University of Twente

EXPLORATION OF NOVEL APPROACHES TO CONVEY 3D EFFECTS IN AUTOMOTIVE USER INTERFACES

Master Thesis

by

Magdalena Szostak

Department of Learning, Data-Analytics, Technology

Faculty of Behavioural, Management and Social Sciences

For the degree of:

Master of Science, Psychology

04.04.2022

Supervisors:

dr. Simone Borsci
dr. Martin Schmettow

External supervisor:
Andrea Narciso, ART S.p.A.

This research has been carried out in collaboration with ART S.p.A. - Italian industry leader of in-car electronic systems.

Abstract

Over a span of the last ten years, the automotive industry has seen a sudden rise of in-vehicle infotainment systems (IVIs). Staying at the forefront of this emerging market requires innovative solutions as the novelty of IVIs is believed to be a source of a vehicle's perceived luxuriousness. One of such advancements is the introduction of 3D effects to the in-vehicle experience. However, the commonly used 3D technologies have multiple shortcomings for in-car use, which begs an exploration of novel approaches to conveying 3D effects in automotive user interfaces. In this context, the camera-based solutions provide opportunities to create novel 3D effects in 2D digital instrument clusters by manipulating the 3D scenes to match the driver's point of view. Considering the novelty of such a safety-critical solution, a key goal of this research was to provide design guidelines and user requirements regarding the implementation of a dynamically simulated 3D instrument cluster for car displays.

A systematic literature review was conducted on the monocular depth cues and their potential for depth recreation on a flat surface. Relevant depth cues were identified and their resulting combination, as well as the remaining questions were addressed in a follow-up experiment. For this purpose, the cues were tested on both depth-related performance and the subjective depth impression. Moreover, interviews were conducted to gather insights on the visual preferences and user acceptance towards the in-car use of the 3D effect.

Overall, the obtained results provide strong support for the observer-produced motion parallax and its effectiveness in creating a simulated 3D effect. Additionally, the 3D effect was received positively and with a high level of acceptance towards the implementation in the automotive context. However, multiple user requirements were mentioned by the interviewees. The key insights gathered from the review and the experiment were therefore used to create practical guidelines for both design and implementation of a 3D digital instrument cluster. Based on these guidelines, several designs have been developed to illustrate their optimal use. Lastly, the remaining questions were outlined and were followed by recommendations for future research directions.

Keywords: Monocular depth cues, depth perception, digital instrument cluster, 2D, 3D, flat display, automotive user interfaces, head tracking, driver monitoring system

Contents

Introduction

It was the first acts of cave painting that have manifested our ability to create flat pictorial representations of our three-dimensional reality. Yet, it was not until the fifteenth century that special interest was given to a realistic representation of depth on a flat surface. Inventions such as linear and aerial perspective have revolutionized the art of painting and have brought us a great step closer to capturing depth on a flat surface (Brooks, 2017). Nowadays, that surface has transformed and diversified its nature including a multitude of digital displays. Furthermore, with technological innovation came new possibilities and the strides to achieve a realistic depth representation have been moved away from a flat-screen into the realm of stereoscopic devices and virtual reality (VR). These solutions are slowly finding their way into the automotive industry giving rise to novel in-car electronic systems, such as head-up displays (Lauber, 2014). They have also been successfully introduced in multiple other fields, ranging from VR gaming headsets (Kongsilp & Dailey, 2017) and 3D movies (Emoto, 2019) to surgical stereoscopic screens, 3D geo-visualizations (Seipel, 2013), or perspective air-traffic displays (Mulligan, 2009). It seems therefore, that the added value of three-dimensional displays lies in their ability to increase both entertainment and the quality and accuracy of information presented to an audience.

These two aspects, information and entertainment, are in turn closely related to what is known as in-vehicle infotainment systems (IVIs) where information and entertainment services are seamlessly integrated to form an embedded platform (Sen & Sener, 2020). The sudden rise of IVI systems in the last ten years has transformed the nature of a car, changing it into an electronic product and leading to a convergence of the automotive and information technology (IT) fields (Berger et al., 2019). Staying at the forefront of this emerging market requires not only constant improvements of the in-car electronic systems but most importantly demands new, innovative solutions (Bolder et al., 2018). Introducing novelty results in a competitive advantage, which is even more important in the context of luxury cars (Sen & Sener, 2020), where the uniqueness of provided options plays a vital role in the vehicle's perceived "premiumness" (Law & Evans, 2007). According to Sen & Sener (2020) novel experiences sparked by state-of-the-art technological advancements to the IVIs, are the main source of the system's perceived luxuriousness.

The introduction of such novelties has nowadays become more easily attainable as a result of the digitalization of IVIs. In this area, one of the most substantial advancements is the replacement of traditional analogue instrument clusters with digital ones, which has brought about immense freedom in terms of displaying information, which is facilitated by adjustable graphics and easily updatable contents (Broy et al., 2014b; Masola et al., 2020). Physical clusters are being transformed from a specified and limited set of physical gauges and signs into a digital screen that can display information of any kind, form and quantity. However, such richness of possible design options and the amount of content necessitates careful adaptation while creating in-vehicle products that are to best serve the driver while performing a task that is already highly demanding (Ostendorp et al., 2016). Additionally, introducing novel solutions requires careful considerations of not only the driver's cognitive ability but also of their already formed mental models. Such concerns sit at the heart of the field of automotive user experience (UX) design, where a fine line for balancing novelty with expectations necessitates a very thorough examination before any features can be introduced to the market.

One of such advancements – where innovation has to be carefully adapted to meet drivers' requirements – is the introduction of 3D effects to the in-vehicle user experience. The most commonly used technology is the augmented reality head-up displays (HUD) (Broy et al., 2014, Xie et al.,2018). HUDs typically present information, such as current speed, by displaying it directly on the windshield, hence in the direct visual field of the driver. This carries the benefit of rapid, safe and effortless information transfer (Broy et al., 2014b). However, a typical HUD will display this information in a static way and on a small, specified area. To enable more dynamic HUDs, a 3D augmented-reality effect can be added, by making use of for instance head-tracking or eye-tracking techniques. This causes the position of the visual elements to dynamically adapt to the driver's viewpoint (Rao et al., 2014), provides a richer user experience by connecting the elements with the driving environment, and allows for presenting multiple information in different depth layers (Broy et al., 2014a). Another approach to countering the static nature of HUDs, is to use head-mounted displays (HMDs) in a form of see-through glasses which make it possible to see the information continuously, regardless of the viewing direction (Lauber et al., 2014). However, such wearable equipment has raised concerns when it comes to drivers' comfort and safety, as they obscure and darken the driver's view of the environment (Broy et al., 2014a). Regardless of the type of the HUDs, their negative effects on performance have been observed in the field of aviation, where the attention of pilots was continuously drifting towards the content of the HUD leading to a decreased attention to outside stimuli (Lauber et al., 2014).

So far, the 3D technology that least obstructs the drivers' view seems to be offered by the advancements in the area of stereoscopic displays, which produce a 3D effect through the use of binocular disparity - sending two slightly offset images to each eye (Li et al., 2013). The regular stereoscopic displays, which also require special glasses to achieve the 3D effect (Broy et al., 2012) are nowadays being replaced by the so-called autostereoscopic screens. With autostereoscopic displays, no glasses are needed to produce a 3D effect, which is in turn created by an advanced technical display design. The most commonly implemented techniques direct two different images to each eye, either by a thin layer with a set of the so-called lenticular lenses, which refract the image (Algorri et al., 2016), or by the use of parallax barriers which occlude parts of the image at different angles (Kakeya et al., 2018;

Lanman et al., 2010). Such a complex mechanism, provides a 3D experience with no need for additional wearable equipment, which creates safer driving conditions and increases comfort.

A successful implementation of a 3D experience can be highly beneficial in the context of automotive interfaces, both when it comes to entertaining and presenting the information. For instance, 3D displays have been shown to improve UX when compared with 2D screens (Broy et al., 2014a), are believed to increase the visual attractiveness of the presented information (Schild et al., 2012) and enjoyment (Broy et al., 2012), as well as decrease visual search times (Huhtala et al., 2011). Moreover, benefits in a form of enhanced depth judgements, higher saliency of prioritized stimuli and better navigation performance have been widely observed (McIntire et al., 2012). These advantages, together with the higher availability of autostereoscopic screens, have sparked the interest of the automotive industry in creating 3D digital instrument clusters (Broy et al., 2015a; Masola et al., 2020;). When compared to a 2D version of the same instrument cluster, an autostereoscopic display has been assessed as more attractive, as well as more usable, as it provided well-structured information grouped at different depths, which enhanced information processing of essential elements (Broy et al., 2014b).

Despite such a wide array of benefits, autostereoscopic screens carry multiple limitations. In fact these types of stereoscopic displays have been reported to cause discomfort in a form of motion sickness (Broy et al., 2015a; Hwang & Peli, 2014) and high levels of eye fatigue due to accommodation-convergence mismatch (Broy et al., 2014a; Wang et al., 2015). Consequently, autostereoscopic screens can increase cognitive workload and result in a decreased driving performance (Broy et al., 2015a). Even though the advantages of 3D for UX and spatial judgements appear highly beneficial in the context of instrument cluster design, the autostereoscopic display is still not an optimal solution. Additionally, autostereoscopic displays are still highly expensive to produce and implement when compared to the more commonplace flat displays. This begs an exploration of novel approaches to conveying 3D effects in the context of automotive user interfaces.

The technology described so far, makes use of the so-called binocular depth cues while producing the 3D effect. The binocular cues allow for depth estimation in a three-dimensional environment (Hendrix & Barfield, 1995) and, as the name suggests, are perceived by both eyes. They include the already mentioned binocular disparity, as well as accommodation and convergence (Li et al., 2013). All of these are physiological processes connected to how visual sensory information is received and processed and are caused by the fact that both eyes register images at a slightly different angle (Emoto, 2019). For the binocular depth cues to work in a simulated environment, two slightly offset images have to be presented separately to each eye and, as discussed above, this is the fundamental mechanism behind the stereoscopic devices. However, the binocular cues are not the only ones that aid our depth perception.

Monocular depth cues are another source of depth judgment and, as opposed to the binocular cues, they can be used to recreate depth on a 2D surface (Brooks, 2017). They include a multitude of effects, such as linear and aerial perspective, shading, occlusion, color, size or motion parallax. Many studies have revealed that a well combined set of strong monocular depth cues can produce 3D experience and depth judgement performance comparable to the one achieved with binocular cues (Emoto, 2019; McIntire et al,. 2012; van Schooten et al., 2010; Seipel, 2013; van Beurden et al., 2010). For instance, van Schooten et al. (2010) have found that in the presence of a strong motion cue, stereoscopy had no added value. Moreover, van Beurden et al. (2010) report that when stereoscopic disparity was added to the motion based cues, it negatively impacted completion time. Despite such results and due to the strong focus on stereoscopic solutions, the monocular cues and the extent to which they can be used to create believable 3D effects on a flat display, have not been sufficiently explored. Considering the disadvantages of stereoscopic technology, this study aims to take a step back and return to a flat surface of a 2D screen, in order to revisit its potential in conveying the impression of depth in the context of automotive user interfaces.

The rationale of the study

This research has been carried out in collaboration with ART S.p.A. – an Italian company leader of in-car electronic systems for the luxury automotive market. The motivation behind this study has been sparked by the already mentioned digitalization of instrument clusters as well as the popularity of 3D solutions and the increased user experience evoked by the novelty of 3D. As reported by ART S.p.A., this prospect is additionally revived by the rise of driver monitoring systems, the introduction of which opens up new doors to system personalization and UX enhancement. In this context, the camera-based solutions provide opportunities to create novel 3D effects in 2D digital instrument clusters. This can be achieved by implementing head tracking to manipulate the graphical elements accordingly to match the driver's point of view. The real-time head position information could serve to simulate a matching transformation of objects inside of a simulated 3D scene, in a way that creates an illusion of actual depth.

To what extent a simulated 3D effect is feasible when using 2D displays and how well 3D simulated information is received by the users are the main questions that drive this exploration. Additionally, since this is a novel and at the same time safety-critical technology, guidelines for the design of such a display have to be carefully created. Therefore, a key goal of this research is to provide design guidelines and user requirements regarding the implementation of a dynamically simulated 3D instrument cluster for car displays.

To arrive at such guidelines, it is firstly essential to understand how the monocular depth cues influence depth perception and which are the most relevant cues in the context of a 3D digital instrument cluster. To achieve that, a systematic literature review will be conducted. The review findings will then be used to identify the most appropriate cues and to provide initial design directions. Based on that, an experiment with a 3D simulated display will be designed to test the combination of the cues derived from the review and to address the potential knowledge gaps. Finally, the results of both will inform the final design guidelines and several potential designs will be created to provide visual examples. Next to that, the remaining uncertainties will be identified and future research directions will be outlined. Therefore, the following paper has been divided into three main sections: (1) a systematic review of monocular depth cues, (2) an experiment testing the display technology and selected cues, and (3) final design guidelines with visual examples followed by future research directions.

A systematic review of monocular depth cues

A systematic literature review was conducted to better understand the influence of monocular depth cues on the design of a 3D digital instrument cluster. The main objective was to identify the most relevant monocular depth cues by classifying them based on their reported influence on depth-related performance and depth impression. While depth performance refers to the accuracy, correctness and speed of assessing spatial relationships, depth impression signifies the subjective believability of the achieved depth. Both of these aspects are important in the context of this research, as accurate and quick depth judgement should be properly balanced with the believability of the achieved 3D effect. Moreover, since most current reviews provide insight on individual cues, and since the design is more often a combination of different visual elements, additional focus was placed on the interaction effects between multiple depth cues. A secondary goal was to better understand the degree to which the monocular cues are effective in creating a realistic depth impression while using actual devices, with a particular focus on existing 3D dashboard solutions and head-tracking based transformation. The review has been conducted in accordance with The Preferred

Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) and its most current guidelines (PRISMA 2020; Page et al., 2020).

Method

Search strategy

The literature search was performed primarily using the EBSCOhost platform. Additionally, the ACM Digital Library was included, as it provides access to the most current technology-related papers. The search was performed from 29th of July till 17th of August 2021 and was limited to records published in English. While using the EBSCOhost platform, the search was narrowed down to psychology and technology-related databases (see Appendix A). For concepts related to depth perception no publication date was set as a requirement for retrieval, but years 1990-2021 were specified for technology-related concepts, such as head-tracking and 3D displays. The main search terms included *monocular depth cues*, *stereoscopic* and *monoscopic displays* as well as *3D instrument cluster*. The full sequence of keywords and the number of yielded works can be found in Appendix B.

Selection Process

The first stage of the selection process consisted of screening the identified records based on their titles and abstracts. This was performed by one reviewer and was followed by a full-text eligibility screening. Foxit Reader and Microsoft Excel were used to facilitate this process. The selection criteria for depth perception related papers included topics that clearly combined monocular depth cues with depth judgement, either when it comes to overall depth impression or depth estimation. When it comes to technology-related topics, papers were included only when combining the information on 3D displays with monocular depth cues. These papers focused mainly on the implementation of monocular depth cues in stereoscopic devices, or where monoscopic cues were used for depth enhancement procedures, such as 2D-to-3D video conversion. As for the exclusion criteria, sources that focused purely on depth perception in the context of the underlying neurological or physiological processes, as well as papers that focused only on technical approaches to producing 3D technology, without including any depth perception context, were excluded from the analysis. Finally, papers assessed as eligible were then divided into two main groups: (1) depth perception related and (2) technology-focused.

Data Collection Process

The study characteristics extracted from all the papers consisted of the authors and the year of publication. This was done by using Mendeley Reference Manager (version 2.59.0). ATLAS.ti (version 9.1.6.0) was used to facilitate collecting relevant data from the gathered literature. Inductive, data-driven coding was employed to extract relevant information. Details were extracted regarding the definitions of monocular depth cues, both as a group and as individual cues. Next to the definitions, the reported effects of individual cues as well as the interaction effects of their combinations were collected. This was done considering both the overall depth impression and depth judgement accuracy of subjects. Lastly, data were collected in the context of existing 3D in-car technology, for both autostereoscopic and if available, head-tracking based solutions, with the purpose of comparing these two techniques.

Data Analysis

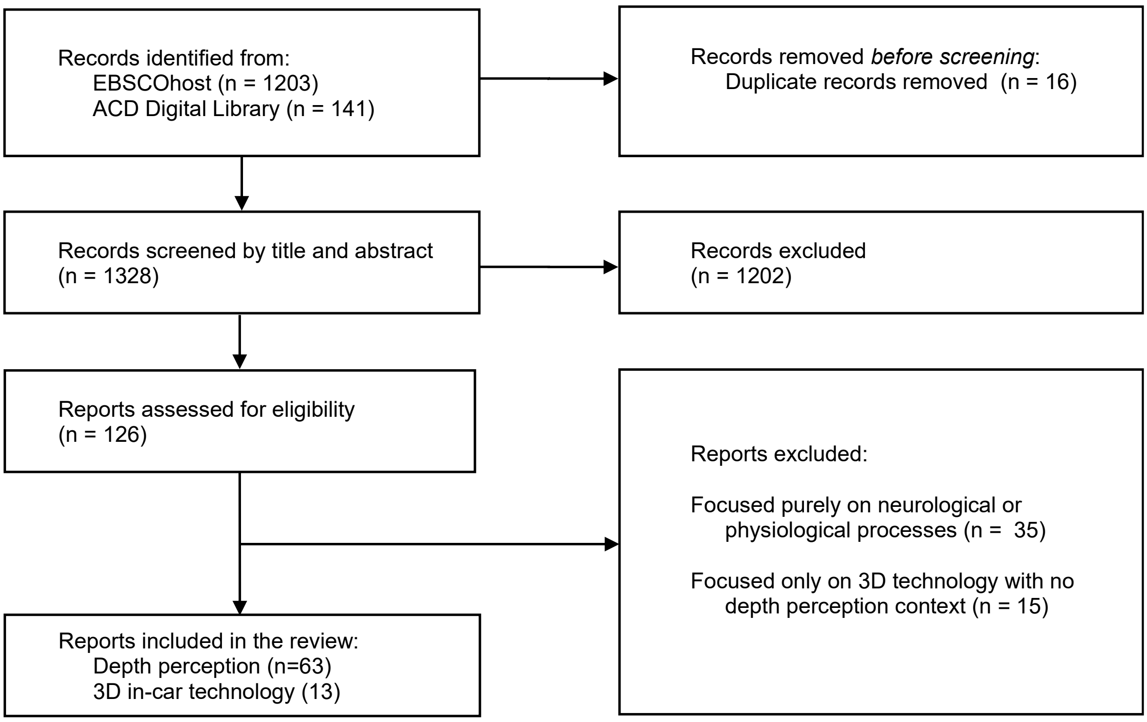
An initial coding was used to further categorize the results. Codes that were referring to the same or highly similar subject were combined and emerging groups formed hierarchies. Such a flexible approach to data extraction allowed for a full exploration of the field and a more thorough analysis. Finally, the data under the same categories were synthesized and the final results underwent an interpretation process. In the case of the reported effects of depth cues, it was for instance important to analyze them both based on the resulting depth judgement performance and the subjective depth impression that they create. This allowed for identifying the most relevant cues for both of these aspects. In the context of using display technology to create a 3D effect, the reports on in-car autostereoscopic and monoscopic displays were compared.

Results

There were 1344 records identified from the databases. Following duplicates removal, 1328 articles were used for the initial title and abstract screening. The analysis of the title and abstract resulted in the exclusion of 1202 items as most of the articles focused either on 3D technology without any mention of its influence on depth judgement or depth impression, or explored purely the underlying physiological and neurological processes of depth perception. This yielded 126 relevant papers which were then used for in-text eligibility review. Subsequently, 50 records were excluded not having met the inclusion criteria. This resulted in 76 articles (see Figure 1), 63 of which regarded depth perception (1956-2021) with the remaining 13 focusing on 3D in-car display technology (2006-2021).

Figure 1

Selection process and the number of records per stage



The section below provides relevant findings in the context of monocular depth cues. It starts with a synthesized definition of monocular depth cues and the reported advantages of their use. Subsequently, an overview is provided of all the unique monocular depth cues identified and grouped by the categories that have emerged from the data. This is accompanied by a review of different depth cue ranking approaches found in the literature, which is then followed by a final synthesis of monocular depth cues relevant for the successful creation of 3D effects on a flat, 2D display. Concurrently, their effects on depth impression and performance are described, with an additional focus on relating these effects to the design of an automotive interface.

Monocular Depth Cues

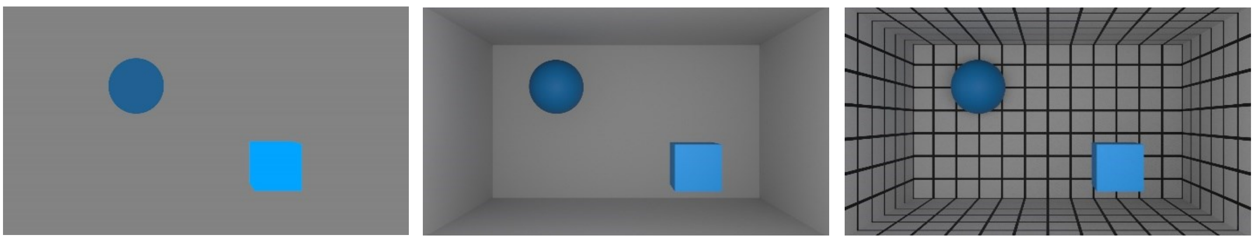
The analyzed studies provide a unified definition of monocular depth cues. These cues, as opposed to binocular cues discussed earlier, can be created in only two dimensions. Therefore, they are often referred to as pictorial cues (Lee & Lee, 2016) and include for example linear and aerial perspective, color or shading. Additionally, they can be observed by only one eye (Brooks, 2017; Hendrix & Barfield, 1995), hence the name *monocular*. Furthermore, in contrast to binocular cues and their underlying physiological processes, such as accommodation or convergence, monocular depth cues are considered to rely on psychological, cognitive responses when perceiving depth (Easa et al., 2013).

Even though monocular depth cues are not the primary source of depth estimation when perceiving a true three-dimensional world (Klinghammer et al., 2016), multiple advantages have been reported of their influence on perceived depth. The main added value of monocular depth cues is that, if used correctly, they can enhance a depth impression on a flat surface (Andersson, 2017; Brooks, 2017). Furthermore, they can easily be created using computer graphics (Hendrix & Barfield, 1995). Consequently, monocular cues can be used to create and display 3D scenes on the ubiquitous, regular 2D displays, which lowers costs and avoids the disadvantages of autostereoscopic displays, such as increased visual fatigue and discomfort (Lambooij et al., 2007).

Additionally, the monocular depth cues very often illustrate additive character, where a combination of various cues increases the overall depth impression (Cutting & Vishton, 1995) as well as depth-related performance (Keefe et al., 2011; Klinghammer et al., 2016; Mather & Smith, 2004; Royden et al., 2016). It is believed that having multiple sources for depth estimation increases the certainty of depth judgement (Keefe et al., 2011), as well as the accuracy and speed of depth ordering (Mather & Smith, 2004). Figure 2. illustrates how the addition of just two cues, shading and texture gradient, increases the impression of depth. However, adding another source of depth information can sometimes harm depth impression, or even result in conflicts between the cues (Dunn & Gray, 1965; Hendrix & Barfield, 1995; Sweet & Kaiser, 2013). This is closely related to the *modified weak fusion* (MWF) model proposed by Landy et al. (1995) which states that multiple depth cues are constantly being weighted based on their quality and location and that the estimated average is used to derive the final depth information. For instance, in a moving scene more weight will be given to depth produced by motion parallax than object overlap. Knowing which cues are given precedent and how they interact with each other is therefore essential for providing the best possible 3D effect.

Figure 2

The effect of increasing the number of monocular depth cues on depth impression



Note. Starting from the left, the used cues are: (1) flat color, (2) shading, (3) texture gradient.

In the analyzed literature, a single, yet widely used approach has been identified (Ni et al., 2007; Kellnhofer et al., 2016; Klinghammer et al., 2016; Li et al., 2013; Pfautz, 2000) to ranking monocular depth cues based on their relative importance. Namely a framework by Cutting and Vishton (1995), who proposed classifying them based on the just noticeable difference (JND), which refers to the smallest difference in the cue level or in terms of intensity needed to detect a change in stimuli. In accordance with their view, the bigger the JND, the less important the cue. Interestingly, they have also proposed that these effects differ depending on a distance from the observer and have identified three distinct zones: (1) *personal* space, which refers to a zone within arm's reach, (2) *action* space, which lies just outside the personal space, and (3) *vista* space, which starts at about 30m from the observer . Table 1 presents the overview of depth cues and their relative importance based on their proposed JND ranking for the *personal* and *action* space, since the *vista* space is not relevant in the context of this research.

Table 1

Ranking of depth information sources based on their JND

		Action space	
Depth cue (source of information)	Personal space	All sources	Pictorial sources
1. Occlusion and interposition	1	1	1
2. Relative size	4	3.5	3
3. Relative density	7	6	4
4. Height in visual field/plane	-	2	2
5. Aerial perspective	8	7	5
6. Motion perspective/parallax	3	3.5	-
7. Convergence	5.5	8.5	-
8. Accommodation	5.5	8.5	-
9. Binocular disparity and stereopsis	2	5	-

Note. Adapted from "Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth," by J. E. Cutting and P. Vishton, 1995, *Perception of Space and Motion*, p. 102.

As shown in Table 1, occlusion and interposition of objects seems to be the best source of depth information, leading to even better depth performance than binocular disparity. It is noteworthy that the distance between the driver's head and an instrument cluster is rather difficult to classify as falling strictly within the personal or the action space. Instead it seems to be located roughly at the edge of the two. Since this exploration focuses on monocular depth cues, the ranking of pictorial sources seems to offer the most relevant information. Furthermore, even though motion parallax is not classified as a pictorial source, it is ranking high compared to other cues. If the binocular cues were excluded from the personal space ranking, motion parallax would rank as a second-best source of depth information. Additionally, occlusion of objects, even though highly beneficial for depth judgement, does not seem an ideal cue for a digital instrument cluster, where the immediate visibility of all the stimuli is essential to drivers' safety. This cue should therefore be used carefully, with no essential information overlapping each other. A solution to this, would be to use an abstract, graphical element in the back of the scene to increase the effect of depth impression.

However, the above ranking contains only six monocular depth cues while many more exist and have been widely studied. Furthermore, it is often a combination of a small number of cues that are studied together by using different tasks and in different specific contexts, which causes difficulties in deriving a unified classification of all the cues. To address these gaps, Table 2 presents a categorized overview of all the monocular depth cues identified in the items included in the review, together with the effects reported by researchers of these cues on depth judgements.

Table 2

The identified depth cues with their categories, frequency of investigation and general effects

Category	Depth Cue	N of studies	Reported strengths and/or limitations
Relative position	Occlusion	12	Rapid, reliable and highly weighted cue
	Intersections	6	Highly effective for precise estimations
	Frame	4	Increases the impression of depth
	Height in the visual field	3	Reference point needed to represent the horizon
Light/Optics	Blur/depth from defocus	8	Inconclusive reports on effectiveness
	Color	4	Effectiveness increases when combined with other cues
	Contrast	6	Highly effective and reliable cue
	Brightness	7	Increases the impression of depth
	Shading	7	Inconclusive reports on effectiveness Has been shown to increase search times
	Cast shadows	5	Potential source of ambiguity
Motion	Motion parallax	13	Observer produced parallax is highly effective Some suggest exaggerating the amount of parallax
Perspective	Linear perspective	9	Very strong depth cue, believed to be a combination of several depth sources
	Aerial perspective	3	Effectiveness increases with distance
	Relative size	5	Objects need to be of common or familiar size Can easily lead to conflicts with other cues
	Texture gradient	6	Regular texture gradient combined with Intersections leads to very good depth performance

Occlusion

Out of the 12 articles that report on the cue of occlusion, only one provides a definition of this cue (Cutting & Vishton, 1995), with the remaining studies relying instead on the intuitive understanding of the term. Cutting and Vishton (1995) referred to occlusion as an overlap of two objects resulting in a complete or partial concealment of the back object (or its part) from view. Despite the lack of such definition in other studies, the use of the cue was consistent with this description. Additionally, this cue is often referred to as *overlap* (Cavanagh, 1987; Easa et al., 2013; Hillstrom et al., 2013) or *interposition* (Canestrari & Farne, 1969; Emoto, 2019; Hendrix & Barfield, 1995; Hsu et al., 2010).

Overall, occlusion provides ordinal information about the relative depth of visual elements, with the occluding object being perceived as a front one (Ni et al., 2007). As such, the cue cannot serve as a source of precise depth estimation and on its own gives no information about the distance between objects (Sweet & Kaiser, 2013). However, multiple studies have reported that occlusion is weighted highly as a source of depth information (Ni et al., 2007; Hillstrom et al., 2010), and takes precedence in case of conflicts with multiple other cues (Cutting & Vishton, 1995). Dynamic occlusion was even observed to be given priority when put in conflict with motion parallax (Ni et al., 2007). Conversely, Ono et al. (1988) report this effect to be true only when the distances between objects are large, while for smaller separations, motion parallax was primarily used to estimate depth. Nonetheless, occlusion is a highly reliable depth cue and its effectiveness is consistent at all perceivable distances (Cutting & Vishton, 1995; Sweet & Kaiser, 2013).

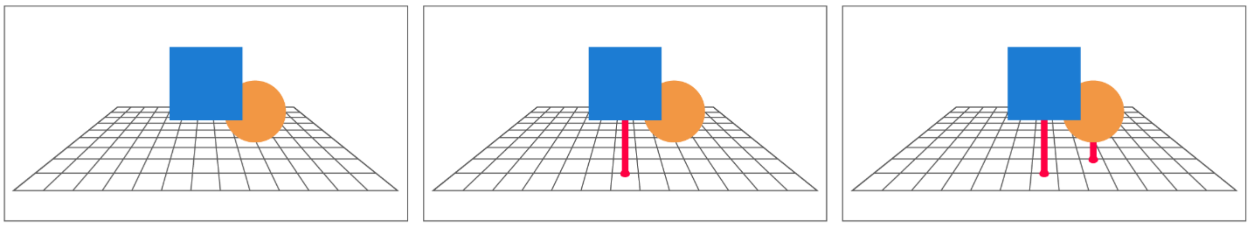
Intersections

The cue of intersections has been conceptualized based on the reviewed literature (9.5%) in order to group differently named, yet very similar sources of depth information. Initially, this cue might be thought of as derived from occlusion, as intersecting edges are also present for occluding shapes. To differentiate between these two cues, the cue of intersections refers to the actual intersecting of the geometry of objects present in 3D space, and not the visual junctions caused by overlap (see Figure 3). Hence, there is a point in space, where the objects intersect, or connect and where the distance between the objects equals zero. In contrast, two overlapping objects are always at an actual distance from each other. As a result, intersections cannot occur between objects that are placed parallel in 3D space. It is when two objects extend in different directions, that intersection or connection points can occur, providing information about their relative position in space. Moreover, this effect can be achieved by using additional visual elements to connect objects with each other. This cue helps disambiguate spatial depth relationships between objects presented on flat surfaces (Hendrix & Barfield, 1995).

As mentioned above, this cue has been utilized and named in different ways. While some (Hendrix & Barfield, 1995; Mulligan, 2009) used the term *dropline* - a line that connected an object to a textured ground plane, Hu and Knill (2011) referred to a similar concept as a *pole*. However, two objects can intersect with each other with no need for additional elements, which has most commonly been achieved by the so-called *ground contact* (Ni et al., 2007; Royden et al., 2016) where an object is directly connected to the ground plane. This has to be made clear, by for instance a thin shadow edge where the two objects meet. Notice how for all of the described measures the addition of a ground plane is a necessary component. Furthermore, a *ground dominance effect* (Bian et al., 2005) seems to exist, so that even if elements in a display are presented in a scene where there are multiple surfaces, such as the walls and the ceiling, the observers will primarily use the ground to derive their estimations.

Figure 3

The difference between occlusion and intersections



Note. The image on the left contains only occlusion of elements, and intersections are added in the other two images through connecting droplines. Notice how just by using occlusion the position of both the objects remains unclear and how the use of linear perspective creates a conflict between these two cues. Adding droplines results in much more precise estimates of the positions of both objects within the scene.

When it comes to the effectiveness of intersections, they have been shown to highly increase observers‘ ability to correctly identify both stationary (Bian et al., 2005) and - what is more relevant in the context of this research - moving objects (Royden et al., 2016). They also appear to provide the most exact information regarding spatial relationships of objects and allow for highly consistent performance in depth judgement (Hendrix & Barfield, 1995). Moreover, intersections (as illustrated in Figure 3) appear to be more effective than cues such as shadows or texture mapping (Hendrix & Barfield, 1995). Considering the high computational power needed to achieve shadows, the cue of intersections seems to serve as a good replacement.

Frame

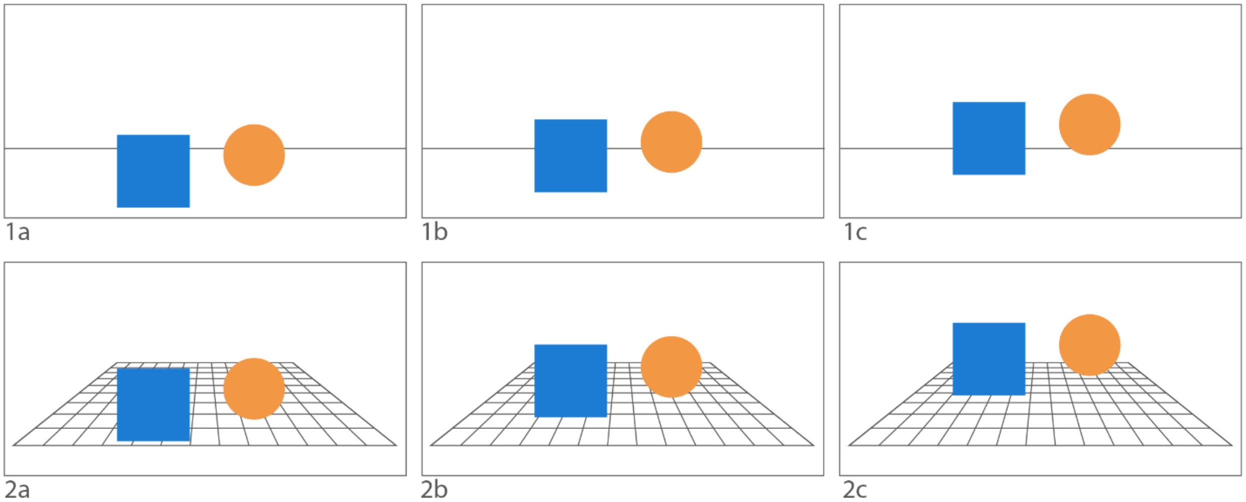
The frame is the next source of depth information on a flat surface that has been conceptualized as a depth cue based on the reviewed literature (6.3%). Despite usually being a flat part of the representation, it provides allocentric (relative to other objects) information and servers as a reference frame for the spatial transformation of objects within it (Klinghammer et al., 2016; Léveillé et al., 2014). This is especially true for dynamic transformations, where the frame around the scene serves as a source of location constancy (Rogers, 2009). Moreover, a frame has also been shown to increase apparent depth in a static photo image, especially when the frame was placed at a distance (Shimono et al., 2021). A possible reason for such an effect, is that the physical depth difference between the frame and the photo causes binocular disparity, which in turn increases the reliability of the monocular depth cues (Shimono et al., 2021). Additionally, a frame, much like a screen, helps to separate its content from the surroundings and serves as a window into another environment, which causes us to expect the same physical laws to apply within it, such as gravity (Pla & Maes, 2013). It is therefore important to make the spatial transformations within the frame physically accurate for a proper depth impression.

Height in the visual field

The last and least reported on (4.7%) monocular cue to depth in the *Relative Position* category is the height in the visual field (see Figure 4). This cue originates in our perception of the 3D environment and the fact that the more distant objects (below our eyesight level) appear higher in our visual field (Cutting & Vishton, 1995). This cue is therefore a good source of ordinal depth information (Royden et al., 2016) and supports recreating an impression of depth (Dunn & Gray, 1965). However, for this cue to work on a flat surface, a reference point is needed to represent the horizon.

Figure 4

Changes in apparent depth of objects caused by manipulation of height in the visual field



Note. The top series of images (1) contains only the horizon line as a reference, while the textured ground plane is included in the bottom row (2). It is clear that the horizon line is enough to serve as a reference point, but the linear perspective present in the texture makes this effect even stronger, speaking again to the additive nature of depth cues.

Light and optics

Light and optics is the second category of monocular cues derived from the review and contains the cues that are present in terms of depth perception due to our interaction with light and its influence on the environment. This category includes: blur or depth from defocus, color, contrast, brightness, shading, and cast shadows.

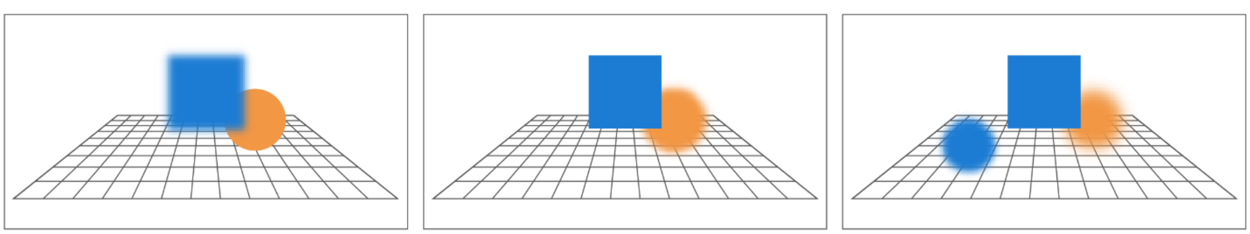
Blur

Blur or depth from defocus was the most commonly reported cue in this category (12.7% of all reviewed studies) and is a well-established depth cue with its importance as a source of depth being highly appreciated in the realms of cinematography and computer graphics (Held et al., 2010). Furthermore, this cue has been successfully used to develop computer vision systems and depth enhancements algorithms (Lee & Lee, 2016; Rößing et al., 2012), as well as for 2D-to-3D video conversion (Zhang et al., 2013). Blur is believed to be a successful depth cue because it closely matches human perception, where accommodation leads to visual focus of a small area, leaving the rest of the image in the visual field blurred (Held et al., 2010). Consequently, blur has been shown to stimulate accommodation responses better than other monocular cues (Busby & Ciuffreda, 2005). By creating a sharp area to focus on, blur is believed to be useful in guiding viewers' attention (see Figure 5). However, despite such common use of blur to compute and recreate depth, the effectiveness of this cue in the context of depth judgment remains unclear. While some declare blur to be a strong depth cue (Held et al., 2010; Rößing et al., 2012), others report on its lack of influence on depth ordering (Easa et al., 2013; Koessler & Hill, 2019). Moreover, when applied to moving elements blur has been shown to negatively impact depth judgement (Sweet & Kaiser, 2013).

Such contrary results and inconsistencies in reported effects of blur point to gaps in knowledge regarding this cue. Considering the high-risk context of a driving task, this cue should best be avoided or applied with careful consideration. Additionally, blur might not be the most optimal cue to apply to an instrument cluster, as the clear visibility of information plays a vital role.

Figure 5

Blur as a cue to guiding attention



Note. Changes in blur should shift attention to the sharp elements. Notice the cue conflict with the height in the visual field – the square is placed higher so should appear further away.

Color

Color is the least commonly mentioned (6.3%) cue in this category, but it is almost always present in the design and is therefore, an interesting cue to explore. As a design element, it has been shown to facilitate visual working memory due to its grouping properties (Qian et al., 2017). Additionally, color can successfully guide and grab viewers' attention, which has been found beneficial for visual warnings in the automotive setting (Broy et al., 2015a; Rao et al., 2014). When exploring color as a depth cue it is important to differentiate it from other cues such as contrast (Cavanagh, 1987), by making sure that two different colors are isoluminant - have no luminance contrast (see Figure 6).

Figure 6

Comparison of an isoluminant color palette within HCL color space with a regular RGB palette

Note. HCL color space stands for Hue, Chroma (saturation sense) and Luminance (lightness). In the case of the color palette above, the luminance is set to 80 (scale 0-100) for all colors, and their chroma value is 90 (scale 0-100). Such an approach results in different color hues (1a) of the same lightness (1b), which is not easily obtainable within other color spaces (Luong et al., 2005), such as RGB (2).

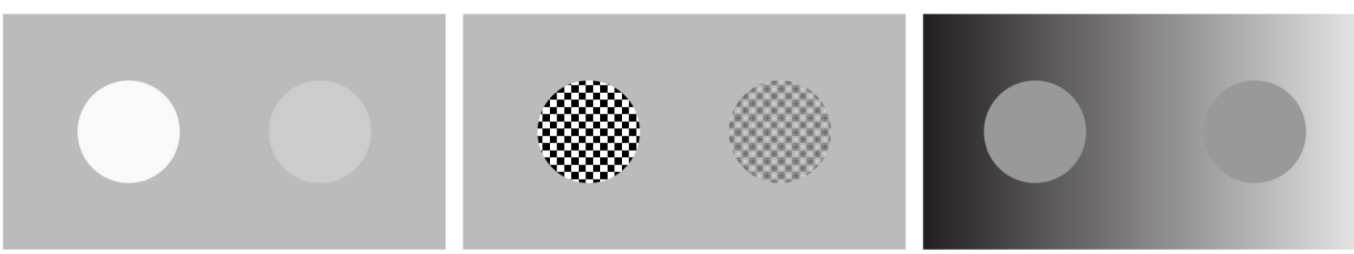
Studies have shown that even with no difference in luminance, color can be used as a cue to depth. For instance, Troscianko et al. (1991) have illustrated that certain color gradients have led to an impression of depth, with for instance a red-to-grey gradient being more effective than red-to-green. The color combinations that were found effective, were similar to what can be found in nature, where the saturation of the color decreases with distance. This is closely connected to another depth cue, namely aerial perspective and points again to the additive nature of the cues. It was indeed found that color's ability to improve performance in a depth ordering task increased in presence of other monocular depth cues (Troscianko et al., 1991). Furthermore, extended color space present in high dynamic range displays has been shown to increase the impression of depth in flat 2D video representations (Emoto, 2019; Rempel et al., 2011). These findings suggest that color can be used to facilitate depth, but the choice of colors should be carefully considered, with brighter and more saturated colors placed in the front.

Contrast

Another depth cue identified in the reviewed items (9.5%) is contrast, which is the perceived difference in luminance between two different visual stimuli (see Figure 7). Contrast diminishes with distance and hence the objects with higher amounts of contrast are judged as closer to us (Ichihara et al., 2007). This is true for both the contrast within the object itself, called texture contrast, and between the object and its background, referred to as area contrast (Ichihara et al., 2007). Additionally, it appears that the size of the object plays a role in the produced area contrast, with bigger objects resulting in higher contrast, than the small objects of the same color (Markov & Tiurina, 2021).

Figure 7

Contrast and its effect on perceived depth



Note: Starting from the left: (1) area contrast between two differently shaded circles, (2) texture contrast differences on the same background, (3) identical shade of grey appears differently depending on its contrast with the background. Notice how in all three images the higher contrast object appears to be the one in front.

The reviewed literature is consistent in reporting contrast as an effective and strong monocular depth cue. It was found that increase in contrast supports the impression of depth of stimuli within the already mentioned high dynamic range displays, where less information is lost as contrast gets higher (Easa et al., 2013; Emoto, 2019; Rempel et al. 2011). Additionally, just increasing the contrast of small elements within the image, such as that of highlights, was beneficial for perceived depth (Rempel et al., 2011). Moreover, contrast appears to be one of the most reliable cues (Easa et al., 2013) as it demonstrates a high ratio of correct depth ordering judgements (91.7%) when compared with other cues, such as blur (58%), transparency (50%) or even shadows (83.3%). Furthermore, high contrast seems to support correct judgement of 3D moving objects and their directions, with low or incorrectly applied contrast leading to confused judgements (Fulvio et al., 2015). Finally, contrast, just like color, is always a part of a display and the design within it, so it is essential to implement it in a way that best supports achieving an impression of depth.

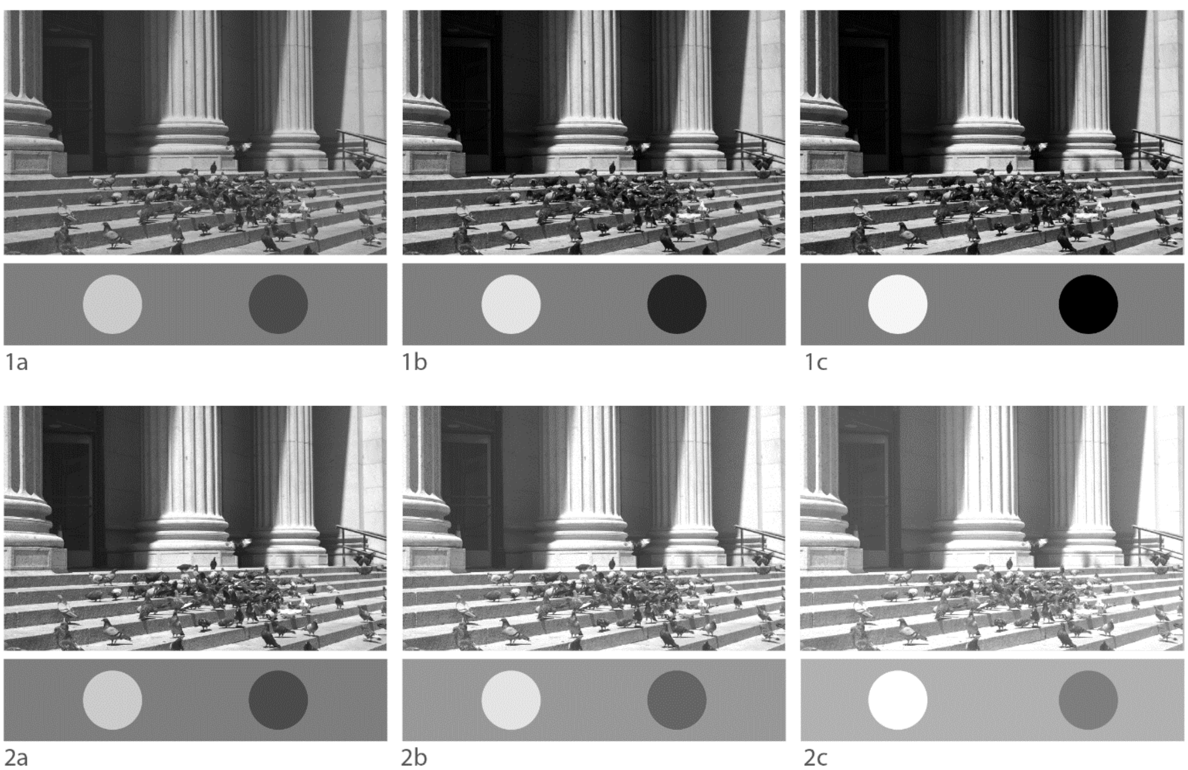
Brightness

Brightness, similarly to color and contrast is a part of every design and the display settings it is presented on. Brightness is a perceptual term and refers to the perceived luminance of stimuli (Gilchrist, 2007). Increasing brightness, will therefore lead to an overall increase of lightness within the image, making the shadows lighter as well. On the other hand, increased contrast causes shadows to become darker while light parts of the image become even lighter (see Figure 8).

The articles that mention this cue (11.1%) are consistent in reporting on the effectiveness of brightness as a source of depth information. It is believed to be a cue to depth, since objects that are closer to us, tend to reflect more light, than distant objects (Easa et al., 2013). This higher brightness of proximal elements is often referred to as *proximity-luminance covariance* (Dosher et al., 1986; Young et al., 1993). It has been shown that gradually increasing a given object's luminance, will make it appear to move forward (Cutting & Vishton, 1995). The reviewed studies highlight the importance of increasing brightness to best support the impression of depth (Emoto, 2019; Hendrix & Barfield, 1995) and depth related performance. Easa et al. (2013) have demonstrated that brightness is one of the most effective depth cues in producing correct depth ordering judgements (ratio of 92.9%). Here again, the high dynamic range of modern displays seems useful for achieving the best depth impression through increased brightness (Emoto, 2019; Rempel et al., 2011).

Figure 8

The difference between contrast and brightness cues



Note. The top row (1) illustrates a gradual increase in contrast. While in 1a the door in the background is well visible, it is lost in the darkened view in 1c. The bottom row (2) shows a gradual increase in brightness. What is worth noticing, is that while the middle grey tone of the background behind the circles in row 1 stays consistent, it gets brighter in row 2.

Shading

Shading is another depth cue found in the reviewed literature (11.1%) and it can be defined as a source of structural information about an object's shape through the use of highlights and shadows on the object's surface (Cutting & Vishton, 1995; Kourtzi & Kanwisher, 2000; Rößing et al., 2012). Shading is also known as *attached shadows* (Brooks, 2017), which distinguishes this cue from that of *cast shadows* – shadows cast by one object onto the surface of another. Despite such a clear distinction these terms are sometimes used interchangeably, which introduces some confusion while interpreting the study results. For the sake of clarity, this paper makes use of the above definitions.

Despite being one of the biggest sources of depth impression in drawings and paintings (Brooks, 2017), the influence of shading on depth-related performance remains unclear. While some evidence was found that shading increases depth ordering judgements (Cavanagh, 1987), others report no advantage of this cue on the perceived depth (Andersson, 2017), as well as no increase of visual search times when compared with flat stimuli (Greene, 2021). The latter is highly important in the context of this project, as shading multiple objects within a digital instrument cluster would be computationally costly, yet appears to offer no improvement over unshaded surfaces.

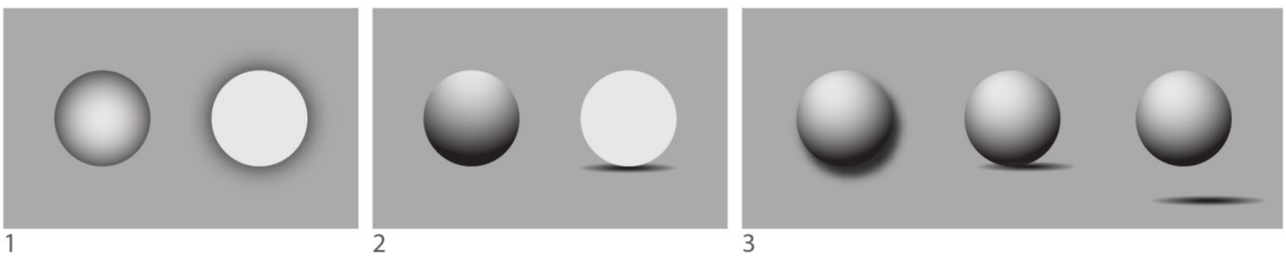
Cast shadows

Another depth cue investigated in the reviewed articles (7.9%) is that of cast shadows. As explained above, cast shadows are different from shading as they are cast onto the surface of another object (see Figure 9). By doing so they provide information of relative positions of objects within a scene and have indeed been found to improve depth judgements (Easa et al., 2013, Hendrix & Barfield, 1995), especially when combined with a textured ground plane

(Hendrix & Barfield, 1995). Additionally, the depth of a moving object can be judged based on the movement of its cast shadow (Katsuyama et al., 2016; Kersten et al., 1997; Ni et al., 2007). However, the motion of a cast shadow has been shown to cause ambiguous depth judgements and even to induce the apparent motion of a static object (Katsuyama et al., 2016). Similarly, when combined with motion parallax, cast shadows can be a source of ambiguity (Ni et al, 2007). Lastly, a cast shadow's movement can override other strong depth cues (Kersten et al., 1997).

Figure 9

Difference between shading (attached shadows) and cast shadows



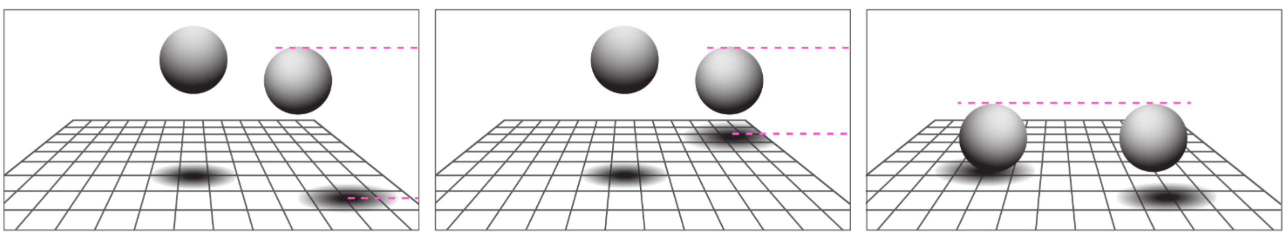
Note. The first two images present shading (left sphere) and cast shadows (right circle) with source of light in front (1) and above (2) . The last image (3) illustrates how different types of cast shadows change the perception of a sphere with the same shading. Starting with the cast shadows (3) from the left: the sphere appears to be either lying on a flat surface and looked at from above; placed on a flat surface and looked at frontally; and finally the sphere appears to be floating with its shadow cast on a surface below it.

Furthermore, shadows have resulted in a smaller depth judgement improvement of around 30%, than the cue of intersections in form of droplines, which caused a 200% improvement (Hendrix & Barfield, 1995). The most likely reason for this difference according to the authors is that cast shadows' position depends on that of the light source. Therefore shadows will never be cast directly below each object (see Figure 10), which requires additional processing of depth by the viewer. This attentional cost, together with the higher computational costs of generating shadows in real-time, have led Hendrix and Barfield (1995) to suggest the use of intersections instead, which can be achieved by abstract representations such as dropline (see

Figure 11). Moreover, Easa et al. (2013) found that shadows, even though overall still successful at aiding depth ordering judgements (success rate = 83.3%), more often than expected lead to false ordering. According to the authors, this might have been caused by confusion between cast shadows and shading of the layered stimuli used. Lastly, another reported disadvantage of cast shadows is the loss of information in the shadowed area (Easa et al., 2013). Overall, the studies cited above show little evidence for effectiveness of cast shadows on performance when estimating relative depth of objects in 3D space. The reported ambiguity this cue causes might be dangerous for a driving task, which requires rapid and correct judgements.

Figure 10

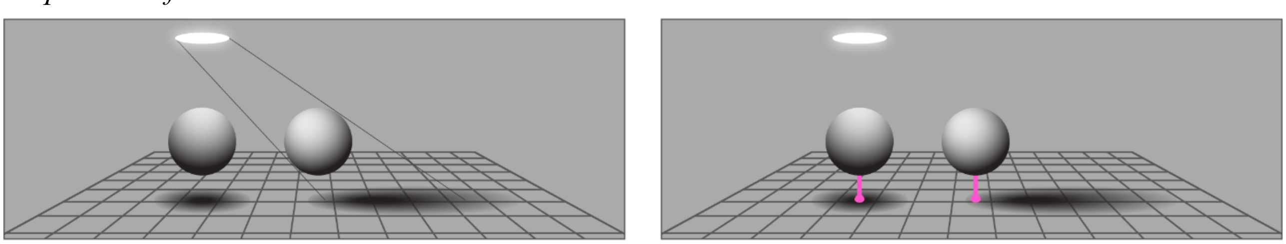
Cast shadows and their influence on perceived depth



Note. In all the images the light source is located centrally above the ground plane which causes different positions of shadows cast. Starting from the left, the first two images illustrate how the change in shadow's position affect the perceived depth of the right sphere, even though its position and scale do not change. The right image shows how a cast shadow might either make a sphere seem to make contact with the ground plane or cause an impression of it floating in space.

Figure 11

Comparison of cast shadows and intersections



Note. The image on the left illustrates how a source of light affects the shadow cast from a floating object. The second image shows how easily and accurately depth information can be derived from the cue of intersections, when compared to the ambiguity of shadows.

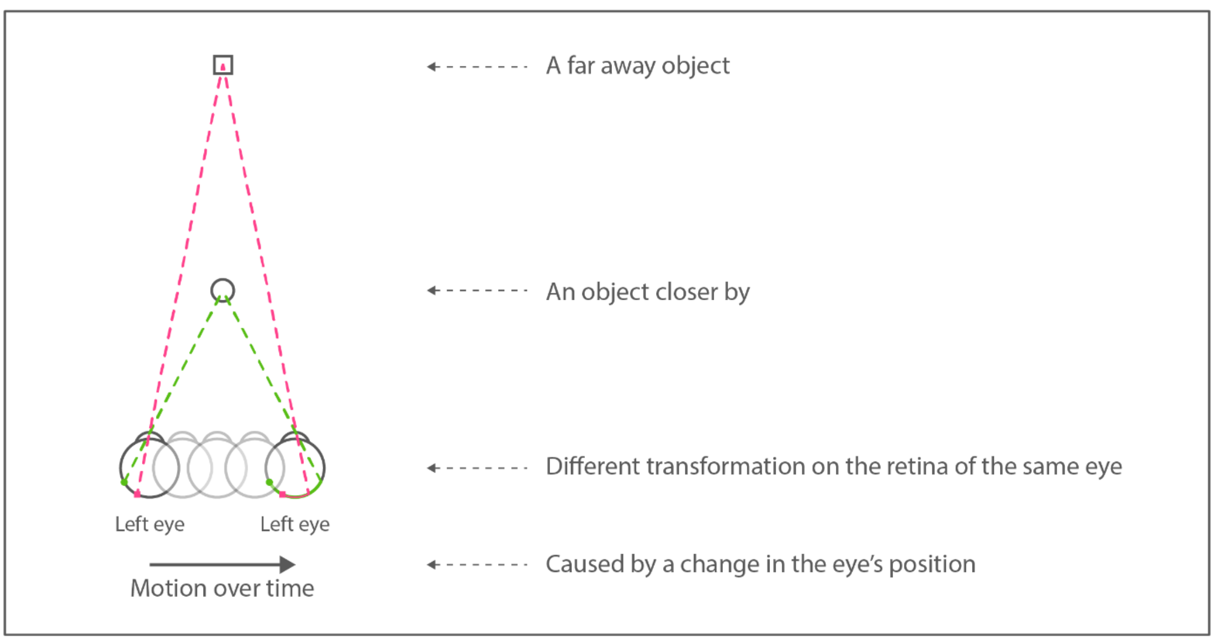
Motion related cues

Motion parallax

Our depth perception processes occur in a world of both static and moving stimuli. Deriving depth information based on an object's movement, be it a wild animal or an approaching car, has always been of high importance to our daily activities (Royden et al., 2016). Additionally, in order to successfully move through the environment, a mental process is needed to constantly update the information on our position in reference to other elements. Here the observer's movements relative to the perceived objects results in a pattern of transformations on the observer's retina (Rogers & Graham, 1979). This pattern, also known as the *radial pattern*, differs depending on the distance at which the objects are located, with the nearby stimuli moving faster on the retina, than the ones far away (Royden et al., 2016), which results in a bigger retinal transformation (see Figure 12). This process is known as motion parallax and is a constant and highly important source of depth information (Rogers et al, 2009).

Figure 12

Motion Parallax as a source of depth information



Given the definition above, motion parallax can be used in a simulated environment to produce an impression of depth. In order to achieve that, the transformation of objects has to correspond to the movement of the observer, also known as observer-produced parallax. This technique was influential to the results of the classic study by Rogers and Graham (1979), who were the first ones to illustrate that "motion parallax can be a sufficient cue to the shape and depth of three-dimensional surfaces, in the absence of all other depth cues." (p. 132). In the reviewed literature, a multitude of studies (20.6%) confirms that motion parallax enhances both the impression of depth (Emoto, 2019; Kongsilp & Dailey, 2017; Uehira et al., 2007) and depth-related performance (Mulligan, 2009; Ni et al., 2007; Parton et al., 1999; Royden et al., 2016; Seipel, 2013). For example, the observer-produced parallax was found beneficial for altitude estimation in air-traffic displays (Mulligan, 2009). Similarly, Seipel (2013) has found that motion parallax increased the accuracy of spatial assessment in the context of geovisualisations, resulting in even better performance than a stereoscopic display. Moreover, Royden et al. (2016) illustrated that in the presence of observer-produced motion parallax, increasing the number of other monocular depth cues resulted in a decreased threshold when detecting moving objects. The additive character of motion parallax is also confirmed by other studies (Ni et al, 2007, Warren & Rushton, 2009). What remains unclear, however, is the degree of parallax that is best received by the viewer. While the studies above measured the physically accurate parallax transformation, others (Hürst et al., 2013; Mulligan, 2009;) report that an exaggerated parallax might be beneficial for certain types of visualizations.

Furthermore, several interaction effects between motion parallax and other cues have been reported. When put in conflict with the cue of occlusion, motion parallax was the main source of depth estimation at small distances between objects, but when the distance increased, occlusion was given precedence (Ono et al., 1988). Additionally, a frame around the screen on which motion parallax occurs has been noted as a highly important element, as its rigid surface provides a reference point, or “location constancy” (Rogers et al, 2009, p. 915), which allows differentiating between the static and moving elements. A frame is therefore seen as a strengthening component, when trying to achieve depth from motion parallax. In the context of automotive displays, the dashboard itself could serve to achieve such a framing effect.

Perspective related cues

The last category of monocular cues is based around the concept of perspective, which can be construed as relative differences in visual stimuli corresponding to their distance from the observer. These differences include attributes such as size, density, color and contrast.

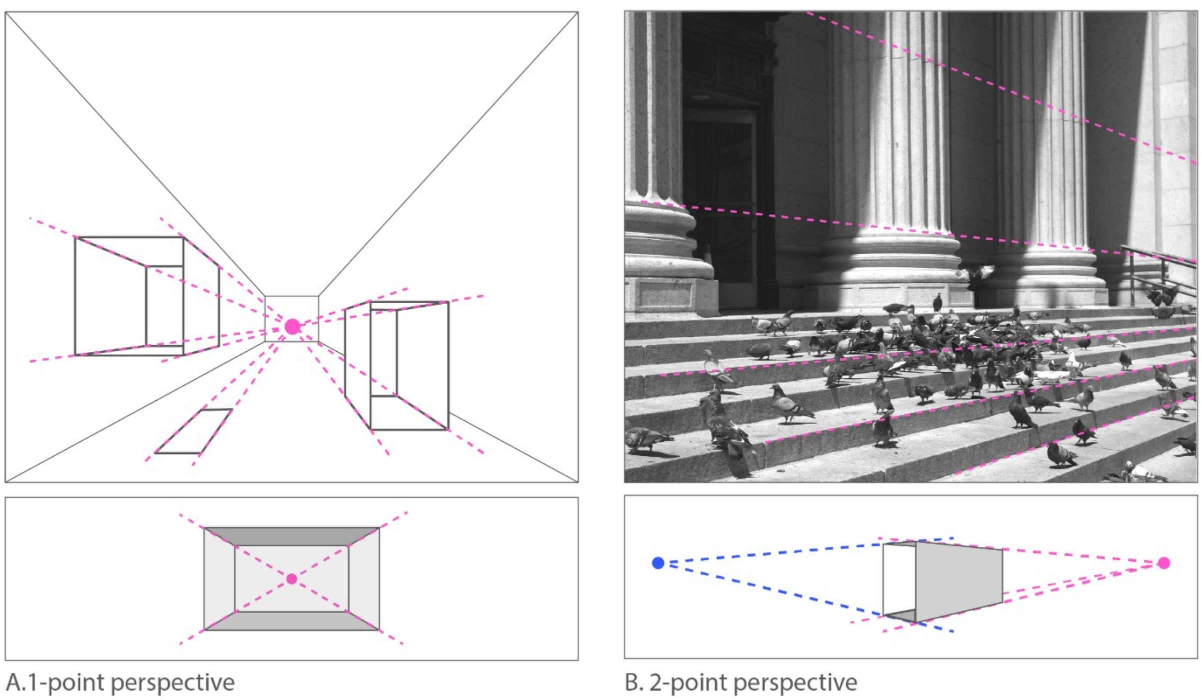
Linear perspective

The most commonly explored (14.3%) source of depth in this category is the linear perspective, which can be explained as the “convergence of each set of parallel lines in the image to a unique vanishing point” (Brooks, 2017, p.4). Its usefulness for recreating depth on a flat surface (see Figure 13) has long been utilized by artists (Goldmann et al., 2012). As a depth cue however, linear perspective seems to pose some theoretical challenges. While some researchers categorize it as a separate cue (Andersson, 2017;), Cutting & Vishton (1995) point out that linear perspective is in reality a combination of several different cues to depth. These include the relative size, texture gradient, and occlusion and according to Cutting & Vishton (1995) the parallel lines of linear perspective serve to simply disambiguate the above cues. This issue arose for instance in the study by Hendrix and Barfield (1995), who used a texture gradient cue located on a ground plane. However, as the authors themselves noted, the regular structure of the gradient (see Figure 14) resulted in parallel lines converging in the distance, adding the linear perspective as a source of depth information.

Regardless of its correct categorization, the reviewed studies confirm the effectiveness of linear perspective in creating an impression of depth (Goldmann et al., 2012; O'Leary & Wallach, 1980; Seipel, 2013; Warren & Rushton, 2009). Linear perspective has also been successfully used for 2D-to-3D video conversion (Zhang et al., 2013). However, linear perspective can easily lead to objects occluding each other (Andersson, 2017). It is therefore important to ensure that the visibility of essential elements is not compromised by this cue. Additionally, as mentioned by Goldmann et al. (2012), linear perspective allows for exact depth estimations, unlike most monocular depth cues, which provide only ordinal information.

Figure 13

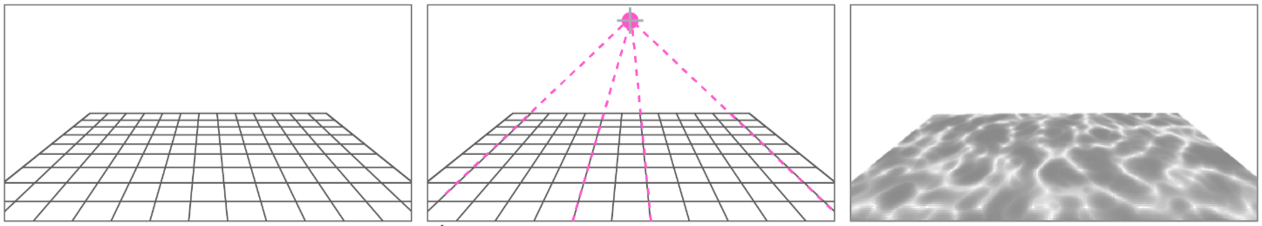
Linear perspective as a source of depth information



Note. The parallel lines can diverge onto one (A), two (B) or even more vanishing points. When one side of an object is seen frontally (A) there is one vanishing point. When an object is viewed from the side a second vanishing point emerges (B). It should be noted that this is a dynamic change, that should be present for perspective transformation based on head movements.

Figure 14

Linear perspective and texture gradient



Note. The slanted texture gradient in the form of a regular grid results in parallel lines (1b) converging with increase in distance. Such a texture gradient strengthens the impression of depth and allows for more accurate depth estimations. Conversely, an irregular gradient (2) does not provide the cue of linear perspective, even though the density changes with distance as well.

Aerial perspective

Another perspective related depth cue is that of aerial perspective, which similarly to the linear perspective, has long been used in the visual arts (Brooks, 2017). Aerial perspective is believed to be a result of the atmosphere and the particles it contains, with their density getting higher with increased distance (see Figure 15), which in turn causes the distant objects to appear blue and desaturated (Brooks, 2017; Cutting & Vishton, 1995). As mentioned before, Cutting and Vishton (1995) argued that linear perspective is a combination of other depth cues. However, they do not seem to make the same conclusion about aerial perspective, despite stating in their description of the cue that "objects in the distance become bluer, decreased in contrast, or both with respect to object in the foreground" (Cutting & Vishton, 1995, p.88). It could potentially be argued, that this cue is again a combination of other monocular depth cues, such as color and contrast. This theoretical issue is again made visible in studies that report on aerial perspective while only manipulating the brightness cue (see for example Canestrari & Farne, 1969).

Figure 15

Aerial perspective in nature and pictorial arts

1 2

Note. Notice how the distant elements of the environment appear more blue and desaturated. 1 presents a picture of a real environment, while 2 (d' Arthois, n.d.) illustrates how painters use the cue of aerial perspective to recreate depth. Notice how in 2 the effect of depth is even more pronounced due to the warm colors in the foreground and their contrast with the cold background.

In the reviewed literature, aerial perspective is mentioned rather scarcely (4.7%). Besides the reported role of aerial perspective in the process of 2D-to-3D video conversion (Zhang et al., 2013), it seems that it is a useful source of depth information for larger distances, and the only cue that demonstrates an increase of effectiveness with distance (Cutting & Vishton, 1995). The cue of aerial perspective can therefore be used to enhance depth when other sources, for instance motion parallax, lose their effect. Lastly, unlike linear perspective, aerial perspective provides only ordinal depth information (Cutting & Vishton, 1995).

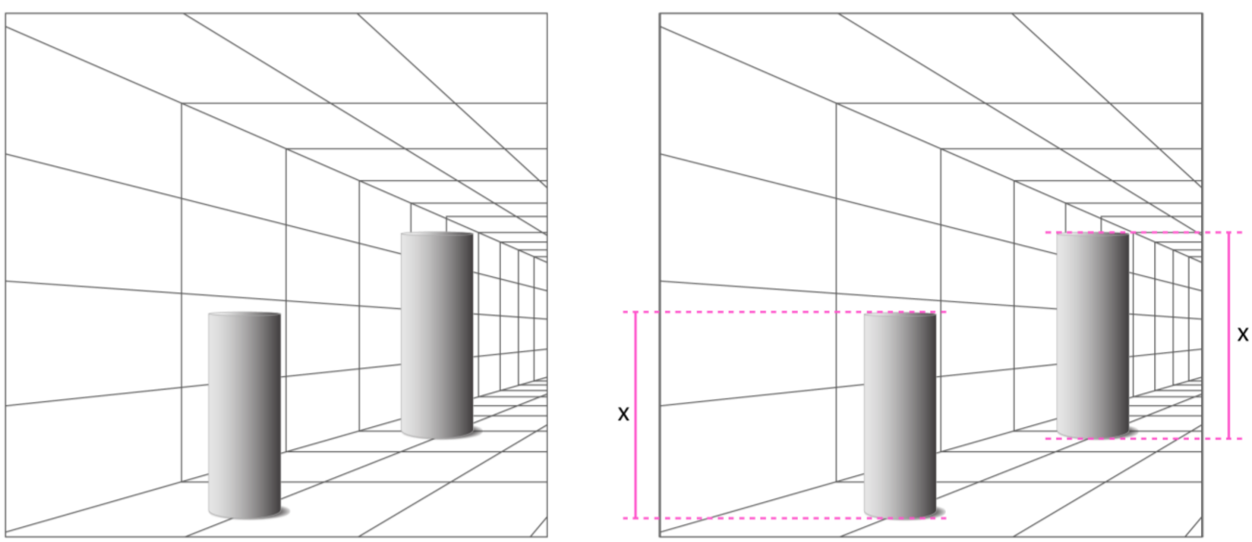
Relative size

Relative size is another perspective cue emerging from the reviewed articles (7.9%). This cue is closely related to perspective as the object's apparent size decreases with depth (Easa et al., 2013). This results in perceived size differences depending on the object's distance from the observer, with larger objects appearing closer than those smaller in size (Cutting & Vishton, 1995). This cue has been shown to increase depth judgement performance in isolation

(Klinghammer et al., 2016), as well as when combined with motion parallax and other monocular depth cues (Warren & Rushton, 2009). Easa et al. (2013) report that the effectiveness of relative size is very high and closely matches that of contrast and brightness. They also mention that this cue is easily recreated on any type of display, yet for the cue of relative size to work, the objects must be of either common or familiar size. Additionally, relative size can result in the so-called "corridor illusion" (Easa et al., 2013) when put in conflict with linear perspective (see Figure 16). Here, the parallel lines converging into the distance will make the object in the back of the corridor appear larger than an object of the same size placed in front. This illustrates the strength of the linear perspective cue and highlights the need for appropriate manipulation of the objects' relative size in the presence of linear perspective.

Figure 16

The corridor illusion caused by the conflict between relative size and linear perspective



Note. The cylinder, despite having the exact same size (x), appears as larger when placed in the back end of the corridor, relative to the one in front.

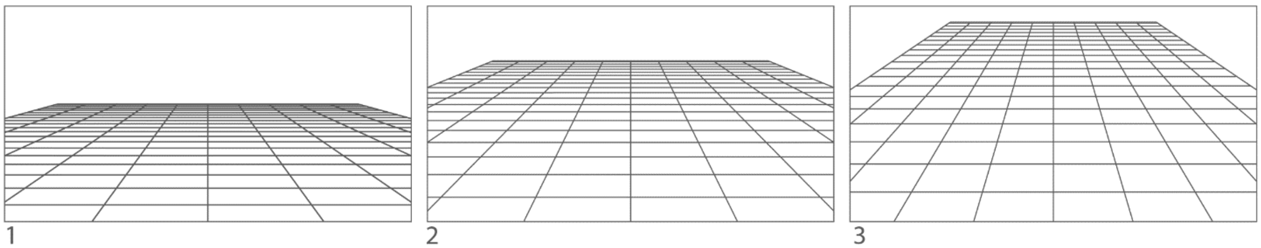
Texture gradient

The last perspective related cue reported in the reviewed literature (9.5%) is the texture gradient, which can be understood as a texture placed on "surfaces receding in depth" (Cutting & Vishton, 1995, p. 94). The word *gradient* refers to the fact, that over distance, the texture's density changes accordingly, informing the observer of the slant of a given surface (see Figure 17). As mentioned above, once the texture has a regular structure that results in converging parallel lines, it can be construed as a linear perspective cue (see Figure 14). Again a theoretical complication arises here, since many reports on the effectiveness of texture gradient are made based on a regular texture (Canestrari & Farne, 1969; Hendrix & Barfield, 1995; Keefe et al., 2011), with a few reporting on a texture of irregular structure (O'Brien & Johnston, 2000; Royden et al., 2016). Moreover, it seems that the studies that implemented the texture gradient in a regular, grid-like form report higher effectiveness of this cue, than the studies that investigated irregular texture gradients, which suggests that the regular gradients carry the added benefit in form of the converging lines.

When it comes to how the texture gradient interacts with other monocular depth cues, it seems that its effectiveness is most optimal when combined with the cue of intersections (Hendrix & Barfield, 1995; Royden et al, 2016). Furthermore, in the presence of motion, the cue of a regular texture gradient seems to increase in effectiveness (Canestrari & Farne, 1969) and can even become the primary source of depth information (O'Brien & Johnston, 2000). Therefore, adding a regular, grid-like texture gradient might be highly beneficial for conveying 3D effects in a dynamically transformed digital instrument cluster.

Figure 17

Texture gradient and its effect on perceived slant



Note. Texture gradient can inform the viewer on a slant of the surface it is applied to. Notice the change in the perspective; in the first image (1) it seems as if the point of view is much lower than in the last one (3), which appears to be observed from a higher perspective.

Summary and relevance in the context of automotive interfaces

Overall, the reviewed literature provides evidence for the effectiveness of monocular depth cues in conveying accurate depth representations. It also supports the depth fusion model, where an increased number of depth cues results in better depth-related performance. The cues of occlusion, intersections, contrast and brightness have consistently been reported as highly reliable and effective sources of depth information. Moreover, the observer-produced parallax and the resulting perspective transformation of a 3D scene has been confirmed as a valid approach to conveying depth on a flat display. Here, the inclusion of a regular, grid-like texture gradient, as well as the addition of a physical frame have proven beneficial. However, the amount of motion parallax that is best received by the observer remains unclear.

On the other hand, commonly used pictorial sources of depth, such as blur, shading, shadows or aerial perspective have been reported as lower in effectiveness than one might expect. Their reported effects are also inconsistent or, as in the case of aerial perspective, simply insufficient. The limited information, as well as the reported depth ambiguity caused by the above cues begs caution when applying them in the high-risk context of driving. Additionally, such cues as blur, shadows or even the highly effective occlusion can easily result in lower visibility of essential information and should therefore be avoided. Moreover, cues such as color, contrast and brightness are always present in the design of a digital instrument cluster. Therefore, their properties should involve careful consideration to best facilitate depth judgement.

Putting aside the theoretical inconsistencies regarding linear and aerial perspective, the more relevant theoretical challenge seems to follow from what constitutes the simulation of an effective depth in a scene. Most of the reviewed studies (76%) have assessed the monocular depth cues and their effectiveness based on the resulting depth-related performance, with a much smaller number of studies investigating the subjective depth impression (25%). Only a single study investigated both the performance and the subjective impression of depth (Rogers & Graham, 1979). While rapid and correct judgements are still highly relevant in an automotive context, the believability of the produced effect should also be ensured. Since this exploration aims at a potential replacement of an autostereoscopic display, the believability of the depth impression seems as important to investigate, as the depth judgement accuracy.

Relevant findings on existing 3D automotive interfaces

Since autostereoscopic displays are still a novelty in the automotive industry, there are relatively few studies on 3D effects in the context of driving. The biggest contribution to the topic is a series of studies by Broy et al. (2012; 2013; 2014a; 2014b; 2015a; 2015b), who investigated 3D in-car solutions both in a controlled and a naturalistic setting (2015a; 2015b). Although stemming from research with stereoscopic devices, their findings and guidelines still offer important insights with regards to the design of spatial relations within a 3D display.

Taken together, their studies highlight the benefit and importance of structuring information into separate depth layers. Such *layered 3D UIs* increase the comprehensibility of the design and, by extension, user performance (Broy et al., 2014a). They suggest using only a few of such layers and placing the main visual elements in the most frontal one, as it “helps the user intuitively recognize the current focus of interaction and consequently the state of the system” (2012, p. 100). Bringing elements to the front in this manner is also useful for communicating urgency (2015b). Additionally, the most important information should be placed at the same level as the screen due to the high visual comfort this position offers (2013). This is also true for textual information as the distortion is minimized (2014b). Moreover, placing visual elements in such a way that they appear to be floating in front of the display should be avoided, as “participants were significantly faster to judge the depth relationship between objects as they are presented behind the screen” (2013, p.66). Furthermore, they have found that the most optimal distance of the farthest depth layer is highly subjective, suggesting to make the depth range adjustable by the user for the most comfortable experience (2013). The complexity of information should also be minimized, as it negatively affects the driving task (2014a). Lastly, the monocular depth cues should be implemented carefully to avoid cue conflicts, as well as such cues as relative size, occlusion or shadows, as they can negatively impact visibility (2014b). This, in particular, highlights the need to carefully consider the monocular depth cues, as their reported strengths might not be beneficial in the context of automotive interfaces.

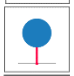
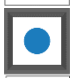
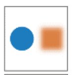
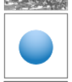
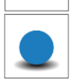
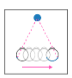
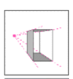
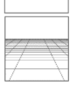
Discussion

The main goal of the systematic review was to determine which monocular depth cues have the highest potential to increase the perception of depth within a simulated 3D digital instrument cluster. For the 3D simulation to work properly, the transformation of the scene needs to follow the point of view of the driver. Here, the observer-produced motion parallax is the crucial cue. Additionally, the unconfirmed benefits of exaggerating the motion parallax are worth investigating, especially that drivers’ head movements tend to be relatively small.

Furthermore, the review illustrated that cues such as shadows, occlusion or relative size, even if beneficial for depth perception, are not necessarily the most optimal for applications where visibility of elements is safety critical. On the other hand, several cues leave little doubt as to their effectiveness for depth judgements. These include texture gradient in combination with intersections, and linear perspective. These cues are beneficial for correct estimations and disambiguating spatial relations, which makes them a promising source of depth information for such infotainment elements as navigation. However, due to their geometric character their sole use might not prompt optimal depth impression in terms of believability. Here, cues such as blur, contrast, brightness, color or aerial perspective are considered more useful. Nonetheless, with comprehensibility of elements being prioritized, blur and aerial perspective seem unsuitable. Instead, the cues such as color, brightness and contrast, which are inherently part of the design and the display settings, should be used to facilitate the subjective impression of depth. Table 3 provides an overview of the identified cues and their relevance in the context of automotive interfaces.

Table 3

The monocular depth cues and their relevance in the context of automotive interfaces

Category		Depth Cue	Relevance (yes/no)	Reasons to include/exclude	Potential application	N of studies
Relative position		Occlusion	No	Threatens the visibility of essential information	-	12
		Intersections	Yes	Highly effective for precise depth estimations	Navigation	6
		Frame	Yes	Increases the impression of depth	Physical frame	4
		Height in the visual field	Yes	Increases the impression of depth	Position of gauges	3
Light / Optics		Blur	No	Threatens the visibility of essential information	-	8
		Color	Yes	Always present in design	Warning system	4
		Contrast	Yes	Highly reliable cue, beneficial for motion parallax	Higher contrast in front	6
		Brightness	Yes	Always present in design	High display brightness	7
		Shading	No	Inconclusive results and increased search times	-	7
		Cast shadows	No	Potential source of ambiguity	-	5
Motion		Motion parallax	Yes	Observer-produced parallax to match driver's perspective	Head-tracking	13
Perspective		Linear perspective	Yes	Benefits both the depth impression and exact estimations	Converging lines	9
		Aerial perspective	No	Insufficient information, useful for large distances	-	3
		Relative size	No	Threatens the visibility of essential information	-	5
		Texture gradient	Yes	High estimation accuracy if combined with intersections	Navigation	6

As seen in Table 3, there are multiple cues that seem beneficial for conveying 3D effects in the automotive interfaces. In accordance with the additive nature of the monocular depth cues, their combinations should result in the most optimal recreation of depth. However, monocular depth cues might easily result in conflicts. It is therefore essential to understand how the relevant cues interact with each other and what is their most optimal combination for both depth-related performance and the subjective depth impression. Additionally, while the reports on some of the cues are consistent and their optimal implementation is clear, several cues raise open questions. Firstly, it remains unclear, whether exaggerating the amount of motion parallax will carry any benefits. Secondly, the cue of contrast was reported to improve depth judgement in the presence of motion, with its incorrect application leading to confused judgements. Testing the implementation of these cues, their interaction and how they affect the other cues used, seems highly relevant. Lastly, while the cue of intersections results in correct and precise depth estimations, its effect on the subjective impression of depth is not sufficiently explored.

Additionally, following the reports of Broy et al. (2013) on the benefits of placing the stimuli behind the screen level, it might be beneficial to further explore the most optimal positioning of visual elements within that space. This could prove useful for structuring visual elements in a 3D digital instrument cluster based on their importance. This can be achieved for both depth and height, through simply splitting the space into quadrants, as illustrated by Hendrix and Barfield (1995).

To address these remaining questions, as well as to test the proposed combination of the relevant cues, an experiment was designed. The following section provides detailed information on the methods used and the obtained results.

Current research

The main goal of this thesis is to provide design guidelines on how to best convey 3D effects in automotive interfaces. To support this, the designed experiment aims to identify the most optimal combinations of the relevant depth cues derived from the review. For this purpose, the cues of intersections, motion parallax, and contrast, together with the depth and height within the 3D space, are tested on both depth-related performance and the subjective depth impression. The combinations that are beneficial for both these aspects will be regarded as the most optimal. Another purpose of this study is to gain deeper understanding of the perception and acceptance of the observer-produced parallax, its potential use in automotive context and the initial user requirements that such an implementation would require. To facilitate this goal, the quantitative methods are accompanied by a qualitative evaluation, where the standardized questionnaire is followed by a semi-structured interview (see Measures).

Method

Design

A mixed factorial design was chosen to conduct the experiment. The design of the study was based on the procedure by Hendrix and Barfield (1995) who used a 2^6 factorial within-subjects design to investigate the influence of both monocular and binocular depth cues on spatial instrument design. The current study employed a 2^5 factorial within subjects design, where participants were additionally randomly assigned to one of the two experimental conditions. Prior to performing the experiment, the experimental group was deceived into thinking that the screen on which the stimuli was presented was a novel autostereoscopic 3D display, while the control group was made aware that that the display was not a 3D but a 2D screen that simulated 3D elements. Such a design allowed for gathering insights on both intra-and inter-subjects level, as well as gaining a better understanding of how prior belief about the nature of the display influences the user interacting with it.

The independent variables that were manipulated within the design were: (1) absence or presence of intersections, (2) absence or presence of contrast, (3) parallax amount – normal or 1.5 times stronger, (4) the depth of the target element – in front or behind the reference element, (5) the height of the target element – above or below the reference cube. To gain understanding of how the given cues and their combinations affect the depth-related performance, three dependent variables were used: (1) time spent on performing the task, (2) depth estimation accuracy, and (3) height estimation accuracy. The subjective impression of depth consisted of two variables: (1) believability and (2) aesthetics. The dependent variables were measured for every task. Additionally, the overall believability of the 3D effect produced by the display was tested after the experiment was completed. This was aimed at differentiating it from the believability of the given cues and their combinations. Next to that, data on the experience with the display and its potential use in the automotive context was gathered by means of a semi-structured interview and a standardized questionnaire (see measures).

Participants

Convenience sampling method was employed to gather data from 40 participants. To be included in the study, participants had to be at least 18 years old, as well as have normal or corrected-to-normal visual- and stereo acuity. All participants met the inclusion criteria, which resulted in a sample of 40 participants, 20 per each experimental condition. The age of respondents ranged from 18 to 38 ($M = 22.15$, $SD = 4.61$), with female constituting more than half the sample ($n = 23$). The two most and equally common nationalities were Dutch (27.5%) and German (27.5%), with the other distinguishable groups being Polish (12.5%) and Chinese

(10%). The nationality of the remaining 22.5% varied strongly. Three male participants were found to be colorblind, but since the experimental setup did not rely on color for depth judgements and their performance during the learning trial was observed as normal, they were included in the sample. 70% of the participants had a driver's license, 50% of which reported a high frequency of driving. Additionally, none of the participants drove luxury cars but two owned executive cars (Audi Q7, BMW 5 Series) and one drove a sports car (Alfa Romeo 4C).

Materials

Prototype software and hardware

In order to easily control the independent variables, a simulation prototype was programmed in *Python* (version 3.7.1), with the 3D functionality being added through the game engine *Panda3D* (version 1.10.09). To transform the virtual perspective according to head movements of participants, tracking functionality was added with the *OpenTrack* software. Tracking was performed with the so-called *ArUco* marker (see for example Li et al., 2021) – a physical marker placed on a baseball cap (see Figure C1), which made it possible to track participants' head movements. To achieve a smooth tracking result the prototype was equipped with a high quality webcam (*Logitech Brio Ultra HD Pro Business Webcam*) which offers a higher than normal framerate (60fps) at full HD resolution (1080p). This made the tracking additionally resilient to the potential change in the lighting conditions. This setup had to be calibrated to each participant once seated in front of the prototype. Here, an adjustable chair proved useful. To not break the 3D effect, the main simulation was displayed on a *Lenovo Thinkpad* laptop and a second screen was used to present the instructions, as well as the input sliders for the experimental task. The continuous interaction between these displays was achieved through wirelessly connecting both the mouse and the *Microsoft Surface Pro* tablet to the laptop running the simulation. To further strengthen the 3D effect and to successfully deceive the experimental group, a wooden black frame was built around the display (see Figure C2). This setup was used for both the control and the experimental group.

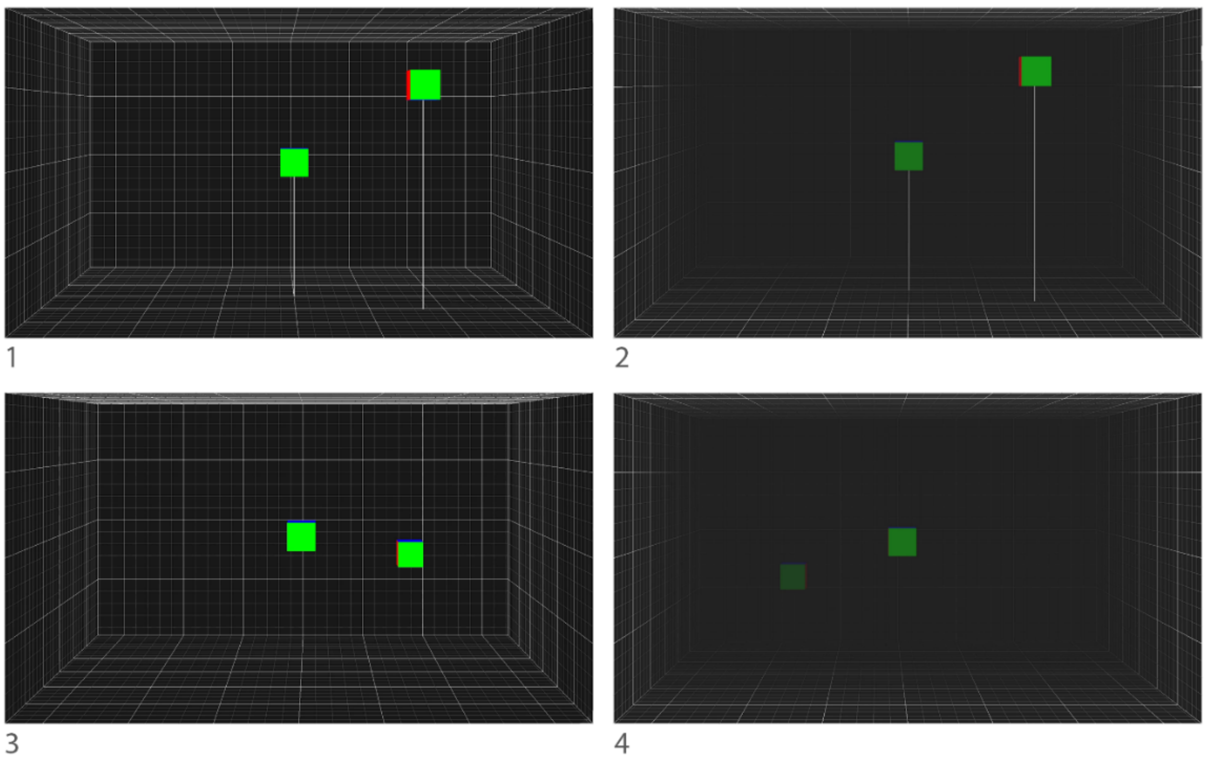
Stimuli

Within a 3D scene, a grid-like texture gradient in the form of a 3D box was used for all the trials to increase both the ease of estimation, as well as the overall depth impression. The texture gradient is a necessary part of the cue of intersections, but having it present only for this condition would result in an advantage over other conditions, as it is an additional source of depth information. To counter this effect, the grid was present for all trials. Two cubes were added to the scene, one of which was a reference cube placed in the middle of both depth and height within the scene. The other was a target cube and its position in the scene varied across trials. To ensure equal representation of positions in all four parts of the screen, as well as to minimize randomness across different conditions, equal increments were created in front and behind-, as well as above and below the reference cube. The cube was then randomly positioned at the specified distance from the reference cube either on its right or left side. The resulting combinations of positions and other factors were always the same, but their order was randomized per participant. For the increased parallax condition, the same amount of head movement corresponded to a 1.5 times higher parallax transformation within the scene. For the condition of intersections, droplines were created to connect the cube to the ground plane and a decrease in contrast over distance was achieved through a fog simulation (see Figure 18 for the examples of possible combinations).

During each one of the 32 trials, participants had to explore the scene from different perspectives by moving their head in order to estimate both the depth and the height of the target cube, as well as provide subjective ratings of believability and aesthetics. To obtain this input, a set of four sliders was created (see Figure 19). To allow for accurate estimations, the input sliders for both depth and height of the target cube included reference lines corresponding to the thick grid lines within the scene (see Figure C3).

Figure 18

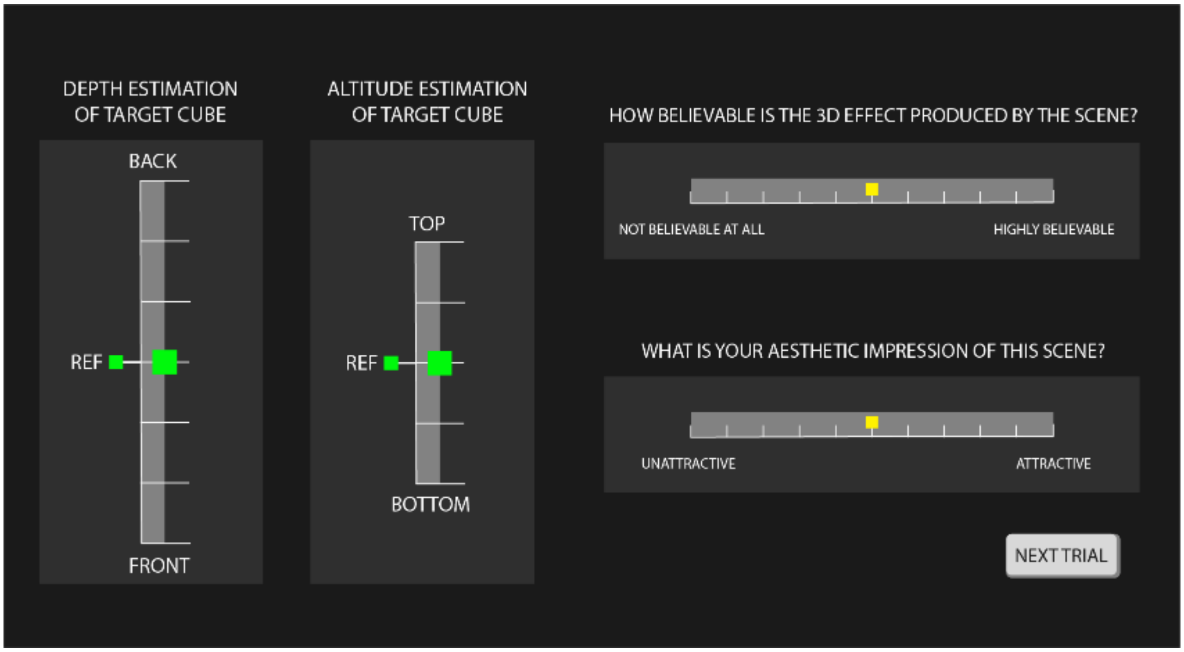
Example stimuli of cue combinations used during the experiment



Note: The first design (1) includes the cue of intersections, and for (2) the fog is added to recreate the proximity dependent change in contrast. Design (3) has no cues but shows how the cube is placed in the lower-back part of the scene. The cube in (4) is in the same quadrant but with the cue of contrast present.

Figure 19

User input screen with sliders for depth and height estimation and the subjective assessment.



Measures

The above mentioned sliders (Figure 19) provided quantitative data for both depth and height estimation. The sliders' length corresponded to the depth and height of the scene accordingly and provided results in millimeters. The input number is represented by how much the estimation was off, so this measure can best be construed as a degree of error. When it comes to the subjective assessment, both the believability and aesthetics sliders were a recreation of an 11-point Likert scale. Similarly, the overall believability of the depth impression within the display was assessed via an 11-point Likert scale. This was to see if the overall believability of the effect differed from those given per trial. Furthermore, the perceived quality of the interaction with the display was assessed by the *User Experience Questionnaire* (Schrepp et al., 2017), which utilizes a 7-point Likert scale and consists of 26 items comprising dimensions of attractiveness, pragmatic-, and hedonic quality (see Appendix D). Lastly, a semi-structured interview was constructed to perform a qualitative assessment of users' visual preferences, their experience with the display, the level of acceptance towards it being used in the automotive context and the potential user requirements (See Appendix E).

Procedure

Prior to performing the experiment, participants signed the informed consent form, which familiarized them with the goal and structure of the study, the involved audio recording of the interview, as well as the confidentiality and anonymity of their data. On obtaining written consent, participants filled in a questionnaire asking for their demographic information, car usage, as well as potential sight or depth vision issues. They then performed the Freiburg Vision Acuity and Contrast Test (FrACT, version 3.10.5) followed by the Ishihara colorblindness test (https://www.colorlitelens.com/ishihara-test). Upon finishing both of these, they entered the experiment room which contained the display setup. Depending on their random assignment, if they were a member of the experimental group, they were then deceived into thinking that the display was a new type of an autostereoscopic screen. The control group was told that the display was a regular 2D screen simulating a 3D view. The baseball cap with a tracker was placed on the participant's head and the setup was calibrated. Participants were familiarized with the dual-screen setup and were presented with instructions displayed on the additional screen. If needed the task was explained and a set of three learning trials was provided. If the task was clear and performed correctly, the participants started the experiment. Upon completing all the 32 trials, they filled in the User Experience Questionnaire. For this assessment, they were asked to focus on their interaction with the display and not the experimental task. Lastly, a spoken interview took place and was audio recorded. The whole procedure took on average 50 minutes.

Data analysis

In the initial stage of analysis, the demographics were examined using SPSS (version 26.0.0.0). The obtained performance data were handled and analyzed in RStudio (version 1.4.1106). The analysis was performed with linear mixed effects models (LMMs), which account for the so-called *random effects*, which makes them effective for data obtained from within-subjects design (Schmettow et al., 2017). However, in the case of this experiment, it was not the participants who were under comparison, but rather the combinations of cues used. This is essential to consider since the study aims to examine design parameters when encountered by people and not the people themselves. Therefore, the analysis follows the *designometric* approach introduced by Schmettow (2021), where the resulting combinations of cues are seen as a unique population of designs and are introduced as a factor. This approach simplifies the selection of the most optimal combination of the monocular depth cues, as it provides clear insights on the effectiveness of the entire design, not the cues in isolation. Such a holistic view in turn makes the derived design guidelines more feasible. All the outcome variables were therefore modeled with the inclusion of both participant- and design random effects dependent on the condition (experimental or control group). In a safety critical context of driving, the design that performs uniformly well on all of the outcomes seems to be the most beneficial. Therefore, a multi-outcome approach was taken, where the designs were presented for all the outcome variables. This was facilitated by radar plots, which allowed to *cherry-pick* (Schmettow, 2021) the best design.

In order to gain an initial understanding of the obtained data, as well as to check if the belief about the nature of the display had any influence on the results, the data for all the outcome variables were initially explored on a descriptive level. The data were also checked on both the participant- and the design-level to assess the differences and levels of variation. The visual exploration was performed with conditional boxplots and density plots. Here the separate cues were also used as factors to visually assess the interaction effects. Such visual exploration allowed for a better understanding of the data and the emerging trends and relations. It was observed that one participant did not provide subjective ratings for any of the trials. The data was removed for these two variables, as well as for time on task (ToT), since the time needed for the task was systematically lower. The depth and height estimation input of this participant was used in the analysis.

To correctly analyze all the outcome variables, proper families had to be chosen to model the data with regards to its type and distribution (see Appendix G). The degree of error for both depth and height estimation, as well as the aesthetics and believability ratings, were modeled through the Beta family with a logit link function applied. Here the data were first rescaled to match the 0-1 bounds. For time on task (ToT), the exgaussian family was used. All the outcome variables were modelled with the fixed-effect condition component, participant-level random effects, as well as the design-level random effect dependent on the condition. The derived posterior distributions and the 95% credibility limits made it possible to select designs that resulted in a statistically worse or better performance or assessment for the given outcome variable. Lastly, the predicted values were used to facilitate the multi-outcome approach and following the proper rescaling and normalization, the results were visualized by radar charts.

Furthermore, to quantify the overall experience and the perception of the observer-produced parallax, the results of the User Experience Questionnaire were analyzed with a designated analysis tool (Schrepp et al., 2017). Next to descriptive outcomes on both item and facet level, the 95% confidence intervals can be obtained for all the mean values. Additionally, the reliability of the obtained results on a dimension level is checked with Guttmans Lambda2, which can provide insights on the different dimensions of the experience and the consistency of the reports within them in the context of this evaluation. Additionally, the tool allows to compare the obtained results with a benchmark data set comprised of 468 unique studies.

Lastly, the data from the semi-structured interviews were analyzed to provide a deeper understanding of how users perceived the interaction with the display and how open they would be to this effect being implemented in the automotive context. Additionally, they were asked if they would describe themselves as *tech-savvy.* This was used to explore the potential relationship between this quality and their understanding and the openness towards the display, especially in the deceived group. The Grounded Theory approach was employed to investigate their visual preferences and potential user requirements. The interview recordings were first automatically transcribed using the Amberscript software (https://www.amberscript.com/en/). Following manual adjustments, the interviews were inductively coded with Atlas.ti (version 9.1.6.0). After initial open coding, the codes were further refined through constant comparison and formed final hierarchies. The final scheme was used to both structure the information, and to identify potential themes when it comes to user acceptance, requirements and design preferences.

Results

The results of depth and height estimation accuracy were based on 1280 unique observations. Following the removal of one participant's data, Time-on-Task, as well as the subjective ratings of believability and aesthetics consisted of 1248 observations. Additionally, the overall believability of the 3D effect was judged by all the 40 participants, resulting in 20 items per each of the experimental groups. For the data gathered from the User Experience

Questionnaire, a single missing value was observed, resulting in 1039 valid observations. Lastly, the interview results were based on 635 meaningful units of analysis that emerged from the data.

Firstly, it seems that the prior beliefs about the display influenced the average performance and subjective ratings resulting in different scores for the two experimental groups. Table 4 presents the population-level effects for all the outcome variables, together with the respective participant-level and design-level random effects standard deviations. The biggest difference between groups is visible for ToT with the regression model estimating the average ToT increase in the experimental group at 3.66 seconds [-0.79, 8.40] $_{CI95}$. Since the lower value is very close to 0 and the upper value is over 8 seconds, there is enough certainty to see this difference as consequential, especially with an average score estimated at 36.94 seconds and considering the high-risk context of driving. The results also suggest that participants are a much bigger source of variance in ToT (SD=6.88) than the different designs (SD=1.36).

Table 4

Population-level coefficients with random effects standard deviations

Measure	Group	Location	CI.025	CI.975	SD_Design	SD_Part
Time-on-Task*	Intercept	36.938910	33.7369023	40.35769	1.3600229	6.875334
	Experimental group	3.663344	-0.7990822	8.40423	0.7150066	NA
Depth estimation error	Intercept	0.0724039	0.0596837	0.0870421	1.528637	1.169532
	Experimental group	0.9638926	0.8290389	1.1226199	1.130676	NA
Height estimation error	Intercept	0.0345228	0.0291328	0.0410872	1.389724	1.288669
	Experimental group	0.8908180	0.7307924	1.0627845	1.107123	NA
Aesthetics	Intercept	1.481226	1.0164072	2.218135	1.470772	2.226313
	Experimental group	1.460323	0.8363639	2.420149	1.177687	NA
Believability	Intercept	2.276695	1.4934406	3.563029	1.213624	2.52585
	Experimental group	1.306812	0.7031064	2.322861	1.198642	NA

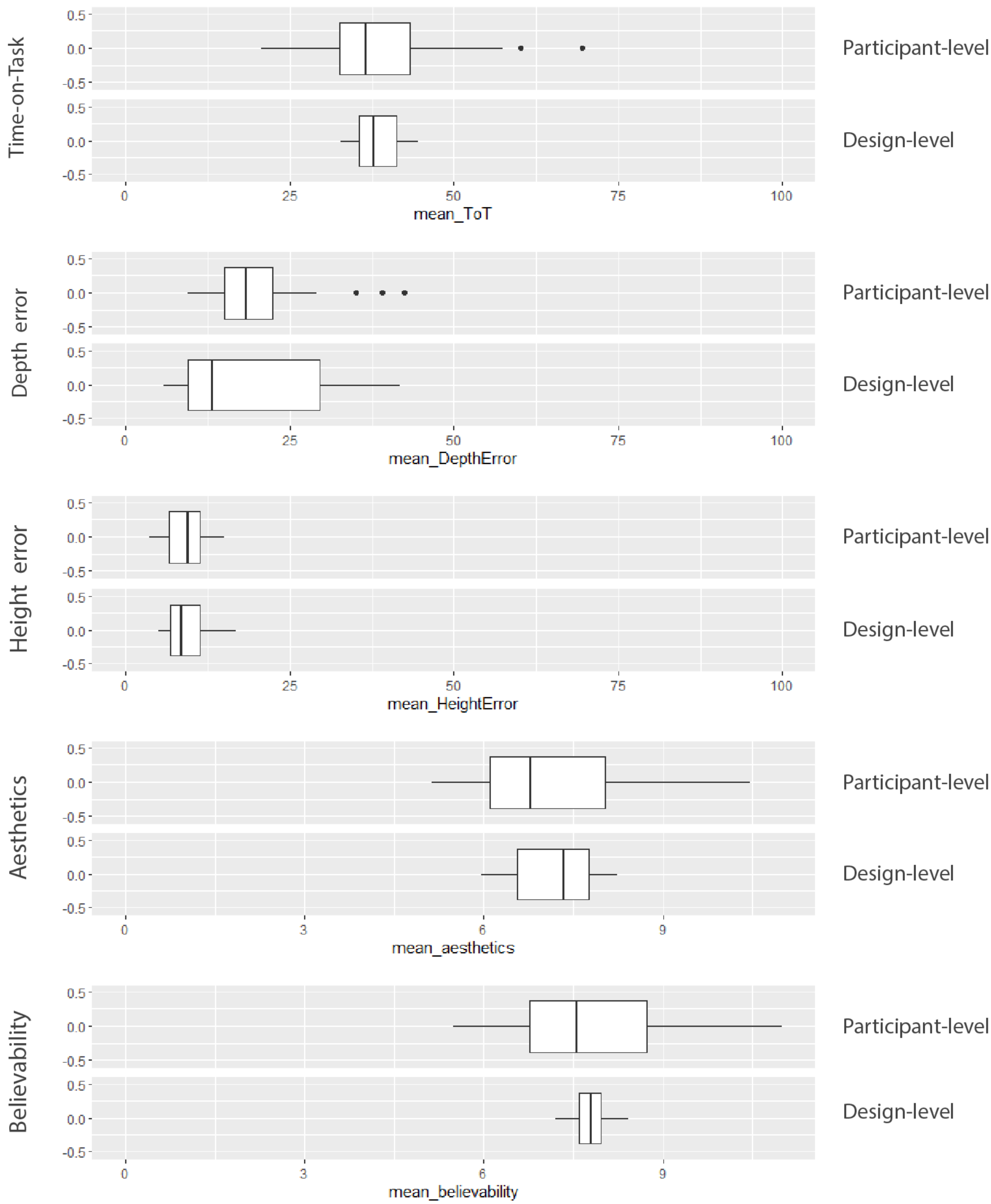
Note: The coefficients for Time-on-Task are in seconds, while all the other coefficients are standardized beta weights and refer to the magnitude of change.

Additionally, it seems that the experimental group gave slightly higher aesthetic ratings with an average increase by the factor of 1.46 [0.84, 2.42] $_{CI95}$. Yet, the credibility interval is rather wide and the lower value suggests a potential decrease by the factor of 0.84. At the same time, there seems to be much more participant-level variation (SD = 2.22) than the variance caused by designs (SD = 1.47), which might be a source of uncertainty. A very similar result is observed for believability, with the experimental group giving on average higher ratings by the factor of 1.30 [0.70, 2.32] $_{CI95}$. Here, again the 95% credibility interval is rather wide and the lower value suggests a potential decrease by the factor of 0.70. Moreover, the participants are again a bigger source of variability (SD = 2.53) than designs (SD = 1.21)

In contrast, there seems to be no clear differences between conditions when it comes to depth ([0.82, 1.12] $_{CI95}$) and height ([0.73, 1.06] $_{CI95}$) estimation accuracy. Even though the average values suggest a slight benefit for the experimental group, it is important to note that the upper limit of estimation error was 230. This puts the average depth estimation error at 16.65 units and the mean error observed for the experimental group at 16.04, which is a rather small difference. A similarly small decrease (less than 1 unit) is visible for the height estimation error, suggesting a lack of substantial differences between the two conditions. However, as shown in Table 4, for both these outcome variables the designs seem to be a bigger source of randomness, than participants. This suggests a strong influence of depth cues on accuracy of depth judgements, especially for the depth estimation, with the variation on the design-level (SD = 1.52) being much higher than on the participant-level (SD = 1.17). For height estimation this effect was less pronounced, with the designs being only a slightly bigger source of variation (SD = 1.39) than the different participants (SD = 1.29). The differences in variance on the participant- and design-levels are illustrated in Figure 20 for all the above outcome variables.

Figure 20

The participant- and design-level variation per each outcome variable

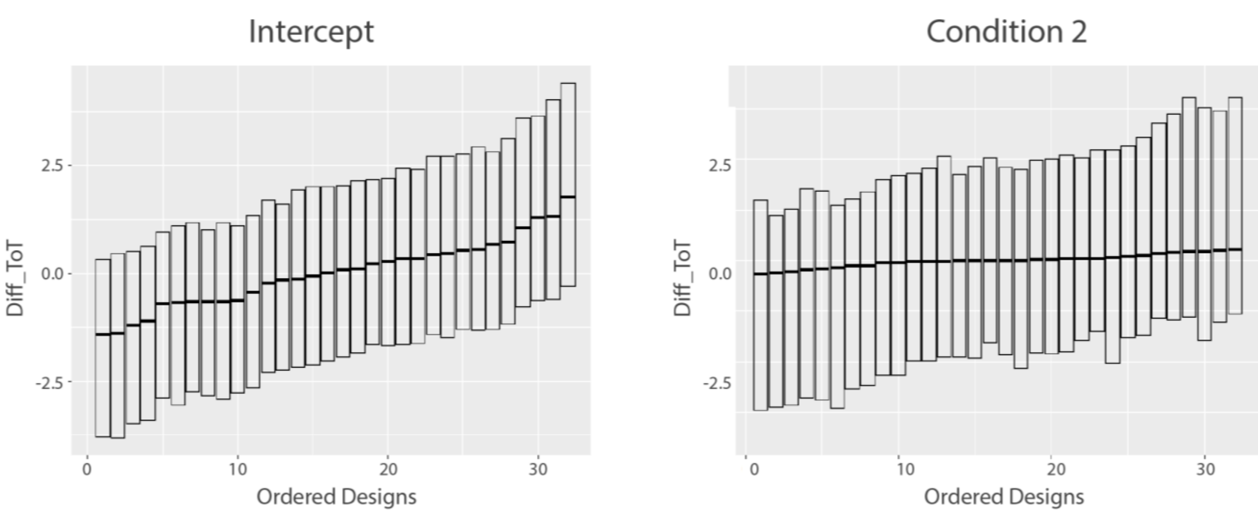


Note. Notice how for ToT and the subjective ratings (aesthetics and believability) the participants are a much bigger source of variation. This is reversed for depth and height estimation accuracy.

As shown in Figure 20, ToT is less varied on the level of designs, than on the level of participants. This lack of differences between designs and their similar influence on ToT values is confirmed on the design-level (see Table F1). However, a few of the combinations stand out as their upper or lower credibility limits are very close to 0 (see Figure 21). For example, ToT had the highest decrease value of -1.44s [-3.81, 0.27]$_{CI95}$ when all the cues were present and when the cube was in the lower-front part of the scene. Similar decrease of -1.43s [-3.80, 0.33]$_{CI95}$ was achieved when the cube was in the lower-back part of the scene in the presence of droplines and increased parallax. On the other hand, the highest ToT increase of 1.77s [-0.19, 4.40]$_{CI95}$ was present for the combination of the cue of contrast and increased parallax, when placed in the upper-front quadrant of the scene. Moreover, figure 21 illustrates that despite the central values being highly similar in the experimental group to the intercept estimates, a bigger intra-design level variation is observed.

Figure 21

Caterpillar graph of individual designs ordered based on their center ToT estimates

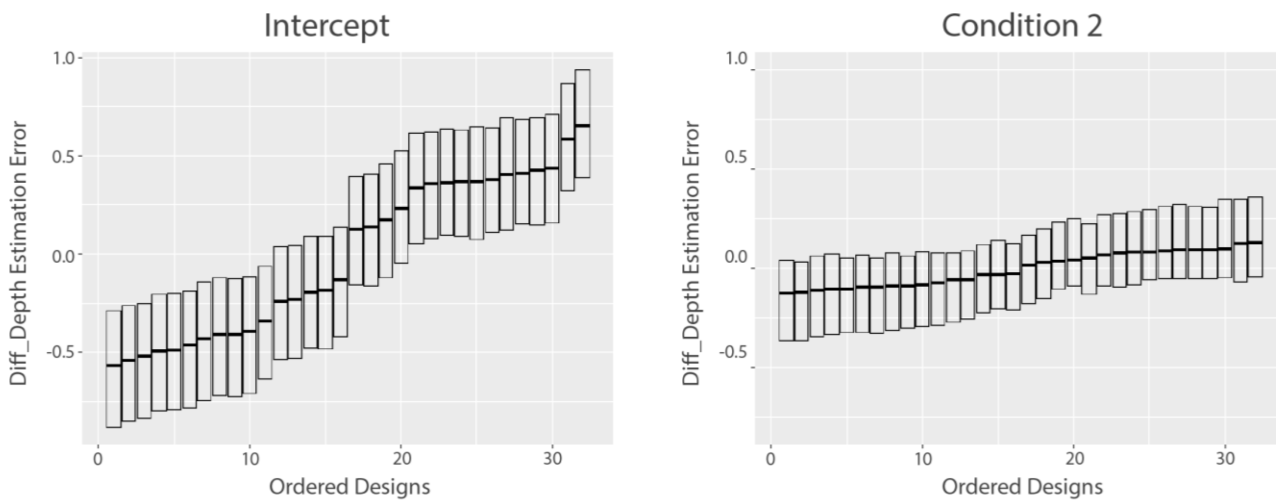


Note. Notice that the decrease in ToT represents a design that results in better performance. The Intercept values show how the designs differed from each other, while the Condition2 values show how ToT differed depending on the condition on the design-level.

When it comes to the error of depth estimation, the designs were a bigger source of variance, than the participants. Figure 22 confirms the differences in depth estimation error on the level of designs.

Figure 22

Caterpillar graph of individual designs ordered based on their center Depth Estimation Error



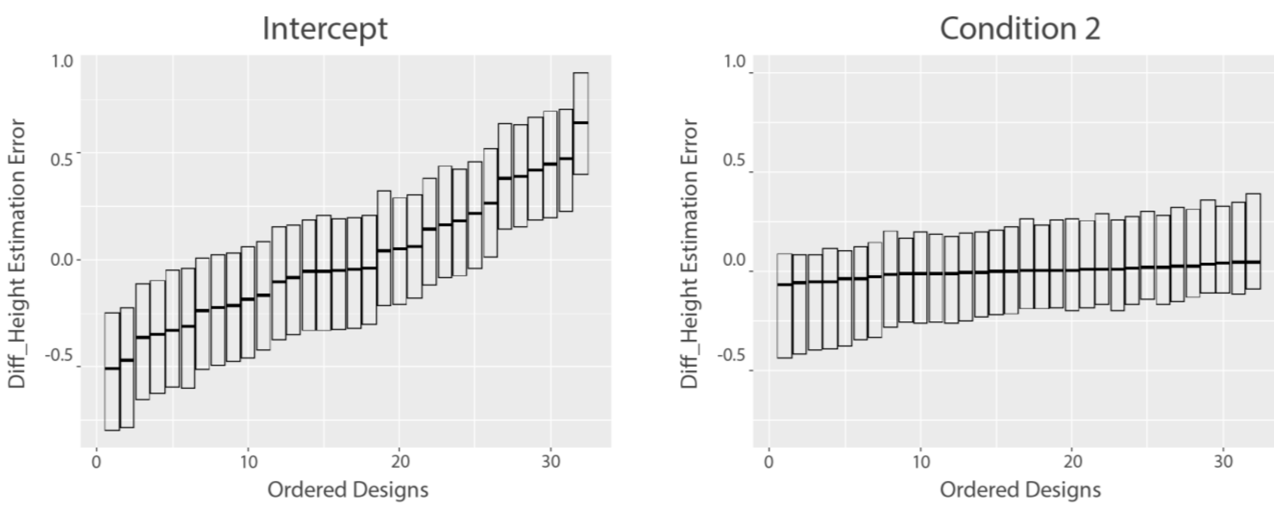
Note. Similarly to ToT, a decrease in depth estimation error represents a better performance. Here, the Intercept values show high inter-design variation. Additionally, there are slight differences in the depth estimation error for the experimental group with several designs showing better or worse performance and with the intra-design level variation being decreased.

Next to Figure 22, the results of the regression analysis (see Table F2) show that the accuracy of depth estimation varied strongly depending on a design. The best design resulted in a degree of error lower by the factor of 0.57 $[0.42, 0.75]_{CI95}$, which represents almost a 50% increase in performance. The elements present for this design were the droplines (cue of intersections) and the fog (cue of contrast) and the cube was placed low in the back of the 3D space. This was followed closely by the same combination of cues but with the cube brought to the front. This design had a decreased degree of error by the factor of 0.58 ($[0.58, 0.77]_{CI95}$). The least effective design for depth accuracy with an increase in a degree of error by the factor of 1.93 $[0.42, 0.75]_{CI95}$, was when the cube was present in the upper-front part of the scene in the absence of the monocular depth cues. Moreover, the intersections were the only cue that in isolation decreased the error by the factor of 0.83 [0.62, 1.10]$_{CI95,}$ but as can be seen in the upper value of the 95% confidence interval, the effect is not certain. Lastly, the posterior distributions of all the designs (see again Table F2) revealed a trend, where the increase in the number of monocular depth cues resulted in improved accuracy of depth judgements.

Similarly to depth estimation accuracy, the differences between designs and their influence on the accuracy of height estimation were confirmed in the results of the regression analysis (see Table F3). Again, no differences were apparent between the two experimental groups, but the performance differed greatly on the level of designs (see Figure 23).

Figure 23

Caterpillar graph of individual designs ordered based on their center Height Estimation Error



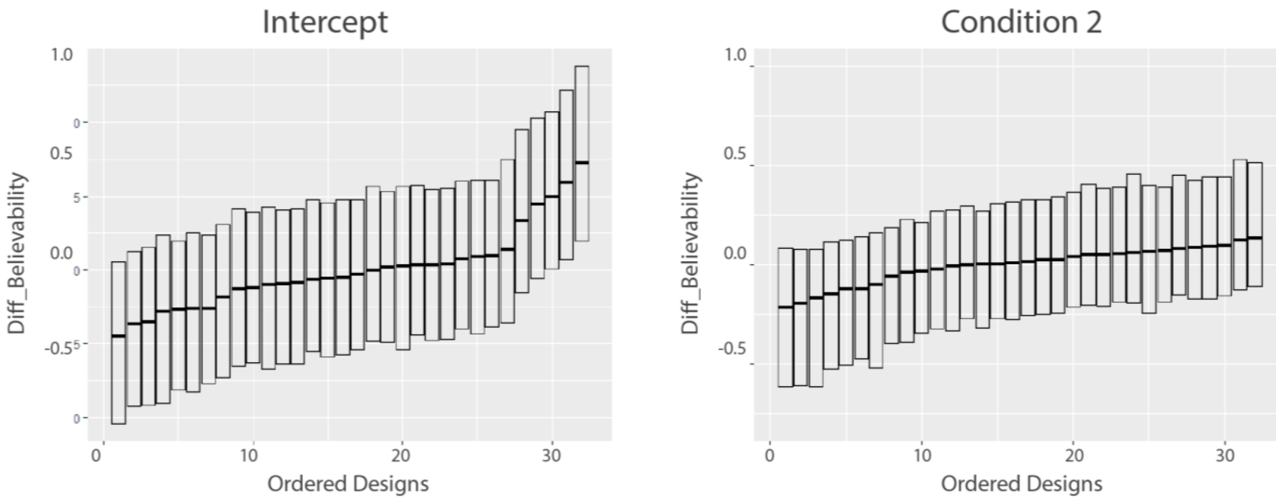
Note. Similarly to depth estimation accuracy, height estimation differed depending on designs. Again, no strong differences are apparent for the experimental groups on the design level.

As can be seen in Table F3, the best design consisted of the increased parallax with the cube positioned in the lower-back part of the scene. This combination decreased the error of height estimation by the factor of 0.59 [0.44, 0.78]$_{CI95,}$ representing over a 40% increase in performance. This was followed closely by the normal amount of parallax and the cube present in the lower-back quadrant, with a decrease of error by the factor of 0.62 [0.45, 0.79]$_{CI95}$. Lastly, a higher amount of parallax was also useful when the cube was in the upper-back part of the simulated scene, with a decrease factor of 0.69 [0.52, 0.88]$_{CI95}$. However, seven designs were found that substantially harm the height estimation. The least effective combination consisted of intersections, contrast, and the increased amount of parallax, with the cube present in the upper-front part of the scene. This design increased the degree of error by the factor of 1.89 [1.48, 2.38]$_{CI95}$. The same combination of monocular depth cues with the cube positioned in the upper-back quadrant resulted in an increase in the degree of error by the factor of 1.58 [1.24, 1.99]$_{CI95}$, Finally, from the posterior distributions (see Table F3), it appears that the contrast cue in the form of fog is present in six out of all the seven badly-performing designs, while not being a part of any of the combinations which benefited the height estimation performance. Additionally, the trend observed for depth estimation, where an increased number of cues was beneficial, is not present for the accuracy of height estimation.

As for the believability of the 3D effect, Figure 24 shows relatively equal ratings between designs, with only several designs increasing or decreasing the believability to a high extent. Furthermore, the believability of unique designs shows high intra variation. These observed results are confirmed by the posterior distributions of the regression model (see Table F4), which shows a very small number of designs that offered a substantial advantage, with no designs that would negatively impact the believability.

Figure 24

Caterpillar graph of individual designs ordered based on their center Believability values



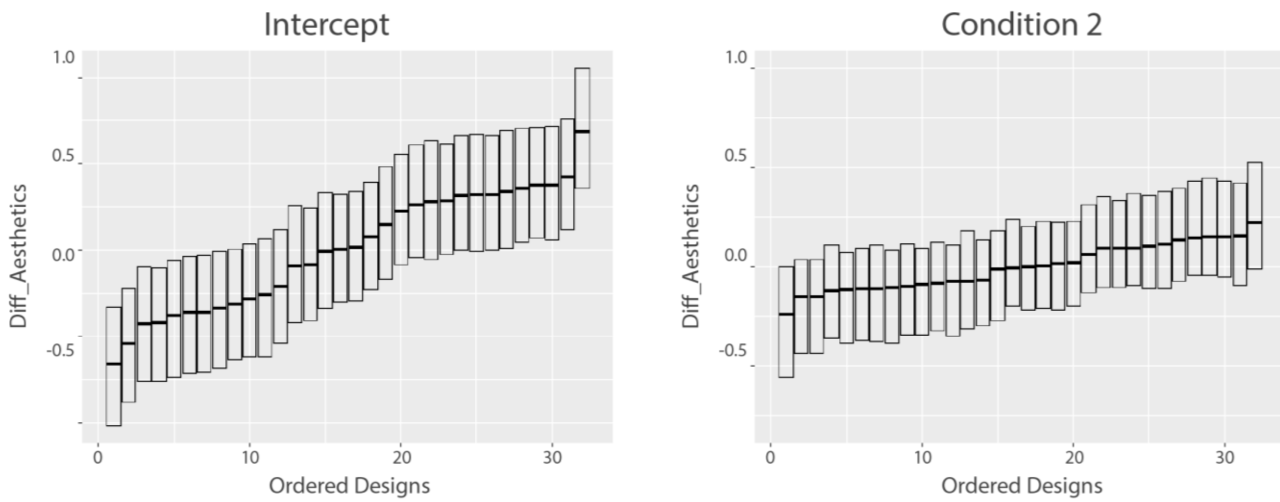
Note. Several designs seem to differ in the resulting believability ratings depending on the experimental condition.

Only two combinations were found to increase the believability of the impression of depth. The most optimal one, with the believability increase by the factor of 1.45 [1.12, 1.98]$_{CI95}$, consisted of droplines and the cube was positioned in the lower-front part of the scene. The second design that improved the believability ratings by the factor of 1.35 [1.05, 1.85]$_{CI95}$, was the combination of droplines with the increased amount of parallax, and with the cube placed in the upper-front quadrant. Surprisingly, the contrast cue in form of fog had a negative effect on the perceived believability when the cube was positioned in the back of the scene and decreased the believability by the factor of 0.83 [0.62, 1.06] $_{CI95}$.

As opposed to believability, the designs appear to be a source of high variability in aesthetic judgements. Additionally, a higher design-level randomness is visible between conditions (see Figure 25). As many as seven designs have been identified as beneficial for aesthetic judgements, with eight combinations greatly harming the rated aesthetics of the scene (see Table F5).

Figure 25

Caterpillar graph of individual designs ordered based on their center Aesthetic values



Interestingly, the most optimal cue combination was also the most beneficial one for the believability of the recreated depth. This design includes the cue of intersection, the regular amount of motion parallax, and the cube is positioned in the lower-front part of the scene and has resulted in an increase in the mean aesthetic judgement by the factor of 1.98 $[1.43, 2.84]_{CI95}$. Similarly, the other cue combination that improved the believability, has also positively affected the aesthetic impression, increasing it by the factor of 1.45 $[1.07, 2.04]_{CI95}$. This combination consists of the intersections with exaggerated motion parallax and the cube placed in the upper-front quadrant. Moreover, out of the seven designs that greatly improved the evoked aesthetics, five of them included the cube present in the front of the scene. Furthermore, the cue of contrast was judged as the least aesthetic one when the cube was present in the upper-back quadrant, decreasing the aesthetic judgement by the factor of 0.51 $[0.35, 0.72]_{CI95}$. Moreover, all of the eight combinations that negatively affected aesthetics ratings contained the cue of contrast. However, it appears that moving the cube to the front in the presence of the contrast cue canceled its negative effect ([0.83, 1.63]$_{CI95}$). This suggests that it was not the contrast cue itself that harmed the aesthetic impression, but its combination with the cube's position.

Multi-outcome analysis

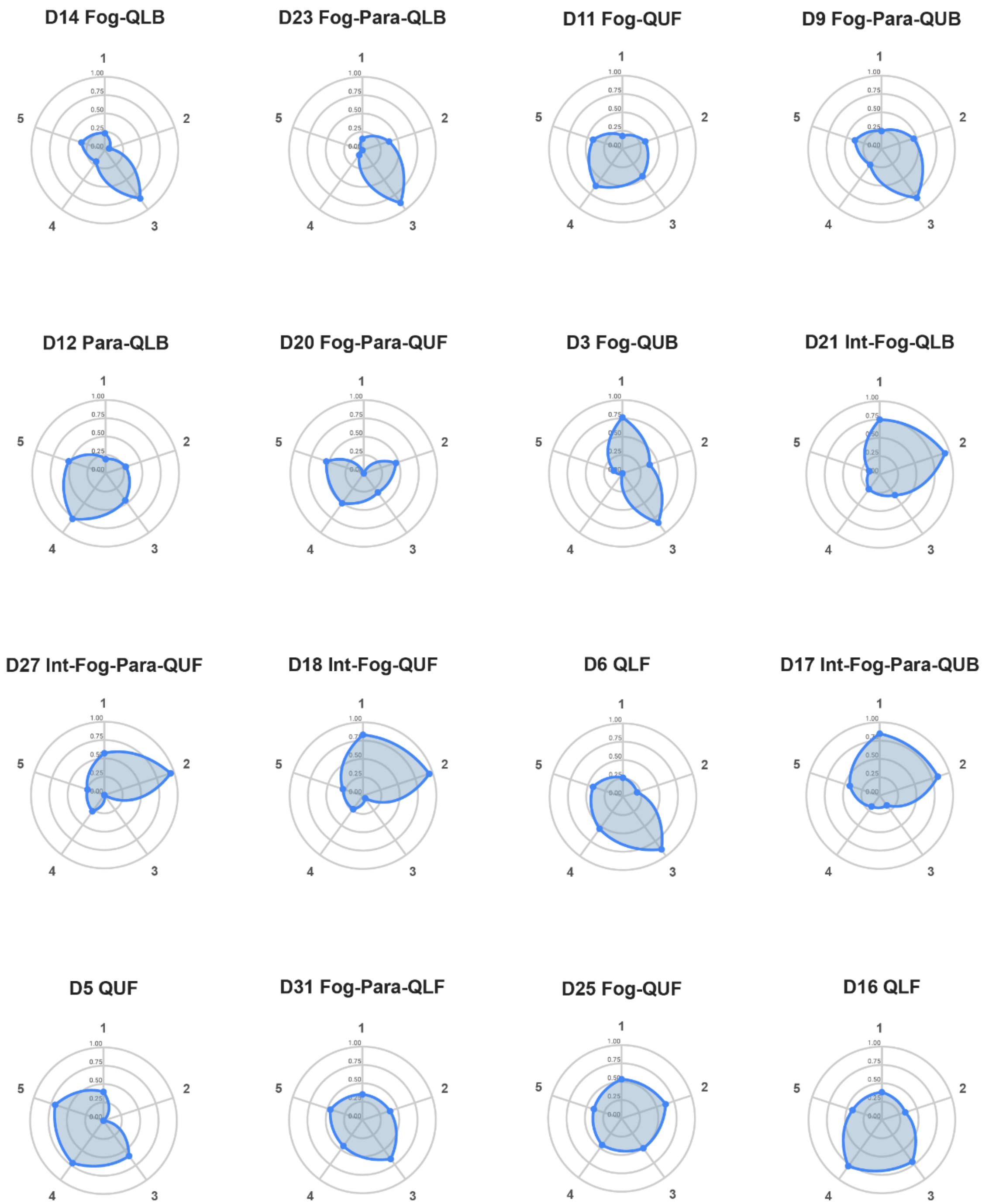
In order to select the design that performs uniformly well across all the above outcome variables, the regression models were used to obtain predicted values per each design for the given outcome. Figure 26 presents the resulting radar-charts. From the charts, it can be seen that the presence of a fog (contrast cue) resulted in the lowest average scores, especially when the target cube was positioned in the back part of the scene (D14 Fog-QLB, D23 Fog-Para QLB, D9 Fog-Para-QUB). However, as can be seen for these designs, the back positioning of the cube resulted in very good height estimation performance.

Furthermore, adding the cue of intersections, even when the fog is present, seems to increase performance. This is especially true for ToT and depth estimation accuracy, as can be seen in designs 21 (Int-Fog-QLB), 27 (Int-Fog-Para-QUF) or 18 (Int-Fog-QUF). Just the addition of intersections to the cube present in the upper-back part of the display (D2 Int-QUB) was already beneficial and stable across all the outcome variables. For the same cue combination, moving the cube to the upper-front quadrant (D10 Int-QUF) shows a very similar performance to D2 Int-QUB, with a slight decrease in the height estimation accuracy.

The decreased height estimation accuracy visible for design 10 (Int-QUF) is most likely caused by the increased distance between the cube and the back wall, the texture of which served as a reference for height estimation. A similar pattern can be seen in design 19 (Int-Para-QUF), where in the presence of intersections and increased parallax, the cube positioned in the upper-front quadrant performs well at all the outcomes except the height estimation.

Figure 26

Designs and their predicted values of all the outcome variables ranked by average score (Part 1)



Factors:
1 - Time on Task
2 - Depth Estimation Accuracy
3 - Height Estimation Accuracy
4 - Aesthetics
5 - Believability

Design abbreviations:
1. Int - Intersections Present
2. Fog - Fog Present
3. Para - Increased Parallax
4. Q - Quadrant in which the target cube was positioned; UB - Upper-back | LB - Lower-back
UF - Upper-front | LF - Lower-front

Designs and their predicted values of all the outcome variables ranked by average score (Part 2)

Factors:
1 - Time on Task
2 - Depth Estimation Accuracy
3 - Height Estimation Accuracy
4 - Aesthetics
5 - Believability

Design abbreviations:
1. Int - Intersections Present
2. Fog - Fog Present
3. Para - Increased Parallax
4. Q - Quadrant in which the target cube was positioned; UB - Upper-back | LB - Lower-back
UF - Upper-front | LF - Lower-front

Note. Designs 30 and 22 seem to be the most stable overall, with the highest subjective impression being achieved through design 24, with design 32 resulting in the most rapid judgements.

Additionally, what stands out in the charts, is how some designs (for example D13 Int-QLB, D28 Int-Fog-Para-QLB and D32 Int-Fog-Para-QLF) are scoring highly on the three performance metrics but have much lower aesthetics and believability scores. All of these three combinations include the cue of intersections and are placed low in the scene. The best performance seems to be evoked by design 22 (Int-Para-QLB), where the cue of intersections is accompanied by the exaggerated parallax with the target cube positioned in the lower-back quadrant. Bringing the cube to the front and adding the fog to this combination seems to result in most rapid judgements (D32 Int-Fog-Para QLF), but a slight decrease in height estimation accuracy. This might suggest that the cue of contrast, when combined with the frontal positioning of elements helps disambiguate spatial relationships and allows for quicker estimations.

When it comes to the subjective ratings, it seems that simply bringing the visual elements to the front of the scene is beneficial (D5 QUF). This is confirmed when looking at all the other designs that scored highly for both aesthetics and believability (D19 Int-Para-QUF, D24 Int-QLF, D26 Para-QLF and 30 Int-Para-QLF), as for all of them the cube was positioned in the front of the scene. The best subjective scores for both aesthetics and believability are achieved by design 24 (Int-QLF), which is composed of the cue of intersections and a normal amount of parallax, while the target cube is positioned in the lower-front part of the 3D scene. In contrast, the worst subjective ratings were given when the cue of contrast was combined with the cube placed in the back of the scene (D3 Fog-QUB, D21 Int-Fog-QLB, D23 Fog-Para-QLB).

It is apparent from the charts that design 22 (Int-Para-QLB) and 30 (Int-Para-QLF) perform uniformly well at all the outcomes, and as such represent the most stable combinations. Both of these designs consist of the cues of intersections and high parallax amount, with the cube positioned in the lower-back quadrant for design 22, and the lower-front part for design 30. Even though, the latter combination shows slightly lower performance results, the subjective rating increases when the elements are brought to the front part of the scene. The lower performance caused by this change could again be explained by the higher difficulty of height estimation, as the separation from the grid-like texture of the back wall increases. The other two valuable combinations seem to be designs 24 (Int-QLF) and 32 (Int-Fog-Para-QLF). Even though design 24 (Int-QLF) seems to result in relatively higher ToT it has the best subjective ratings. In contrast design 32 has the most rapid judgements (Int-Fog-Para-QLF). These two designs differ only by the addition of the cue of contrast and increased parallax in design 32. Adding those elements, even though it decreases the subjective ratings, might be beneficial where the speed of information transfer should be prioritized, for instance in an emergency warning. However, when it comes to achieving a balance between the performance and subjective ratings, design 30 (Int-Para-QLF) seems to be the most optimal combination.

Overall believability

The average overall believability of the 3D effect produced by the device (M=9.25, SD=1.06) was relatively high compared to the mean believability estimated per trial (M=7.8, SD=1.96) and had more consistent results. Additionally, no significant differences were found in the results of the two experimental groups ($t_{(38)} = 0.91$, $P = 0.769$), with the control group having a slightly higher average (M=9.3, SD=1.17), than the experimental group (M=9.2, SD=0.95).

The User Experience Questionnaire

Overall, the results of the User Experience Questionnaire (UEQ) indicate a positive experience with the simulated 3D effect. The Attractiveness of the display was the highest scoring dimension (M=1.61), followed by the Pragmatic Quality (M=1.54), with the Hedonic Quality having the lowest, but still a relatively good score (M=1.24). The positive evaluation is in a way unexpected, considering that the experimental task was highly repetitive and, as often reported by participants, at times difficult, boring or even annoying. It is therefore surprising that the scale of Perspicuity (ease of use) was the highest scoring one (see Table 5).

Table 5

Results obtained from the UEQ and the confidence intervals for all scales

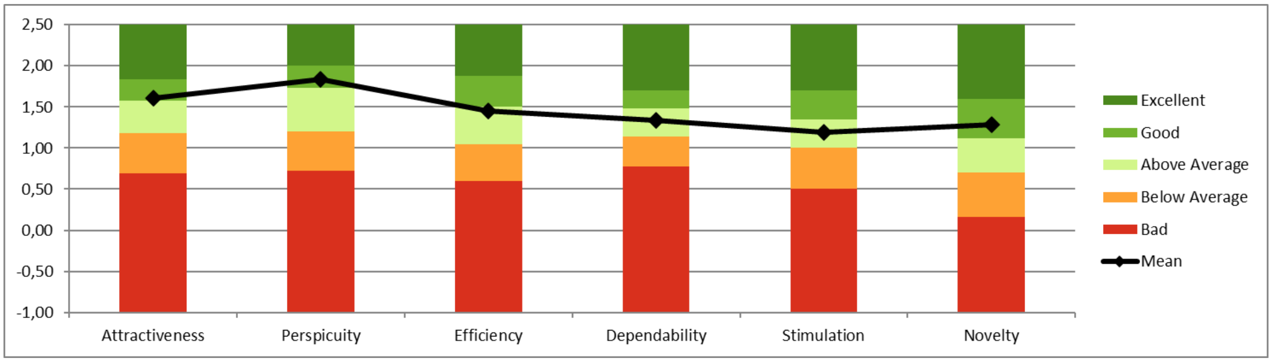
UEQ Scale	Mean	Sd.	N	Confidence	95% CI	
Attractiveness	1.613	0.686	40	0.213	1.400	1.825
Perspicuity	1.838	0.802	40	0.248	1.589	2.086
Efficiency	1.450	0.829	40	0.257	1.193	1.707
Dependability	1.338	0.609	40	0.189	1.149	1.526
Stimulation	1.188	0.751	40	0.233	0.955	1.420
Novelty	1.288	0.722	40	0.224	1.064	1.511

Note. The UEQ transforms the obtained results to the bounds of -3 and 3, with the values below -0.8 considered a negative assessment, while a positive evaluation is observed above 0.8, with average values above 2 considered extremely rare.

However, as we can see in Table 5, the scale of Stimulation scored the lowest, which might suggest the influence of the experimental task. This scale included items such as *boring-exciting*, or *motivating-demotivating*, which participants might have attributed to the performed task. Nonetheless, all of the scales were assessed positively and according to the benchmark assessment (see Figure 27) the relative experience with the display was good or above average. Considering the simplicity of the simulation and the potential influence of the task, these results are promising.

Figure 27

Benchmark assessment of the obtained results

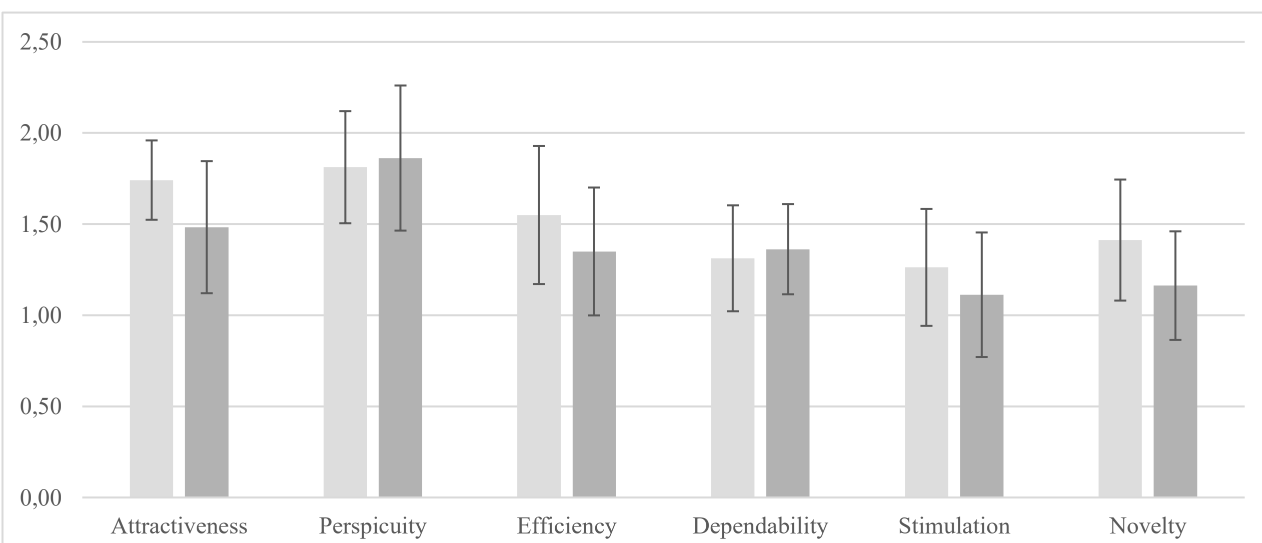


Note. The results are compared with the results of 468 studies with different types of products.

Furthermore, there were no statistically significant differences between the two experimental groups and their assessments on all the facets. However, an interesting trend was observed (see Figure 28), where the experimental group gave slightly higher ratings for the attractiveness, efficiency, stimulation and novelty of the display while judging the perspicuity (ease of use) and dependability dimensions slightly lower than the control group. This might suggest the influence of prior beliefs on the perception of the device.

Figure 28

Results per each scale for the experimental (left) and the control (right) group



Interview results

The interview data revealed that the experimental task was predominantly perceived as *difficult at first, but easy once you start*, as well as *a bit too long and repetitive*. However, subjects also enjoyed how it made them interact with the produced 3D effect, for example: *I liked it. I think it's a great task to encourage someone to use it, so to work with the depth it includes*. Additionally, 50% said they would describe themselves as tech-savvy (60% in the control group and 40% in the experimental group). These results were used to better understand the potential influence of both the experimental task and the technical knowledge on the assessment.

Overall experience with the display

The interview data revealed that the overall experience evoked by the interaction with the display was positive and no differences were found between the two experimental groups. While most participants referred to it as *really interesting* (27.5%) other common descriptions included: *really good* (25%), *cool* (17.5%), *unique* (17.5%) , *fun* (12.5%), *enjoyable* (10%), *exciting* (5%), or even *futuristic* (2.5%). Even though the visual stimuli was simplified for experimental purposes, the observer-produced parallax provided a rich experience, as illustrated by one of the accounts: *It was fun. It was good. It was definitely good. Yeah, it was great. It felt like it was something completely new, something that I've never actually seen before. (...) It's like, oh, it's like another world in a way. Yeah, it's like a whole other world*. Among the contributors to their positive experience, the participants often named the novelty of the interaction (37.5%) and the display's responsiveness together with its smoothness and consistency (25%), as in: *It was responsive in the way I expected it, you could get predictable results. You knew that moving to the right would make you see more of the object's side. So, yeah, I think it was very pleasant and predictable*. Some subjects even claimed that the setup was *thought-stimulating* (5%), *creative*

(5%), and *smart* (2.5%). They also appreciated that it was *easy to use* (7.5%), *user-friendly* (5%), *pleasant to look at* (5%) and felt *natural* (10%) in the way it reacted to the user's movement. No negative terms were used to describe the experience, which also shows that users were able to differentiate it from the experimental task.

The believability of the 3D effect

When it comes to the overall believability of the 3D effect, the majority of participants (77.5%) reported that the 3D effect was never hindered, nor broken. The remaining group mentioned several factors that harmed the 3D effect. Most commonly it was the limited camera angle at which the tracking worked (15%), as well as the glitches that sometimes occurred while tracking the movement (5%). Additionally, at the start of one session, the tracking stopped completely with a sudden change in lighting conditions (2.5%). Nonetheless, one participant reported that the glitches did not diminish the produced depth: *But it was still there. The 3D was always there. I constantly had this feeling of 3D in there, even if it was lagging I still saw the cube from the side at that point. So I could see that it was 3D*. This was confirmed by other participants: *There were, of course, some parts that were more believable than others. But I could see the three dimensional effect in all of them and I can't think about any cases where I didn't have the feeling that it was 3D. So for me, it worked I think all the time*. It appears that, even though the three-dimensionality was almost always perceived, the believability of the 3D effect differed depending on the manipulated visual aspects. One participant reported that in the presence of higher contrast and the objects positioned in the front, the 3D effect was more convincing, yet: *It was always convincing enough. I never had the idea that I was staring at a 2D picture, so that's good*. These reports speak to a potential separation of something being perceived as a 3D object, or a 3D scene, from the believability of the produced depth. In the light of these reports, the latter might pertain to the speed and ease with which depth is perceived. They could also indicate the strength of the observer-produced parallax as a cue to depth, which made the spatial 3D relations visible even in the absence of other monocular depth cues. This possibility is supported by the account of one participant who stated that the presence of droplines, which made the estimation quite easy and did not require additional movements, *felt like it gave it a bit away, so that you look less at the 3D effect*. Lastly, it was often mentioned by the participants that the realism of the produced depth was not important to them and that the 3D effect and the observer-produced parallax were enough to make the depth believable. This can be illustrated by the following comment made by one participant: *I kind of expect it to be some kind of high-tech stuff and you know when you play the VR games, I hate VR games with motion blur, whatever. I don't need that kind of level of reality, you know*.

Furthermore, the majority (N=14, 70%) of the participants in the experimental group, were successfully deceived and were surprised to hear that the 3D effect was simulated. Interestingly, only two of those participants described themselves as *tech-savvy*. In contrast, those that stated to be *tech-savvy* were often quickly able to realize it was a simulated effect, as in: *I also just assumed it from the set up with a camera, that it filmed me and then I had this tracker on top, which made me assume that it's somehow a computer calculating it*. Next to that, the participants who did not get deceived, often reported that the nature of the display did not matter to them: *As I was getting used to it. I think I realized that's, ok, maybe it's just like a screen, but I didn't think about it much. It wasn't important for me*. Furthermore, no differences were observed between the two experimental groups in their reports on the experience or the visual preferences. However, an interesting difference emerged in the accounts of discomfort. Only one participant in the control group (5%) expressed slight dizziness towards the end, but attributed that to being too short for the experimental setup which resulted in their feet not touching the ground, which when combined with constant movements: *was very much like being in a boat*. In contrast, in the experimental group, five subjects (25%) claimed to have experienced discomfort. While one of them had a headache that started before the experiment, the four remaining ones said that using the display was straining for their eyes. They all claimed that it was caused by the autostereoscopic display technology and that they *don't have that watching a normal screen*. This strongly suggests that their beliefs on the nature of the display influenced their perceived comfort. Yet, as other reasons could be involved, the degree to which the simulated 3D effect causes eye-strain should be further investigated.

Visual preferences

When it comes to the visual preferences regarding the monocular depth cues, the reports suggest a high level of subjectivity in what constitutes a believable 3D effect. Despite the varied and often opposite reports, no differences were observed between the two experimental groups (see Table 6).

The highest consensus was found in the accounts on the cue of intersections, with the majority (70%) stating that the droplines were helpful for depth estimation and made it more instantaneous*: I especially liked the task when the wire went down, because then it's very easy to indicate how far the depth is, especially when you have the ones with the lines, the depth is very easy to see*. The positive influence of droplines on believability was echoed by another participant: *I think the dropline always made it a bit more 3D for me, which was maybe because it was also more surface, so the reference got better*. This could explain why the cue of intersection was a part of the most aesthetic and believable design, according to the subjective ratings provided per trial. However, 35% of those that appreciated the cue of intersection, mentioned that, despite being useful, it made the effect less aesthetic. Lastly, only two subjects disliked the cue altogether and preferred when the objects were floating in space.

Table 6

The overview of visual preferences and their distribution within the experimental groups

		Reported preference (N participants)		
Depth Cue	Grouped by	Disliked	Neutral	Preferred
Contrast	Overall	14 (35%)	4 (10%)	22 (55%)
(fog)	Control group	6	2	11
	Experimental group	8	2	11
Intersections	Overall	2 (5%)	10 (25%)	28 (70%)
(droplines)	Control group	1	4	14
	Experimental group	1	5	14
Parallax	Overall	6 (15%)	22 (55%)	12 (30%)
(higher)	Control group	4	11	6
	Experimental group	2	11	6

As shown in Table 6, the cue of contrast achieved through the fog effect resulted in the most varied reports. While over a half of the interviewees (55%) preferred having it on, as much as 35% disliked it. Participants that appreciated its presence mentioned several reasons for its effectiveness. Firstly, it appears that the decreased contrast in the back disambiguated the spatial relationships: *I liked the fog better than the ones that were so bright because it was easier for me to figure out where the cube was exactly*. This was supported by another participant: *Well, I would like the fog as an initial indication of how far the object is, and the line to be more precise*. Another reason mentioned was the natural impression evoked by this cue, as in: *I think when the fog was on, since it's about the sole depth perception and so on and if you look at something further away you can't see it*, which improved the perception of the effect: *The fog was very aesthetically pleasing and helped with making the illusion more believable*.

Additionally, some reports suggest that the fog made the scene more pleasant to look at, as described by one participant: *I liked it way more when it was covered by fog and it was just easier on your eyes*. This was also reflected by another subject: *If the whole scene was very bright, I think you can orient well in it but that might be very distracting to attend to. I think it's exhausting to look at it a long time*. Additionally, one participant who did not like this cue, provided an insightful comment: *So it's actually conflicting because just in general, I'd like it more if the contrast is higher, so the fog effect was annoying me. But at the same time, sometimes it could create a little bit of a calm picture because it's less detailed*. This connects to the ability of this cue to successfully guide attention, as noted by one participant: *I think having the fog can benefit here because it further gives a signal to what's important and what isn't. So if you have something irrelevant in the background and it's foggy then you don't pay attention too it*. However, the effectiveness of this cue also differed depending on the position of the target cube, as one participant put it: *That depended on the position of the cube, so when it all was in the back it seemed really dark, and not so attractive. But when it was in the front it was really bright and that was much better*. The need to have the essential objects in the front in the presence of the contrast cue was recognized by many other participants (29%) that favoured the effect.

Lastly, the participants that disliked the contrast cue, most commonly stated that they appreciated the clarity that was achieved in its absence, as in: *As long as I can see everything clearly, that is more visually appealing to me. Because in real life I expect to see everything. The 3D effect is still there, with or without the fog*. Additionally, the change in contrast often resulted in an impression of lower brightness overall, which was another reason it was considered unsuitable: *But for the fog I just, kind of got the feeling you don't really see anything*. This led one participant to claim that: *When the light was really dim, it was very confusing*, and one more subject even added that: *I had to recheck several times and I can imagine if you're driving, it's distracting, it's just distracting*. These views are in strong opposition to the accounts of those who preferred the cue of contrast, as they often found that: *When it was so bright, it was kind of distracting. It was so bright and so saturated. I couldn't sometimes realize I had to move so much to see if it's closer or further, yeah, I liked the darker feel better*.

In the context of motion parallax, a surprising number of participants (45%) did not notice the exaggerated condition, with one subject stating that *The change in movement, I didn't really notice so it's apparently not very important to me*. This is also reflected in the fact that over half of the participants did not have a preference for this cue (55%). Moreover, those that claimed to have noticed, were often stating that they subconsciously knew something was different but did not realize it was a bigger parallax transformation. The main reason provided for preferring the higher parallax amount was that it was easier to be accurate: *I liked being able to see more when I moved, then I could definitely see more and check the depth of the cube against the sides*. Another added value of the exaggerated parallax was the smaller head movements it required, which one participant saw as beneficial for an in-car display: *I didn't really notice it but as a driver you're kind of fixed to a seat. I think you don't move around as much as what I had to do for the experiments, so I think that you should be able to move around with sufficient change in views, with very small movements, I guess, so more sensitive would be better I think*. In contrast, those that stated to prefer the normal parallax amount said it should feel natural: *The movement just has to be realistic. I don't want to nudge and then just go completely off or have no movement at all*, and that the increased amount felt uncomfortable: *I had a little bit less comfort with it, so it was a bit too fast*.

What was also interesting, is that when asked for potential improvements to the display, multiple participants (20%) said that they would like the parallax amount to be much higher, for example: *Even if there is a little bit more scope in how far you can see it. I would like to see all of it and not necessarily the back, so just the front and all the sides*. Several participants also expressed the need for a higher up-and-down parallax in particular, for example: *Mostly when I was looking up and down with the device, I think I needed to make quite big movements to make it move and I think it's maybe easier if it focuses more on microlevel changes, that I move my head just a bit up and down and that it moves more*. However, these requests might have been closely related to the experimental task.

With regards to positioning within the 3D space, a common view amongst the participants was that, for the most optimal 3D effect, the elements should be placed in the front part of the 3D space, close to the level of the screen. However, two interviewees mentioned that having the object too close diminishes the impression of depth. As one of them accounted: *I'd say once it wasn't very convincing, because probably it was very close so I couldn't see around it. I could only see like a sliver of the sides. So then it's almost as if I'm watching a flat cube, a square*. This is noteworthy, as such a close positioning could then be beneficial for elements which would suffer from a perspective distortion, for example textual information. Moreover, no preferences were given for the height of the elements. Another aspect related to positioning within the 3D space, was the distance between the two cubes, as one participant put it: *It was also about the difference in position. So the greater the difference in position, the more realistic it seemed to be, because if the difference in position was smaller it was harder to see right at the beginning*. Indicating again that the rapid assessment offered by the large separation in space might have increased the realism of the perceived depth.

Overall, the reported visual preferences varied greatly and often represented opposing views. As one participant said: *I find this more of a subjective thing, it really depends on every person*. Even though some trends have emerged, such as the added value of placing elements in the front, addressing the potential differences in how the design is received seems essential in light of the above findings.

User acceptance towards in-car use

The interview data revealed a high level of acceptance towards in-car implementation of the 3D effect, with only one participant (2.5%) voicing strong concerns: *I don't think so because it is kind of an eye catcher. And I think that the kind of displays that we already have, I cannot work with them, it is too much for me. My focus drifts towards that instead of the roads. So for me, it would not work because there are things that are moving and my eye catches those things. When I'm driving and there's a little fly in sight I get very uncomfortable*. The rest of the participants ranged in their openness to the idea, with some being highly positive about the prospect: *Yeah, I think it will work, and I think it's pretty useful too. It's like a step forward in technology, it's really innovative. I imagine really expensive cars having this*, while others voiced some concerns about the display being a potential source of distraction: *I think overall it would be a very good idea. I just don't know with safety. It depends, because some scenes were a bit harder to assess and if you're focusing too much on that, like on the screen, that could cause some difficulties with security. But overall, I think that's a very nice idea to do*.

Interestingly, some participants had doubts about the advantages of the display if they were to use it themselves, but strongly agreed that it would add value if implemented in a luxury car. One participant mentioned that: *I'm wondering if that would make a big difference for me. I think if it wasn't necessarily navigation I don't know if I would really care. I wouldn't know how* *it would really add too much value to my car, because all of the monitors are flat and I don't really care too much about it*, but then went on to say that: *Oh yeah, in luxury cars it would be a good idea! I think in terms of functionality I don't know if it would add too much, but I think in terms of it being a gimmick and in terms of it being like the newest technological advancement and improvement, I think that if you were to add a label luxury car, then having this type of display would make more people want to buy it, but only if it's branded towards luxury cars*. This view was reflected by other participants who all agreed that the 3D effect would add value to a luxury car, mostly because it is: *so futuristic, it's probably also expensive. So I guess in a luxury car, you would feel even more exclusive or it would make it more what others don't have. So I guess the first car brand that would put those 3D screens in would be very futuristic or seen as a modern company. Leading-edge, yeah*.

However, almost one-third of the participants (32.5%) said that the 3D digital instrument cluster would have to offer practical value. Most commonly, they saw this in relation to the driving environment and how the 3D representation of it would benefit their spatial judgement. This included a 3D navigation, which was also the most frequently reported (65%) desired functionality (see Table 7). One participant provided a good summary of the ways in which 3D would add to the spatial assessment while driving: *I think for example in navigation it really adds clarity to a map. If you can see in 3D over the 3D parallax effect where the next exit is going to be that helps. Moreover, if you have, for example, a 360 camera, view of your car and you have it behind your steering wheel, it's obviously a lot more practical when it's in 3D, so you can, not just hear the beeps or see the lines on your backup camera, you can actually get a feeling of the depth. I think that would really help*. Next to the functionalities that connected the 3D with the driving environment, the participant named multiple functions related to the visual aspects of the display itself. The most important reported functionality (15%) seems to be the option to switch back to the 2D version of the display: *For that screen it should just be like a feature that I can use, but I don't really need to if I don't want to. So it shouldn't be something automatically turned on maybe*.

Table 7

The overview of desired functionalities and their frequency of mention

Category	Desired functionality	Reported by N (%) participants
Driver-assistance	Navigation	26 (65%)
	Parking assistance	10 (25%)
	Omniview technology (360° viewing systems)	5 (12.5%)
	Narrow road assistant	4 (10%)
	Adaptive cruise control	2 (5%)
	Backup camera	2 (5%)
	Driver monitoring system	1 (2.5%)
	Intersection assistance	1 (2.5)
Visual aspects	Option to switch back to a 2D view	6 (15%)
	Use depth to support the warning effectiveness	5 (12.5%)
	Use depth to prioritize and structure information	4 (10%)
	Customizable design (e.g. the parallax amount)	3 (7.5%)
	Design matches the interior of the car	3 (7.5%)
	Use depth to reflect the speed of the car	3 (7.5%)
	Use depth to make the fuel level clear	1 (2.5%)
Other	Musical playlist	2 (5%)
	Voice assistance	2 (5%)
	Short tutorial on how the display works	1 (2.5%)

It was also commonly mentioned that the depth should be used to support the visual structure of the elements within it (10%) and that it could be used to support warnings (12.5%) and their saliency: *If the caution light is on for the check engine warning and then it comes more to the front to draw my attention*. Additionally, three participants mentioned that they would like the option to adjust the design elements, such as colors, but also the depth within the display, as well as the amount of motion parallax, for example*: Maybe if you could adjust the depth that you* *want the icons to be on, the colors and maybe it would be kind of smart to, you know how a computer mouse can have like high responsiveness, the same thing with this, because maybe you think that it reacts too much to what you do*. Next to the desired functionalities, the participants provided multiple requirements (see Table 8) that they believed were essential for a successful implementation of the 3D effect in the automotive context.

Table 8

User requirements for successful implementation

User requirement	Reported by N (%) participants
Adding practical value (e.g. navigation, saliency of proximal elements)	13 (32.5%)
No tracking issues	10 (25%)
Does not hinder the normal control of the car	8 (20%)
Does not distract from driving (the effect should be subtle)	7 (17.5%)
Option to switch back to a 2D view	6 (16%)
Clear visibility of essential stimuli	5 (12.5%)
Reliable in all driving conditions (e.g. strong sunlight, night)	4 (10%)
Simplistic design	4 (10%)
Attractive design	3 (7.5%)
Customizable design (e.g. the parallax amount)	3 (7.5%)
Easy to use	3 (7.5%)
Does not require big movements to see the 3D effect	2 (5%)
Accounts for sudden, unrelated movements	2 (5%)
Easy to use	3 (7.5%)
Easily adjusts to the height of the driver	1 (2.5%)

As shown in Table 8, adding practical value was most commonly reported as a requirement (32.5%): *That it adds practical value. That I cannot easily find a way to do better with existing stuff. So in this case, it would be that the effect works, is consistent, and that it allows me to control something with it. If it's not controlling something, if I'm not using it to interact with something, then I feel it's very gimmicky. It might be nice in the beginning, but after a while you get used to this*. This was followed by the reliability of the display and how the 3D effect could potentially threaten safety, which was summarized well by one participant: *Yes, I as* *a driver think it has to be very responsive, the real product, and like I said, I probably don't really care about the practicality but it cannot hinder my normal control of the car, so if it feels normal and the control is smooth and it's just like a cool visual effect. I think that just makes the driving experience more pleasing, so I wouldn't mind that, but yeah, the control has to be good*. The visual aspects of the display were also considered important for safety. Besides the aforementioned option to switch between the 2D and 3D view, the requirements were mostly related to the clear structure of the design, as one participant put it: *I'd like to have all the information presented to me in a very clear way. So, those speed meters and stuff like that, they should be at the front and then if you move around then you could see the map in the background and I guess I don't want it to lose functionality over a normal screen in the car*. Lastly, some participants (10%) valued the simplicity and the attractiveness of the design over the realistic depth representation: *So somewhere realistic, but somewhere better looking, so I would prefer it more better looking because it's not something that you especially need I guess, it's only something extra, I would say so it should also look believable but also attractive I guess, so more attractive than believable*. Another participant even stated that the realism of the 3D effect could be too attention grabbing: *At the same time, you might get too immersed into it in a way. There's lots of sides of it which are, well, very beneficial, for example, for the maps and being able to have a bit more of a real-life situation when driving, but we shouldn't be too realistic so that you can still pay attention to what is around you*. Therefore, as shown in the reports above, putting the drivers at ease and increasing their trust in the display should be achieved by a combination of the optimal design with a flawless technical setup.

Discussion

Two main goals were driving this investigation. Firstly, this exploration aimed at creating guidelines on the most optimal design and implementation of a simulated 3D digital instrument cluster, through improved understanding of the monocular depth cues and their potential to recreate depth on a flat surface. Secondly, the study set out to gain insights into user acceptance of the observer-produced parallax, the believability of the achieved effect and the overall feasibility of this approach. The above experiment aimed at furthering both goals and the findings relevant for each goal are discussed below.

Findings relevant for design

When it comes to the design guidelines, several visual design implications can be derived from the obtained results. However, before discussing the findings, it is important to note how the depth-related performance was often not correlated with the subjective ratings of the impression of depth. This is an important insight, as the believable and aesthetically pleasing recreation of depth on a flat surface seems to differ from the one that produces the most accurate and rapid depth judgements. Understanding which one of these two aspects is more important for a given 3D design and its application, seems essential to ensure its success. As stated in the introduction, in the context of the automotive interfaces, both of these elements are highly important and therefore a balance between all the outcome variables should be achieved. However, considering the interview results which indicate a strong separation between the visual aspects and the practical functionality, it might still be useful to consider them in separation. For example, for some infotainment elements such as navigation or parking assistance, the accuracy and speed of depth judgements should be prioritized, while the aesthetics and believability might play a smaller role. On the other hand, achieving the most believable and aesthetic 3D design of the digital instrument cluster might take precedence over the rapid and accurate depth estimations. This prioritization should remain flexible and if needed the order of importance should be reversed, for instance when a warning is displayed, and the quick and correct transfer of the presented information is more important than the subjective perception of the 3D effect.

Depth-related performance

Firstly, in the context of the most optimal depth-related performance, the exaggerated amount of motion parallax seems beneficial for both accuracy and speed of judgements. This confirms the suggestions of Mulligan (2009) who claimed that higher than normal motion parallax might potentially aid the observer in assessing spatial relationships. Moreover, it is in line with the interview findings, where many participants expressed a desire for an even bigger parallax transformation and perceived it as useful for the estimation task. Additionally, while some saw it as potentially beneficial for in-car use, where drivers should not be required to make big movements, others expressed their needs for a natural transformation. Since many participants did not notice the increased transformation, it could be argued that a slight increase can still feel natural. By how much the parallax amount can be increased before it disturbs viewers, and how the increased value affects drivers' performance remains an open question.

Similarly, the cue of intersections carries strong benefits for performance, especially for ToT and depth estimation accuracy. This is in line with both the review and interview findings and confirms the added value of this cue. Moreover, the effectiveness of intersections for depth estimation accuracy was consistently reported by the majority of the participants and the clear spatial relationships resulting from this cue were highly appreciated. Based on both the regression results and the interview findings, it appears that this cue is the most beneficial one for facilitating spatial judgements and is therefore highly useful for applications such as 3D navigation or parking assistance. However, it should be noted that the dropline used in the experiment does not equal the cue of intersections, which can be created in many different ways, as long as the connection or the intersection point between two objects is visually clarified. The most optimal approaches to achieving this in the automotive UI should be explored.

Surprisingly, the cue of contrast, which was mostly expected to benefit the subjective ratings, highly decreased the time spent on performing the task. This was further confirmed in the interview results, with many participants stating that the fog disambiguated the cubes' relative positions and made the judgement quicker. However, as seen in the regression results and as stated by many participants, the target cube has to be positioned in the front part of the scene to ensure the most optimal use of the contrast cue. When both the cubes were in the back part of the scene, and hence the contrast differences decreased, the cue was not effective and many participants reported difficulties with providing exact estimations. Moreover, such application led some participants to perceive the entire design as darker and less clear. This speaks to the importance of the proper implementation of this cue and is in line with the review findings, that, in the presence of motion parallax, *well-applied* contrast differences help disambiguate spatial relations. This cue should therefore be applied carefully to not decrease the visibility of the important elements and to not darken the design as a whole. Lastly, the interview results revealed that the cue of contrast was often seen as a means to guiding attention and prioritizing the essential stimuli, which could benefit the driving task.

Another observation regards the height estimation and its increased accuracy when the cube was present in the back part of the scene. As mentioned above the most likely reason is the proximity of the back wall with a regular, grid-like texture gradient, the lines of which served as a reference for estimation. Next to confirming the relevance of the regular texture gradients, this finding has two implications: 1) if judgements of relative positions of objects are to be made, their proximity in space should be small; 2) if the cube was in front, the side walls could just as easily be used to estimate its height, yet it appears that the back wall was preferred. This is similar to the *ground dominance effect* for depth estimation reported by Bian et al. (2005) and might suggest that in the presence of other surfaces, the backdrop is a dominant source for height estimation. Putting this finding aside, it seems that the height estimation is less relevant than the speed and accuracy of depth judgement, so bringing the essential elements to the front part of the scene might be more beneficial. This is even more true in light of the higher subjective ratings for elements placed in front of the scene and is highly supported by the interview results, which suggest a strong visual preference for such positioning. However, the added value of a proximal textured backdrop for height estimation could still prove useful for some visual elements, for example, those that indicate levels of fuel or oil pressure, as indicated by one interviewee.

Lastly, it was observed that increasing the number of monocular depth cues improved depth estimation accuracy. While this trend was not observed for height estimation, it is also visible in the posterior distributions of ToT values. To a certain degree, this finding confirms the additive nature of monocular depth cues and by extension - the depth fusion model.

Subjective impression of depth

From the obtained trial results it appears that the ratings of aesthetics and believability were highly correlated. For both of these aspects, the same combinations were rated as the most- and the least optimal. The best impression of depth was achieved when the cue of intersections was present, with the target cube placed in the lower-front quadrant. The success of intersections was unexpected in light of both the review and interview findings, as this cue was established as primarily optimal for depth-related performance, not depth impression and was often mentioned by participants as less aesthetic. However, the ratings suggest that depth impression might be improved simply by introducing more geometric objects to the scene, even such basic and abstract as the droplines, which was also noted by one participant. Therefore, a certain level of complexity within the design of a digital instrument cluster could aid the perception of the 3D effect but, as the interview results suggest, it should be well balanced with the simplicity and subtleness of the design.

Additionally, from the literature, it seemed that more pictorial cues, such as contrast, would be more beneficial for subjective depth judgement. Surprisingly, this cue produced the lowest trial ratings for both aesthetics and believability of the simulated depth. However, as already mentioned this was true only if the target cube was positioned in the back part of the scene, and was hence less visible. Moving the cube to the front in the presence of the fog, negated its detrimental effects, which was also supported by the interviewees. Moreover, the majority of them mentioned that the cue of contrast made the scene aesthetically pleasing and more believable. It was also reported as more natural and pleasant to look at. However, even though this cue was appreciated by more than half of the participants, a large group believed that the decreased clarity of the scene made the 3D effect less believable, which again highlights the need for careful implementation of this cue in the final design.

Furthermore, the trial ratings suggest that placing the elements in the front part of the simulated depth evokes the best depth impression, which finds strong confirmation in the interview results. However, it was also reported that when an object is positioned too close to the surface of the screen, the depth is more difficult to perceive due to its small parallax. As already mentioned, this position might be useful for textual or any other type of information, readability of which could suffer from a perspective transformation. However, this stands in opposition to the review findings (Broy et al., 2013), which indicated that the most essential elements should be placed at screen level, as this position offers the highest visual comfort. Further exploring the most optimal positioning of the gauges could therefore be beneficial.

Moreover, it appears that participants gave highly consistent believability ratings, which suggests that the cue combinations were not the main source of their subjective depth impression. This could also explain the high intra-design variation, which illustrates the high subjective differences in how the participants perceived the overall believability of the produced 3D effects. This was confirmed by the interview findings, with participants describing highly opposing visual preferences. Overall, this might indicate that allowing users to adjust the visual parameters of the design to their liking, would best support their subjective impression of depth.

Lastly, the overall believability of the 3D effect produced by the display was much higher than the average believability ratings obtained per trial. With the mean score of 9.25 out of 11, this result suggests that, although not realistic, the simulated depth was perceived as highly believable. This is confirmed by the interview findings, which revealed that the 3D effect was almost always present, with some participants reporting the 3D to still be visible even in presence of technical issues, such as glitches in tracking. Furthermore, the separation between the three-dimensionality of the scene and the believability of the produced depth is an interesting finding. This disparity is most likely caused by the 3D effect being strongly supported by the observer-produced motion parallax, while the higher believability followed from the addition of other monocular depth cues. This could again be seen as a confirmation of the depth fusion model. Additionally, achieving a high realism of depth, for instance in the 3D navigation, is not necessary for the 3D effect to be believable and appreciated by the users. Especially that some participants even voiced concerns that a realistic representation could be too immersive in the context of driving.

The most optimal and stable designs

Despite the apparent need for separating the depth-related performance from the subjective depth impression, it is still important to find the combination of cues that performs uniformly well for both these aspects. The regression results indicate that the cue of intersections was not only beneficial for depth estimation and the speed of judgement but also highly increased the subjective ratings, even if the latter was not consciously noticed by most of the participants. Similarly, the often undetected increase in the motion parallax is present in the designs that are most optimal and stable for all outcome variables. Consequently, combining the cue of intersections with an exaggerated parallax and the lower-front positioning seems to be the most stable combination (design 30). However, this design is not the most optimal one for any of the outcomes and the performance might benefit from adding the contrast cue (design 32), which increases all the three performance-related ratings, while the subjective impression might improve with a change to a normal parallax amount (design 24). Understanding the changing priorities of the driver in its interaction with the 3D effect and manipulating the design accordingly seems to be the most optimal approach.

Findings relevant for user acceptance

The influence of prior beliefs

Despite the small certainty of the differences between the two experimental groups, the observed discrepancies are worth reflecting upon. This is especially true for the time the participants took to perform the task, with the deceived group taking on average 4 seconds longer, and since the lower value of the 95% confidence interval was close to 0 ([-0.86, 20.16]). This might indicate that being told that the screen is a novel autostereoscopic display caused a certain decrease in trust towards the perceived depth or an expectation of higher difficulty that follows from using novel 3D technologies. This is also reflected in the UEQ results, with the experimental group perceiving a higher level of difficulty of use and lower dependability of the device, as well as the reported eye-strain by four of the deceived participants. However, another reason for their higher ToT could be that they valued the device more and wanted to interact with the produced 3D effect. Especially, that the experimental group also judged the effect as both more believable and aesthetically pleasing. Even though the causes remain unknown, these differences might suggest the influence of prior beliefs about the nature of the display on how it is received. The higher ToT values are especially alarming and should be further investigated in the context of a driving task. Moreover, exploring these issues might be helpful in understanding the most optimal approach to introducing this technology to the user in a way that best facilitates their perception of- and hence their performance with the product.

Acceptance towards in-car use

The results of both the UEQ and the interviews suggest a positive experience with the simulated 3D effect, mostly due to its novelty and the high responsiveness of the perspective transformation. The positive experience with the observer-produced parallax resulted in a high level of acceptance towards its use in the automotive context and particularly in the luxury sector. However, many requirements for successful implementation have emerged mostly when it comes to the reliability of the display. Next to that, participants largely reported that their acceptance of the 3D effect is dependent on the practical value it adds to the driving task and distinguished themselves from a luxury car driver who, according to them, might be satisfied with the purely visually entertaining character. Yet, using the 3D effect to support the desired functionality such as navigation, might be highly appreciated in the luxury sector as well, where the uniqueness and novelty of the infotainment functions increase the vehicle’s perceived premiumness (Sen & Sener, 2020). According to the interviewees’ reports, the 3D effect should not be treated as purely a visual addition, but should instead be used to facilitate the driving task through the most optimized use of the third dimension. This holds true for both the connection with the driving environments, as well as using the depth to effectively guide drivers’ attention. It appears that in order to increase the driver’s acceptance of the product, the flawless technical setup has to meet a simple, yet attractive design, while making use of the possibilities offered by the addition of the third dimension.

Conclusion

This study set out to explore the potential of the observer-produced parallax to convey 3D effects in the context of 2D automotive interfaces. This was investigated from both the perspective of design elements needed to best recreate the impression of depth, as well as the overall feasibility of the approach and the user acceptance towards the in-car use. To better understand the initial design directions, a systematic review was performed to identify the most relevant monocular depth cues and to ensure their most optimal implementation. The review findings have illustrated that multiple depth cues, even though highly effective, might not be beneficial in the context of driving, where the comprehensibility of the visual elements should be prioritized. The relevant depth cues and the remaining open questions were further investigated in a follow-up experiment. When it comes to the most optimal design, the results of the experiment provide strong support for the conceptualized cue of intersections, as well as the benefits of the exaggerated motion parallax, both of which facilitated depth judgements when it comes to performance and the subjective impression. The combination of these two monocular depth cues with a frontal positioning of the essential elements within the scene appears to form the most stable design. However, adding the cue of contrast further increases performance, while changing the parallax amount to a realistic one benefits the subjective depth impression. In light of the evident separation between the performance measures and the subjective ratings, remaining flexible in applying the monocular depth cues seems to be the most optimal approach.

Furthermore, the experiment findings provide strong support for the observer-produced motion parallax and its effectiveness in creating a correct and believable impression of depth on a flat surface. The simulated 3D effect was received positively by the viewers and, despite some technical issues, it resulted in a high level of acceptance towards its implementation in the automotive context. Considering the disadvantages of the autostereoscopic screens, the simulated approach might introduce in-car 3D effects at a lower cost, with higher resolution and without sacrificing the drivers' comfort. However, the final acceptance is dependent on multiple user requirements, which have to be met to ensure a successful implementation of the 3D effect. Taken together, the review and the experiment findings provide multiple insights on the visual and technical aspects relevant for creating and implementing a 3D digital instrument cluster. The section below provides an overview of the key insights, the design guidelines together with the identified user requirements, and outlines the future research directions.

Key insights to drive design and user research on simulated 3D displays

Overall, the study provides support for the monocular depth cues and their effectiveness in reproducing depth on a flat surface highlighting, in particular, the strength of the observer-produced motion parallax. It also shows the importance of exploring the monocular depth cues when it comes to their influence on both depth-related performance and depth impression, as these two aspects appear to be separated. Including both of them in the evaluation of the final design, but also in any research on visual sources of depth judgement seems necessary to ensure the most thorough understanding. Additionally, while testing design solutions in any field it is important to separate the effects of the design options from those caused by the differences in participants. Moreover, it appears that the geometric three-dimensionality of objects differs from the overall depth believability and, as the findings indicate, the 3D effect, even if not entirely realistic, is enough to evoke a positive user experience. The design should therefore prioritize smoothness of parallax transformation and the visual simplicity, clarity, and attractiveness over a realistic depth impression. Instead, the depth should be used to benefit drivers both in their tasks, such as navigating or parking and to effectively guide their visual attention.

Moreover, the study shows how important it is to take a holistic, human-centered approach in investigating a novel in-car solution, as the interview findings strongly benefit the understanding of the obtained trial results, as well as the level of acceptance towards in-car use and the user requirements that have to be met for a successful implementation. The interview reports were invaluable, for example for disentangling the influence of the contrast cue and the actual visual preferences most viewers had. This is especially beneficial in light of the high differences in subjective assessment of depth impression and the small differences in the believability trial ratings. Overall, the study provides multiple design- and implementation guidelines for a simulated 3D effect in the automotive context. Both the design guidelines and user requirements, together with the future research directions for conveying 3D effects in automotive interfaces are outlined below.

Design guidelines

Multiple design insights have been gathered from the literature review, the experiment results and the interview reports. Firstly, based on the review several monocular depth cues were identified as threatening to the visibility of the essential stimuli and therefore not relevant for further exploration of their potential for recreating depth in a simulated 3D display. However, since these cues are widely applied in the design of modern 3D applications, such as VR, it seems important to be aware of how they can hinder the comprehensibility of the essential stimuli. However, their careful implementation might still be useful for certain elements. For instance, aerial perspective, which is considered effective for large separations, might be applied to a navigation map to suggest large distances. Therefore, Table 9 again outlines these depth cues and the considerations that surround their use.

Table 9

The overview of monocular depth cues threatening the visibility of essential information

Depth cue	Considerations
Occlusion	Essential stimuli should not occlude each other to suggest depth Some level of occlusion will always be present The occluded elements should be abstract and not relevant to the driving task Useful for some elements, such as the music playlist while browsing through albums
Blur	Detrimental for comprehensibility of the visual stimuli Should never be applied to essential elements If used, the effect should be applied subtly and only to the abstract elements, such as grid lines
Shading	Increases visual search time when compared with flat stimuli High computational costs
Cast shadows	Potential source of ambiguity Requires more attention to assess the spatial relationships High computational costs Should be replaced by the cue of intersections with abstract representations that connect objects Is computationally costly but seems to offer no advantage
Aerial perspective	Insufficiently explored Effective for large distances Can be used to suggest high separations in the driving environment, for example in a 3D map
Relative size	The essential stimuli should never be made smaller to suggest depth The cue will be partly present in the lines of a texture gradient or through the use of perspective

Additionally, next to the findings on the monocular depth cues, the positioning of the elements in the scene, as well as the distance between them were analyzed and the relevant insights are provided in Table 10. The key insight here is the difference between the obtained results and the reports of (Broy et al., 2013), who claimed that the essential elements, such as gauges should be placed at the screen level, as this position offers the highest comfort. However, many participants mentioned that such close positioning often harmed the believability of the produced depth. The most optimal positioning for the gauges is therefore broadly considered as the front part of the scene, but whether they should be at the screen-level or slightly moved to the back should be further explored.

Table 10

The overview of design guidelines on structuring information in depth

Design guidelines	Source
Do not place any elements in front of the screen level	Review/ratings/interviews
Place textual information at the screen level for the highest comfort	Review
Place the gauges and other essential stimuli in the front of the 3D scene	Ratings/interviews
Structure information into a small number of separate depth layers	Review/interviews
Use depth to support saliency of warnings by bringing them to the front	Review/ratings/interviews
For the most exact estimations objects should be close to each other	Review/ratings/interviews
The optimal amount of depth within the display should be easily adjustable by the driver	Review/interviews

Furthermore, multiple insights have been gathered on other monocular depth cues, which were considered relevant for conveying 3D effects in automotive interfaces. Table 11 provides an overview of the design guidelines that have emerged on their use. These guidelines are a collection of insights from the literature review, the obtained trial ratings and finally the interviewees' accounts.

Table 11

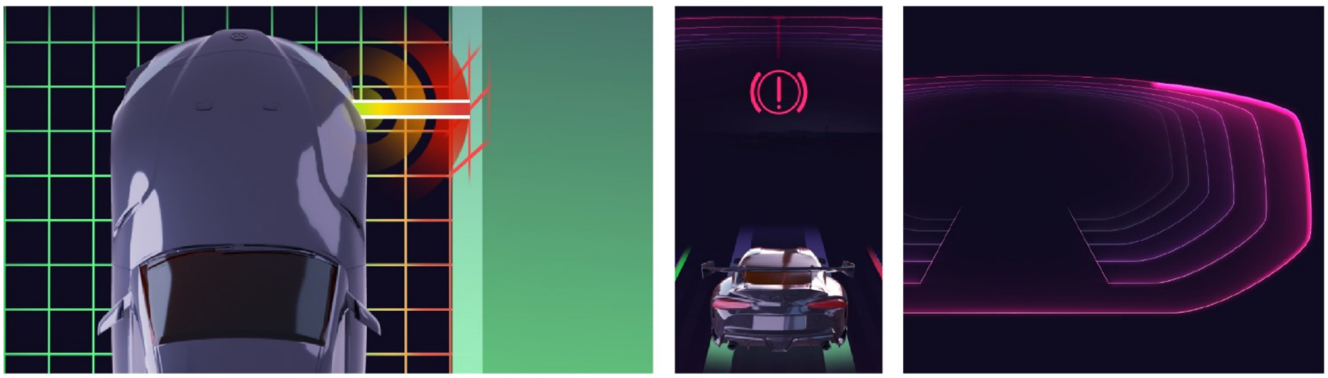
The relevant monocular depth cues and their optimal application

Category	Depth Cue	Design guidelines
Relative position	Intersections	Connect two objects that extend in different directions to disambiguate their relative position Use to support exact depth estimation (for example in 3D navigation, parking assistance) Connect objects to the ground (here the regular texture gradient is a necessary component) Explore the best approaches to introducing this cue into the design, as it also increases the impression of depth
	Frame	Do not use a flat, tablet-like display that is separated from the dashboard Place a 3D digital instrument cluster inside of the dashboard to make it seem as if actual depth extends inside When framing place the display slightly deeper than the dashboard level
	Height in the visual field	Introduce a horizon line to facilitate this cue Elements that are in front should either be placed higher or lower in the visual field This might be difficult since drivers are used to the central positioning of the gauges Make sure that the elements that are further away, are placed more closely to the horizon line
Light	Color	Use color gradients to facilitate the impression of depth Place warmer, brighter and more saturated colors in the front and cooler and desaturated colors in the back Do not use diagonal or circular gradients Use color to communicate urgency (saturated red) and to group elements that are connected with each other
	Contrast	Make sure that the essential elements have high area and texture contrast Do not decrease contrast of any elements that have to stay readable Slightly decrease contrast of the abstract elements in the back of the scene to facilitate depth impression
	Brightness	Ensure sufficient brightness of the display Apply in accordance with the proximity-luminance covariance, where the frontal objects are more luminant Avoid making the objects in the back too dark
Motion	Motion parallax	Increasing the parallax transformation might be beneficial, explore the increase that still feels natural Let the drivers adjust the parallax amount to match their comfort Let the drivers turn the parallax off and switch back to a static, 2D view Consider implementing nested parallax effects (smaller transformation of the gauges but bigger for navigation) Explore the possibility of manipulating the display only when it is looked at
Perspective	Linear perspective	Use converging lines to support the impression of depth, these lines can be suggested by corners of other objects Make sure that the use of this cue does not lead to objects occluding each other
	Texture gradient	Use a regular texture gradient to support both depth impression and estimation accuracy The gradient does not have to be a grid but can be comprised out of parallel lines, be it horizontal or vertical

Several designs have been developed to illustrate the optimal use of the guidelines presented above. Firstly, the cue of intersections can be used not only for spatial assessment, such as parking assist but also to support warnings and to strengthen the impression of depth through the use of intersecting geometry (see Figure 29).

Figure 29

The potential use of the cue of intersections



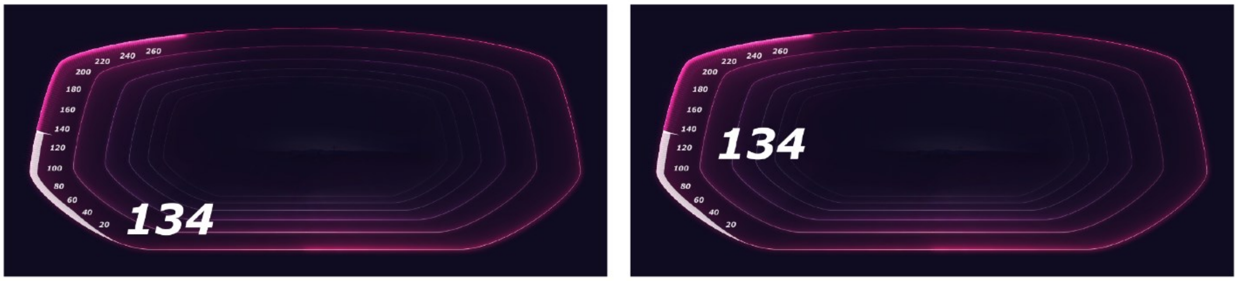
1 2 3

Note. 1 shows how the cue of intersections could be helpful for the exact positioning of the car, 2 shows how it can facilitate the frontal positioning of a warning, while 3 presents how the lines of a texture gradient intersect with the road and clarify the spatial relationship between these elements.

Additionally, the manner in which the display is placed within the dashboard is of high importance and framing should be used in such a way that it suggests potential depth behind the screen level, as well as a physical depth in front. Here, separated and tablet-like screens should be avoided due to their flatness and the display should be seamlessly integrated into the dashboard. Moreover, to best facilitate the impression of depth, the cue of the height in the visual field should be properly implemented. Here, the usual central placement of gauges conflicts with the cue's main assumption, that the objects that are near appear more separated in height from the horizon line than the objects that are far away. A potential solution to this issue could be a different implementation of the meters (see Figure 30) but the novelty of such positioning and the resulting lack of prototypicality needs further investigation.

Figure 30

Height in visual field and the design of meters



1 2

Note. 1 shows how the height in the visual field could be implemented by moving the speed meter to the bottom side of the screen. However, as seen in 2, the central positioning is more prototypical and thus potentially safer. Notice the subtle horizon line in both images.

As shown in Figure 30, the prototypicality of the display is in conflict with creating the most optimal impression of depth. However, the height in the visual field cue could still be implemented without having to readjust the drivers' existing mental models. This could be achieved through a change in perspective of the 3D scene in such a way that it is seen slightly from above (see Figure 31). By doing so, the horizon line moves up and the frontal elements appear lower in comparison. Such a perspective would also match the drivers' position better than a completely frontal view since drivers most often have to look down at the instrument cluster. The higher point of view would also benefit any 3D spatial assessment, such as navigation or Adaptive Cruise Control.

Furthermore, digital instrument clusters provide new opportunities for structuring information, for example replacing circular gauges with different types of meters. This might be beneficial in the case of a 3D design, where regular, circular gauges placed in front would block most of the scene from view (see Figure 32), consequently diminishing the potential of 3D representation. The degree to which novel design structures affect driving performance requires further investigation. A deeper understanding of how drivers adapt to new layouts and how they affect performance would be highly beneficial for unlocking the design possibilities of digital instrument clusters.

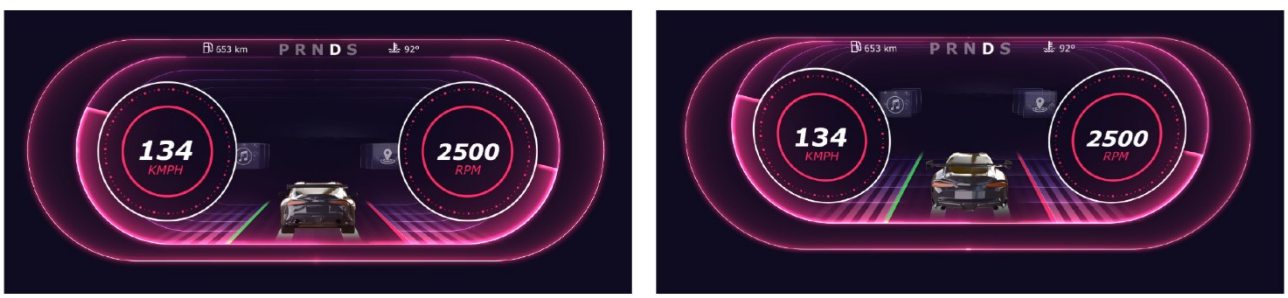
Figure 31

Matching the 3D scene with the point of view of the driver

Note. While the image on the left shows how the 3D scene looks like when viewed frontally, the right image presents a perspective that better matches the driver's point of view. Such a view also supports the cue of height in the visual field, as well as such functionalities as 3D navigation.

Figure 32

The circular meters and their occlusion of the 3D scene



1 2

Note. Both the images show that the circular meters block a considerable amount of the 3D scene from view. Additionally, as shown in 2, the top perspective view causes strange distortions in how the round objects are perceived.

It appears that the circular shapes within a 3D design are more difficult to work with. Nonetheless, the issue of the gauges occluding most of the 3D space could be solved by bringing the gauges closer to the screen-level and adjusting their design to have them take on less screen space (see Figure 33).

Figure 33

The potential new representation of the circular meters

1 2

Note. The outer white circle of the meter, as well as the dashed line, are only partly visible to reveal the 3D scene behind the meters. 2 represents the upper perspective, which again results in strange spatial relations between the circular shapes.

Figures 32 and 33 show that repeating round shapes in depth leads to a rather strange effect, which might be amplified even further by the parallax transformation. Therefore, if the circular shapes are to be used, the scene needs to be further simplified to avoid their repetition in depth with the gauges placed at the screen-level. Such simplification is possible, but when compared with the angular layout, the scene is stripped of several other depth cues (see Figure 34). These include the cue of intersections, relative size and to a high degree the cue of perspective, with their absence harming the impression of the recreated depth.

Figure 34

The difference between round and angular layouts

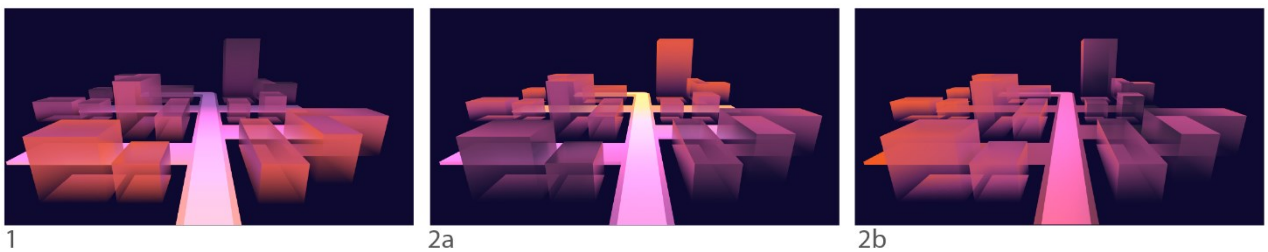
1 2

Note. Although the circular layout appears to work in such a simplified design (1) multiple other sources of depth are missing when compared with the angular design (2), which additionally includes the cues of intersections, relative size and a higher number of converging perspective lines.

The impression of depth can also be successfully created by the use of color gradients. However, this has to be applied properly, with the color's hue becoming cooler and the saturation and luminance decreasing with distance (see Figure 35). Additionally, gradients should be used in depth and not horizontally or in a circular form, which is sometimes used for meters as an indication of an increasing value, for example, speed.

Figure 35

The use of color gradients to support depth impression



Note. 1 shows the correct application of the cue of color, with warmer, brighter and more saturated colors in the front. This is reversed for 2a, notice how the brightest part in the back grabs attention. 2b represents another use of color that should be avoided, where color gradients are used not in depth but horizontally.

However, it might be the case that a cold hue, such as blue or cyan must be used as a primary color in the design. This can still be successfully applied, as long as the color gradient transitions from the most luminant and saturated colors in the front to the darkest and desaturated colors in the back (see Figure 36).

Figure 36

The cue of color and its different potential application

Similarly, the cue of contrast should be applied carefully, with the contrast decreasing in depth in such a way that the essential elements are not negatively affected. Therefore, even if the meters are not positioned in the most frontal depth layer, their contrast should not decrease, which should only happen to elements in the back that are not considered important to the driving task and are not currently used by the driver (see Figure 37).

Figure 37

The cue of contrast and its change over distance

Note. It is also clear that the hue of the color changes to blue with distance, which is a result of the color cue applied accordingly to the cue of aerial perspective.

The cue of brightness could be added to the design above by increasing the overall brightness of the image (see Figure 38). However, the increase should be subtle as the contrast of the elements should remain sufficiently high and the brightness should not be eye-straining.

Figure 38

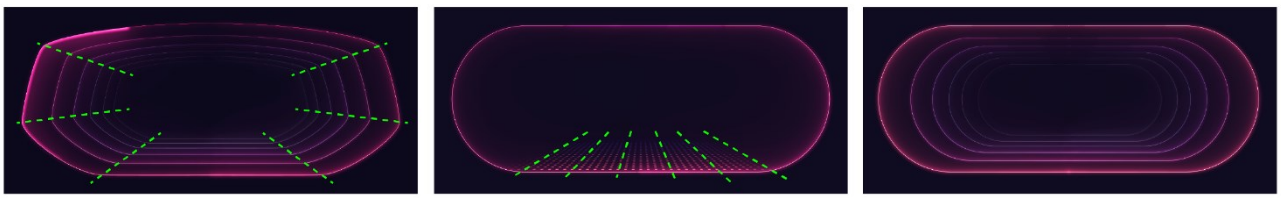
The gradual increase of brightness of the design

Note. While 1 presents the design on a black background, dark blue is used in 2 which is made even brighter for 3. As long as contrast is not negatively affected, the background color of the display could be made brighter.

Furthermore, the linear perspective cue should be applied to facilitate the correct and believable impression of depth. This cue can most easily be achieved through converging lines. However, these lines can also be suggested through corners of objects or the texture gradient (see Figure 39). Moreover, it appears that through the use of an angular design the perspective cue is achieved more successfully, than in the presence of circular shapes.

Figure 39

The cue of perspective and the converging lines



1 2 3

Note. The converging lines can become apparent from the repetition of the same angular object placed at equal depth increments (1), as well as from a regular structure of a texture gradient (2). In contrast, the repetition of round shapes in depth (3) does not seem to result in clear perspective lines.

Additionally, different types of texture gradients can be used to achieve the impression of depth, as long as the texture is regular in its structure (see Figure 40). A regular texture gradient is beneficial not only for the impression of depth but also for the accuracy of spatial judgements. It should therefore be used both as a design element and in support of such functionalities as parking assist (see Figure 41).

Figure 40

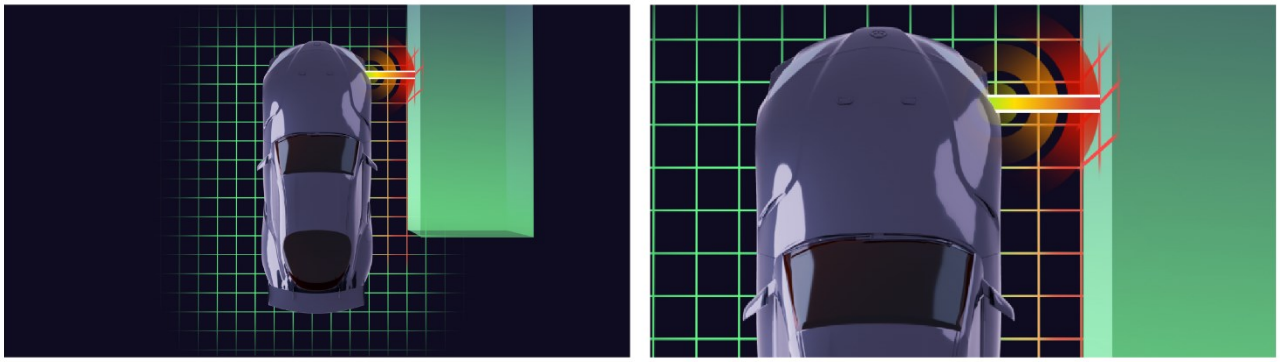
Different types of texture gradients

1 2

Note. 1 illustrates a texture gradient in the form of straight vertical lines, while 2 presents a gradient made out of dots that form a regular grid.

Figure 41

The added value of the texture gradient for spatial assessment



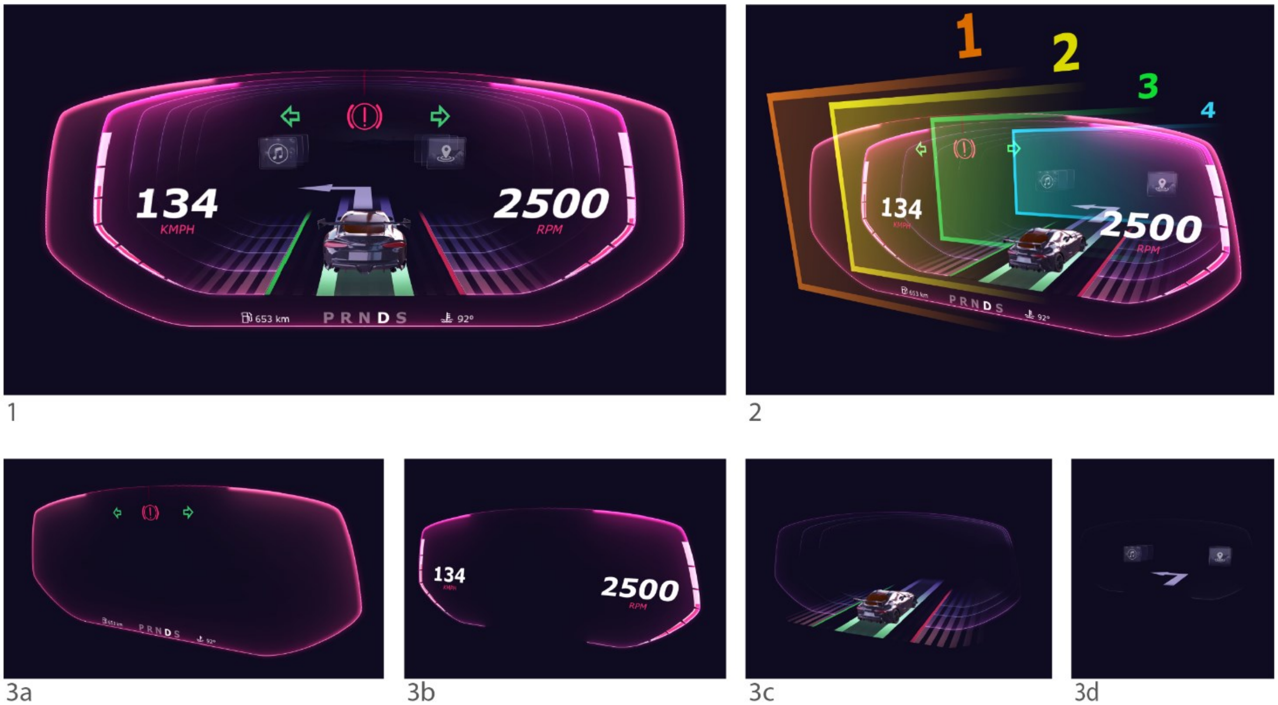
1 2

Note. 1 shows the overview of the environment that could assist the driver with tasks such as parking. 2 shows how the cue of intersections is combined with the texture gradient to facilitate depth assessment. Drivers could also be made aware of the exact distance in centimeters that corresponds to one square within the grid.

Next to the monocular depth cues, structuring information in depth has to be considered while creating a 3D digital instrument cluster (see Figure 42). The most frontal positioning, that of the screen-level, should be reserved for textual information, icons, as well as warnings. As already mentioned, a round-shaped design requires the gauges to be placed at the screen level as well, which makes the first depth layer rather crowded and the structuring of information in depth is not supported. Therefore, employing an angular design, where the meters can be moved slightly to the back seems to better promote such structuring. This information can then be separated into its own depth layer, but should still be placed in the front of the 3D scene. Additionally, any functionality related to spatial assessment such as 3D navigation or Adaptive Cruise Control, should be placed on yet another depth layer (see Figure 42). Lastly, the infotainment elements that are not currently used by the driver should be separated and placed all the way in the back. While they are being used, they should move forward and elements in the third depth layer should be readjusted accordingly.

Figure 42

The structuring of information into several depth layers



Note. 1 presents a frontal view of all the depth layers with 2 showing the numbered layers and their structure from the side. 3 iterates through all the four depth layers, starting with the closest one (3a) and ending with the one that is most far away from the viewer (3d). Notice how 3c is much wider than other layers, as it is reserved for spatial assessment functionality and should fit a 3D model of the car or the car's surroundings, as well as other infotainment elements while they are being used.

As illustrated above, through the continuous consideration of the developed guidelines, the creation process of a 3D digital instrument cluster is facilitated and new insights are emerging once the theoretical framework is put into practice. This includes for example the advantages offered by the angular shapes over the round ones or the apparent issues with the prototypical circular meters. Such meters, when combined with the motion parallax and the resulting perspective transformation, would occlude even more of the 3D scene. This shows that the design of a 3D digital instrument cluster is much more than simply adding different monocular depth cues. Instead, it requires novel approaches to structuring information in a way that best supports recreation of depth, as well as a continuous investigation of the proposed design solutions and their influence on the driving performance. Figure 43 provides the final design developed based on the proposed guidelines where the 3D effect was given priority. Overall, it appears that static 2D and dynamic 3D layouts require distinct design approaches and the differences should be further investigated in the context of automotive interfaces.

Figure 43

The exemplary design of a 3D digital instrument cluster based on the proposed guidelines

User requirements

While designing 3D in-car displays attention should be paid not only to the visual elements but also to the user requirements. Hence, the insights gathered from the interview reports have been combined with the review findings in order to create the final user requirements. Table 12 presents the identified functionalities that users would like the design to support, together with the level of desirability of these functions. This overview is based purely on the interview accounts. On the other hand, Table 13 outlines the requirements derived from both the interviews and the review findings. Meeting these requirements is essential for a successful implementation of the simulated 3D effect in the automotive context.

Table 12

The overview of desired functionalities

Category	Requirements	Desirability
Support driving tasks	3D navigation	High
	Omniview technology (360° viewing systems)	High
	Narrow road assistant	High
	Adaptive cruise control	High
	Parking assistance	High
	Backup camera	Medium
	Driver monitoring system	Medium
	Intersection assistance	Medium
Visual aspects	Option to switch back to a 2D view	High
	Warning system that communicates urgency through the depth	High
	Customizable design (e.g. the parallax amount)	Medium
Other	Musical playlist	Medium
	Voice assistance	Medium
	Short tutorial/instructions on how the display works	Medium

Table 13

The user requirements for in-car implementation of the simulated 3D effect

Category	The reported requirement	Source	Importance
Technical factors	Complete smoothness of tracking and the resulting parallax transformation	Interviews	Necessary
	Wider tracking angle, preferably with no clear boundaries	Interviews	Necessary
	Reliable in all driving conditions and resilient to sudden changes in lighting	Interviews	Necessary
	Tracking accounts for differences between drivers (e.g. height, racial differences)	Interviews	Necessary
	Tracking accounts for sudden, unrelated head movements	Interviews	Necessary
	Does not require big movements to see the 3D effect	Interviews	Necessary
	Higher contrast range of the display	Interviews/review findings	Desirable
	Higher brightness of the display	Interview/review findings	Desirable
Driving task	Does not hinder the normal control of the car	Interviews	Necessary
	Does not distract from driving (the effect should be subtle)	Interviews	Necessary
	Adds practical value by 3D functionality (navigation, parking, omniview camera)	Interviews	Desirable
	Minimized complexity of information	Interviews/review findings	Desirable
	Ease of use and interaction	Interviews	Desirable
Visual aspects	Option to easily switch back to a 2D view	Interviews	Necessary
	Ensure clear visibility of the essential stimuli	Interviews	Necessary
	Prioritize attractive visual design over the realistic depth representation	Interviews	Necessary
	Make the design of the display match the car's interior	Interviews	Desirable
	Structure information into a small number of separate depth layers	Review findings	Desirable
	Use depth to support saliency of warnings by bringing them to the front	Interviews/review findings	Desirable
	Customizable design options (parallax, depth within the display, colors)	Interviews/review findings	Desirable
Other	Support other infotainment elements (playlist, placing a call, voice assistance)	Interviews	Desirable
	Provide clear, well-framed instructions about the display's nature and interaction	Interviews	Desirable

Future research directions

Overall, the above exploration provides evidence for a high level of acceptance towards the use of observer-produced parallax to convey 3D effects in the automotive setting. Yet, the major limitation of this study is the separation of the experimental task and the driving context, which makes the reported level of acceptance hypothetical and highly dependent on the correct implementation of the derived design guidelines and the user requirements. Therefore, this thesis lays the groundwork for the design of the 3D digital instrument cluster and the resulting design should further be investigated in the high-risk context of a driving task.

The designed 3D display should first be tested in the presence of a simulated driving task. To understand how the 3D effect affects driving performance the 3D parallax-based simulation could be compared with a static version of the same design. In the 3D condition, the amount of the parallax transformation should be manipulated and most preferably several different levels ranging from natural to highly exaggerated should be tested. Such an approach would allow to further investigate the influence of increased parallax on driving performance and identify the level of transformation that feels natural. As already mentioned, it is, therefore, important to investigate not only the performance but also the users' subjective impressions evoked by their interaction with the 3D effect. Especially understanding how users experienced the parallax transformation and the true acceptance towards in-car use would benefit from a qualitative approach.

Additionally, as the findings of this thesis indicate, the visual preferences are highly subjective and should further be investigated before the final design is implemented. To support this goal, users could be asked to customize the design elements to their liking and the resulting combinations could be used to derive the most attractive visual layouts. However, the influence of the 3D effect on the driving performance should be tested with the same version of the design for all participants. If each participant used a different layout for the experimental task, it would be impossible to differentiate between participant- and design-level effects.

Furthermore, 3D layouts seem to require novel design approaches which result in low prototypicality of the display. Investigating how the new layouts affect driving performance while examining the learning effects would help increase the current understanding of the potential use of 3D effects in the automotive industry. Considering the high-risk nature of the driving task, continued efforts are needed to introduce such effects in a safe and responsible way.

.

References

Algorri, J. F., Urruchi, V., García-Cámara, B., & Sánchez-Pena, J. M. (2016). Liquid crystal microlenses for autostereoscopic displays. *Materials*, *9*(1). https://doi.org/10.3390/ma9010036

Andersson, U. (2017). Effect of depth cues on visual search in a web-based environment [Master's thesis, School of Computer Science and Communication]. DiVA - Academic Archive Online. http://www.diva-portal.org/smash/record.jsf?dswid=-333&pid=diva2%3A1084759

Berger, M., Pfleging, B., & Bernhaupt, R. (2019). A tactile interaction concept for in-car passenger infotainment systems. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*, 109–114. https://doi.org/10.1145/3349263.3351914

Bian, Z., Braunstein, M.L. & Andersen, G.J. (2005). The ground dominance effect in the perception of 3-D layout. *Perception & Psychophysics, 67,* 802–815. https://doi-org/10.3758/BF03193534

Bolder, A., Grünvogel, S. M., & Angelescu, E. (2018). Comparison of the usability of a car infotainment system in a mixed reality environment and in a real car. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology.* https://doi.org/10.1145/3281505.3281512

Brooks, K. R. (2017). Depth perception and the history of three-dimensional art: Who produced the first stereoscopic images? *I-Perception*, *8*(1). https://doi.org/10.1177/2041669516680114

Broy, N., Alt, F., Schneegass, S., Henze, N., & Schmidt, A. (2013). Perceiving layered information on 3D displays using binocular disparity. *Proceedings Of The 2Nd ACM International Symposium On Pervasive Displays - Perdis '13*. https://doi.org/10.1145/2491568.2491582

Broy, N., Alt, F., Schneegass, S., & Pfleging, B. (2014a). 3D Displays in Cars. *Proceedings Of The 6Th International Conference On Automotive User Interfaces And Interactive Vehicular Applications*. https://doi.org/10.1145/2667317.2667319

Broy, N., André, E., & Schmidt, A. (2012). Is stereoscopic 3D a better choice for information representation in the car?. *Proceedings Of The 4Th International Conference On Automotive User Interfaces And Interactive Vehicular Applications - Automotiveui '12*. https://doi.org/10.1145/2390256.2390270

Broy, N., Guo, M., Schneegass, S., Pfleging, B., & Alt, F. (2015a). Introducing novel technologies in the car. *Proceedings Of The 7Th International Conference On Automotive User Interfaces And Interactive Vehicular Applications*. https://doi.org/10.1145/2799250.2799280

Broy, N., Schneegass, S., Alt, F., & Schmidt, A. (2014b). FrameBox and MirrorBox. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems*. https://doi.org/10.1145/2556288.2557183

Broy, N., Schneegass, S., Guo, M., Alt, F., & Schmidt, A. (2015b). Evaluating Stereoscopic 3D for Automotive User Interfaces in a Real-World Driving Study. *Proceedings Of The 33Rd Annual ACM Conference Extended Abstracts On Human Factors In Computing Systems*. https://doi.org/10.1145/2702613.2732902

Broy, N., Zierer, B., Schneegass, S., & Alt, F. (2014c). Exploring virtual depth for automotive instrument cluster concepts. *CHI '14 Extended Abstracts On Human Factors In Computing Systems*. https://doi.org/10.1145/2559206.2581362

Busby, A., & Ciuffreda, K. J. (2005). The effect of apparent depth in pictorial images on accommodation. *Ophthalmic and Physiological Optics*, *25*(4), 320–327. https://doi.org/10.1111/j.1475-1313.2005.00305.x

Canestrari, R., & Farne, M. (1969). Depth Cues and Apparent Oscillatory Motion. *Perceptual And Motor Skills*, *29*(2), 508-510. https://doi.org/10.2466/pms.1969.29.2.508

Cavanagh, P. (1987). Reconstructing the third dimension: Interactions between color, texture, motion, binocular disparity, and shape. *Computer Vision, Graphics, And Image Processing*, *37*(2), 171-195. https://doi.org/10.1016/s0734-189x(87)80001-4

Cutting, J. E., & Vishton, P. (1995). Perceiving layout and knowing distances: The interaction, relative potency, and contextual use of different information about depth. In W. Epstein, & S. Rogers (Eds.), *Perception of Space and Motion* (pp. 69-117). New York: Academic Press

d' Arthois, J. (n.d.). *Landscape* [Painting]. The Royal Museum of Fine Arts, Antwerp, Belgium. https://kmska.be/en/masterpiece/landscape-12

Dosher, B. A., Sperling, G., & Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision research*, *26*(6), 973–990. https://doi.org/10.1016/0042-6989(86)90154-9

Dunn, B. E., Gray, G. C, & Thompson, D. (1965). Relative Height on the Picture-Plane and Depth Perception. *Perceptual And Motor Skills*, *21*(1), 227-236. https://doi.org/10.2466/pms.1965.21.1.227

Easa, H. K., Mantiuk, R. K., & Lim, I. S. (2013). Evaluation of monocular depth cues on a high-dynamic-range display for visualization. *Acm Transactions on Applied Perception*, *10*(3). https://doi.org/10.1145/2504568

Emoto, M. (2019). Depth perception and induced accommodation responses while watching high spatial resolution two-dimensional TV images. *Displays*, *60*, 24–29. https://doi.org/10.1016/j.displa.2019.08.005

Fulvio, J. M., Rosen, M. L., & Rokers, B. (2015). Sensory uncertainty leads to systematic misperception of the direction of motion in depth. *Attention, Perception, & Psychophysics, 77*(5), 1685–1696. https://doi.org/10.3758/s13414-015-0881-x

Gilchrist, A. L. (2007). Lightness and brightness. *Current Biology*, *17*(8), 267–9. https://doi.org/10.1016/j.cub.2007.01.040

Goldmann, L., Ebrahimi, T., Lebreton, P., Raake, A. (2012). Towards a Descriptive Depth Index for 3D Content: Measuring Perspective Depth Cues. *Sixth International Workshop on Video Processing and Quality Metrics for Consumer Electronics – VPQM,* Scottsdale, Arizona, United States.

Greene, C. M., Broughan, J., Hanlon, A., Keane Seán, Hanrahan, S., Kerr, S., & Rooney, B. (2021). Visual search in 3d: effects of monoscopic and stereoscopic cues to depth on the validity of feature integration theory and perceptual load theory. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.596511

Held, R. T., Cooper, E. A., O'Brien, J. F., & Banks, M. S. (2010). Using blur to affect perceived distance and size. *Acm Transactions on Graphics, 29*(2), 1–16. https://doi.org/10.1145/1731047.1731057

Hendrix, C., & Barfield, W. (1995). Relationship between monocular and binocular depth cues for judgements of spatial information and spatial instrument design. *Displays*, *16*(3), 103-113. https://doi.org/10.1016/0141-9382(96)81210-8

Hillstrom, A. P., Wakefield, H., & Scholey, H. (2013). The effect of transparency on recognition of overlapping objects. *Journal of Experimental Psychology. Applied*, *19*(2), 158–70. https://doi.org/10.1037/a0033367

Hsu, L. C., Kramer, P., & Yeh, S. L. (2010). Monocular depth effects on perceptual fading. *Vision Research*, *50*(17), 1649–1655. https://doi.org/10.1016/j.visres.2010.05.008

Hu, B., & Knill, D. C. (2011). Binocular and monocular depth cues in online feedback control of 3D pointing movement. *Journal of Vision, 11*(7), 23–23. https://doi.org/10.1167/11.7.23

Huhtala, J., Karukka, M., Salmimaa, M., & Häkkilä, J. (2011). Evaluating depth illusion as method of adding emphasis in autostereoscopic mobile displays. *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, 357–360. https://doi-org/10.1145/2037373.2037427

Hürst, W., Beurskens, J., & van Laar, M. (2013). An experimentation environment for mobile 3D and virtual reality. *MobileHCI 2013 - Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 444–447. https://doi.org/10.1145/2493190.2494659

Hwang, A. D., & Peli, E. (2014). Instability of the perceived world while watching 3D stereoscopic imagery: A likely source of motion sickness symptoms. *i-Perception, 5*(6), 515–535. https://doi-org/10.1068/i0647

Ichihara, S., Kitagawa, N., & Akutsu, H. (2007). Contrast and depth perception: effects of texture contrast and area contrast. *Perception, 36*(5), 686–95.

Kakeya, H., Hayashishita, A., & Ominami, M. (2018). Autostereoscopic display based on time-multiplexed parallax barrier with adaptive time-division. *Journal of the Society for Information Display*, *26*(10), 595–601. https://doi.org/10.1002/jsid.717

Katsuyama, N., Usui, N., & Taira, M. (2016). Activation of the human mt complex by motion in depth induced by a moving cast shadow. *Plos One*, *11*(9), 0162555. https://doi.org/10.1371/journal.pone.0162555

Keefe, B. D., Hibbard, P. B., & Watt, S. J. (2011). Depth-cue integration in grasp programming: No evidence for a binocular specialism. *Neuropsychologia, 49*(5), 1246–1257. https://doi.org/10.1016/j.neuropsychologia.2011.02.047

Kellnhofer, P., Didyk, P., Ritschel, T., Masia, B., Myszkowski, K., & Seidel, H. P. (2016). Motion parallax in stereo 3D: Model and applications. *ACM Transactions on Graphics*, *35*(6). https://doi.org/10.1145/2980179.2980230

Kersten, D., Mamassian, P., & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception, 26*(2), 171–192. https://doi.org/10.1068/p260171

Klinghammer, M., Schütz, I., Blohm, G., & Fiehler, K. (2016). Allocentric information is used for memory-guided reaching in depth: A virtual reality study. *Vision Research*, *129*, 13–24. https://doi.org/10.1016/j.visres.2016.10.004

Koessler, T., & Hill, H. (2019). Focusing on an illusion: Accommodating to perceived depth? *Vision Research*, *154*, 131–141. https://doi.org/10.1016/j.visres.2018.11.001

Kongsilp, S., & Dailey, M. N. (2017). Motion parallax from head movement enhances stereoscopic displays by improving presence and decreasing visual fatigue. *Displays*, *49*, 72–79. https://doi.org/10.1016/j.displa.2017.07.001

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *The Journal of Neuroscience*, *20*(9), 3310–8. https://doi.org/10.1523/JNEUROSCI.20-09-03310.2000

Landy, M.S, Maloney, L.T., Johnston, E.B. & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research, 35(3),* 389-412. https://doi.org/10.1016/0042-6989(94)00176-M

Lanman, D., Hirsch, M., Kim, Y., Raskar, R., & ACM SIGGRAPH Asia (2010). Content-adaptive parallax barriers: optimizing dual-layer 3d displays using low-rank light field factorization. *Acm Transactions on Graphics*, *29*(6). https://doi.org/10.1145/1866158.1866164

Lauber, F., Böttcher, C., & Butz, A. (2014). You’ve got the look: Visualizing infotainment shortcuts in head-mounted displays. *AutomotiveUI 2014 - 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, in Cooperation with ACM SIGCHI - Proceedings*. https://doi.org/10.1145/2667317.2667408

Law, B., & Evans, S. (2007). Understanding Luxury in the Premium Automotive Industry. *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces,* 168-179. *https://doi-org/10.1145/1314161.1314176*

Lee, K., & Lee, S. (2016). A new framework for measuring 2d and 3d visual information in terms of entropy. *Ieee Transactions on Circuits and Systems for Video Technology*, *26*(11), 2015–2027. https://doi.org/10.1109/TCSVT.2015.2477915

Léveillé, J., Myers, E., & Yazdanbakhsh, A. (2014). Object-centered reference frames in depth as revealed by induced motion. *Journal of Vision*, *14*(3), 1-11. https://doi.org/10.1167/14.3.15

Li, C., Lai, Y., Wu, C., Tsai, S. F., Chen, T. C., Chien, S. Y., & Chen, L. G. (2013). Brain-inspired framework for fusion of multiple depth cues. *IEEE Transactions on Circuits and Systems for Video Technology*, *23*(7), 1137–1149. https://doi.org/10.1109/TCSVT.2012.2223874

Luong, T., Seth, A., Klein, A.W., & Lawrence, J. (2005). Isoluminant color picking for non-photorealistic rendering. *Proceedings of Graphics Interface 2005,* 233-240.

Markov, Y. A., & Tiurina, N. A. (2021). Size-distance rescaling in the ensemble representation of range: Study with binocular and monocular cues. *Acta Psychologica*, *213*. https://doi.org/10.1016/j.actpsy.2020.103238

Masola, A., Gabbi, C., Castellano, A., Capodieci, N., & Burgio, P. (2020). Graphic Interfaces in ADAS: From requirements to implementation. *Pervasive Health: Pervasive Computing Technologies for Healthcare*, 193–198. https://doi.org/10.1145/3411170.3411259

Mather, G., & Smith, D. R. (2004). Combining depth cues: effects upon accuracy and speed of performance in a depth-ordering task. *Vision Research*, *44*(6), 557–62.

McIntire, J. P., Havig, P. R., & Geiselman, E. E. (2012). What is 3D good for? A review of human performance on stereoscopic 3d displays. *Proceedings of SPIE - the International Society for Optical Engineering*. https://doi.org/10.1117/12.920017

Mulligan, J. B. (2009). Late-news poster: motion parallax enhances depth in a perspective air-traffic display. *Digest of Technical Papers*, *40*, 1231–1233.

Ni, R., Braunstein, M. L., & Andersen, G. J. (2007). Scene layout from ground contact, occlusion, and motion parallax. *Visual Cognition, 15*(1), 48–68. https://doi.org/10.1080/13506280600646657

O'Brien, J., & Johnston, A. (2000). When texture takes precedence over motion in depth perception. Perception, 29(4), 437–452. https://doi.org/10.1068/p2955

Ono, H., Rogers, B. J., Ohmi, M., & Ono, M. E. (1988). Dynamic occlusion and motion parallax in depth perception. *Perception*, *17*(2), 255–266. https://doi.org/10.1068/p170255

O'leary, A., & Wallach, H. (1980). Familiar size and linear perspective as distance cues in stereoscopic depth constancy. *Perception & Psychophysics*, *27*(2), 131-135. https://doi.org/10.3758/bf03204300

Ostendorp, M. C., Feuerstack, S., Friedrichs, T., & Lüdtke, A. (2016). Engineering automotive HMIs that are optimized for correct and fast perception. *Proceedings of the 8th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, 293–298. https://doi.org/10.1145/2933242.2935869

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson A, Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). Prisma 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *Bmj (Clinical Research Ed.)*, *372*, 160. https://doi.org/10.1136/bmj.n160

Parton, A. D., Bradshaw, M. F., & de Bruyn, B. (1999). The Design of Telepresence Systems: The Task-Dependent Use of Binocular Disparity and Motion Parallax. *International Journal Of Cognitive Ergonomics, 3*(3), 189-202.

Pla, P., & Maes, P. (2013). Display Blocks: A Set of Cubic Displays for Tangible, Multi-Perspective Data Exploration. *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction,* 307-314. https://doi-org/10.1145/2460625.2460677

Pfautz, J.D (2000). *Depth Perception in Computer Graphics* [Doctoral dissertation]. University of Cambridge. https://doi.org/10.17863/CAM.31733

Rao, Q., Grünlery, C., Hammori, M., & Chakraborty, S. (2014). Design methods for augmented reality in-vehicle infotainment systems. *Proceedings - Design Automation Conference*. https://doi.org/10.1145/2593069.2602973

Rempel. A. G., Heidrich, W., & Mantiuk, R. (2011). The Role of Contrast in the Perceived Depth of Monocular Imagery. *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization,* 115. https://doi-org/10.1145/2077451.2077478

Rogers, B. (2009). Motion parallax as an independent cue for depth perception: a retrospective. *Perception*, *38*(6), 907–11.

Rogers, B., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception, 8*(2), 125–134. https://doi.org/10.1068/p080125

Royden, C. S., Parsons, D., & Travatello, J. (2016). The effect of monocular depth cues on the detection of moving objects by moving observers. *Vision Research*, *124*, 7–14. https://doi.org/10.1016/j.visres.2016.05.002

Rößing, C., Hanika, J., & Lensch, H. (2012). Real-time disparity map-based pictorial depth cue enhancement. *Computer Graphics Forum*, *31*(2), 275–284. https://doi.org/10.1111/j.1467-8659.2012.03006.x

Schild, J., LaViola, J., & Masuch, M. (2012). Understanding User Experience in Stereoscopic 3D Games. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 89-98. https://doi-org/10.1145/2207676.2207690

Schmettow, M. (2021). *New Statistics for design researchers*. Springer International Publishing. Chapter 6.8. https://doi.org/10.1007/978-3-030-46380-9

Seipel, S. (2013). Evaluating 2D and 3D geovisualisations for basic spatial assessment. *Behaviour and Information Technology*, *32*(8), 845–858. https://doi.org/10.1080/0144929X.2012.661555

Sen, G., & Sener, B. (2020). Design for Luxury Front-Seat Passenger Infotainment Systems with Experience Prototyping through VR. *International Journal of Human-Computer Interaction*. https://doi.org/10.1080/10447318.2020.1785150

Shimono, K., Higashiyama, A., Kihara, K., & Matsuda, Y. (2021). A frame at a different depth than a photograph enhances the apparent depth in the photograph. *Attention, Perception, & Psychophysics, 83*(8), 3216–3226. https://doi.org/10.3758/s13414-021-02345-7

Sweet, B. T., & Kaiser, M. K. (2013, June 03-06). *Choosing your poison: optimizing simulator visual system selection as a function of operational tasks* [Paper presentation]. IMAGE Society, Scottsdale, Arizona, United States.

Troscianko, T., Montagnon, R., Clerc, J., Malbert, E., & Chanteau, P. (1991). The role of colour as a monocular depth cue. *Vision Research*, *31*(11), 1923-1929. https://doi.org/10.1016/0042-6989(91)90187-a

Uehira, K., Suzuki, M., & Abekawa, T. (2007). 3-D Display using Motion Parallax for Extended-Depth Perception. *2007 IEEE International Conference on Multimedia and Expo*, 1742-1745. https://doi.org/10.1109/ICME.2007.4285007.

Qian, J., Li, J., Wang, K., Liu, S., & Lei, Q. (2017). Evidence for the effect of depth on visual working memory. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-06719-6 van Beurden, M. H. P. H., Kuijsters, A., & IJsselsteijn, W. A. (2010). Performance of a path tracing task using stereoscopic and motion based depth cues, *2010 Second International Workshop on Quality of Multimedia Experience (QoMEX)*, 176-181, https://doi.org/10.1109/QOMEX.2010.5516268

van Schooten, B.W., van Dijk, E. M. A. G., Zudilova-Seinstra, E., Suinesiaputra, A., & Reiber. J. H. C. (2010). The effect of stereoscopy and motion cues on 3D interpretation task performance. In *Proceedings of the International Conference on Advanced Visual Interfaces*. Association for Computing Machinery, New York, NY, USA, 167–170. https://doi-orgC /10.1145/1842993.1843023

Warren, P. A., & Rushton, S. K. (2009). Perception of scene-relative object movement: optic flow parsing and the contribution of monocular depth cues. *Vision Research*, *49*(11), 1406–1419. https://doi.org/10.1016/j.visres.2009.01.016

Xie, X., Shin, K. G., Yousefi, H., & He, S. (2018). Wireless CSI-based head tracking in the driver seat. *Proceedings of the 14th International Conference on Emerging Networking EXperiments and Technologies*, 112–125. https://doi.org/10.1145/3281411.3281414

Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision research*, *33*(18), 2685–2696. https://doi.org/10.1016/0042-6989(93)90228-o

Zhang, Z., Zhou, C., Wang, Y., & Gao, W. (2013). Interactive stereoscopic video conversion. *IEEE Transactions on Circuits and Systems for Video Technology*, *23*(10), 1795–1808. https://doi.org/10.1109/TCSVT.2013.2269023

Appendix A – Searched databases

Table A

Databases included in the EBSCOhost search

Database	Focus
Business Source Elite	Business
Library, Information Science & Technology Abstracts	Technology
APA PsycArticles	Psychology
Psychology and Behavioral Sciences Collection	Psychology
APA PsycInfo	Psychology
Regional Business News	Business
OpenDissertations	Dissertations and Theses
eBook Open Access (OA) Collection (EBSCOhost)	Curated collection of eBooks

Appendix B

Search keyword sequences

Table B

Search keyword sequences with the number of yielded results per database

Keyword Sequence	Database	Records Identified
Monocular depth cues	EBSCOhost	171
Stereoscopic display	EBSCOhost	419
Autostereoscopic display	EBSCOhost	96
Stereoscopic display AND monoscopic	EBSCOhost	21
3D AND flat display	EBSCOhost	81
Motion parallax AND depth	EBSCOhost	394
Head tracking AND 3D display	EBSCOhost	21
Depth impression AND monocular cues	ACM Digital Library	49
Depth impression AND flat display	ACM Digital Library	61
3D instrument cluster AND automotive	ACM Digital Library	19
3D dashboard AND automotive	ACM Digital Library	12

Note. EBSCOhost was used primarily for depth perception related concepts and ACM Digital Library was searched mostly for technology records. However, to ensure that such division did not cause records to be omitted, both included one or two additional searches for the other topic.

Appendix C

Experimental setup

Figure C1

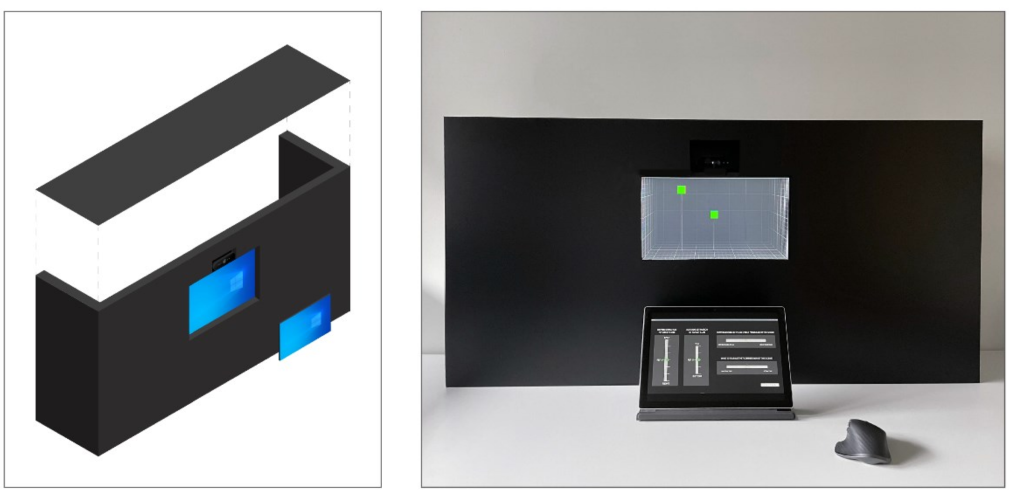
The ArUco tracker used for observer-produced parallax

1 2

Note. The tracker was placed on a hard cardboard and was connected to a baseball cap in such a way that its position could easily be re-adjusted (1). The cap was then placed on the participant's head in such a way that the tracker faced the camera (2).

Figure C2

The experimental setup with the main display placed in a hardwood frame

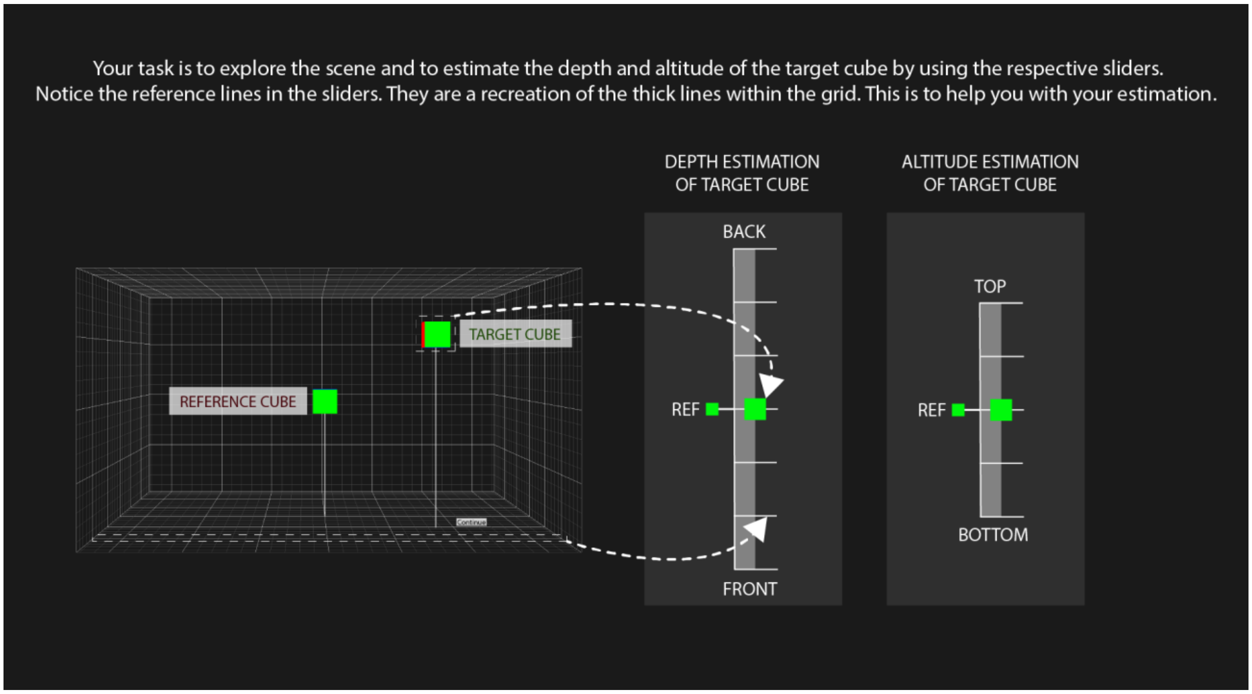


1 2

Note. To ensure easy access the top lid was made removable (1) and there was no back wall so that the cables could be connected while remaining hidden from the participant's view. Both the second display and the mouse were wirelessly connected to the laptop inside the frame (2).

Figure C3

Sliders used for height and depth estimation in reference to the stimuli



Note. The above image is part of the instructions participants were given before the task. It shows how the thick lines of the grid correspond to the lines within the sliders.

Appendix D

The User Experience Questionnaire

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Appendix E

Interview questions

- Would you describe yourself as a tech-savvy person?
- If you were to describe this 3D display to someone with similar computer experience as yours, what would you say?
- Shortly describe what you liked about the device.
- Shortly describe what about the device could be improved.
- What did you think of the task itself?
- Putting the task of the experiment aside, how would you assess your experience with the 3D display?
- Would you be interested in owning such a 3D display? Explain your answer.
- Can you imagine any potential use for it?

Showing the example image to the participants:

Note. An example of a digital instrument cluster. Reprinted from Behance (Sams, 2021). *Automotive Digital Cockpit.* https://www.behance.net/gallery/122017093/Automotive-Digital-Cockpit

- Here you can see an example of a regular flat instrument cluster inside of a car. Do you think adding the 3D effect like the one you just saw is a good idea? Explain why or why not.
- Do you think installing such displays in luxury cars is a good addition to the car's entertainment services? Explain why or why not.
- Imagine that you were to buy a car with such a display in it, what would your requirements be?
- While using the device just now you saw a few different visual elements, like the fog or the dropline coming down, and sometimes there was also a different amount of parallax, so the same movement of your body corresponded to a bigger transformation within the scene, did you notice that?

- Could you maybe tell me a little bit about your preferences when it comes to those visual elements?
- While using the device, was the 3D effect at any point hindered or broken? If so, what do you think caused that?
- While using the device, did you at any point experience any discomfort? If so what kind and what do you think was the reason?

For the deceived group:

- In the beginning you were informed that this display is a new type of an autostereoscopic 3D display, while in reality it’s a regular 2D laptop screen and the 3D effect was simulated. What are your thoughts on that?

For all:

Do you have any other comments or questions? Anything you would like to add?

Appendix F

Regression results

Table F1

Design-level coefficients for Time-On-Task

Design (cue combinations and positioning)	Location	CI.025	CI.975
1. Normal parallax	0.4634222	-1.4111337	2.5925951
2. Droplines	-0.0374178	-1.9779815	2.1686148
3. Fog	-0.1501861	-2.2790978	1.6216534
4. Parallax	0.3958791	-1.654585	2.4549799
5. Depth (front)	-0.4592045	-2.6623091	1.3876155
6. Height (below)	1.066549	-0.7331492	3.514954
7. Droplines:fog	-0.67351	-2.7743427	1.1482488
8. Droplines:parallax	0.6742503	-1.2072904	2.9505754
9. Fog:parallax	0.5954741	-1.3179912	2.9767247
10. Droplines:depth (front)	-0.064172	-1.9981014	1.8516123
11. Fog:depth (front)	0.5728697	-1.3742475	2.8317374
12. Parallax:depth (front)	0.4988317	-1.3693972	2.757479
13. Droplines:height (below)	-0.6343276	-2.8519193	1.1059261
14. Fog:height (below)	0.1039187	-1.8031269	2.0611705
15. Parallax:height (below)	1.3221196	-0.5154733	3.6628404
16. Depth (front):height (below)	0.0933932	-1.9269092	2.0824969
17. Droplines:fog:parallax	-0.6996818	-2.8644416	1.0220236
18. Droplines:fog:depth (front)	-1.1193013	-3.420873	0.7015426
19. Droplines:parallax:depth (front)	-0.6832001	-2.9058531	1.0302406
20. Fog:parallax:depth (front)	1.7769153	-0.1890201	4.3956946
21. Droplines:fog:height (below)	-0.705915	-2.8816834	1.2087256
22. Droplines:parallax:height (below)	-1.4282464	-3.8012	0.3265965
23. Fog:parallax:height (below)	0.3535953	-1.5331212	2.3208248
24. Droplines:depth (front): height (below)	0.7371038	-1.1711711	3.1182747
25. Fog:depth (front):height (below)	0.2082367	-1.7172445	2.126065
26. Parallax:depth (front):height (below)	-0.6933287	-3.0545991	1.0223745
27. Droplines:fog:parallax:depth (front)	-0.2286954	-2.3303498	1.6993124
28. Droplines:fog:parallax:height (below)	-1.1685892	-3.5721379	0.5268309
29. Droplines:fog:depth (front):height (below)	0.2580755	-1.7135036	2.3364454
30. Droplines:parallax:depth (front):height (below)	-0.1039952	-2.2181085	1.8786816
31. Fog:parallax:depth (front):height (below)	1.3650741	-0.5569043	4.0265684
32. Droplines:fog:parallax:depth(front):height (below)	-1.4409426	-3.8111651	0.2695293

Note: Posterior distributions of the parameters are presented as a mean (location) and its 95% CI. The parameters are estimated in seconds, with the minus value meaning a decrease in ToT.

Table F2

Design-level coefficients for depth estimation accuracy

Design (cue combinations and positioning)	Location	CI.025	CI.975
1. Normal parallax	1.4056012	1.0786111	1.8474975
2. Droplines	0.8331133	0.6225441	1.1027934
3. Fog	1.4474306	1.103471	1.9348651
4. Parallax	1.552815	1.182445	2.0517611
5. Depth (front)	1.9308312	1.4632795	2.5586333
6. Height (below)	1.5442383	1.1840188	2.0379356
7. Droplines:fog	0.6721037	0.4978148	0.89754
8. Droplines:parallax	0.6810089	0.5013777	0.8984477
9. Fog:parallax	1.2769602	0.9766711	1.6703689
10. Droplines:depth (front)	0.7970321	0.5910859	1.0499973
11. Fog:depth (front)	1.4327426	1.0866577	1.8751528
12. Parallax:depth (front)	1.4694012	1.1127342	1.9230386
13. Droplines:height (below)	0.6139564	0.4469987	0.8098518
14. Fog:height (below)	1.7983059	1.3804768	2.3490004
15. Parallax:height (below)	1.4511672	1.0876868	1.8776418
16. Depth (front):height (below)	1.5252164	1.1661622	2.0029573
17. Droplines:fog:parallax	0.7211668	0.5269756	0.9495652
18. Droplines:fog:depth (front)	0.6167702	0.4517696	0.8227547
19. Droplines:parallax:depth (front)	0.882816	0.6588192	1.1553596
20. Fog:parallax:depth (front)	1.136464	0.860482	1.4879595
21. Droplines:fog:height (below)	0.5702168	0.4159111	0.7590863
22. Droplines:parallax:height (below)	0.6328185	0.4597879	0.8459451
23. Fog:parallax:height (below)	1.5202472	1.1590862	2.0068132
24. Droplines:depth (front): height (below)	0.5980147	0.4340888	0.7920695
25. Fog:depth (front):height (below)	1.1490848	0.8643462	1.5074092
26. Parallax:depth (front):height (below)	1.1990899	0.9023068	1.5804705
27. Droplines:fog:parallax:depth (front)	0.6705776	0.4915885	0.8813875
28. Droplines:fog:parallax:height (below)	0.6549532	0.4805695	0.8633437
29. Droplines:fog:depth (front):height (below)	0.5876509	0.4269964	0.7795754
30. Droplines:parallax:depth (front):height (below)	0.8332055	0.6200641	1.0906608
31. Fog:parallax:depth (front):height (below)	1.4421868	1.1056018	1.9058327
32. Droplines:fog:parallax:depth(front):height (below)	0.7956206	0.589338	1.0463036

Note: Posterior distributions of the parameters are presented as a mean (location) and its 95% CI. The parameters are estimated in factors by which the odds change, with the values smaller than 1 signifying a decrease in the degree-of error in depth estimation.

Table F3

Design-level coefficients for height estimation accuracy

Design (cue combinations and positioning)	Location	CI.025	CI.975
1. Normal parallax	0.7206731	0.5465553	0.9573935
2. Droplines	0.9167568	0.6979625	1.168598
3. Fog	0.946044	0.7248146	1.2097078
4. Parallax	0.6935593	0.5253359	0.8857743
5. Depth (front)	1.0348721	0.8019207	1.3391752
6. Height (below)	0.6220867	0.4544224	0.7946942
7. Droplines:fog	1.5060915	1.1707061	1.9197519
8. Droplines:parallax	0.894718	0.6873952	1.1456074
9. Fog:parallax	0.9458829	0.7324403	1.2008664
10. Droplines:depth (front)	1.1450115	0.8970137	1.4427888
11. Fog:depth (front)	1.2943509	1.0028452	1.6663972
12. Parallax:depth (front)	1.1740651	0.9155977	1.5059234
13. Droplines:height (below)	0.8034789	0.6025603	1.0323434
14. Fog:height (below)	0.9408676	0.7155117	1.1994207
15. Parallax:height (below)	0.5946256	0.4439825	0.7818423
16. Depth (front):height (below)	0.8419008	0.6470164	1.0800062
17. Droplines:fog:parallax	1.5885842	1.24452	1.9985011
18. Droplines:fog:depth (front)	1.5560029	1.2096742	1.9711569
19. Droplines:parallax:depth (front)	1.445383	1.1386552	1.859452
20. Fog:parallax:depth (front)	1.4699614	1.1698825	1.8661714
21. Droplines:fog:height (below)	1.1893607	0.9203097	1.5108693
22. Droplines:parallax:height (below)	0.7095844	0.5380114	0.9296224
23. Fog:parallax:height (below)	0.8283181	0.632203	1.0618711
24. Droplines:depth (front): height (below)	0.793651	0.6000734	1.0200635
25. Fog:depth (front):height (below)	1.2313381	0.9547672	1.5566971
26. Parallax:depth (front):height (below)	0.7013946	0.5330084	0.8957832
27. Droplines:fog:parallax:depth (front)	1.891979	1.4785827	2.3849775
28. Droplines:fog:parallax:height (below)	1.0471421	0.8089255	1.3261501
29. Droplines:fog:depth (front):height (below)	1.0564918	0.8160268	1.3384555
30. Droplines:parallax:depth (front):height (below)	0.7838615	0.5957808	0.9943794
31. Fog:parallax:depth (front):height (below)	0.9389954	0.7124629	1.1933768
32. Droplines:fog:parallax:depth(front):height (below)	0.9444272	0.7358336	1.2031052

Note: Posterior distributions of the parameters are presented as a mean (location) and its 95% CI. The parameters are estimated in factors by which the odds change, with the values smaller than 1 signifying a decrease in the degree-of error in height estimation.

Table F4

Design-level coefficients for believability

Design (cue combinations and positioning)	Location	CI.025	CI.975
1. Normal parallax	1.0471012	0.8205478	1.3574928
2. Droplines	1.0055303	0.7668522	1.3056721
3. Fog	0.832667	0.6256005	1.0653158
4. Parallax	0.9384877	0.7276357	1.2194338
5. Depth (front)	1.1879873	0.940935	1.5717107
6. Height (below)	0.9860187	0.760449	1.2738274
7. Droplines:fog	0.9480982	0.7194553	1.2305478
8. Droplines:parallax	0.9645641	0.7426284	1.2548
9. Fog:parallax	0.9733379	0.753143	1.2645353
10. Droplines:depth (front)	1.0205386	0.7926916	1.3261934
11. Fog:depth (front)	1.0769395	0.8284034	1.469989
12. Parallax:depth (front)	1.0116414	0.7603562	1.3356686
13. Droplines:height (below)	0.9402769	0.7258925	1.2236051
14. Fog:height (below)	1.00559	0.7764615	1.3350381
15. Parallax:height (below)	0.976451	0.7599922	1.2655001
16. Depth (front):height (below)	1.0176966	0.8072802	1.334561
17. Droplines:fog:parallax	0.9496562	0.7161761	1.2251246
18. Droplines:fog:depth (front)	0.8659648	0.643733	1.108937
19. Droplines:parallax:depth (front)	1.3560443	1.0452348	1.8548166
20. Fog:parallax:depth (front)	1.0505054	0.8352091	1.3679558
21. Droplines:fog:height (below)	0.8763337	0.6698162	1.1334351
22. Droplines:parallax:height (below)	1.047106	0.8102275	1.3575462
23. Fog:parallax:height (below)	0.7933395	0.5981828	1.0306872
24. Droplines:depth (front): height (below)	1.4502318	1.1157695	1.9841664
25. Fog:depth (front):height (below)	0.9459549	0.7300179	1.2154897
26. Parallax:depth (front):height (below)	1.2788415	0.9918552	1.7211311
27. Droplines:fog:parallax:depth (front)	0.8749996	0.6574513	1.1209963
28. Droplines:fog:parallax:height (below)	0.8739816	0.6573728	1.1112024
29. Droplines:fog:depth (front):height (below)	0.8358961	0.6251834	1.066815
30. Droplines:parallax:depth (front):height (below)	1.2476201	0.9699961	1.6931136
31. Fog:parallax:depth (front):height (below)	1.0205954	0.7907263	1.3359937
32. Droplines:fog:parallax:depth(front):height (below)	0.9123762	0.6930005	1.1623374

Note: Posterior distributions of the parameters are presented as a mean (location) and its 95% CI. The parameters are estimated in factors by which the mean changes, with the values smaller than 1 signifying a decrease in the believability ratings.

Table F5

Design-level coefficients for aesthetics

Design (cue combinations and positioning)	Location	CI.025	CI.975
1. Normal parallax	1.2437081	0.9144914	1.7282767
2. Droplines	1.4339801	1.0508866	1.9913788
3. Fog	0.5119536	0.3558119	0.7192242
4. Parallax	1.5290868	1.1108376	2.1651085
5. Depth (front)	1.2981633	0.955881	1.8203033
6. Height (below)	1.0187316	0.7542716	1.3975965
7. Droplines:fog	0.7773231	0.5489604	1.0623123
8. Droplines:parallax	1.3174097	0.9508398	1.8738159
9. Fog:parallax	0.7101518	0.5034777	0.9886243
10. Droplines:depth (front)	1.3580412	0.9891865	1.9020656
11. Fog:depth (front)	1.1611109	0.8333307	1.6399712
12. Parallax:depth (front)	1.404901	1.0141135	1.9550136
13. Droplines:height (below)	1.0036268	0.7346013	1.3720979
14. Fog:height (below)	0.6532522	0.469072	0.9042098
15. Parallax:height (below)	1.0767402	0.7820928	1.480762
16. Depth (front):height (below)	1.4586814	1.0538936	2.0644133
17. Droplines:fog:parallax	0.6569769	0.463539	0.9045408
18. Droplines:fog:depth (front)	0.6773571	0.4717175	0.9389154
19. Droplines:parallax:depth (front)	1.4538537	1.0703569	2.0434012
20. Fog:parallax:depth (front)	0.9898091	0.7183052	1.3831844
21. Droplines:fog:height (below)	0.6936318	0.488568	0.9672035
22. Droplines:parallax:height (below)	1.3737938	0.988695	1.9126641
23. Fog:parallax:height (below)	0.5812898	0.4132615	0.8099899
24. Droplines:depth (front): height (below)	1.9819698	1.4259579	2.8386483
25. Fog:depth (front):height (below)	0.9121423	0.6523637	1.2758707
26. Parallax:depth (front):height (below)	1.3802666	1.0042125	1.9554922
27. Droplines:fog:parallax:depth (front)	0.7322027	0.515407	1.0167865
28. Droplines:fog:parallax:height (below)	0.7512809	0.5391591	1.0296019
29. Droplines:fog:depth (front):height (below)	0.8067182	0.5826993	1.1294398
30. Droplines:parallax:depth (front):height (below)	1.3254985	0.9664399	1.841137
31. Fog:parallax:depth (front):height (below)	0.9170508	0.6639744	1.2778668
32. Droplines:fog:parallax:depth(front):height (below)	0.6982031	0.4925456	0.9716782

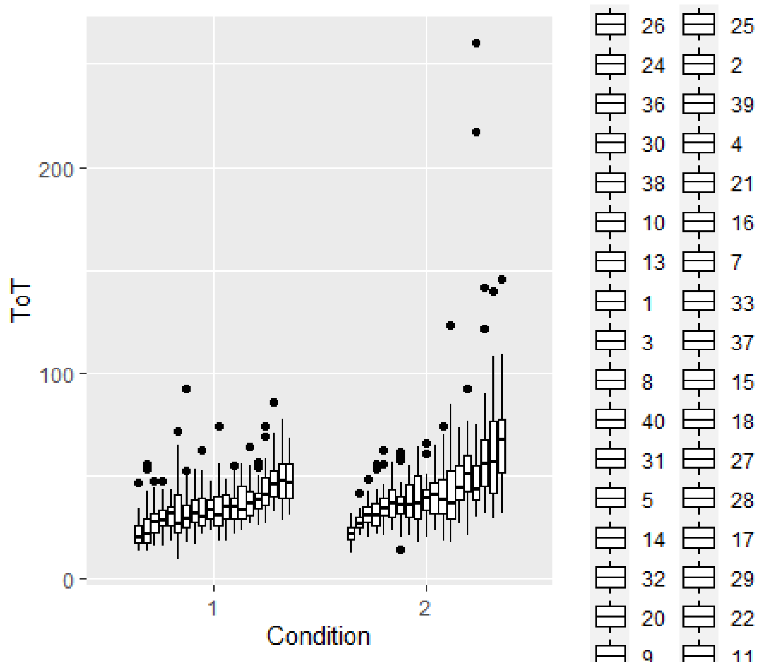
Note: Posterior distributions of the parameters are presented as a mean (location) and its 95% CI. The parameters are estimated in factors by which the mean changes, with the values smaller than 1 signifying a decrease in the aesthetic ratings.

Appendix G

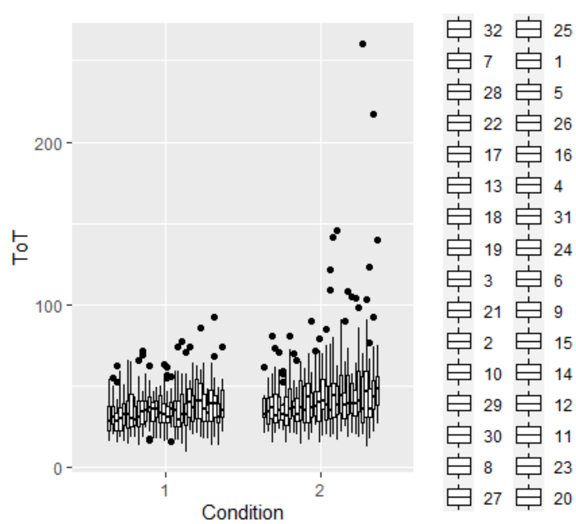
R Syntax

```
E1_data <- read_csv("New_exp_data10_v4.csv",
                    col_types = cols(
                                  Part = col_character(),
                                  Condition = col_character(),
                                  Obs = col_double(),
                                  droplines = col_logical(),
                                  fog = col_logical(),
                                  height = col_double(),
                                  depth = col_double(),
                                  parallax = col_character(),
                                  Design = col_character(),
                                  heightInput = col_double(),
                                  depthInput = col_double(),
                                  believability = col_double(),
                                  aesthetics = col_double(),
                                  ToT = col_double(),
                                  accuracyHeight = col_double(),
                                  accuracyDepth = col_double(),
                                  absAccuracyHeight = col_double(),
                                  absAccuracyDepth = col_double(),
                                  depthDistance = col_double(),
                                  heightFactor = col_character(),
                                  depthFactor = col_character()
                                )) %>%



E2_data_cleanToT <-
  E1_data %>%
    filter(Part != 23, !is.na(ToT))



E2_data_cleanToT %>%
    ggplot(aes(x = Condition,
               y = ToT,
               color = reorder(Part, ToT))) +
    geom_boxplot() +
  scale_color_discrete(labels = order)+
    scale_colour_manual(values = rep("black",40))
```

Visual Exploration – example code

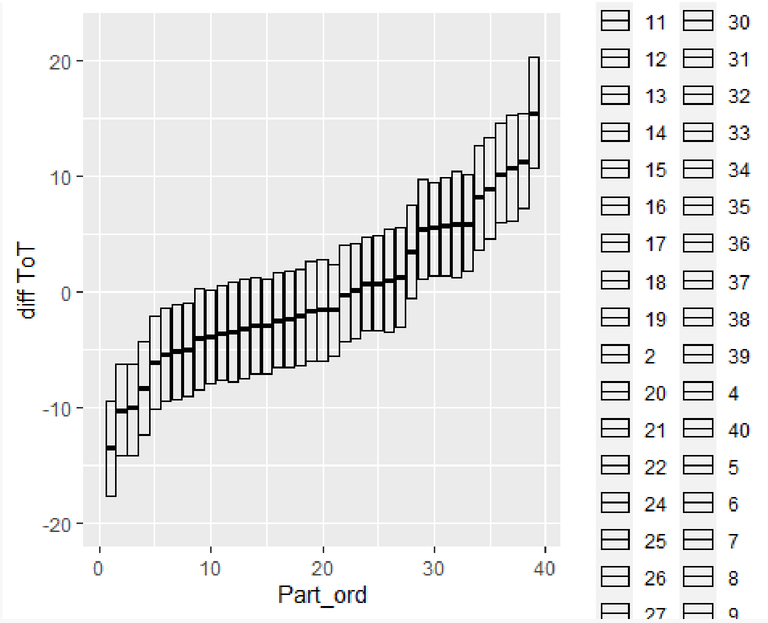



```
E2_data_cleanToT %>%
    ggplot(aes(x = Condition,
               y = ToT,
               color = reorder(Design, ToT))) +
    geom_boxplot() +
  scale_color_discrete(labels = order)+
    scale_colour_manual(values = rep("black",40))
```

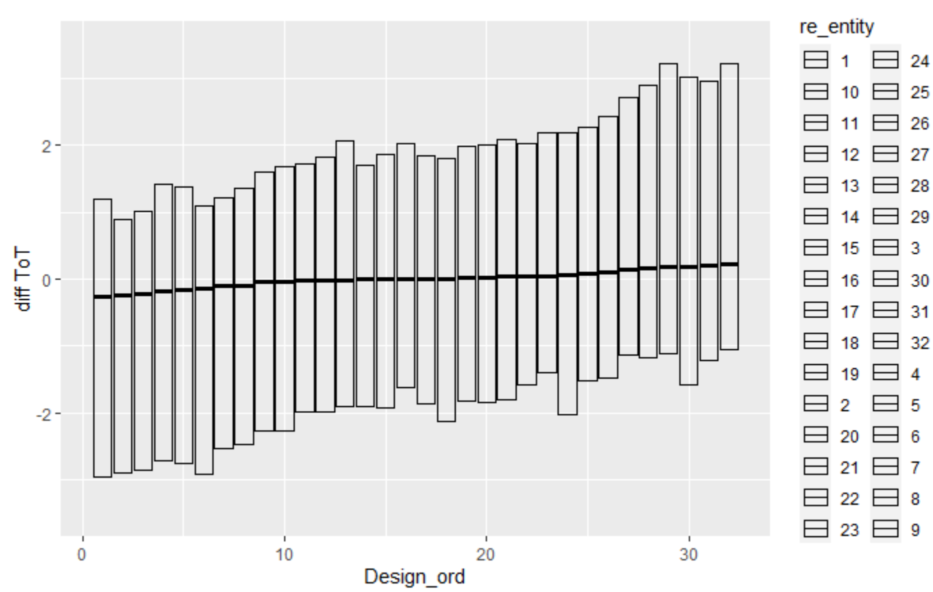


ToT

```
M1_ToT_new <-
  E2_data_cleanToT %>%
  brm(ToT~ 1 + Condition + (1|Part) + (1 + Condition|Design),
      family = exgaussian(),
      data = .,
      inits = 0)

E2_data_cleanToT %>%
  select(Part, Design, Condition, ToT) %>%
  bind_cols(., predict(M1_ToT_new)) %>%
  group_by(Design) %>%
  summarize(mean_ToT = mean(center, na.rm = TRUE))

P1 <- posterior(M1_ToT_new)

D_cue8_x <-
  E2_data_cleanToT %>%
  unite(Part, Design:Part, remove = F) %>%
  as_tbl_obs()

D_cue8_x

P1 %>%
  filter(type %in% c("ranef", "fixef", "grpef")) %>%
  clu()

ranef(P1) %>%
  filter(fixef == "Intercept", re_factor == "Part") %>%
  mutate(Part_ord = rank(center)) %>%
  ggplot(aes(x = Part_ord, ymin = lower, y = center, ymax = upper, color = re
_entity)) +
  geom_crossbar() +
  scale_color_discrete(labels = order) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff ToT") +
  ylim(-20,22)
```




```
ranef(P1) %>%
  filter(fixef == "Condition2", re_factor == "Design") %>%
  mutate(Design_ord = rank(center)) %>%
  ggplot(aes(x = Design_ord, ymin = lower, y = center, ymax = upper, color =
re_entity)) +
  geom_crossbar() +
  scale_color_discrete(labels = order) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff ToT")
```



```
ranef(P1) %>%
  filter(fixef == "Condition2" | fixef == "Intercept",re_factor == "Design")
 %>%
  group_by(fixef) %>%
  mutate(Design_ord = rank(center)) %>%
  ggplot(aes(x = Design_ord, ymin = lower, y = center, ymax = upper, color =
re_entity, na.rm = TRUE)) +
  geom_crossbar() +
  facet_wrap(~fixef) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff ToT") +
  ylim(-20,22)
```

```
fixef_ml(M1_ToT_new)
```

Population-level coefficients with random effects standard deviations

fixef	center	lower	upper	SD_Design	SD_Part
Intercept	37.041891	33.847343	40.192948	1.3202741	6.860395
Condition2	3.686489	-0.944734	8.220153	0.7180166	NA

```
grpef(M1_ToT_new)
```

Coefficient estimates with 95% credibility limits

fixef	re_factor	center	lower	upper
Intercept	Design	1.3202741	0.2789404	2.322510
Condition2	Design	0.7180166	0.0378217	2.235332
Intercept	Part	6.8603950	5.4555702	8.769026

```
ranef(M1_ToT_new)
```

Accuracy of depth estimation

```
#mutating accuracy to fit 0-1
E1_data_2_depth <-
  E1_data %>%
  mutate(absAccuracyDepth = mascutils::rescale_unit(absAccuracyDepth, lower=0,
 upper=280)) %>%
  mutate(absAccuracyDepth = mascutils::rescale_centered(absAccuracyDepth, sca
le=.999))

M1_depth <-
  E1_data_2_depth %>%
    brm(absAccuracyDepth ~ 1 + Condition + (1|Part) + (1 + Condition|Design),
    family = Beta(link = "logit"),
    data = .,
    inits = 0)

ranef(M1_depth, mean.func = exp)

fixef_ml(M1_depth, mean.func = exp)
```

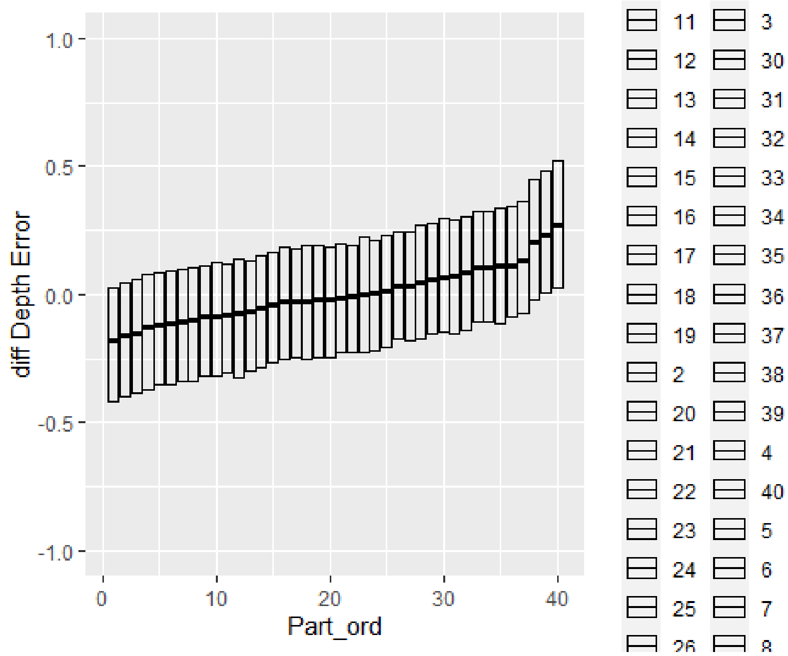
Population-level coefficients with random effects standard deviations

fixef	center	lower	upper	SD_Design	SD_Part
Intercept	0.0727177	0.0599436	0.0865417	1.530169	1.170009
Condition2	0.9605756	0.8322766	1.1074982	1.131237	NA

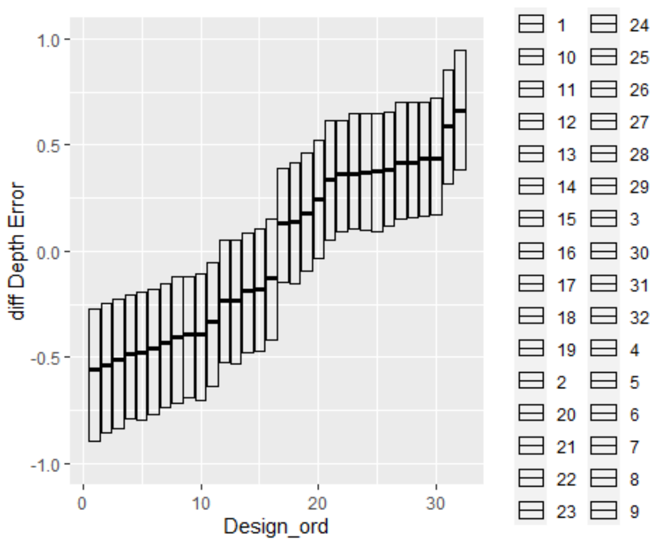
```
P2_depth <- posterior(M1_depth)

P2_depth %>%
  filter(type %in% c("ranef", "fixef", "grpef")) %>%
  clu()
```

Parameter estimates with 95% credibility limits

```
ranef(P2_depth) %>%
  filter(fixef == "Intercept", re_factor == "Part") %>%
  mutate(Part_ord = rank(center)) %>%
  ggplot(aes(x = Part_ord, ymin = lower, y = center, ymax = upper, color = re
_entity)) +
  geom_crossbar() +
  scale_color_discrete(labels = order) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff Depth Error") +
  ylim(-1,1)
```

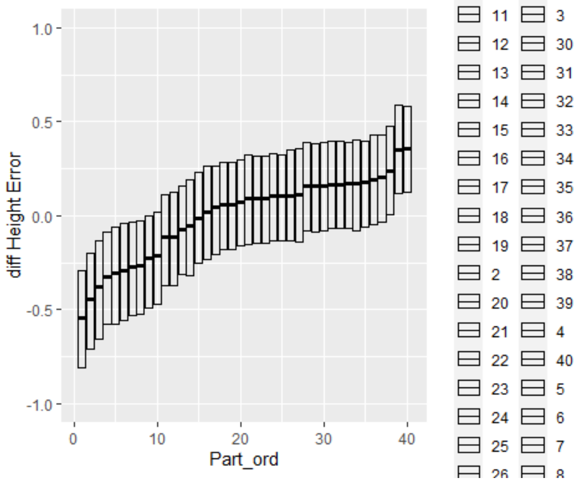



```
ranef(P2_depth) %>%
  filter(fixef == "Intercept", re_factor == "Design") %>%
  mutate(Design_ord = rank(center)) %>%
  ggplot(aes(x = Design_ord, ymin = lower, y = center, ymax = upper, color =
re_entity)) +
  geom_crossbar() +
  labs(y = "diff Depth Error") +
  scale_colour_manual(values = rep("black",40))+
  ylim(-1,1)
```

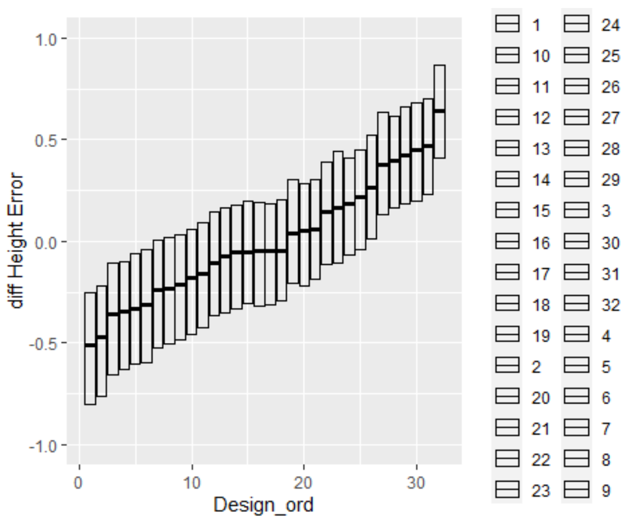


Accuracy of height estimation

```
#mutating accuracy to fit 0-1
E1_data_2_height <-
  E1_data %>%
  mutate(absAccuracyHeight = mascutils::rescale_unit(absAccuracyHeight, lower
=0, upper=280)) %>%
  mutate(absAccuracyHeight = mascutils::rescale_centered(absAccuracyHeight, s
cale=.999))

M1_Height <-
  E1_data_2_height %>%
    brm(
      formula = absAccuracyHeight ~ 1 + Condition + (1|Part) + (1 + Condition
|Design),
    family = Beta(link = "logit"),
    data = .,
    inits = 0)

ranef(M1_Height, mean.func = exp)

P3_height <- posterior(M1_Height)

P3_height %>%
  filter(type %in% c("ranef", "fixef", "grpef")) %>%
  clu()

ranef(P3_height) %>%
  filter(fixef == "Intercept", re_factor == "Part") %>%
  mutate(Part_ord = rank(center)) %>%
  ggplot(aes(x = Part_ord, ymin = lower, y = center, ymax = upper, color = re
_entity)) +
  geom_crossbar() +
  scale_color_discrete(labels = order) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff Height Error") +
  ylim(-1,1)
```




```
ranef(P3_height) %>%
  filter(fixef == "Intercept", re_factor == "Design") %>%
  mutate(Design_ord = rank(center)) %>%
  ggplot(aes(x = Design_ord, ymin = lower, y = center, ymax = upper, color =
re_entity)) +
  geom_crossbar() +
  labs(y = "diff Height Error") +
  scale_colour_manual(values = rep("black",40))+
  ylim(-1,1)
```



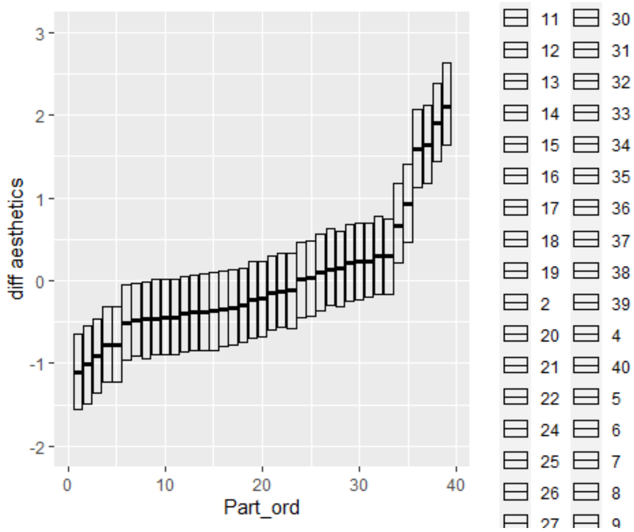

```
fixef_ml(M1_Height, mean.func = exp)
```

Population-level coefficients with random effects standard deviations

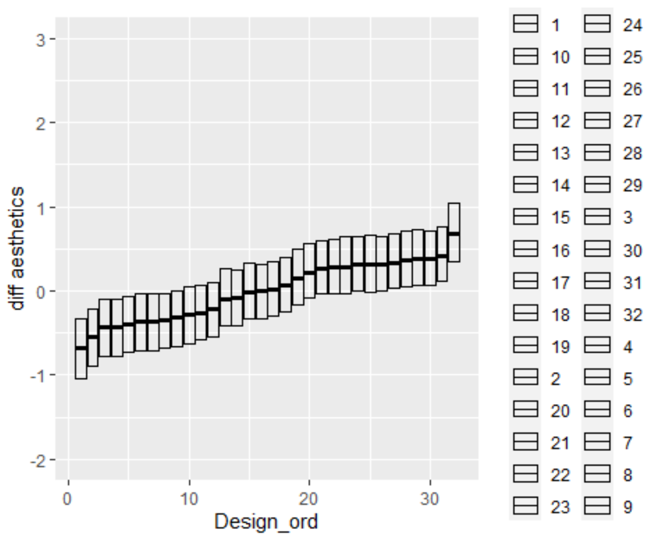
fixef	center	lower	upper	SD_Design	SD_Part
Intercept	0.0345372	0.0290321	0.0412276	1.388059	1.290004
Condition2	0.8941790	0.7426748	1.0691754	1.108826	NA

Models for aesthetics and believability

```
#mutating aesthetics to fit 0-1 for beta
E1_data_beta_aes2 <-
  E2_data_cleanToT %>%
  mutate(aesthetics = mascutils::rescale_unit(aesthetics, lower=1, upper=11))
 %>%
  mutate(aesthetics = mascutils::rescale_centered(aesthetics, scale=.999))

#mutating aesthetics to fit 0-1 for beta
E1_data_beta_bel2 <-
  E2_data_cleanToT %>%
  mutate(believability = mascutils::rescale_unit(believability, lower=1, uppe
r=11)) %>%
  mutate(believability = mascutils::rescale_centered(believability, scale=.99
9))

M1_aes <-
  E1_data_beta_aes2 %>%
  brm(aesthetics ~ 1 + Condition + (1|Part) + (1 + Condition|Design),
            family = Beta(link = "logit"),
           data = .,
      inits = 0)

P4_aes <- posterior(M1_aes)

P4_aes %>%
  filter(type %in% c("ranef", "fixef", "grpef")) %>%
  clu()

ranef(P4_aes) %>%
  filter(fixef == "Intercept", re_factor == "Part") %>%
  mutate(Part_ord = rank(center)) %>%
  ggplot(aes(x = Part_ord, ymin = lower, y = center, ymax = upper, color = re
_entity)) +
  geom_crossbar() +
  scale_color_discrete(labels = order) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff aesthetics") +
  ylim(-2,3)
```

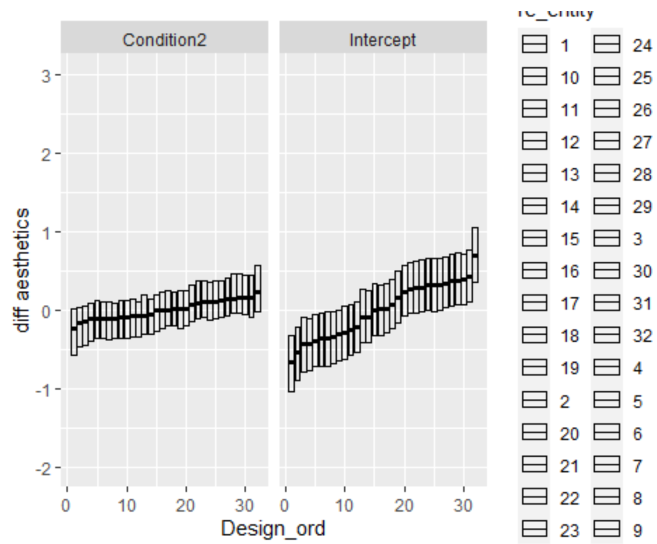



```
ranef(P4_aes) %>%
  filter(fixef == "Intercept", re_factor == "Design") %>%
  mutate(Design_ord = rank(center)) %>%
  ggplot(aes(x = Design_ord, ymin = lower, y = center, ymax = upper, color =
re_entity)) +
  geom_crossbar() +
  labs(y = "diff aesthetics") +
  scale_colour_manual(values = rep("black",40))+
  ylim(-2,3)
```




```
ranef(P4_aes) %>%
  filter(fixef == "Condition2" | fixef == "Intercept",re_factor == "Design")
 %>%
  group_by(fixef) %>%
  mutate(Design_ord = rank(center)) %>%
  ggplot(aes(x = Design_ord, ymin = lower, y = center, ymax = upper, color =
re_entity, na.rm = TRUE)) +
  geom_crossbar() +
```

```
  facet_wrap(~fixef) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff aesthetics") +
  ylim(-2,3)
```




```
fixef_ml(M1_aes, mean.func = exp)
```

Population-level coefficients with random effects standard deviations

fixef	center	lower	upper	SD_Design	SD_Part
Intercept	1.495351	1.0041532	2.241267	1.473924	2.218557
Condition2	1.416133	0.8099478	2.355647	1.179808	NA

```
M1_bel <-
  E1_data_beta_bel2 %>%
  brm(believability ~ 1 + Condition + (1|Part) + (1 + Condition|Design),
            family = Beta(link = "logit"),
           data = .,
      inits = 0)

ranef(M1_bel, mean.func = exp)

grpef(M1_bel, mean.func = exp)
```

Coefficient estimates with 95% credibility limits

fixef	re_factor	center	lower	upper
Intercept	Design	0.1965886	0.1069893	0.3033589
Condition2	Design	0.1834668	0.0285456	0.3398815
Intercept	Part	0.9305843	0.7497441	1.2210726

```
fixef_ml(M1_bel, mean.func = exp)
```

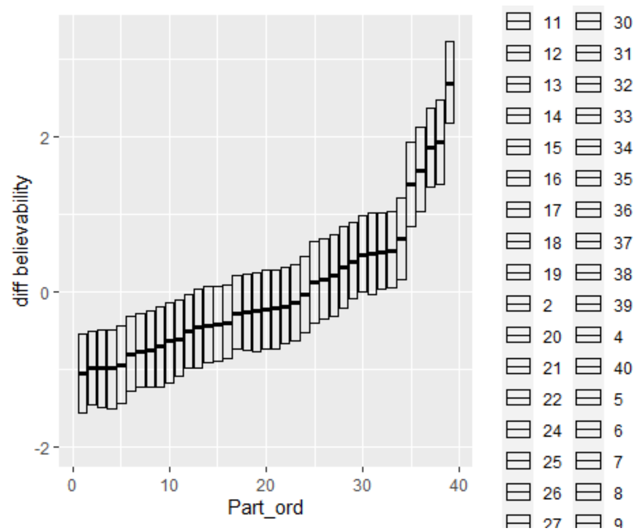
Population-level coefficients with random effects standard deviations

fixef	center	lower	upper	SD_Design	SD_Part
Intercept	**2.284137**	**1.4975334**	**3.440692**	**1.217243**	**2.535991**
Condition2	**1.268874**	**0.7027161**	**2.353035**	**1.201375**	**NA**

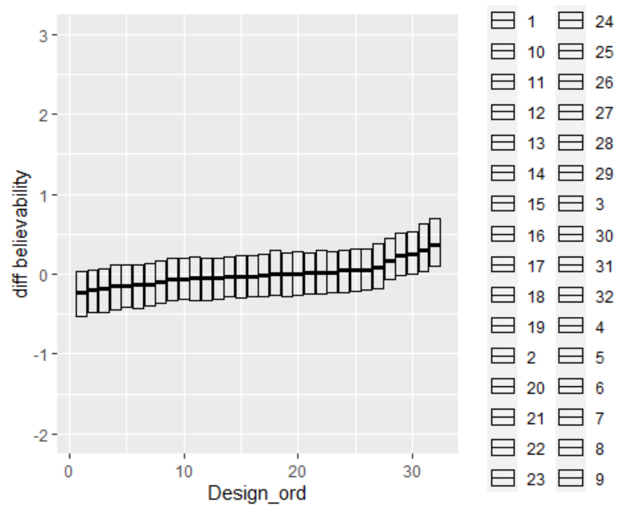
```
P5_bel <- posterior(M1_bel)

P5_bel %>%
  filter(type %in% c("ranef", "fixef", "grpef")) %>%
  clu()
```

Parameter estimates with 95% credibility limits

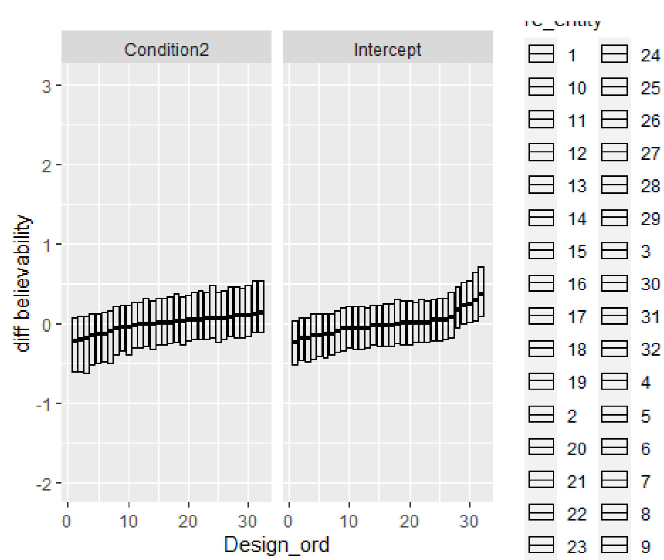
```
ranef(P5_bel) %>%
  filter(fixef == "Intercept", re_factor == "Part") %>%
  mutate(Part_ord = rank(center)) %>%
  ggplot(aes(x = Part_ord, ymin = lower, y = center, ymax = upper, color = re
_entity)) +
  geom_crossbar() +
  scale_color_discrete(labels = order) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff believability") +
  ylim(-2,3.3)
```




```
ranef(P5_bel) %>%
  filter(fixef == "Intercept", re_factor == "Design") %>%
  mutate(Design_ord = rank(center)) %>%
  ggplot(aes(x = Design_ord, ymin = lower, y = center, ymax = upper, color =
re_entity)) +
  geom_crossbar() +
  labs(y = "diff believability") +
  scale_colour_manual(values = rep("black",40))+
  ylim(-2,3)
```

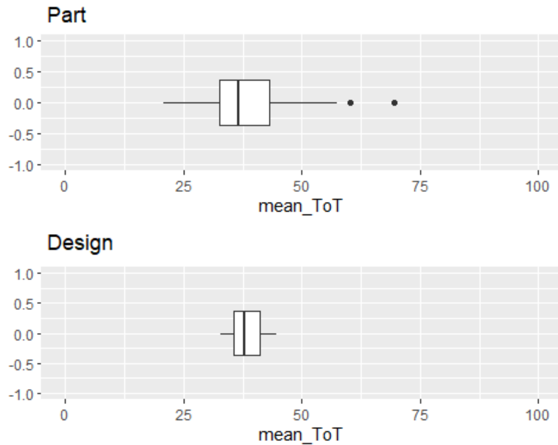



```
ranef(P5_bel) %>%
  filter(fixef == "Condition2" | fixef == "Intercept",re_factor == "Design")
 %>%
  group_by(fixef) %>%
  mutate(Design_ord = rank(center)) %>%
  ggplot(aes(x = Design_ord, ymin = lower, y = center, ymax = upper, color =
re_entity, na.rm = TRUE)) +
  geom_crossbar() +
  facet_wrap(~fixef) +
  scale_colour_manual(values = rep("black",40))+
  labs(y = "diff believability") +
  ylim(-2,3)
```

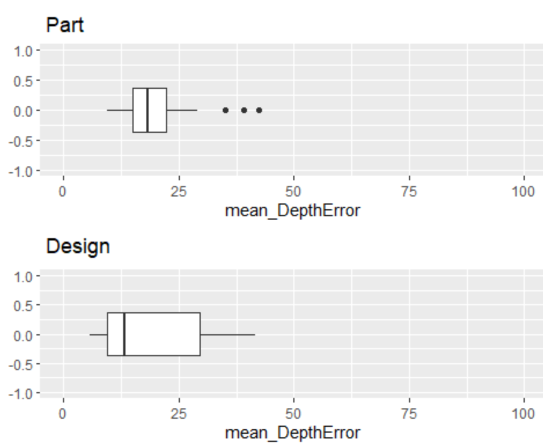



```
E1_data_beta_aes2%>%
  select(Part, Design, Condition, aesthetics) %>%
  na.omit() %>%
  bind_cols(., predict(M1_aes, mean.func = exp)) %>%
```

```
  group_by(Design) %>%
  summarize(mean_aes = mean(center, na.rm = TRUE))

E1_data_beta_bel2%>%
  select(Part, Design, Condition, believability) %>%
  na.omit() %>%
  bind_cols(., predict(M1_bel, mean.func = exp)) %>%
  group_by(Design) %>%
  summarize(mean_bel = mean(center, na.rm = TRUE))

E1_data_2_depth %>%
  select(Part, Design, Condition, absAccuracyDepth) %>%
  na.omit() %>%
  bind_cols(., predict(M1_depth, mean.func = exp)) %>%
  group_by(Design) %>%
  summarize(mean_depthAccu = mean(center, na.rm = TRUE))

E1_data_2_height %>%
  select(Part, Design, Condition, absAccuracyHeight) %>%
  na.omit() %>%
  bind_cols(., predict(M1_Height, mean.func = exp)) %>%
  group_by(Design) %>%
  summarize(mean_depthAccu = mean(center, na.rm = TRUE))

plot_level <- function(Level) {
  level <- enquo(Level)
  out <-
  E1_data_2_depth %>%
  group_by(!!level) %>%
  summarize(mean_ToT = mean(ToT)) %>%
  ggplot(aes(x = mean_ToT)) +
  geom_boxplot () +
  labs(title = quo(!!level)) +
  xlim(0, 100)+
    ylim(-1,1)
}


grid.arrange(
plot_level(Part),
plot_level(Design)
)
```

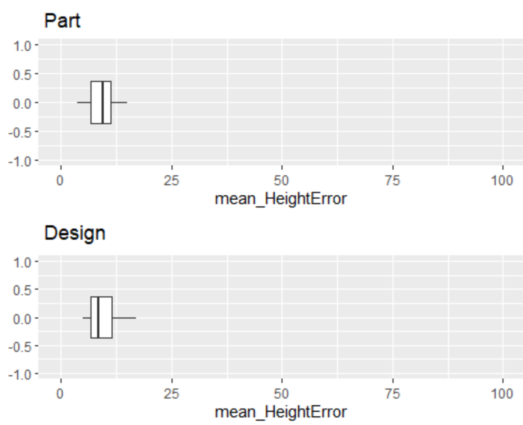



```
plot_level <- function(Level) {
  level <- enquo(Level)
  out <-
  E1_data %>%
  group_by(!!level) %>%
  summarize(mean_DepthError = mean(absAccuracyDepth)) %>%
  ggplot(aes(x = mean_DepthError)) +
  geom_boxplot() +
  labs(title = quo(!!level)) +
  xlim(0, 100) +
    ylim(-1,1)
}


grid.arrange(
  plot_level(Part),
  plot_level(Design)

)
```

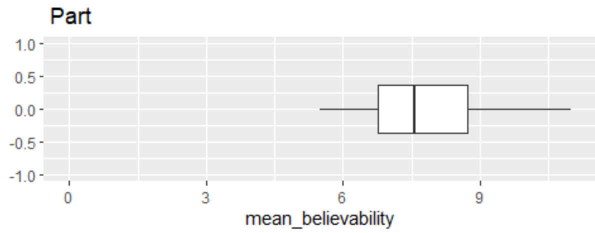


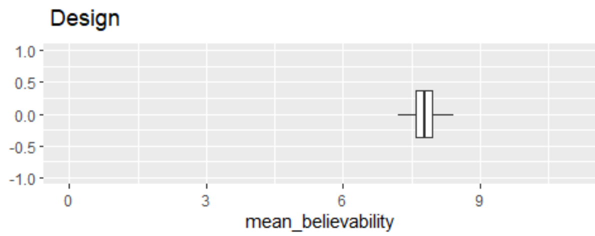

```
plot_level <- function(Level) {
  level <- enquo(Level)
```

```
  out <-
  E1_data %>%
  group_by(!!level) %>%
  summarize(mean_HeightError = mean(absAccuracyHeight)) %>%
  ggplot(aes(x = mean_HeightError)) +
  geom_boxplot() +
  labs(title = quo(!!level)) +
  xlim(0, 100) +
  ylim(-1,1)

}


grid.arrange(
plot_level(Part),
plot_level(Design)
)
```

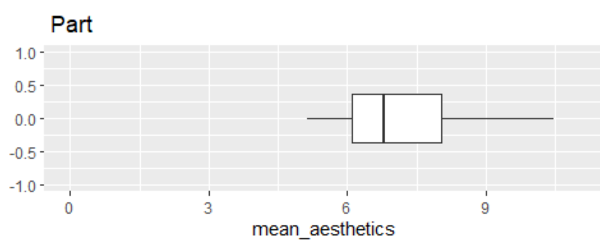


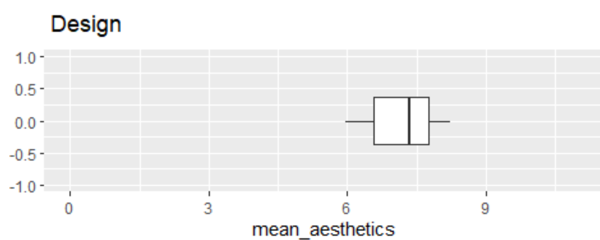

```
plot_level <- function(Level) {
  level <- enquo(Level)
  out <-
  E2_data_cleanToT %>%
  group_by(!!level) %>%
  summarize(mean_believability = mean(believability)) %>%
  ggplot(aes(x = mean_believability)) +
  geom_boxplot() +
  labs(title = quo(!!level)) +
  xlim(0, 11)+
    ylim(-1,1)
}


grid.arrange(
plot_level(Part),
plot_level(Design)
)
```






```
plot_level <- function(Level) {
  level <- enquo(Level)
  out <-
  E2_data_cleanToT %>%
  group_by(!!level) %>%
  summarize(mean_aesthetics = mean(aesthetics)) %>%
  ggplot(aes(x = mean_aesthetics)) +
  geom_boxplot() +
  labs(title = quo(!!level)) +
  xlim(0, 11) +
    ylim(-1,1)
}


grid.arrange(
plot_level(Part),
plot_level(Design)
)
```





```
clu_cor <-
  function(model){
    model %>%
      posterior() %>%
      filter(type == "cor") %>%
      mutate(parameter = str_remove_all(parameter, "cor_")) %>%
      group_by(parameter) %>%
      summarize(center = median(value),
                lower = quantile(value, .025),
                upper = quantile(value, .975)) %>%
      separate(parameter, into = c("re_factor", "between", "and"),
               sep = "__")
  }

M1_ToT_new %>%
  clu_cor()
```

re_factor	between	and	center	lower	upper
Design	Intercept	Condition2	0.0085511	-0.9080967	0.9404518