (location: $F_{1,23} = .249, p = 0.622, \eta_p^2 = 0.011$; angle: $F_{2,46} = 1.464, p = .242, \eta_p^2 = 0.06$). The interaction between projection type and discrepancy parallax was significant ($F_{3.378,77.683} = 2.977, p = 0.031, \eta_p^2 = 0.115$), and therefore, we performed post-hoc tests separately on parallel and perspective conditions (Figure 22).

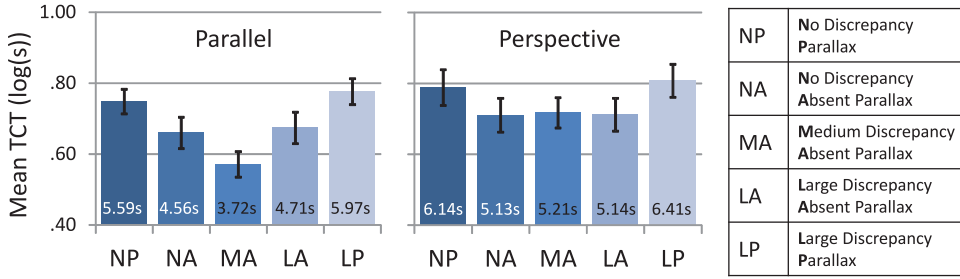


Fig. 22. Average TCTs (log-transformed) for the orientation estimation task. Numbers at the bottom of the column represent the untransformed value in seconds for the average of the log-transformed completion times. Error bars represent standard error.

H1 is partially contradicted by the completion time data, since no-discrepancy conditions resulted in shorter completion times than the medium discrepancy condition for the parallel projection and resulted in similar completion times to the large discrepancy conditions in both parallel and perspective projections.

H2.A is not supported since comparisons of no-discrepancy conditions with perspective and parallel projections do not show an advantage for perspective—in fact, the averages seem to contradict this. H2.B is also not supported since large discrepancy conditions are symmetric to no-discrepancy conditions for both types of projection.

H3.A is partially contradicted, because the parallax condition takes longer to judge, which is exactly what H5 predicts. The results are consistent with H3.B, but it seems more plausible that the source of the increased completion time is parallax itself rather than a difference in difficulty of perceiving large discrepancy conditions with parallax. The statistical support for the H3 hypotheses is only significant for one of the four tests (although the remaining three tests all approach significance). The results do not strongly support H3.

H4 is supported by the time data for parallel projection, since the medium discrepancy parallel condition is, as shown clearly by the post-hoc tests, the fastest to judge (more than one second faster than the next fastest), and the medium discrepancy perspective condition is no slower than any of the other conditions with perspective projections.

5.7.4. Subjective Assessment. For the subjective measures, the omnibus ANOVAs of the rank-transformed data showed that the discrepancy-parallax factor had a main effect on the participants' ratings of difficulty ($F_{4,92} = 4.459, p = 0.002, \eta_p^2 = 0.162$) and preference ($F_{4,92} = 4.364, p = 0.003, \eta_p^2 = 0.159$). However, no interaction was found between discrepancy-parallax and projection geometry for either difficulty ($F_{3,4,78.3} = 1.907, p = 0.13, \eta_p^2 = 0.077$) or preference ($F_{3,4,78.3} = 1.907, p = 0.13, \eta_p^2 = 0.022$). The results of the planned post-hoc analyses are displayed in Table II.⁸

As expected, the statistical analyses of the subjective responses are less powerful than their objective measure counterparts and show fewer significant results. Nevertheless, the post-hoc tests provide some support for H1: larger discrepancy conditions were considered more difficult or were less preferred than the medium-discrepancy condition (Table II). H2 and H3 are generally not supported. The medium-discrepancy condition had highest preference scores, but the difference was not significant.

⁸The participant responses are displayed in tables in Appendix B, Experiment 1.

Table II. Planned Comparisons for Subjective Likert Results of the Orientation Estimation Task

	<i>Difficulty</i>	<i>Preference</i>
NA to MA	$p = 0.979$	$p = 0.366$
MA to LA	$p = 0.128$	$p = 0.002$
NP to NA	$p = 0.759$	$p = 0.930$
LA to LP	$p = 0.055$	$p = 0.511$
MA to NP	$p = 0.324$	$p = 0.323$
MA to LP	$p = 0.005$	$p < 0.001$

Since the interaction between projection geometry and discrepancy was not significant, parallel and perspective responses were averaged and analyzed together.

5.8. Summary and Discussion

The results of this experiment indicate that reducing discrepancy is generally beneficial for the accuracy of orientation estimation of objects with respect to the world. The evidence does not support a significant advantage of parallax for angle estimation; on the contrary, providing parallax results in longer estimation times. Perspective and parallel projections result in similar levels of error, with two notable exceptions: the medium discrepancy parallel projection (CoP fixed above the table) resulted in the lowest errors, and the parallel projections had larger side-to-side (transverse) error in the no-discrepancy conditions. A more detailed analysis showed that part of the general advantage in longitudinal error from the parallel medium discrepancy condition comes from how it is unaffected by the location of the stimuli in the table. TCTs were also lowest for the parallel medium discrepancy condition. Additionally, the time data indicates that the presence of parallax results in longer judgment times, in exchange for no significant advantage in accuracy.

In practical terms, the results of this experiment suggest that: a) the medium discrepancy parallel condition is preferable if both time and accuracy are important; b) dynamic parallax takes extra time, but does not seem to help for angle perception; c) large discrepancy has a clear detrimental effect in accuracy with respect to medium discrepancy and no-discrepancy conditions; and d) the parallel projection seems to have an advantage, especially for medium discrepancy, in terms of time, resilience to the position of the stimuli, and general accuracy, but also results in slightly larger transversal errors in the no-discrepancy condition. Since the medium discrepancy with parallel projection did not seem to be affected by this problem, this orientation task seems best performed by a parallel medium discrepancy projection, which also has the advantage that it does not require sensing the position of viewers around it.

6. EXPERIMENT 2: INTERNAL STRUCTURE ESTIMATION

This experiment was designed to determine how discrepancy and projection type affect perception of object structure—participants were asked to estimate the internal spatial relationships of a simple triangular object. In this study, the estimation does not involve the physical space around the table, but only internal relationships within the table space.

6.1. Participants

Twenty-four participants (13 female) completed this experiment (ages 18 to 33, $Mdn = 26$, $SD = 5.2$) and were recruited in pairs (four all-female, three all-male, and five mixed). Participants performed a six-item Purdue spatial rotation test [Bodner and Guay 1997] previous to the experiment (correct answers, out of 6: $M = 4.44$, $Mdn = 5$).

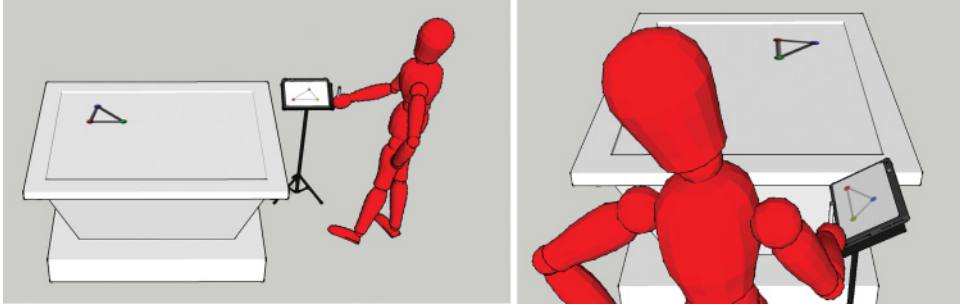


Fig. 23. 3D simulation of the apparatus for experiment 2. The participant provided input about their perception of objects through a pen on a tablet held by a tripod.

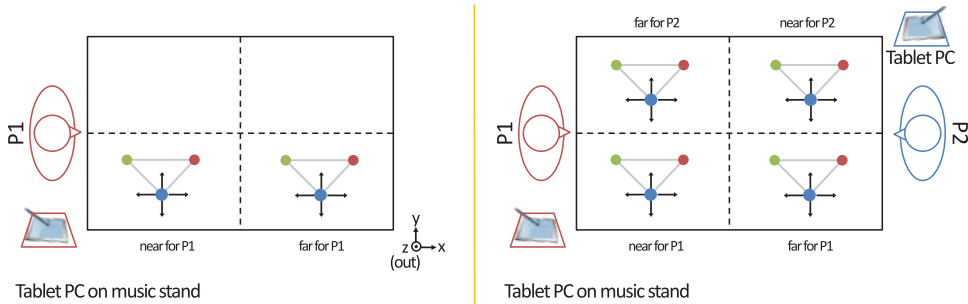


Fig. 24. Diagram of the experimental setup for Experiment 2. The left diagram displays elements that concern only the task of participant 1. The right diagram shows the positions of all objects as the participants were tested in pairs.

This study was conducted together with Experiment 3, and so upon completing this experiment, the same pair proceeded to Experiment 3.

6.2. Apparatus

We used the same table as described for Experiment 1 (see Section 5.2), but this time participant responses were entered using a stylus on a tablet PC (one per participant, 25 cm × 18 cm, resolution of 1024 × 768). Tablets were placed next to a participant on a music stand, tilted at 30° to match the stimuli on the table (see Section 6.3 and Figure 23). Participants were free to alter the height or location of the stand, but did not change the tilt.

6.3. Task

Participants had to match the relationships in a triangular shape representing a simple molecule constructed with three spheres (nodes) and three gray long cylinders (links) connecting the spheres (see Figures 25 and 26). The model to be matched appeared represented in 3D on the table, and it had to be replicated by adjusting a 2D model on the tablet, next to the participant (Figure 23). To reduce complexity, the green and red nodes of the molecule were always in the same position. In different trials, the position of the blue node would vary, which effectively changed all the angles in the triangle. Consequently, participants just had to adjust the position of the blue node on the tablet, to try to reproduce what they perceived to be the shape of the molecule on the table.

The molecule could appear in one of two positions: near and to the left, or far and to the right (see Figure 24). The 3D triangular molecule plane was tilted 30° from

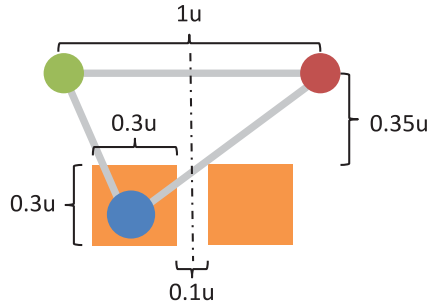


Fig. 25. The position of the blue node was randomly selected within the two squares in the figure. These squares were not visible in the experiment. Distances are indicated relative to the distance between the green and red nodes, which was kept constant across all trials.

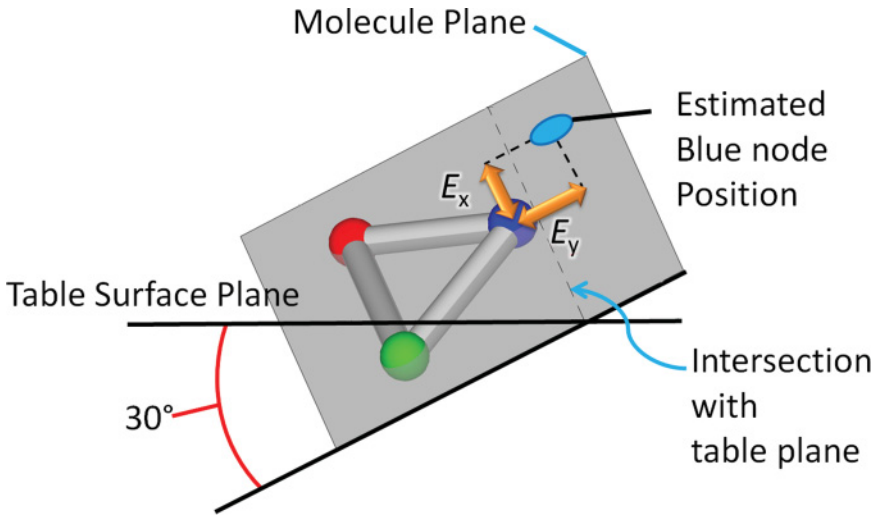


Fig. 26. The position of the molecule in virtual space and the measures of error (E_x and E_y).

the plane of the table, rotated along the x -axis (parallel to the long side of the table, Figure 26). Correspondingly, the tablet computer was tilted to match the 3D model (Figure 23).

The position of the blue node was altered randomly along the x and y axis, but to avoid potential perceptual effects of symmetry and square angles, we limited the possible positions to two square areas that would result in no symmetry in the molecule (Figure 25).

To provide their answers, participants tapped on the tablet PC to indicate the relative position of the blue node with respect to the red and green nodes, following similar methods described by Goldstein [1991]. The representation of the molecule on the tablet was planar (i.e., 2D geometrical shapes), which, combined with the tilting of the tablets and explicit verbal instructions to the participants, conveyed that the shapes on the response tablet's screen represented the spatial shape of the object rather than a 3D view.

Once both participants had entered their answers, the next trial would begin. The distance from the red to green nodes on the table's surface in the medium-discrepancy condition with parallel geometry was 462 pixels (24.3 cm). The distance from the red

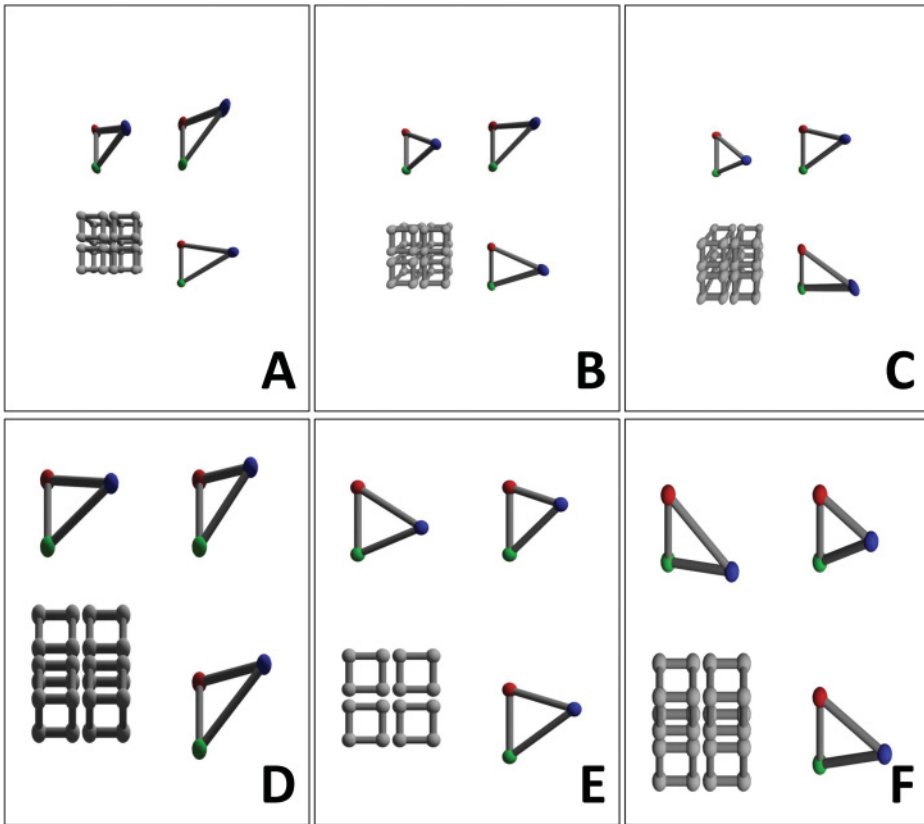


Fig. 27. Table renderings of target objects in each of the main conditions. Notice that in a real trial, only two molecules would be visible, and another set of gray cubes would take its place. Figures (A)–(F) follow the same convention as in Figure 6.

to green circles on the tablets was 323 pixels (7.9 cm). This scaling was necessary due to the large difference in size between the tablets and the objects on the table. As in the first experiment, the quadrants of the table not used to display target objects were occupied by reference objects (sets of four cube molecules) using the same representation as the corresponding trial in order to provide some context (one example of these objects is also represented in Figure 27). As in Experiment 1, targets were always at the participant's right, so they never estimated the structure of the same object simultaneously.

6.4. Procedure and Experimental Design

Targets could appear either in the “near” or “far” halves of the table (Figure 24). Participants carried out two practice trials (one near and one far) and eight testing trials (four repetitions of near and far) for each *discrepancy-parallax* \times *projection* combination, for a total of 100 trials (including practice). These ten combinations were counterbalanced between participants using a random Latin Square. Therefore, the experiment tested a 5 discrepancy-parallax \times 2 projection geometry \times 2 location fully crossed within-participants design.

Table III. Planned Comparisons for Experiment 2 (Internal Structure Estimation)

	E_x		E_y	TCT
	Parallel	Perspective	Parallel+Perspective	Parallel+Perspective
NA to MA	$p < \mathbf{0.001}$	$p = \mathbf{0.004}$	$p = 0.183$	$p = \mathbf{0.006}$
MA to LA	$p < \mathbf{0.001}$	$p < \mathbf{0.001}$	$p = 0.192$	$p = \mathbf{0.004}$
NP to NA	$p = 0.433$	$p = 0.415$	$p = 0.391$	$p = 0.384$
LA to LP	$p = .032$	$p = .533$	$p = .805$	$p = .053$
MA to NP	$p < \mathbf{.001}$	$p = 0.032$	$p = 0.567$	$p = \mathbf{0.009}$
MA to LP	$p < \mathbf{0.001}$	$p < \mathbf{0.001}$	$p = 0.402$	$p < \mathbf{0.001}$

6.5. Measures

Both error and TCT were measured. Errors were considered along the x - and y -axes:

E_x : the error measured along the x -axis of the table, calculated as the orthogonal distance between the measured and target positions of the blue node, expressed as a fraction of the distance between the green and red nodes (Figure 26).

E_y : the error measured along the y -axis of the molecule (at 30° with respect to the y -axis of the table—see Figure 26), calculated as above.

6.6. Experiment 2 Specific Hypothesis

The study was designed to isolate errors along the x -axis of the table (E_x) due to the change in discrepancy. Thus, hypotheses referring to discrepancy (H1–H4, see Section 4.5) apply to the E_x measure, and H5 still refers to the time measure. Additionally, the H6 hypothesis from Experiment 1 (that the type of projection geometry will affect E_y) applies to E_y error.

6.7. Results

We performed a 5 discrepancy-parallax \times 2 projection geometry \times 2 location RM-ANOVA, and planned comparisons for pairwise differences. We report on transversal errors (error in x), longitudinal error (error in y), time, and subjective measures.

6.7.1. Error in x (E_x). The omnibus ANOVA of the main error measure (E_x —longitudinal error) revealed significant effects of discrepancy-parallax ($F_{3,20,73.50} = 49.334, p < 0.001, \eta_p^2 = 0.682$) and an interaction between discrepancy-parallax and projection type ($F_{3,36,77.22} = 16.836, p < 0.001, \eta_p^2 = 0.423$), but not of projection type ($F_{1,23} = 2.944, p = 0.1, \eta_p^2 = 0.113$) or stimuli location ($F_{1,23} = 2.365, p = 0.138, \eta_p^2 = 0.093$). Additionally, the interaction of all factors was also significant ($F_{3,358,77.224} = 16.836, p < 0.001, \eta_p^2 = 0.423$). Because several of the interactions (notably projection type vs. discrepancy-parallax) were significant, we ran post-hoc comparisons separately for parallel and perspective projections (see Table III), and we display longitudinal error in Figure 28.⁹

H1 is not supported by the longitudinal error data: although errors are slightly larger with large discrepancy, especially without parallax, than with no discrepancy, the surprisingly low error for the parallel medium discrepancy condition breaks the pattern. If there is a relationship between discrepancy and longitudinal error, this relation is not monotonic or, more likely, the medium discrepancy is a special case (H4).

H2.A is partially supported by the error data, since the error was larger in the low-discrepancy conditions for the parallel projections than for the perspective projection. However, the error was also larger for parallel projections in large-discrepancy conditions, which contradicts H2.B. Overall, H2 is not supported.

⁹A comprehensive table of the statistical tests and all their details for all measures can be found in Appendix B, Experiment 2.

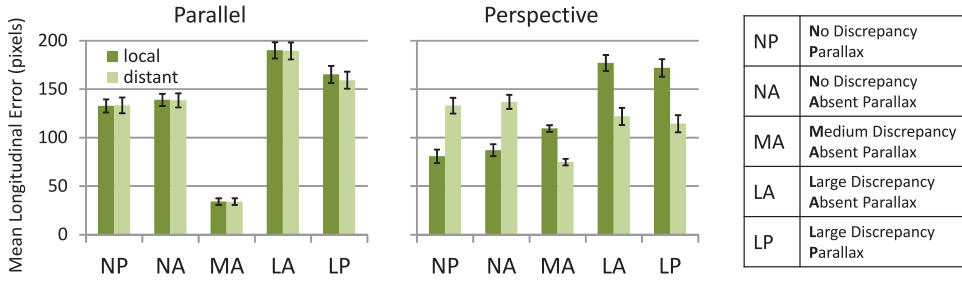


Fig. 28. Average errors within the longitudinal plane for the internal structure estimation task, split by the location of the stimuli on the table. Error bars indicate standard error.

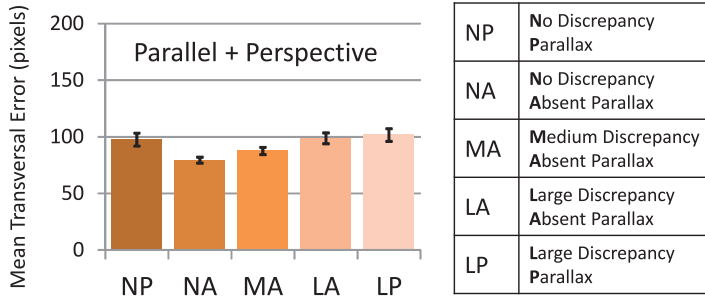


Fig. 29. Average errors within the transversal plane (E_y) for the internal structure estimation task. Error bars indicate standard error.

H3.A and B do not find any support in the data since none of the post-hocs comparing parallax with non-parallax conditions were significant and the figures show how, if these differences exist, they are very small. H3 is, thus, also not supported.

H4 (whether the medium-discrepancy case is special) is strongly supported by the data. Almost all post-hoc comparisons that include the medium discrepancy condition show how the error was significantly smaller in this condition.

A further analysis of the effects on error of the location of the stimuli shows an interesting difference between the parallel and perspective projections that is clearly visible in Figure 28. Each of the different discrepancy-parallax conditions in the parallel projection have very similar average errors for stimuli that appeared in the close and distant parts of the table. However, when projected in perspective, closer stimuli resulted in smaller errors for low discrepancy and larger errors for large discrepancy, an effect that is inverted for stimuli that are located further away.

6.7.2. Error in y (E_y). The measures in the transversal direction of error (Figure 29) did not show any significant effect of discrepancy-parallax ($F_{1,312,30,167} = 0.899$, $p = 0.378$, $\eta_p^2 = 0.038$), projection type ($F_{1,23} = 0.121$, $p = 0.731$, $\eta_p^2 = 0.005$), or any of the interactions (all $p > 0.46$), although it showed a small effect of stimuli location ($F_{1,23} = 5.708$, $p = 0.025$, $\eta_p^2 = 0.199$), which indicates that, overall, there is more transversal error when judging stimuli in close locations ($M_{\text{local}} = 98.03$ pixels) than when judging stimuli in distant locations ($M_{\text{distant}} = 87.83$ pixels).

H6 is not supported by the analysis of E_y since projection geometry did not make a significant difference in transversal error nor did it interact with any of the other factors.

6.7.3. TCT. The omnibus ANOVA on the log-transformed TCTs showed a main effect of discrepancy-parallax ($F_{3,198,73,555} = 6.212$, $p = 0.001$, $\eta_p^2 = 0.213$), but none of the

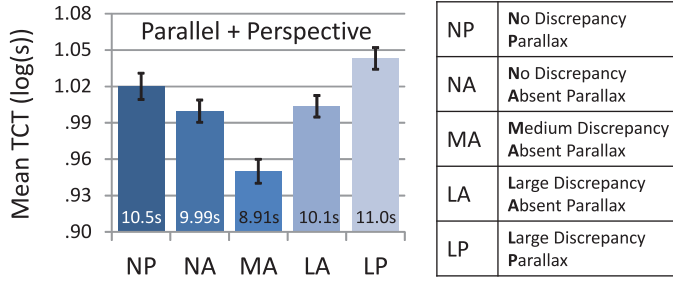


Fig. 30. Average TCTs (log-transformed) for the internal structure estimation task. Numbers at the bottom of the column represent the untransformed value in seconds for the average of the log-transformed completion times. Both projection types are aggregated because the interaction between discrepancy-parallax and projection geometry was not significant. Error bars represent standard error.

other main effects were significant (projection type: $F_{1,23} = 3.932$, $p = 0.059$, $\eta_p^2 = 0.146$; stimuli location: $F_{1,23} = 0.681$, $p = 0.418$, $\eta_p^2 = 0.029$), neither were any of the interactions between factors (all $p > 0.064$). The bar graphs in Figure 30 and the post-hoc comparisons in Table III show that the shortest times correspond to the medium discrepancy condition (both perspective and parallel projections were aggregated).

H1 is contradicted by the data, since completion times for no discrepancy and large discrepancy are pretty much equivalent, and the medium discrepancy condition clearly shows shorter completion times that are also backed up by the post-hoc tests.

H2.A and H2.B are not supported: perspective does not make it faster in the low discrepancy condition nor slower in the large discrepancy conditions.

H3.B is consistent with the average TCTs, but the effects were too small to show significant differences. H3.A is not supported by the values or the tests.

H4 is strongly supported since the medium discrepancy conditions were the fastest, and the post-hoc tests show statistically significant differences between the medium discrepancy and all other discrepancy conditions.

H5 is supported by the average trial durations, which were longer for conditions with parallax; however, these differences did not show up as statistical differences in the post-hoc tests, and therefore, the results about H5 are inconclusive.

6.7.4. Subjective Assessments. The omnibus ANOVAs of the rank-transformed data for perceived difficulty revealed a main effect of discrepancy-parallax ($F_{4,92} = 5.39$, $p = 0.001$, $\eta_p^2 = 0.19$), but no effect of projection type ($F_{1,23} = 0.76$, $p = 0.39$, $\eta_p^2 = 0.032$) nor interaction between the two ($F_{4,92} = 1.51$, $p = 0.20$, $\eta_p^2 = 0.062$). For the preference measure, discrepancy-parallax was again significant ($F_{3,0,67.2} = 4.17$, $p = 0.009$, $\eta_p^2 = 0.159$), and projection type was not ($F_{1,22} = 4.46$, $p = 0.002$, $\eta_p^2 = 0.068$); however, the interaction between the two was significant ($F_{4,88} = 3.18$, $p = 0.017$, $\eta_p^2 = 0.126$), which leads to a separate post-hoc analysis for the parallel and perspective conditions.¹⁰

The post-hoc analyses (Table IV) provide partial support for some of the objective findings. Participants judged the no-discrepancy no-parallax condition as easier than the no-discrepancy with parallax condition (a perceived negative effect of parallax). They also found the medium discrepancy condition easier than the large-discrepancy with and without parallax conditions and easier than the no discrepancy with parallax condition. These results do not provide strong support for H1; although large discrepancy is generally found to be worse than medium and no discrepancy, medium is better than no discrepancy with parallax and equivalent to no discrepancy without parallax.

¹⁰Degrees of freedom in the preference tests differ slightly from the difficulty rating and other experiments because one participant failed to rate two conditions and can, therefore, not be included in the RM-ANOVA.

Table IV. Planned Comparisons for the Subjective Likert Results of Experiment 2 (Internal Structure Estimation Task)

	<i>Difficulty</i>	<i>Preference</i>	
	Parallel and Perspective	Parallel	Perspective
NA to MA	$p = 0.703$	$p = 0.053$	$p = 0.393$
MA to LA	$p = \mathbf{0.001}$	$p = \mathbf{0.004}$	$p = 0.032$
NP to NA	$p = \mathbf{0.007}$	$p = 0.018$	$p = 0.776$
LA to LP	$p = 0.357$	$p = 0.908$	$p = 0.464$
MA to NP	$p = \mathbf{0.011}$	$p < \mathbf{0.001}$	$p = 0.583$
MA to LP	$p = \mathbf{0.004}$	$p = \mathbf{0.002}$	$p = 0.065$

For the difficulty question, the interaction between projection type and discrepancy-parallax was not significant, and therefore, parallel and perspective ratings were aggregated prior to analysis. For the preference question, the interaction between the two main factors was significant and therefore the two levels of projection type were analyzed separately.

The subjective difficulty ratings contradict H3.A—participants thought that parallax made the task more difficult with no discrepancy—and partially support H4, since the medium discrepancy showed some significant advantages.

The preference ratings are similar to the difficulty ratings, but only for the parallel geometry: medium discrepancy is preferred over no discrepancy with parallax and over large discrepancies with and without parallax. A preference for the medium discrepancy condition provides some support for H4, although not all comparisons are significant, most likely due to the limited power of the conservative post-hoc non-parametric tests.¹¹

6.8. Summary

Unlike in Experiment 1, this experiment does not show a general advantage of reducing discrepancy to improve accuracy. Results indicate that placing the CoP above the table provides the best accuracy when perceiving the geometrical relationships internal to objects represented on the table, which is noticeably better than providing a PoV-adapted CoP. This advantage in accuracy is more pronounced for parallel projections, but it is present also in perspective projections. The time measurements further support this, since medium discrepancy trials were also the fastest to make. The subjective results partially support the objective measurements.

Additionally, the analysis including the location of the stimuli shows an interesting difference between parallel and perspective projections: parallel projections do not seem to be vulnerable to the location of the stimuli, whereas in perspective projections, we can observe that distant stimuli result in larger errors for no-discrepancy conditions and local stimuli result in larger errors for large discrepancy conditions.

7. EXPERIMENT 3: OBJECT STRUCTURE RECOGNITION (I)

This experiment tested people's ability to recognize 3D objects represented on the tabletop—in particular, how accurately people can determine that two representations correspond to the same 3D structure. This task involves estimation of both internal structure and external orientation.

7.1. Participants and Apparatus

The same 24 participants (13 female, 11 male) of Experiment 2 completed this experiment (this study was conducted second). The apparatus for this experiment is also the same as for Experiment 2, including the use of the tablets to provide input, except that

¹¹Detailed tables of subjective ratings for this experiment are available in Appendix B, Experiment 2.

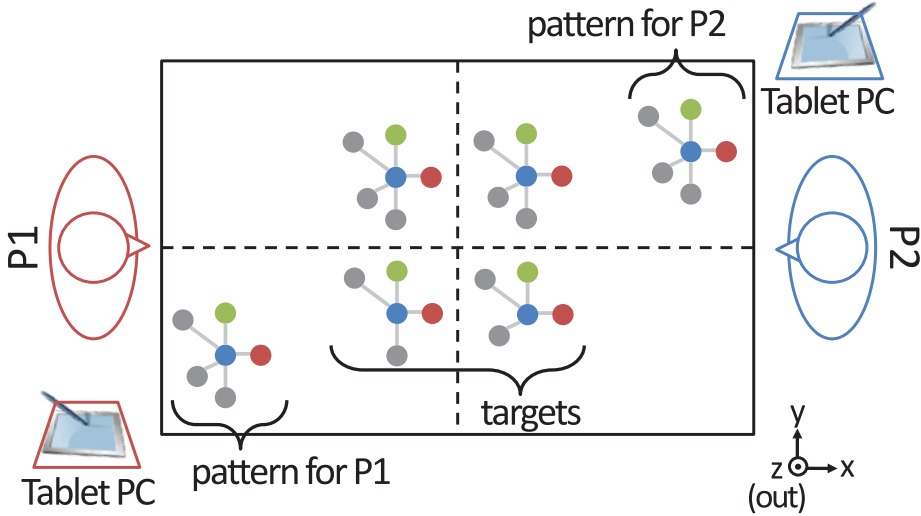


Fig. 31. Diagram of the experimental setup for Experiment 3. In this experiment, the stand holding the tablet PC could be adjusted in height and angle. Representations of the molecules are diagrammatic; for a representation of a molecule as in the experiment, see Figure 32.

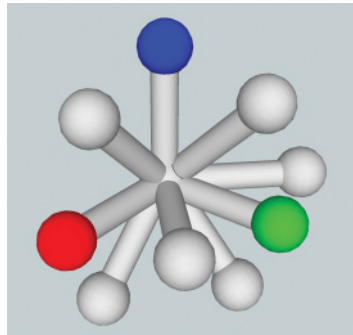


Fig. 32. The basic molecule of Experiments 3 and 4. The red, blue, and green nodes were always present. Each target or pattern molecule could have 3–5 randomly chosen gray nodes of the six possible gray nodes shown in this figure.

in this experiment, the participants were free to alter both the height and tilt angle of the stand. Experimental setup is shown in Figure 31.

7.2. Task

Each participant was presented with a pattern molecule at their end of the display and four target molecules in the center of the display (Figure 31). Participants were asked to compare their pattern molecule to the target molecules and indicate how many of the target molecules had the same shape as the pattern. The molecules were built from spherical nodes connected to a central node, at different angles (Figure 32). In each molecule the axis nodes (red, x ; green, y ; and blue, z) were all present. Molecules differed in the number (3–5) and choice of gray nodes.

To provide their answer, the tablet had four squares that corresponded to the four target molecules; participants had to tap any and all quadrants that matched their pattern. Once both participants had answered, the correct answers were shown.

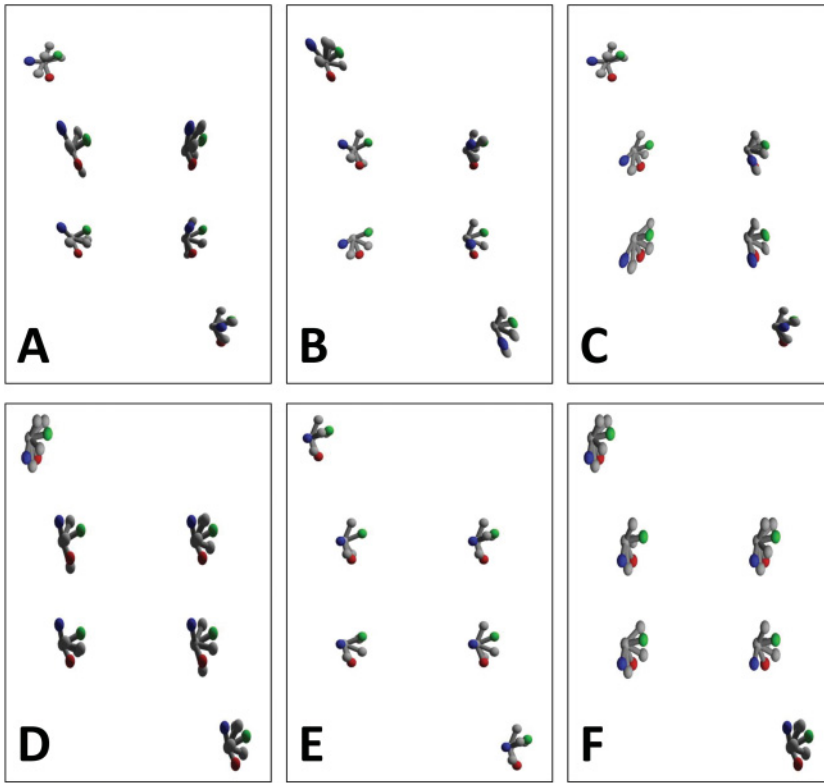


Fig. 33. Table renderings of target and pattern objects in each of the main conditions for Experiments 3 and 4. Figures (A)–(F) represent the same conditions as in Figure 6.

Table V. Centers of Projection used for the Different Conditions

Condition	CoP of Target Molecules	CoP of Local Molecule (Pattern)
Low Discrepancy	PoV of viewer	PoV of viewer
Medium Discrepancy	Above center of table	Above center of table
Large Discrepancy	PoV of opposite Person	PoV of viewer

In parallel projections the representations are analogous, but with directions of projection and *direction* of view instead of center of projection (CoP) and point of view (PoV).

The presentation of conditions for Experiment 3 is slightly different to that of experiments 1 and 2 due to the presence of a 3D-rendered local pattern. In the no-discrepancy and medium discrepancy conditions of this experiment, the pattern and the targets were rendered using the same projection and CoP. That is, CoP would coincide with PoV of the participant in the no-discrepancy condition, and CoP would be above the table for the medium discrepancy condition. However, for the large discrepancy condition, the pattern molecule would be rendered with no discrepancy, which results in not all molecules in the table being rendered in the same way (unlike in Experiments 1 and 2). The different representations of the local and target molecules are summarized in Table V. We decided to test large-discrepancy conditions in this way because it is unlikely that a *local* object would, in the real world, be displayed in a way that is most difficult to see by its owner. We also anticipated that this experimental choice would better represent real-life scenarios where viewers are trying to match an object of which they already understand its 3D structure but is not visually present at that time. The

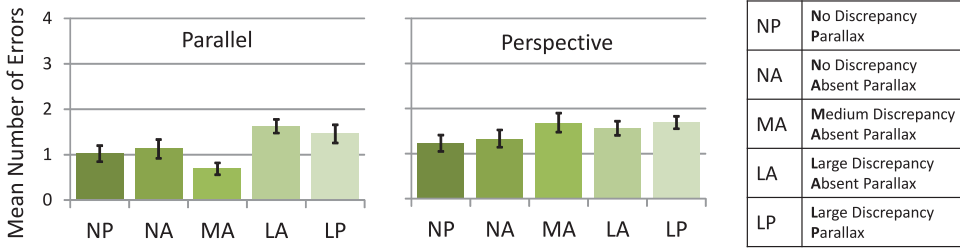


Fig. 34. Mean number of errors in recognizing patterns for the object structure recognition task (Experiment 3). Error bars indicate standard error.

participant would presumably look at a pattern that was rendered for her, understand its 3D structure, and use this understanding when trying to find it in the table. This design also made it harder for participants to directly compare the projections without understanding the 3D structure.

7.3. Procedure and Design

In each of the ten condition-projection combinations, participants performed all combinations of the remaining factors in random order. Participants carried out one practice trial and two testing trials for each combination, for a total of 30 trials. The ten condition-projection pairs were counterbalanced between participants using a random Latin Square.

The study, therefore, used a 5 discrepancy-parallax condition \times 2 projection geometry fully crossed within-participants design. Notice that stimulus location cannot be a factor in this experiment (or Experiment 4) because all the molecules on the table served as stimuli simultaneously for each trial.

Error and TCT were both measured. The error was defined as Error in Matches (E_g)—the number of incorrectly matched patterns. This measure included both false positives and false negatives and was in the range 0–4.

This experiment tested the general hypotheses (H1–H5) described above (Section 4.5).

7.4. Results

As in previous experiments, we provide error, time, and subjective assessment results.

7.4.1. Error in Matches (E_g). Bar graphs of errors are represented in Figure 34. The omnibus ANOVA revealed main effects of discrepancy-parallax ($F_{4,92} = 3.524$, $p < 0.01$, $\eta_p^2 = 0.133$), projection type ($F_{1,23} = 13.395$, $p < 0.001$, $\eta_p^2 = 0.368$), and an interaction between the two ($F_{4,92} = 2.706$, $p = 0.035$, $\eta_p^2 = 0.105$). Because the interaction was significant, we ran the post-hoc tests separately for parallel and perspective conditions (see Table VI).

H1 is only partially supported by the results. In the parallel conditions, our tests do show that MA has significantly fewer errors than LA and LP, but if H1 were true, we would expect to see NP and NA with fewer errors than MA, since they have the least discrepancy. This is not the case and, in fact, the averages show exactly the opposite: the parallel medium discrepancy condition showed the least amount of errors (0.68 errors). In the perspective projections, post-hoc tests did not reveal any effect of discrepancy.

H2.A and H2.B are not supported by the results, since perspective was generally unaffected by discrepancy and the parallel discrepancy conditions result, on average, in fewer or similar errors than their perspective counterparts. H2 is, therefore, unsupported.

Table VI. Planned Post-Hoc Comparisons for Experiment 3 (Object Structure Recognition), Errors and Task Completion Time

	<i>Errors</i>		<i>TCT(log10)</i>	
	Parallel	Perspective	Parallel	Perspective
NA to MA	$p = 0.09$	$p = 0.17$	$p = 0.151$	$p = 0.24$
MA to LA	$p = \mathbf{0.001}$	$p = 0.623$	$p < \mathbf{0.001}$	$p = 0.516$
NP to NA	$p = 0.633$	$p = 0.61$	$p = 0.218$	$p = 0.945$
LA to LP	$p = 0.547$	$p = 0.401$	$p = 0.664$	$p = 0.258$
MA to NP	$p = 0.057$	$p = 0.076$	$p = 0.033$	$p = 0.298$
MA to LP	$p < \mathbf{0.004}$	$p = 0.986$	$p < \mathbf{0.001}$	$p = 0.06$

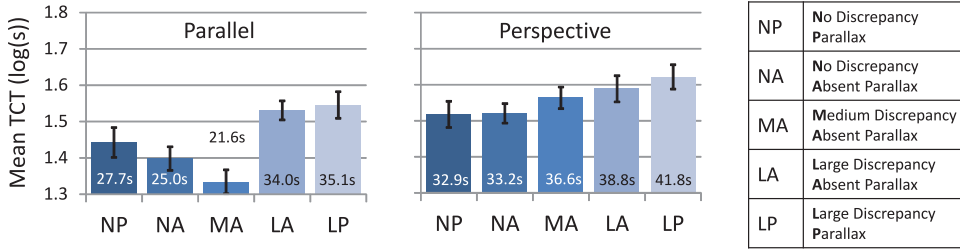


Fig. 35. Mean TCT (log10-transformed) for the object structure recognition task. Error bars indicate standard error. The numbers within or above the bars indicate the un-transformed value of the log mean in seconds. Note that the y-axis does not represent an absolute scale (it starts at 1.3).

H3.A and H3.B are not supported; self-parallax does not seem to affect perspective projection conditions (NP-NA and LA-LP in perspective are not significant and have very close average values). H3 is, therefore, unsupported.

H4 is partially supported, but only for parallel geometry conditions and with respect to large discrepancy only. Two significant post-hoc comparisons support an advantage of the parallel MA condition against the LP condition but the comparisons between NA and the low discrepancy conditions are not significant (although they approach significance).

7.4.2. TCT. Figure 35 shows bar graphs of the TCTs. The omnibus ANOVA of the log-transformed TCTs revealed main effects of discrepancy-parallax ($F_{4,92} = 11.235$, $p < 0.001$, $\eta_p^2 = 0.328$), projection type ($F_{1,23} = 18.39$, $p < 0.001$, $\eta_p^2 = 0.444$), and an interaction between the two ($F_{4,92} = 2.61$, $p = 0.04$, $\eta_p^2 = 0.102$). The interaction led to separate post-hoc tests for parallel and perspective conditions (see Table VI).

H1 is not supported by the time data. Although looking at the average values only for the perspective condition one can see a clear monotonic increase in time, the selected post-hoc tests were not powerful enough to show significant differences. In contrast, the parallel conditions do contradict H1. For example, the lowest average time corresponds to MA (21.6s), albeit without strong evidence from the post-hocs, which were only able to statistically separate MA from the large discrepancy conditions. Results for H1 are then ambivalent as to the role of discrepancy in the completion time of these tasks.

H2.A is not supported, since the times for parallel projections were generally lower than those for perspective conditions. H2.B might still hold, but we do not have individual explicit tests to support it. H2 is, therefore, not supported.

H3.A and H3.B are unsupported by the data, since none of the NP-NA or LA-LP comparisons were significant. H3 is, therefore, also unsupported by the time data.

H4 is supported by Figure 35, and only in the Parallel projection, but none of the post-hoc tests provides conclusive evidence.

Table VII. Planned Comparisons for Subjective Likert Results of Experiment 3 (Structure Recognition Task)

	Difficulty		Preference	
	Parallel	Perspective	Parallel	Perspective
NA to MA	$p = 0.767$	$p = \mathbf{0.005}$	$p = 0.844$	$p = 0.036$
MA to LA	$p < \mathbf{0.001}$	$p = 0.220$	$p < \mathbf{0.001}$	$p = 0.507$
NP to NA	$p = 0.031$	$p = 0.602$	$p = 0.200$	$p = 0.832$
LA to LP	$p = 0.355$	$p = 0.272$	$p = 0.964$	$p = 0.592$
MA to NP	$p = 0.068$	$p = \mathbf{0.003}$	$p = 0.430$	$p = 0.037$
MA to LP	$p < \mathbf{0.001}$	$p = 0.100$	$p = \mathbf{0.001}$	$p = 0.271$

7.4.3. Subjective Assessment. The non-parametric omnibus tests of the difficulty questions showed a significant discrepancy-parallax main effect ($F_{2.78,64.01} = 13.229$, $p < 0.001$, $\eta_p^2 = 0.365$), as well as projection type ($F_{1,23} = 27.585$, $p < 0.001$, $\eta_p^2 = 0.545$), and an interaction ($F_{4.92} = 9.508$, $p < 0.001$, $\eta_p^2 = 0.292$). The tests for the preference question mirror the results of the difficulty question¹²: discrepancy-parallax was significant ($F_{2.587,56.9} = 9.438$, $p < 0.001$, $\eta_p^2 = 0.30$), as well as projection type ($F_{1,22} = 15.229$, $p = 0.001$, $\eta_p^2 = 0.41$), and their interaction ($F_{4.88} = 2.474$, $p = 0.05$, $\eta_p^2 = 0.101$). Table VII shows a post-hoc comparison result summary.¹³

H1 is not supported by the subjective data. However, a separate look at parallel and perspective projection geometries shows interesting different patterns. In perspective projections, the conditions were considered more difficult and less preferable as the discrepancy increased, except that the medium discrepancy condition that projects from the top of the table was recognized as the worst. This result is statistically supported in the case of difficulty. For parallel conditions, discrepancy also makes things worse, except that the medium discrepancy condition appears statistically equivalent to the no-discrepancy conditions, and shows a statistically significant advantage to the LP large discrepancy conditions in both difficulty and preference.

H2.A is not supported, since perspective was generally rated harder than parallel in the no-discrepancy conditions. As in the objective measures, perspective seemed to make large-discrepancy conditions worse on average. This trend could support H2.B but the statistical tests only indirectly support this. The support for H2 is thus, ambivalent at best.

H3.A and H3.B are not supported since none of the NP-NA or LA-LP post-hoc tests were significant. Therefore, H3 is also not supported.

H4 is not supported by the subjective data, since MA was not rated better than the lower discrepancy conditions. However, the difference between parallel and perspective is revealing: for perspective projections, the medium discrepancy was rated badly on average, with very similar scores to the large discrepancy conditions, whereas for parallel projection, the result is the opposite (ratings for MA are best or close to best).

7.5. Discussion and Summary

The results for the object structure recognition task do not provide strong support that large discrepancies between PoV and CoP are detrimental for this task. Instead, the data support a benefit of the parallel medium-discrepancy condition, which is the best for this task, resulting in the fewest average errors and the shortest completion times; however, the statistical tests were not able to single it out from the lower discrepancy

¹²As in Experiment 2, the degrees of freedom in the preference tests differ slightly from the difficulty rating and other experiments because one participant failed to rate two conditions and can, therefore, not be included in the RM-ANOVA.

¹³Frequency tables of participant answers for the two questions, are available in Appendix B, Experiment 3.

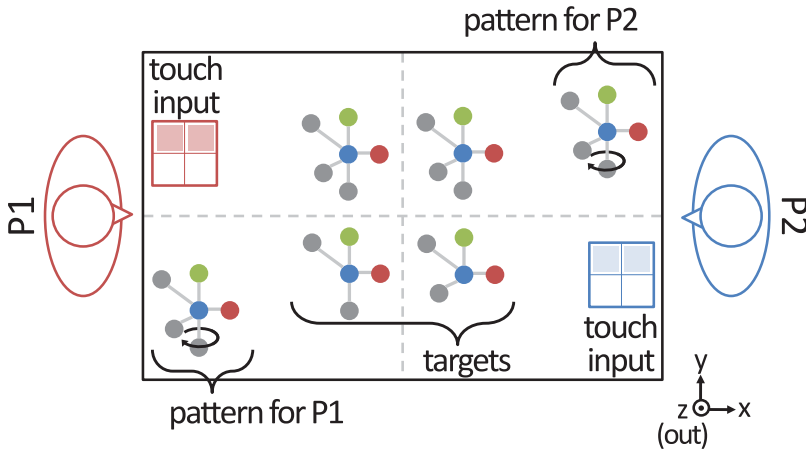


Fig. 36. Diagram of the experimental setup for Experiment 4.

conditions, and neither did the participants in their subjective assessments. When considered overall, the parallel geometry results in fewer errors and shorter completion times than the perspective geometries, and are recognized as better by participants as well.

8. EXPERIMENT 4: OBJECT STRUCTURE RECOGNITION (II-INTERACTIVE)

When interacting with 3D content, it is often possible to alter the position of the objects to see them from a different point of view. This is complementary to self-motion parallax, but instead of moving the head we move the object (see Section 3). The main motivation for this experiment was to generalize perception effects in structure recognition to situations where interactive manipulation of the pose of objects is allowed. A secondary goal was to explore differences in performance and accuracy due to interaction (although this requires cross-experiment analysis that is described in Section 9).

8.1. Participants

The same 24 participants of Experiment 1 took part in this experiment (Exp 1 was performed first). Since the participant pools for Experiments 2 and 3 and 1 and 4 were separate, none of the participants in this study took part in its twin, Experiment 3.

8.2. Apparatus

The apparatus for this experiment was the same as for Experiment 3, with two exceptions: participants could rotate their own pattern molecule through direct touch, and participants gave their answers on the table instead of through tablets (Figure 36). Touching and dragging the pattern molecule made the molecule rotate around the axis perpendicular to the direction of rotation and parallel to the table plane [Hancock et al. 2007]. The central gray node (Figure 32) was always in the axis of rotation, so the center node was never translated.

8.3. Task, Procedure, Design, Measures, and Hypotheses

The task, procedure, design, measures, and hypotheses were identical to the task in Experiment 3, except that participants were allowed to rotate the local pattern molecule.

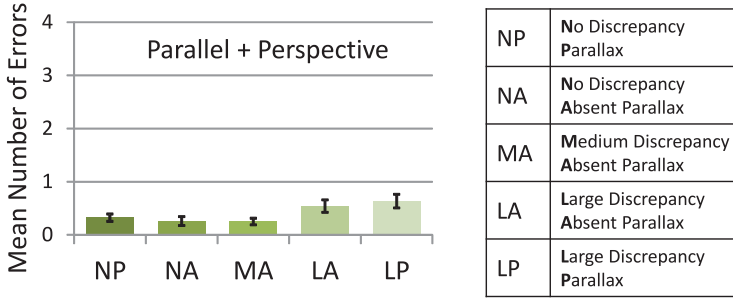


Fig. 37. Mean number of errors in recognizing patterns for the interactive object structure recognition task (Experiment 4). Error bars indicate standard error.

Table VIII. Planned Post-Hoc Comparisons for Experiment 4, Errors and TCT

	Errors	TCT (\log_{10})	
	Parallel and Perspective	Parallel	Perspective
NA to MA	$p = 0.816$	$p = 0.65$	$p = 0.016$
MA to LA	$p = 0.043$	$p = \mathbf{0.001}$	$p = 0.332$
NP to NA	$p = 0.193$	$p = 0.53$	$p = 0.113$
LA to LP	$p = 0.496$	$p = 0.512$	$p = 0.032$
MA to NP	$p = 0.389$	$p = 0.158$	$p < \mathbf{0.001}$
MA to LP	$p = \mathbf{0.005}$	$p = \mathbf{0.001}$	$p < \mathbf{0.001}$

8.4. Results

8.4.1. Error in Matches (E_g). Figure 37 shows number of errors in the aggregated parallel and perspective condition. The omnibus ANOVA was significant for discrepancy-parallax ($F_{4,92} = 4.277, p = 0.003, \eta_p^2 = 0.157$), although neither projection type ($F_{1,23} = 0.594, p < 0.449, \eta_p^2 = 0.025$) nor the interaction between discrepancy-parallax and project type ($F_{4,92} = 0.837, p = 0.505, \eta_p^2 = 0.035$) were significant. Due to the lack of significant interaction, post-hoc tests compare discrepancy-parallax conditions aggregated across projection (parallel and perspective together, see Table VIII). Figure 37 shows results aggregating the parallel and perspective conditions.

H1 is partially supported by the difference in errors between medium discrepancy without parallax (0.25 errors/trial) and large discrepancy with parallax (0.63 errors/trial), which is statistically significant. However, although not significant, medium discrepancy without parallax had the lowest average rates (0.25 errors/trial) also compared with no discrepancy without parallax (0.26) or with parallax (0.32). The lack of significant results in the error measures is attributable to the small number of errors.

H2.A and H2.B cannot be supported, since the omnibus test did not show a difference between projection types. H2 is, therefore, also not supported.

H3.A and H3.B are unsupported, since none of the differences between parallax and non-parallax conditions are significant or even show big differences in average values. H3 is, therefore, also not supported.

H4 is not supported, but there seems to be a pattern, with the medium discrepancy condition showing fewer errors than any of the others. As in H1, the significance comparisons between NA and the other conditions are limited due to the small number of errors.

8.4.2. TCT. Figure 38 shows completion time. The omnibus ANOVA showed main effects of discrepancy ($F_{4,92} = 5.021, p = 0.001, \eta_p^2 = 0.179$), projection type ($F_{1,23} = 55.85, p < 0.001, \eta_p^2 = 0.708$) and their interaction ($F_{4,92} = 8.932, p < 0.001, \eta_p^2 = 0.28$).

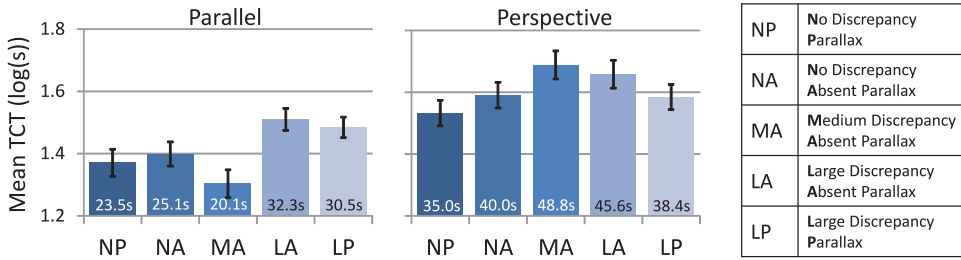


Fig. 38. Mean TCT (log10-transformed) for the interactive object structure recognition task. Error bars indicate standard error. The numbers within the bars indicate the un-transformed value of the log mean in seconds. Note that the y-axis does not represent an absolute scale (it starts at 1.2).

Figure 38 shows the difference between parallel and perspective conditions, with perspective taking longer for every condition. The difference between the medium discrepancy times is also obvious: medium discrepancy without parallax is fastest for parallel and slowest for perspective. Since the interaction was significant we analyze and display the parallel and perspective geometries separately. The results of the post-hoc tests are displayed in Table VIII.

H1 is not supported, since the only significant differences were between medium and large discrepancies in the parallel conditions, and not between the low discrepancy and large discrepancy conditions. Moreover, the medium discrepancy is, on average, the fastest for parallel and the slowest for perspective, although the post-hoc tests were only able to support the poorer performance of large discrepancy in the parallel projection with respect to medium discrepancy, and the longer times for the medium discrepancy in perspective.

H2.A is not supported since parallel conditions were generally faster than perspective. H2.B seems plausible, but rather than due to the effect of discrepancy on projection, this seems just an effect of the poorer performance of perspective in this task. Evidence supporting H2 is, therefore, at best contradictory.

H3.A and H3.B are also not supported by any post-hoc tests, although the values of the averages seem to indicate exactly the opposite—the averages of parallax conditions were, in all cases, lower than their non-parallax counterparts. H3 is, therefore, not supported.

H4 is supported, since the medium discrepancy condition is indeed special, only in very different ways for the parallel and the perspective conditions. Although the post-hocs are only significant for the medium and large discrepancy conditions without parallax and the medium discrepancy without parallax and large discrepancy with parallax conditions, the task took the shortest time for the parallel medium discrepancy condition and the longest for the perspective medium discrepancy condition.

H5 is not supported, the differences between parallax and no parallax were not significant; moreover, the average values might indicate the opposite.

8.4.3. Subjective Assessment. The subjective omnibus tests of the difficulty responses showed a main effect of discrepancy-parallax ($F_{4,92} = 13.881, p < 0.001, \eta_p^2 = 0.376$), but not of projection type ($F_{1,23} = 1.96, p = 0.175, \eta_p^2 = 0.076$), nor their interaction ($F_{3,107,71.4} = 0.867, p = 0.465, \eta_p^2 = 0.036$). However, for the preference question, all tests were significant: a main effect of discrepancy-parallax ($F_{4,92} = 15.847, p < 0.001, \eta_p^2 = 0.408$), projection type ($F_{1,23} = 5.951, p = 0.023, \eta_p^2 = 0.206$), and their interaction ($F_{4,92} = 4.093, p = 0.004, \eta_p^2 = 0.151$). Post-hoc tests are displayed in Table IX.¹⁴

¹⁴Tables with response frequencies are available from Appendix B, Experiment 4.

Table IX. Planned Comparisons for the Subjective Likert Results of Experiment 4 (Structure Recognition Task—Interactive)

	Difficulty	Preference	
	Parallel and Perspective	Parallel	Perspective
NA to MA	$p = 0.862$	$p = 0.302$	$p = 0.027$
MA to LA	$p = \mathbf{0.002}$	$p = \mathbf{0.001}$	$p = 0.272$
NP to NA	$p = 0.248$	$p = 0.216$	$p = 0.176$
LA to LP	$p = 0.445$	$p = 0.724$	$p = 0.749$
MA to NP	$p = 0.374$	$p < 0.684$	$p = 0.091$
MA to LP	$p = \mathbf{0.001}$	$p < \mathbf{0.001}$	$p = 0.163$

For the difficulty question, the interaction between projection type and discrepancy-parallax was not significant, and therefore, parallel and perspective ratings were aggregated prior to analysis. For the preference question, the interaction between the two main factors was significant, so projections were analyzed separately.

For H1, the results for the subjective tests do not contradict a general increase of difficulty with discrepancy as in the time measures (although the average ratings follow this order), but they do not provide strong evidence either, since only the post-hoc comparisons of medium discrepancy to large discrepancy were significant.

H2.A and B are not supported by the difficulty subjective results because the interaction was not significant. The preference question results do not provide support for H2 either, since no-discrepancy and large-discrepancy perspective results were generally equivalent to their parallel counterparts. H2 is, therefore, not generally supported.

H3.A and B are also not supported; none of the parallax-to-no-parallax comparisons were significant. H3 is, therefore, not supported.

H4 cannot be supported, since significant post-hocs only concern medium discrepancy conditions in relationship with the large discrepancy conditions.

8.5. Discussion and Summary

Errors in interactive object structure recognition were low, resulting in few conclusive tests. This suggests that explicit interaction such as through touch can have a large positive impact in this kind of task. The differences between the conditions become most evident through the time measure, where we can see different results for the parallel and perspective projections. Interestingly, the task took longer with perspective conditions, and the medium discrepancy condition was worst in perspective, but best in parallel. Self-generated parallax did not seem to have an effect in any measure. The subjective measures are only useful to confirm that large discrepancy is recognized as worse than medium or no-discrepancy, but participants did not seem to recognize the disadvantages of the medium discrepancy condition in the perspective geometry that were evident in the time measures.

9. COMBINED ANALYSIS OF EXPERIMENTS 3 AND 4

Experiments 3 and 4 are very similar, and their main difference is whether participants could interact with the visual stimuli. As discussed earlier, interactivity and parallax are distinct ways to see an object or scene from a different angle. Although the two experiments were performed by a disjoint set of participants, it is still possible to analyze the data jointly by modeling interactivity as a between-subjects factor. With a between-subjects (but fully crossed) interactivity factor, we can learn more about the effects of interactivity.

The main aggregate analysis is a 5 discrepancy-parallax condition (within) \times 2 projection geometry (within) \times interactivity (between) fully crossed mixed design. Except

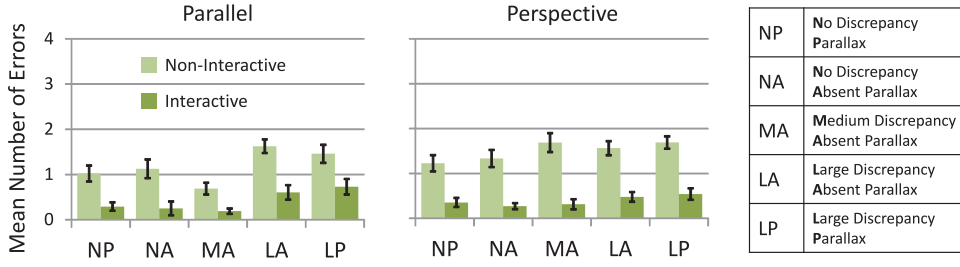


Fig. 39. Mean number of errors in recognizing patterns for the object structure recognition task (Experiments 3 and 4 combined). Error bars represent standard error.

for the addition of the between-subjects interactivity factor, the analysis was equivalent to those of Sections 6.8 and 8. Subjective tests were not analyzed aggregately.

The inclusion of interactivity in experiment 4 also allows the following hypotheses.

H8: Interactivity will reduce error.

H9: Interactivity will increase the time to complete the task.

9.1. Results

We performed a 5 discrepancy-parallax condition \times 2 projection geometry \times 2 interactivity (interactive, non-interactive) repeated measures ANOVA.

9.1.1. Error in Matches (E_g). Figure 39 shows the bar graphs of the errors, comparing the interactive and non-interactive condition. The omnibus ANOVA, which this time includes interactivity as a factor, shows significant main effects of discrepancy-parallax ($F_{4,184} = 8.858, p < 0.001, \eta_p^2 = 0.161$), projection type ($F_{1,46} = 15.165, p < 0.001, \eta_p^2 = 0.248$), and interactivity ($F_{1,46} = 90.592, p < 0.001, \eta_p^2 = 0.663$). From the interactions, Projection Type \times Discrepancy-Parallax was not significant ($F_{4,184} = 0.686, p = 0.603, \eta_p^2 = 0.015$) and Discrepancy-parallax \times Interactivity was not significant either ($F_{4,184} = 0.883, p = 0.475, \eta_p^2 = 0.019$). The other interactions (Projection Type \times Interactivity, and the three-way interaction) were all significant (Respectively: $F_{1,46} = 10.686, p = 0.002, \eta_p^2 = 0.189$; $F_{4,184} = 3.439, p = 0.01, \eta_p^2 = 0.07$).

H8 is strongly confirmed by the significance of the main effect of interactivity. It is clear from Figure 39 that interactive conditions greatly improved accuracy. The average number of errors, across all conditions was 1.342 for the non-interactive condition and 0.402 for the interactive condition: an improvement due to interaction of 0.94 errors per trial (in a task that can have a maximum of four errors). The interactions above highlight that the magnitude of the improvement is not uniform across conditions, but Figure 39 is enough to see that the advantage of interactivity is consistent across all cells.

9.1.2. TCT. For the time measures, the results are redisplayed in Figure 40. The omnibus ANOVA of log-transformed completion times shows a different picture than for the error measures: the Projection Type factor still has a main effect ($F_{1,46} = 69.406, p < 0.001, \eta_p^2 = 0.601$), as does Discrepancy-parallax, ($F_{3,778,173.80} = 12.91, p < 0.001, \eta_p^2 = 0.219$). However, interactivity does not have a significant effect on duration ($F_{1,46} = 0.028, p = 0.867, \eta_p^2 = 0.001$). Of the interactions, Projection Type \times Discrepancy-Parallax is significant ($F_{3,767,173.288} = 9.835, p < 0.001, \eta_p^2 = 0.176$) as is Projection Type \times Interactivity ($F_{1,46} = 5.30, p = 0.026, \eta_p^2 = 0.103$), but neither Discrepancy-Parallax \times Interactivity ($F_{4,184} = 2.145, p = 0.077, \eta_p^2 = 0.045$) nor the three-way interaction ($F_{4,184} = 0.7, p = 0.593, \eta_p^2 = 0.015$) were.

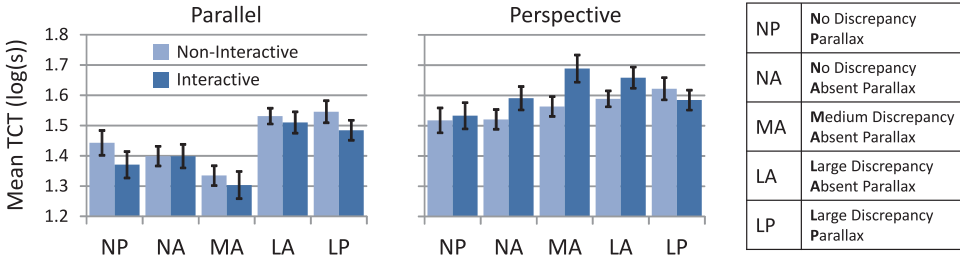


Fig. 40. Mean TCTs (log10-transformed) for the object structure recognition task (Experiments 3 and 4 combined). Error bars indicate standard error. Note that the y-axis does not represent an absolute scale (values start at 1.2).

H9 is not generally supported by our tests, since the main effect of interactivity was not significant, and the actual completion time averages were very similar for the interactive and non-interactive experiments (38.3s vs. 35.8s—a difference of 6.5%). However, the interactions point to a difference in timing present only in the perspective projections; more specifically, the perspective neutral condition took extra time with the interactive system.

9.2. Discussion and Summary

Our results indicate a clear advantage of enabling interactivity in terms of accuracy, and that this advantage does not come at a cost in completion time, except for specific conditions within the perspective projection geometry. Whereas parallax does not seem to make a difference in terms of accuracy in any of the previous experiments, enabling interaction does, even though both support what is visually the same action: changing the point of view. This discrepancy between the two could be due to parallax requiring more physical movement to achieve the same change in perspective (full body movement) and to differences in sensory inputs (e.g., ventral system) as discussed in Section 3.3.

10. SUMMARY OF RESULTS

Figures 41, 42 and 43 summarize the results across all experiments in terms of hypotheses (Figure 41) and in terms of projection and discrepancy-parallax conditions (Figure 43). In this section, we discuss the results presented in Sections 5–9, grouped by topic, followed by the limitations of the study and the main lessons for practitioners.

11. DISCUSSION

11.1. Geometrically Accurate ≠ Perceptually Accurate

It is very tempting, in the absence of knowledge, to assume that a graphical system that reproduces the properties of the represented object most accurately will allow viewers to be less error prone. Our experiments tested this assumption in two ways: first, we tested three levels of discrepancy between the PoV and the CoP, which result in increasing levels of dissimilarity between the retinal projections of a representation and a real object; second, we tested perspective and parallel projection, the latter presenting a similar, if more subtle, geometrical disagreement between the retinal images of real and projected objects.

11.1.1. PoV-CoP Discrepancy. The results of our study do not support the assumption described above: although in the orientation task, error generally increases with discrepancy, this is the exception rather than the rule. In the internal structure estimation

	Hyp.	Description	Measure	Exp1	Exp2	Exp3	Exp4	Exp3/4
Common	H1	+discrepancy → worse estimations	Error	S	I	I	S	N/A
			Time	A	A	C	I	N/A
			Subjective	S	I	I	I	
	H2	perspective better or worse according to projection	Error	I	I	I	I	N/A
			Time	I	I	I	C	N/A
			Subjective	I	I	C	I	
	H2.A	perspective & no discrepancy → better estimations	Error	I	S	I	I	N/A
			Time	I	I	I	I	N/A
			Subjective	I	I	I	I	
	H2.B	perspective & large discrepancy → worse estimations	Error	I	I	I	I	N/A
			Time	I	I	I	S	N/A
			Subjective	I	I	S	I	
	H3	parallax better or worse according to projection	Error	I	I	I	I	N/A
			Time	C	I	I	I	N/A
			Subjective	I	C	I	I	
	H3.A	self parallax → better estimations	Error	I	I	I	I	N/A
			Time	A	I	I	I	N/A
			Subjective	I	A	I	I	
	H3.B	other's parallax → worse estimations	Error	I	I	I	I	N/A
			Time	S	I	I	I	N/A
			Subjective	I	S	I	I	
H4	medium discrepancy is special → better estimations	Error	S	SS	S	I	N/A	
		Time	SS	SS	I	S	N/A	
		Subjective	I	S	I	I		
H5	parallax → longer judgement time	Error	N/A	N/A	N/A	N/A	N/A	
		Time	S	I	I	I	N/A	
		Subjective	N/A	N/A	N/A	N/A		
Exp 1/2	H6	projection-geometry affects transversal error	Error	SS	I	N/A	N/A	N/A
			Time	N/A	N/A	N/A	N/A	N/A
			Subjective	N/A	N/A	N/A	N/A	
H7	long. error reduced for 0 and 90 degree angles (DRE)	Error	SS	N/A	N/A	N/A	N/A	
		Time	N/A	N/A	N/A	N/A	N/A	
		Subjective	N/A	N/A	N/A	N/A		
Exp 3/4	H8	interactive → fewer errors	Error	N/A	N/A	N/A	N/A	SS
			Time	N/A	N/A	N/A	N/A	N/A
			Subjective	N/A	N/A	N/A	N/A	
H9	interactive → longer judgement time	Error	N/A	N/A	N/A	N/A	N/A	
		Time	N/A	N/A	N/A	N/A	I	
		Subjective	N/A	N/A	N/A	N/A		

Fig. 41. Table summary according to hypothesis.

Strongly Supports	SS
Supports	S
Inconclusive/ Not-supported	I
Contradictory/Ambivalent	C
Against	A
Strongly Against	SA

Fig. 42. Hypothesis table summary legend.

task, the medium discrepancy condition showed better accuracy than low discrepancy conditions, and low discrepancy conditions were barely better than large discrepancy conditions. In the structure recognition tasks, results do not generally support effects of discrepancy on error, but the time measure shows that, in the parallel condition, the medium discrepancy condition is actually faster than the low discrepancy condition.

These results suggest that technological efforts to make the PoV and CoP coincide (e.g., fish-tank VR), are not only more expensive, but likely detrimental for accurate

				Parallel					Perspective				
		Parallel	Perspective	NP	NA	MA	LA	LP	NP	NA	MA	LA	LP
Internal Orientation Estimation	Accuracy Longitudinal	N	N	VG	VG	VG	VB	VB	VG	VG	G	VB	VB
	Accuracy Transversal	N	N	VB	VB	VG	VG	VG	VG	VG	VB	VG	VG
	Time	VG	VB	VB	G	VG	G	VB	VB	G	G	G	VB
	Perceived Difficulty	N	N	G	G	G	VB	VB	G	G	G	VB	VB
	Preference	N	N	G	G	G	VB	VB	G	G	G	VB	VB
Internal Structure Estimation	Accuracy Longitudinal	N	N	VB	VB	VG	VB	VB	B	B	VG	VB	VB
	Accuracy Transversal	N	N	N	N	N	N	N	N	N	N	N	N
	Time	N	N	VB	VB	VG	VB	VB	VB	VB	VG	VB	VB
	Perceived Difficulty	N	N	B	VG	VG	VB	VB	B	VG	VG	VB	VB
	Preference	N	N	VB	B	VG	B	B	B	B	G	VB	B
Object Structure Recog.	Accuracy	G	B	G	G	G	B	B	B	N	N	N	N
	Time	VG	VB	G	G	VG	B	B	N	N	N	N	N
	Perceived Difficulty	VG	VB	G	G	G	VB	VB	N	N	B	B	B
	Preference	VG	VB	G	G	G	VB	VB	N	N	B	B	B
	Accuracy	N	N	G	G	G	B	B	G	G	G	B	B
Interactiv e Structure Recog.	Time	VG	VB	G	G	VG	B	B	B	B	VB	B	B
	Perceived Difficulty	N	N	G	G	G	B	B	G	G	G	B	B
	Preference	G	B	G	G	G	B	B	G	G	B	B	B
	Accuracy	N	N	G	G	G	B	B	G	G	G	B	B

Very Good/Best	VG
Good	G
Neutral/Inconclusive	N
Bad but not worst	B
Very Bad/Worst	VB

Fig. 43. Table summary according to condition. *For these measures, there was no interaction between the Parallax-discrepancy and projection, therefore parallel and perspective conditions were analyzed together, and the corresponding values are copied across the parallel and the perspective projection columns.

perception, at least in the tasks that we tested, which involved binocular observation of planar horizontal single-image surfaces. This does not necessarily discard the fish-tank VR approach altogether, and there might be many other reasons and situations where fish-tank VR is preferable, but it does provide a warning to designers that assume this is generally the case.

11.1.2. Projection Geometry. In terms of projection geometry, the results are also inconsistent with a direct relationship between geometric accuracy and perceptual accuracy. As we have discussed above, parallel projections produce an artificial distortion of the shapes, making them project a different retinal image than what a real object would project. However, in all tasks, the perspective did not show statistically significant advantages over parallel projections in terms of accuracy. In fact, the evidence points in the opposite direction: accuracy was somewhat better for parallel projections in the non-interactive object structure recognition task, and it took less time to complete the tasks with parallel projections for all tasks except for the internal structure estimation task.

This could explain why parallel projections are used widely in engineering and architecture. Parallel representations maintain properties that are useful for these professionals. For example, two lines that are parallel in the model will be parallel in the projection; similarly, the lengths of two lines that are at the same orientation will keep the same length relative to each other regardless of their position on the scene or the position of the point of view. Our analysis of Experiments 1 and 2 also shows how some of the advantages of the no-discrepancy parallel conditions, and the parallel medium-discrepancy condition in particular, come from the fact that parallel projections are more resilient to the placement of the object; that is, accuracy for local and distant objects is equivalent within each condition for parallel projections, whereas for

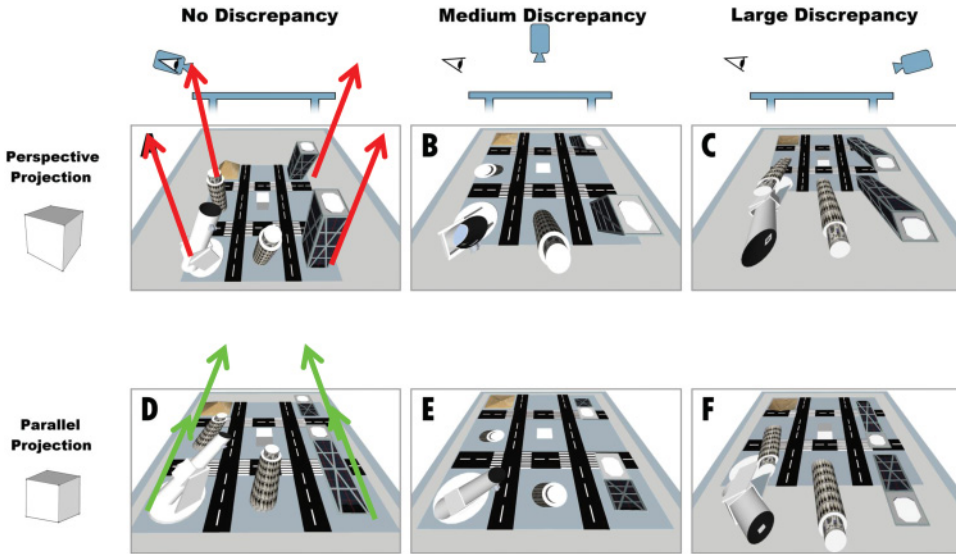


Fig. 44. The tilting-in effect: in a large display, a parallel no-discrepancy projection (D) gives the objects a transversal (left-to-right or vice versa) tilt when the objects are away from the central line of sight.

perspective projections, the location of the stimulus makes a difference in the error (see Figures 20 and 28).

Finally, the parallel projections did result in some small but noticeable distortion in the transversal direction for the orientation experiment (E_r , see Figure 21). Specifically, no-discrepancy parallel projections result in larger orientation estimation errors in the transversal direction. We believe this is due to the fact that a no-discrepancy parallel projection can only adapt the projection to one specific direction (all projection lines are parallel). When the display is large, as is the case in our setup, objects that are furthest from the central area of the display tend to look *tilted-in* (see Figure 44).

Notice, however, that this tilting-in effect did not affect orientation perception for the medium discrepancy condition, or the large-distortion projections. This lack of tilting-in effect in the projection from above is somewhat surprising and serves to further single out the parallel mid-discrepancy condition, which we discuss in more detail in Section 10.2.

11.1.3. Subsidiary Awareness of the Projection Surface. A plausible explanation for the effects described in Sections 7.4 and 8.4 is that it is easier to compare objects, and judge their internal structure and external orientation when the projection is identical, in terms of pixels, regardless of the location on the table. This might be particularly true for the pattern matching tasks: if the images that participants compare are identical pixel by pixel, the 3D comparison task might revert to a 2D image comparison task, where participants compare 2D shapes rather than 3D structures. However, although the shape of the stimuli might be identical in these cases, the different positions on the table force the projection on the retina of the viewer to be different. Figure 45 illustrates how, although the parallel table projections of the stimuli are identical regardless of location, the retinal projections are not, precisely due to the different location of the stimuli, which is exacerbated by the natural obliqueness of a tabletop display.

These results, therefore, indicate that the two types of visual differences between objects (in projection, and in retinal image) are treated differently by the brain. This

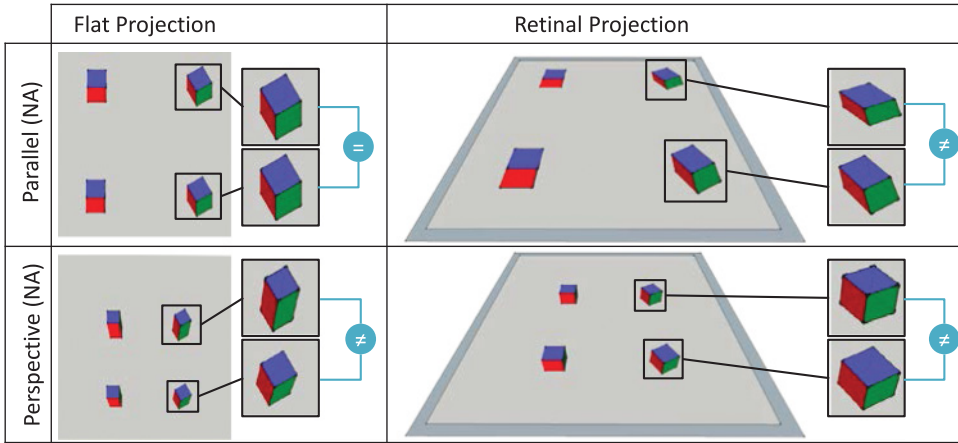


Fig. 45. Even when the objects' pixels are identical (the case for all parallel conditions), different locations on the table will make the retinal images different, due to the oblique position of the person and the effect of optics in the human eye. Note: Retinal projections here are simulated using rectilinear optics that are similar but not identical to the curve projections in the human eye due to the non-planar nature of the retina and the presence of a lens.

supports the hypothesis, first stated by Pirenne [1970], that we have a subsidiary awareness of the surface onto which a 3D representation is projected, which is corrected for by the brain before the 3D shape is interpreted from the planar representational cues (e.g., foreshortening and occlusion). In other words, the position and orientation of the surface that contains the image can be obtained from other visual cues such as vergence, accommodation, the existence of a frame, or binocular disparity; this indicates to our visual system that the content represented in that surface has to be interpreted as a picture or a photograph, rather than directly as the effect of an object floating in space. Then, the visual system corrects for the position and orientation of the surface in order to interpret the projected image, which results in a conflict of cues if the projection has been designed to function as if the image was not projected in a surface, as seen in fish-tank VR approaches. This mismatch in interpretation leads to the unintuitive result that more realistic projections (i.e., perspective) lead to more errors and a longer judgment time.

These findings seem to contradict some of the conclusions and hypotheses from previous literature, namely those that predict perception accuracy performance according to the geometry of the retinal projection (e.g., Cutting [1988]). Although our experiment was not explicitly designed with the purpose of solving this controversy from the perception literature, it does provide support for previous research, which found that the exclusive geometrical approach to perception does not account correctly for the perception data (e.g., Koenderink et al. [1993], Perkins [1973], Goldstein [1991], and Ellis et al. [1991]). Similarly, it provides support for more current accounts of perception, in which cue conflicts are posited to be the cause of degraded perception [Vishwanath 2014].

11.2. The Parallel From the Top Projection (PTP)

The results from our experiments show that the *frontoparallel* or *parallel from the top projection* (PTP; the parallel medium-discrepancy condition) is a special projection. Considered across all tasks, it resulted in the highest accuracy: in the internal structure estimation task, it was the projection with the fewest errors; in the orientation

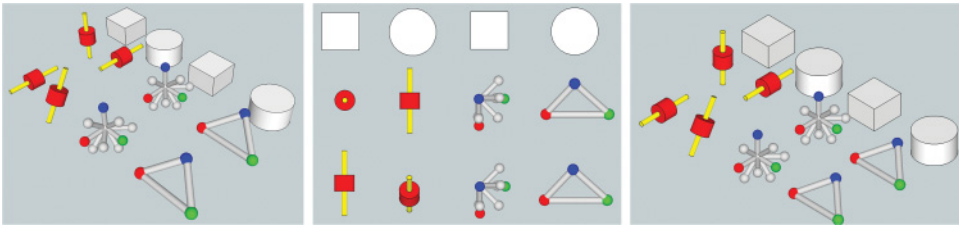


Fig. 46. A collection of experimental objects and regular objects displayed using a perspective (left), parallel from the top (center), and parallel at an angle (right) perspective geometries.

estimation task, it is equivalent to best in terms of average errors and likely responsible for the significant main effect of projection geometry;¹⁵ and in the object structure recognition experiments, it was consistently faster (and equivalent or better in number of errors to the other conditions). The causes of these differences can be found in the special geometry of this projection and the nature of our experiments.

The PTP is the most similar to a map or a satellite image taken from a perfectly perpendicular angle; in this projection, object sides that are perpendicular to the horizontal surface become lines and are, therefore, invisible. Cubes and cuboids that are placed horizontally or vertically are easily recognizable due to their simple projection, which looks like a regular 2D shape (see Figure 46). Although the DRE [Goldstein 1987] predicts that perceived orientation errors will be maximal for objects that point out of the screen, this is not the case for this condition; viewers recognize the simplicity of the shapes as a parallel from the top projection and judge accordingly, without any instruction.

Because the PTP is most accurate and viewer-independent (the image is the same regardless of the number and position of collaborators around a table), one might think that it should be the unqualified projection of choice. We believe that this will be the case in many situations; however, there are several caveats that complicate the picture and make this choice more difficult. We discuss the two main caveats in the following subsections.

11.2.1. Angle Dependence in Orientation Estimation. As we discussed above, the PTP makes it easy for people to recognize the external orientation of objects with parallel or perpendicular angles to the table; however, intermediate angles seem to have similar accuracy to its perspective equivalent, somewhere between the error rates of no-discrepancy and large discrepancy conditions (see Figure 19). Although this might not be relevant for applications where the 3D models are very regular, such as city models, where most faces are orthogonal or parallel to the ground plane, it might be problematic for models that are very irregular, such as a model of the human body. This dependence of accuracy on the angle being judged could create surprising anomalies such as non-monotonic estimations of the angle, where ordering of angular estimations are reversed when the angles are close to perpendicular.

It is fair to say that our external orientation estimation experiment is biased in favor of regular model situations—which are probably also the most common in real scenarios. Note, however, that the overall advantage of PTP due to straight angle regularity can only explain the results for the orientation estimation task; although

¹⁵We did not perform a statistical analysis of the comparison between the medium discrepancy parallax conditions due to the *a priori* restriction in the selection of pairwise comparisons; however, the differences in average accuracy seem evident from Figure 20.

this regularity did not exist in the other experimental tasks, since objects were always at an angle and without planar faces, we also observed advantages of PTP in these tasks.

11.2.2. Losing the Sides. In the discussion above, we highlight the important advantage for external orientation estimation of PTP in regular models with mostly orthogonal faces. Due to this regularity, however, a PTP projection also results in a projection that hides all information about object surfaces that are perpendicular to the table plane. If this information is important for the application (e.g., if it is important for an architectural application to see the façade of buildings), PTP will not be a good option. Moreover, PTP also results in a reduced set of 3D cues, which make it more difficult to perceive depth; PTP, unlike other projections, does not provide relative size or relative density cues, although it can provide occlusion, and perhaps, a virtual version of aerial perspective.¹⁶ Practitioners would have to consider if a projection like this, which might be closer to a map than to a 3D representation, will fulfill the requirements of the application.

11.3. Self-Generated Motion Parallax Vs. Object Orientation Manipulation

Time measures across the experiment indicate that when motion parallax is available, viewers try to make use of it by moving their heads, which unsurprisingly, results in longer trial times. However, the accuracy measures of our experiment could not detect any significant benefits of self-generated motion parallax, regardless of the extra information shown and the sometimes longer viewing times. This does not necessarily mean that the self-generated motion parallax is worthless, but it does provide a strong indication that, for the tasks that we tested, continuously tracking PoVs of viewers provides very little performance benefit, at least compared to the effect of changing the projection from perspective to parallel or statically determining the level of discrepancy. Naturally, it is possible that the value of self-generated motion parallax is relevant for other measures (e.g., sense of immersion, realistic perception), or in other tasks.

Interestingly, when participants were allowed to manipulate the local object through touch in the object structure recognition task, this resulted in large improvements in accuracy (see Figure 39). This is surprising since head movement and object orientation manipulation have a similar visual effect, and the results partially contradict the findings in Rogers and Graham [1979], Simons et al. [2002], which suggests that the value of the extra information gathered through the ventral system (head-position and orientation awareness) is not very important compared with the ability to effortlessly rotate the object through touch manipulation, at least in this task. One possible explanation for the difference here is that head position changes require much larger movements and effort than touch interaction to achieve equivalent changes; moreover, some views of the object might not be achievable with self-movement (e.g., from the bottom).

11.4. Different Projections for Different Tasks

One expected, although important, conclusion derived from our results is that the way in which different projections affect perceptual performance is heavily influenced by the task. This has important implications for the design of tabletop applications and systems that use 3D graphics; systems that are focused on different tasks will need to use different projections to avoid large perceptual errors. Our study provides

¹⁶We are not aware of any system that uses aerial perspective (the loss of air transparency due to light scattering from the air molecules) to simulate distance for parallel projections, but we do not see any reason why it could not contribute to depth perception.

measurements of accuracy and performance for each of the three tasks but, most importantly, it offers strong evidence that the origin of these differences is in the tasks themselves. Unlike most of previous research, we studied multiple tasks and kept most of the elements of the experiment design unchanged; this makes the comparison across tasks stronger, and validates some of the differences highlighted by separate pieces of previous research that used different tasks (e.g., the different findings from Sedgwick [1993], Goldstein [1987], and Hagen and Elliot [1976]).

11.5. Limitations and Future Work

As any empirical study, our experiments are necessarily limited in scope and range. To our knowledge, the four experiments described here present the most comprehensive study of the effect of representation types of 3D objects on tabletops to date; however, the covered tasks are only a subset of all possible types of tasks in this emerging field of applications, and other tasks, such as 3D object manipulation and 3D abstract data representation, will need to be investigated before we can understand how they are affected by discrepancy, projection, self-generated parallax, and interaction. Similarly, we investigated mostly performance related measures of the perception of 3D content on tabletops (accuracy, completion time); there are several other important aspects that are important for tabletop applications that need to be considered in the future. For example, it will be valuable to study how the different projections affect the sensation of immersion as experienced by viewers.

To avoid possible confounds and make sure that the power of the experiment was reasonable, we also made experimental design decisions that could interact with perception. For example, the figures used as stimuli for the four experiments are shaded, but do not project shadows. Similarly, we provided context for the projected figures in the form of cubic molecules (except for the structure recognition task, see Figure 31), but the backgrounds and surrounding areas of the objects to judge were mostly clear of other objects. Our results are valuable to understand the effects of discrepancy, projection, and self-generated parallax independently, but they should be further validated for denser content and alternative rendering techniques.

Our experiments are also limited in that they cover what happens when CoP and PoV coincide or are located across the longitudinal axis of a table (*longitudinal discrepancy*). In tabletops, there will be situations when CoP and PoV are not aligned with this axis and have therefore varying degrees of *oblique discrepancy* with respect to the table. Although further research is needed to assess what is the impact in those cases, related experiments (e.g., Wigdor et al. [2007] and our own experiences when using and preparing 3D representations on large tabletops indicate that large oblique discrepancies could be even more disruptive than the large longitudinal discrepancies that we tested).

Binocular rendering of 3D content for collaborative environments is currently economically and computationally expensive; moreover, binocular rendering does come with added problems due to conflicting depth cues (e.g., accommodation and binocular disparity will result in different estimations of depth [Vishwanath 2014]). Nevertheless, future development of display technologies and systems that enable binocular observation of 3D scenes is plausible. Our studies can serve as a template to study the interaction of discrepancy, projection type and self-generated motion parallax with binocular display systems.

We believe that most of our results could apply to non-horizontal surfaces. Although vertical surfaces, and everything in between, are not observed from oblique angles as often as tabletops, collaborative scenarios are now common, which result in a similar conflict between reducing discrepancy and providing images for all viewers. There

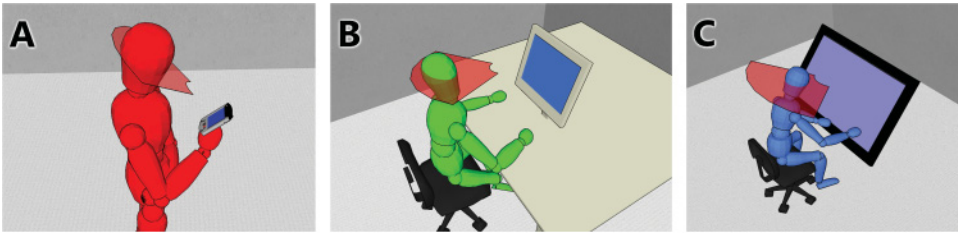


Fig. 47. Typical location of the point of view for personal displays (red transparent wedge) in three canonical situations: (A) mobile device, (B) monitor, and (C) large personal display.

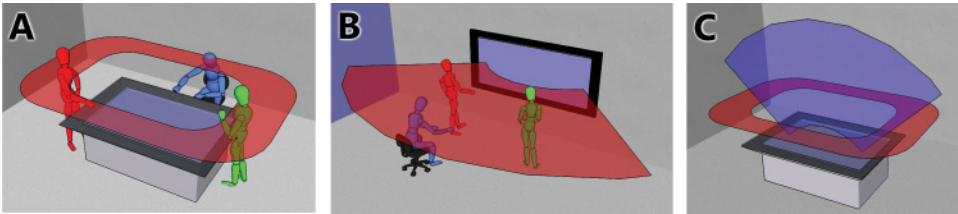


Fig. 48. Typical location of the point of view for large shared displays (red transparent wedge): (A) tabletop display, (B) vertical display, (C) comparison of the two when adjusting for orientation.

are, however, at least two notable differences between horizontal and vertical surface scenarios: binocular alignment and angle of approach.

11.5.1. Horizontal Vs. Vertical Surface Perception: Binocular Alignment. We know that the binocular disparity produced by the distance between the eyes is an important depth cue [Cutting 1997]. Horizontal surfaces, however, result in a type of binocular disparity that is different from the one produced by a vertical display because unlike in vertical display viewing, the display surface is parallel to the horizontal alignment of the eyes. Although our apparatus does not project binocular images, the projected images and the 3D objects that hold them are still subject to disparity, which in turn might result in different degrees of conflict between the 3D cues provided by the environment and the projected image (i.e., subsidiary awareness of the projection plane and the projected content). This issue requires further empirical research to determine if the orientation of a surface is a significant factor in how 3D information is perceived.

11.5.2. Angle of Approach. Another important difference between horizontal and vertical displays is the position and angle of approach to the display or, in other words, the typical PoV that observers take when using the different displays. Figure 47 represents three canonical display orientation types; the red semi-transparent shapes represent the planes close to which we can expect to find the point of view of people using the display.

Personal displays generally allow viewers to configure their own position or orientation, and manipulate the position and orientation of the device, especially for mobile devices and small monitors. This allows them to maximize the perpendicularity of their line of sight to the plane of the display. The likely PoVs in Figure 47 (A)–(C) are therefore quite similar to each other. On the larger displays of Figure 48, the variation of likely points of view is much larger due to their size and orientation. For example, on a large vertical display the variation is larger than on a monitor, but the position alternatives are mostly constrained to a rotation around the vertical axis—viewers will typically see the screen from different sides—east to west—but mostly at the same height. However, a horizontal display is typically seen from consistently off-axis PoVs,

and the relative position of PoV with respect to the display varies both east-west and north-south.

These differences in angle of approach will likely result in different effects of discrepancy and projection type for different display positions depending on the point of view. Although the likely positions of the PoV for a large vertical display are simpler and more similar to the better known personal and monitor displays, they will still require further research to validate the difference in 3D perception between different discrepancies, projections and points of view. To our knowledge, this research has not yet been conducted.

11.5.3. Multiple and Alternative Projections. Another area that has not yet received sufficient attention is the use of multiple different projections within the same display. This can be done by explicitly splitting different parts of the display into different views, or by modifying the projection of individual objects. Although there are obvious problems that can arise from the merging of multiple points of view within the same plane, these effects are largely unstudied, and there might be situations that can still benefit from their use. The same question extends to situations in which visual projections are distributed across displays (c.f., Karnik et al. [2013]).

Additionally, it might be worth considering setups with centers of projection that do not correspond to the viewer's point of view or to an infinite distantly located point of view (parallel projection), but instead use an intermediate location. This approach has some empirical support; Hagen and Elliot [1976] found that participants judged projections at an intermediate distance as "more natural."

Finally, there are non-linear ways of projecting 3D objects that could have additional benefits, such as the ability to project wider fields of view while preserving other properties [Brosz et al. 2007]. Although we can expect these to produce even more inconsistent effects on perception, they might be appropriate for certain tasks.

11.6. Recommendations for Practitioners

The findings of our studies paint an interesting and somewhat unexpected picture of the effects of different ways of representing 3D objects on tabletops. Recommendations from comprehensive empirical evaluations such as ours are rarely straightforward or universal; we cannot recommend a single way to project information on the table in every task and every situation. However, the results do provide valuable information for the design of tabletop applications with 3D content for specific tasks.

If vertical information is not important, choose a parallel from the top projection. The parallel-from-the-top projection (parallel medium discrepancy) is a special case that increases accuracy and reduces time for most of the tasks we tested; moreover, this type of projection affects all viewers equally and so it is independent from the PoV, and does not require viewer tracking and so is cheaper to implement. In exchange, this projection significantly reduces the information displayed in vertical planes. Applications that have irregular models, or where information of vertical planes is not crucial can benefit from this approach.

If vertical information is important, reduce PoV-CoP discrepancy. As discussed above, eliminating discrepancy does not guarantee the most accurate or efficient perception of the content. However, in some situations, a parallel from the top projection might be out of the question due to the necessity of representing vertical information, including distances in the z -axis. In this case, it is preferable to adapt the CoP to the PoV to each viewer than to provide a random location point, which could result in the impaired accuracy of the LA and LP conditions of our study. Providing low discrepancy is easy enough for a single-user display, especially if the PoV does not change significantly

throughout the use session (see Figure 47; in these cases, it is enough to calibrate the position at the beginning, otherwise, real-time tracking with devices such as a Kinect would be required). However, multiple users in displays such as those of Figure 48 would require presenting a different image to each viewer, which is technically harder—although not impossible—and significantly more expensive.

Do not go out of your way to provide self-generated motion parallax. Since we could not find any significant beneficial effect of self-generated motion parallax, and it has a cost in time in some cases, we suggest that, unless there are other reasons to include it (see Section 10.3), real-time self-generated parallax can be safely avoided for the tested tasks.

Enable direct manipulation of object orientation if object recognition is important. Unlike self-generated motion parallax, the manipulation of object orientation through touch does increase significantly the ability of people to recognize these objects. This happens at no significant cost in time for most conditions. Even if the objects in the application are static, it is worth considering an interaction technique that would allow rotation of the object temporarily while a viewer is trying to recognize the object.

If objects have to appear in different locations of a large display, use parallel projections to achieve a consistent performance regardless of location.

One of the strong advantages of parallel projections seems to be that they are robust with respect to the location where an object is being displayed. If the display is large, and it is important to not give advantage to central locations in the display with respect to other locations, then perspective neutral projections should be avoided.

Consider the task. Different tasks will require different types of perceptual estimations, and these are affected differently by projection. When choosing a projection, it is important to analyze which perceptual information will be most relevant for the supported activity.

12. CONCLUSION

Displaying 3D objects on tables presents several problems in the way that the 3D virtual scene is presented on the 2D surface, and as a result, different choices in the way the projection is designed can lead to distorted images and can alter the interpretation of represented objects. However, little is known about how projection geometry affects perception on large horizontal displays, and about which projection is best in which circumstances.

To investigate these problems, we studied people's ability to perceive and recognize 3D objects under different projection conditions. We carried out four experiments to test perception of the structure of visual objects, perception of orientation of objects, and ability to recognize objects with and without motion parallax. We systematically varied three variables of projection geometry: the type of projection (perspective vs. parallel), the separation between the point of view of the observer and the center of projection used for rendering (discrepancy), and the presence of motion parallax (with and without parallax).

We found strong effects of projection geometry, but that these effects depend on the perceptual task. Overall, our studies show that optimizing the perspective for one person at a shared table is likely to cause major problems for the others in the group—especially in tasks that involve estimation of object orientation—but that providing a neutral projection above the table and using parallel projection geometry can mitigate the problems.

We contribute a broad analysis of how the choice of projection will fundamentally affect perception of 3D content in tabletop displays. The results have far-reaching

implications for the design of 3D views on tables, in particular, for multi-user applications where projections that appear correct for one person will not be perceived equally by another, and can therefore create communication and coordination conflicts.

ELECTRONIC APPENDIX

The electronic appendix for this article can be accessed in the ACM Digital Library.

STATEMENT OF PREVIOUS RESEARCH

A small part of the results of this article were presented at ITS 2009 (citation below). The ITS paper presented a single study; the present article is a much-expanded version that incorporates three additional studies and substantial additional analysis and discussion. The ITS paper was an initial look at the problem; the present article is a comprehensive treatment.

Mark Hancock, Miguel Nacenta, Carl Gutwin, and Sheelagh Carpendale. 2009. The effects of changing projection geometry on the interpretation of 3D orientation on tabletops. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'09)*. ACM, New York, NY, 157–164. DOI=<http://dx.doi.org/10.1145/1731903.1731934>

ACKNOWLEDGMENTS

The authors would like to thank the participants that took part in our studies, and the members of the Interaction lab at the University of Saskatchewan, the Interactions lab at the University of Calgary, and the Innovis group at the University of Calgary.

REFERENCES

- M. Agrawala, A. C. Beers, I. McDowall, B. Fröhlich, M. Bolas, and P. Hanrahan. 1997. The two-user responsive workbench: support for collaboration through individual views of a shared space. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH97)*. ACM Press/Addison-Wesley Publishing Co., 327–332. DOI: 10.1145/258734.258875
- R. Arsenault and C. Ware. 2004. The importance of stereo and eye-coupled perspective for eye-hand coordination in fish tank VR. *Presence-Teleop. Virt.* 13, 5, 549–559. DOI: 10.1162/1054746042545300
- K. W. Arthur, K. S. Booth, and C. Ware. 1993. Evaluating 3D task performance for fish tank virtual worlds. *ACM Trans. Inf. Syst.* 11, 3, 239–265. DOI: 10.1145/159161.155359
- C. Bichlmeier, S. M. Heining, L. Omary, P. Stefan, B. Ockert, E. Euler, and N. Navab. 2009. MeTaTop: a multi-sensory and multi-user interface for collaborative analysis. Video. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'09)*. ACM. DOI: 10.1145/1731903.1731950
- I. Biederman and P. C. Gerhardstein. 1993. Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *J. Exp. Psychol. Hum. Percept. Perform.* 19, 6, 1162–1182. DOI: 10.1037/0096-1523.19.6.1162
- G. M. Bodner and R. B. Guay. 1997. The Purdue visualization of rotations test. In *The Chemical Educator*. Springer-Verlag, 1–17. DOI: 10.1007/s00897970138a
- J. Brosz, F. F. Samavati, S. Carpendale, and M. Costa Sousa. 2007. Single camera flexible projection. In *Proceedings of the 5th International Symposium on Non-photorealistic Animation and Rendering (NPAR'07)*. ACM, 33–42. DOI: 10.1145/1274871.1274876
- I. Carlbom and J. Paciorek. 1978. Planar geometric projections and viewing transformations. In *ACM Comput. Surv.* 10, 4, 465–502. DOI: 10.1145/356744.356750
- J. E. Cutting. 1987. Rigidity in cinema seen from the front row, side aisle. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 3, 323–334. DOI: 10.1037/0096-1523.13.3.323
- J. E. Cutting. 1988. Affine distortions of pictorial space: Some predictions for Goldstein (1987) that La gournerie (1859) might have made. *J. Exp. Psychol.* 14, 2, 305–311.
- J. E. Cutting. 1997. How the eye measures reality and virtual reality. *Behav. Res. Methods* 29, 1, 27–36.
- J. E. Cutting and P. M. Vishton. 1995. Perceiving layout and knowing distances: the integration, relative potency, and contextual use of different information about depth. *Percept. Space Motion* 5, 69–117.

- G. de Haan, R. Molenaar, M. Koutek, and F. H. Post. 2007. Consistent viewing and interaction for multiple users in projection-based VR systems. *Comput. Graph. Forum* 26, 695–704.
- J. de la Gournerie. 1859. *Traité de perspective linéaire contenant les tracés pour les tableaux, plans & courbes, les bas reliefs & les décorations théâtrales, avec une théorie des effets de perspective*. Dalmont et Dunod.
- V. A. Diwadkar and T. P. McNamara. 1997. Viewpoint dependence in scene recognition. *Psychol. Sci.* 8, 4, 302–307. DOI:10.1111/j.1467-9280.1997.tb00442.x
- S. R. Ellis, S. Smith, A. Grunwald, and M. W. McGreevy. 1991. Direction judgement error in computer generated displays and actual scenes. In *Pictorial Communication in Virtual and Real Environments* (2nd ed.). Taylor & Francis, Inc, 504–526.
- J. Geng. 2013. Three-dimensional display technologies. *Adv. Opt. Photon.* 5, 456–535. DOI:10.1364/AOP.5.000456
- E. B. Goldstein. 1987. Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *J. Exp. Psychol. Hum. Percept. Perform.* 13, 2, 256–266.
- E. B. Goldstein. 1991. Perceived orientation, spatial layout and the geometry of pictures. In *Pictorial Communication in Virtual and Real Environments*. Taylor & Francis, 480–485.
- T. Grossman and D. Wigdor. 2007. Going deeper: a taxonomy of 3D on the tabletop. In *Proceedings of the 2nd Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*. IEEE, 137–144. DOI:10.1109/TABLETOP.2007.18
- M. A. Hagen and H. B. Elliott. 1976. An investigation of the relationship between viewing condition and preference for true and modified linear perspective and adults. *J. Exp. Psychol. Hum. Percept. Perform.* 2, 4, 479.
- M. Hancock, S. Carpendale, and A. Cockburn. 2007. Shallow-depth 3D interaction: design and evaluation of one-, two- and three-touch techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*. ACM, 1147–1156. DOI:10.1145/1240624.1240798
- M. Hancock, T. Ten Cate, S. Carpendale, and T. Isenberg. 2010. Supporting sandtray therapy on an interactive tabletop. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems (CHI'10)*. ACM, 2133–2142. DOI:10.1145/1753326.1753651
- T. Höllerer, S. Feiner, T. Terauchi, G. Rashid, and D. Hallaway. 1999. Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system. *Comput. Graph.* 23, 6 (Dec. 1999), 779–785. DOI:10.1016/S0097-8493(99)00103-X
- S. Holm. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian J. Statistics* 6, 2, 65–70.
- K. F. T. Holografika. 2011. Holovizio Displays. Retrieved Jul. 19, 2011 from <http://www.holografika.com/>.
- R. K. Jones and M. A. Hagen. 1978. The perceptual constraints on choosing a pictorial station point. *Leonardo* 11, 3, 191–196. DOI:10.2307/1574141
- A. Karnik, T. Grossman, and S. Subramanian. 2013. Comparison of user performance in mixed 2D-3D multi-display environments. In *Proceedings of INTERACT 2013. Lecture Notes in Computer Science 8117*. Springer, Berlin, 260–277. DOI:10.1007/978-3-642-40483-2_18.
- A. Karnik, D. Martinez Plasencia, W. Mayol-Cuevas, and S. Subramanian. 2012. PiVOT: Personalized view-overlays for tabletops. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST'12)*. ACM, New York, NY, 271–280. DOI:10.1145/2380116.2380151
- Y. Kitamura, T. Konishi, S. Yamamoto, and F. Kishino. 2001. Interactive stereoscopic display for three or more users. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'01)*. ACM, 231–240. DOI:10.1145/383259.383285
- Y. Kitamura, T. Nakashima, K. Tanaka, and T. Johkoh. 2007. The illusionhole for medical applications. In *IEEE Virtual Reality Conference (VR'07)*. IEEE, 231–234. DOI:10.1109/VR.2007.352488
- K. Kiyokawa, M. Billingham, B. Campbell, and E. Woods. 2003. An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR'03)*. IEEE Computer Society, Washington, DC, 133–141.
- J. J. Koenderink, A. J. van Doorn, and A. M. L. Kappers. 1994. On so-called paradoxical monocular stereoscopy. *Perception* 23, 583–594.
- J. J. Koenderink, A. J. van Doorn, A. M. L. Kappers, and J. T. Todd. 2004. Pointing out of the picture. *Perception* 33, 513–530.
- W. Kruger, C. A. Bohn, B. Frohlich, H. Schuth, W. Strauss, and G. Wesche. 1995. The responsive workbench: A virtual work environment. *Computer* 28, 7, 42–48. DOI:10.1109/2.391040
- A. Kulik, A. Kunert, S. Beck, R. Reichel, R. Blach, A. Zink, and B. Froehlich. 2011. C1x6: A stereoscopic six-user display for co-located collaboration in shared virtual environments. In *Proceedings of the 2011*

- SIGGRAPH Asia Conference (SA'11)*. ACM, New York, NY, Article 188, 12 pages. DOI: 10.1145/2024156.2024222
- M. Matsushita, M. Iida, T. Ohguro, Y. Shirai, Y. Kakehi, and T. Naemura. 2004. Lumisight table: a face-to-face collaboration support system that optimizes direction of projected information to each stakeholder. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (CSCW'04)*. ACM, New York, NY, 274–283. DOI: 10.1145/1031607.1031653
- M. Mauderer, S. Conte, M. A. Nacenta, and D. Vishwanath. 2014. Depth perception with gaze-contingent depth of field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'14)*. ACM, New York, NY, 217–226. DOI: 10.1145/2556288.2557089
- M. Möllers, P. Zimmer, and J. Borchers. 2012. Direct manipulation and the third dimension: Co-planar dragging on 3D displays. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS'12)*. ACM.
- M. A. Nacenta, S. Sakurai, T. Yamaguchi, Y. Miki, Y. Itoh, Y. Kitamura, S. Subramanian, and C. Gutwin. 2007. E-conic: a perspective-aware interface for multi-display environments. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology (UIST'07)*. ACM, New York, NY, 279–288. doi:10.1145/1294211.1294260
- D. N. Perkins. 1973. Compensating for distortion in viewing pictures obliquely. *Percept. Psychophys.* 14, 13–18.
- M. H. Pirenne. 1970. *Optics, Painting & Photography*. Cambridge University Press.
- B. Rogers and M. Graham. 1979. Motion parallax as an independent cue for depth perception. *Perception* 8, 2, 125–134. DOI: 10.1068/p080125
- S. Sakurai, Y. Kitamura, S. Subramanian, and F. Kishino. 2008. Visibility control using revolving polarizer. In *Proceedings of the 3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems (TABLETOP'08)*. 161–168. DOI: 10.1109/TABLETOP.2008.4660200
- J. A. Saunders and B. T. Backus. 2006. The accuracy and reliability of perceived depth from linear perspective as a function of image size. *J. Vis.* 6, 9, 933–954. DOI: 10.1167/6.9.7
- S. D. Scott, K. D. Grant, and R. L. Mandryk. 2003. System guidelines for co-located, collaborative work on a tabletop display. In *Proceedings of the 8th European Conference on Computer Supported Cooperative Work (ECSCW'03)*. Kluwer Academic Publishers, 159–178.
- H. A. Sedgwick. 1993. The effects of viewpoint on the virtual space of pictures. In *Pictorial Communication in Virtual and Real Environments* (2nd ed.). Taylor & Francis, Inc., 460–479.
- A. Simon and S. Scholz. 2005. Multi-viewpoint images for multi-user interaction. In *Proceedings of IEEE Virtual Reality*. IEEE, 107–113. DOI: 10.1109/VR.2005.1492760
- D. J. Simons, R. F. Wang, and D. Roddenberry. 2002. Object recognition is mediated by extraretinal information. *Percept. Psychophys.* 64, 4, 521–530. DOI: 10.3758/BF03194723
- F. Steinicke, G. Bruder, and S. Kuhl. 2011. Realistic perspective projections for virtual objects and environments. *ACM Trans. Graph.* 30, 5, Article 112. DOI: 10.1145/2019627.2019631
- H. J. Sun, G. S. W. Chan, and J. L. Campos. 2004. Active navigation and orientation-free spatial representations. *Memory Cognit.* 32, 1, 51–71. DOI: 10.3758/BF03195820
- D. Valkov, A. Giesler, and K. Hinrichs. 2012. Evaluation of depth perception for touch interaction with stereoscopic rendered objects. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS'12)*. ACM, 21–30. DOI: 10.1145/2396636.2396640.
- D. Vishwanath, A. R. Girshick, and M. S. Banks. 2005. Why pictures look right when viewed from the wrong place. *Nat. Neurosci.* 8, 10, 1401–1410. DOI: 10.1038/nn1553
- D. Vishwanath. 2014. Toward a new theory of stereopsis. *Psychol. Rev.* 121, 2, 151–178. DOI: 10.1037/a0035233
- U. von Zadow, S. Buron, K. Sostmann, and R. Dachselt. 2013. The simmed experience: Medical education on interactive tabletops. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces (ITS'13)*. ACM, New York, NY, 297–300. DOI: 10.1145/2512349.2514587
- C. Ware, K. Arthur, and K. S. Booth. 1993. Fish tank virtual reality. In *Proceedings of the INTER-ACT'93 and CHI'93 Conference on HUMAN Factors in Computing Systems (CHI'93)*. ACM, 37–42. DOI: 10.1145/169059.169066
- M. Wexler, F. Panerai, I. Lamouret, and J. Droulez. 2001. Self-motion and the perception of stationary objects. *Nature* 409, 6816, 85–88.
- D. Wigdor, C. Shen, C. Forlines, and R. Balakrishnan. 2007. Perception of elementary graphical elements in tabletop and multi-surface environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'07)*. ACM, 473–482. DOI: 10.1145/1240624.1240701

- A. D. Wilson, S. Izadi, O. Hilliges, A. Garcia-Mendoza, and D. Kirk. 2008. Bringing physics to the surface. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST'08)*. ACM, 67–76. DOI: 10.1145/1449715.1449728
- J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*. ACM, New York, NY, 143–146. Retrieved from <http://doi.acm.org/10.1145/1978942.1978963>.

Received November 2014; revised October 2015; accepted November 2015