# The Effect of Interior Bezel Presence and Width on Magnitude Judgement

James R. Wallace [*]
Wilfrid Laurier University

Daniel Vogel [†]
University of Waterloo

Edward Lank [‡]
University of Waterloo

## ABSTRACT

Large displays are often constructed by tiling multiple small displays, creating visual discontinuities from inner bezels that may affect human perception of data. Our work investigates how bezels impact magnitude judgement, a fundamental aspect of perception. Two studies are described which control for bezel presence, bezel width, and user-to-display distance. Our findings form three implications for the design of tiled displays. Bezels wider than 0.5cm introduce a 4–7% increase in judgement error from a distance, which we simplify to a 5% rule of thumb when assessing display hardware. Length judgements made at arm's length are most affected by wider bezels, and are an important use case to consider. At arm's length, bezel compensation techniques provide a limited benefit in terms of judgement accuracy.

**Index Terms:** H.5.2 [User Interfaces]—Graphical user interfaces (GUI);

## 1 INTRODUCTION

Large, interactive displays have been shown to support activities such as navigation and wayfinding [4, 27], the physical navigation of large datasets [3], and co-located, collaborative work [29]. However, when constructing physically large displays one must choose between two imperfect approaches. First, one may buy a single, contiguous display. However, these displays are costly, and are available in limited sizes, aspect ratios, and resolutions of 1080 or 4K that are inadequate at large scales. Alternatively, one may tile many smaller displays together, offering lower cost, more flexible display sizes, and significantly higher resolution, but with the trade-off of introducing visual discontinuities created by individual display frames, called *interior bezels*. These bezels are aesthetically unpleasing, and potentially disruptive to users [7].

Previous work has explored aspects of large display interaction and perception when interior bezels are present (e.g. [7, 6]) and proposed mapping techniques to compensate for bezels [13]. However, the extent to which data crossing interior bezels interferes with magnitude judgement remains unexplored. It has been suggested that existing research has lacked control over many confounds in the design and use of tiled displays [7], and that a more careful investigation of these design issues would assist in developing software that takes full advantage of their capabilities [31]. Such an understanding is very practical. For example, display manufacturers now offer panels with "ultra-thin" bezels, intended for use in video wall installations and priced at a premium compared to their desktop counterparts. Based on our current understanding of the interactions between bezels and human perception, it remains unclear how much of a benefit, if any, these displays provide.

In this work, we contribute practical results regarding the presence and width of bezels, motivated by the construction of our own

---

[*]e-mail:jwallace@wlu.ca

[†]e-mail:dvogel@uwaterloo.ca

[‡]e-mail:lank@uwaterloo.ca

Figure 1: Tiling many independent displays provides a means of creating large workspaces that can support activities such as data visualization and sensemaking, but introduces interior bezels.

4m wide interactive display. We present the results of two empirical studies that investigate human perception of data across bezels at a distance and at arm's length. Our results show that in most cases, introducing interior bezels has an impact on human ability to gauge the relative size of data, but once bezels are introduced, larger interior bezel width has minimal effect. However, for interactions at arm's length, we found that wider bezels more significantly impacted magnitude judgement. We also tested a 'French Window' [13] bezel compensation technique and found no reduction of overall user error. These results lead directly to practical design guidelines for large, tiled displays: (1) bezels introduce approximately 5% additional error into user magnitude judgements; (2) length judgements made at arm's length are affected by wide bezels; and (3) bezel compensation techniques have a limited effect on judgement, but may provide aesthetic improvements.

## 2 TILED DISPLAYS

Tiling multiple, high-resolution monitors enables researchers to create displays that offer a number of advantages when compared to projected or traditional desktop displays. For example, tiled displays can support resolutions that surpass human visual acuity [40], and allow users to view images and data an order of magnitude larger than is possible on traditional displays [38]. The large surfaces created by tiled displays provide a shared workspace for the analysis of data that can benefit users both within arms reach, and those nearby [23]. And the combination of touch interaction, and the ability to physically navigate large data sets provides opportunities to leverage new modalities of interaction such as proxemics [19, 17] to enhance interactions with these displays. In line with these opportunities, tiled displays have been constructed to support work in many fields, including analytics [31], information visualization [28, 38], and command and control [26, 24].

Despite these advantages, the literature questions whether interior bezels may interfere with tasks such as visual search [14], stereoscopic vision [18], or target acquisition [22], and whether

bezels may negatively impact user experience and aesthetic [2]. For example, can users effectively interpret data when it is divided across many small screens? To mitigate these potential issues, researchers have contributed guidelines for the development of software on tiled displays [1], such as placing materials on tiled displays to avoid crossing bezels, as well as techniques to alleviate their impact on human perception and interaction. For example, de Almeida et al. [13] simulated a 'French Window' effect via head tracking data, and allowed users to 'look behind' bezels to improve their performance at a cross-display tracing task.

However not all research supports the claim that bezels negatively impact user performance. Moreland [28] notes that there is little research supporting a negative impact on user perception. McNamara et al. [27] studied their impact on navigation and wayfinding and found no difference between conditions with and without a bezel. Robertson et al. [30] reported both positive and negative effects in their investigation of visual search on large, tiled displays. Other research has suggested that bezels have little or no negative impact on tiled display usability (e.g. [34, 39, 9]), and may even be leveraged by users to *improve* their performance [4, 6, 25].

Bi et al. [7] note the difficulties in drawing strong conclusions due to the lack of experimental control in the literature and argue that it would be beneficial to further explore the impact of bezels on user interaction with tiled displays. In their work, Bi et al. explored the impact of bezels on visual search, target selection, and tunnel steering tasks using a mouse and keyboard at a distance, and controlled for the number of interior bezels present on a user's display. However, they did not experimentally control for the width of interior bezels, or investigate their impact on human perception. Grüninger and Krüger [18] explored the impact of bezel width, colour, and tile size on depth perception and found that bezels smaller than 1.2cm provided little benefit to users, however their investigation did not provide guidance for 2D displays. We aim to build on their work and determine if interior bezel presence and width may impact human perception of displayed data, both at a distance and at arm's length.

## 2.1 Magnitude Adjustment Tasks

To investigate any distortions introduced by the presence of interior bezels, we leveraged an approach developed by psychophysicists to explore the impact of physical phenomena on human perception: the magnitude adjustment task [15, 16]. The magnitude adjustment task consists of a series modulus and stimulus shapes, and asks participants to make estimates of their relative sizes. These tests primarily evaluate performance according to Stevens' Power Law [32], and rely on a user's ability to estimate or reproduce a displayed stimulus. Typically, stimuli representative of graphical representations such as bar, line, and pie graphs are used, as they provide a measure of how effectively users will be able to interpret common visualizations of data, as found in the scientific literature [11, 10]. A comprehensive summary of the development of magnitude judgment tasks is beyond the scope of this work. For a more recent summary of magnitude judgement studies, and meta-analysis of visual phenomenon that may impact user perception for these tasks, we direct the reader to Wagner [35].

HCI researchers have used magnitude judgement tasks to identify scenarios in which human perception of data is compromised. For example, Wigdor et al. [37] identified cases where perception of shapes on a shared, digital tabletop may be compromised. Similarly, in the context of large, tiled displays, Bezerianos et al. [6] found that the distortions caused by viewing a display at extreme angles impacted a user's ability to perceive data. By controlling the conditions under which participants engage with data, these studies are able to isolate the impact of factors of interest from confounds such as display size [31], position and angle [33], or the presence and position of other users [21], and contribute towards a funda-

mental understanding of how human perception is influenced by factors such as the presence of bezels. Magnitude adjustment tasks are particularly useful in the context of this research because they not only can help to identify potential impacts of bezels on human perception, but also provide a means of quantifying those impacts. Such quantification is useful in developing design guidelines for interfaces that take into account the capabilities of their users.

In this work, we investigated the impact of bezels from the perspective of human perception. Our goal was to build on the existing literature, and to use carefully controlled studies to quantify the effect of interior bezels on human perception. By better understanding this effect, we are able to suggest under which settings the guidelines and techniques developed to compensate for interior bezels most benefit users. This work consists of two studies. First, we studied the impact of bezels on magnitude judgement at a distance. Second, we compared the effect of bezel compensation techniques on magnitude judgement when users were able to interact with a display via touch. In both studies, we introduced bezel width as a controlled variable, allowing us to investigate the impact that bezel size has on both types of tasks. We now describe the two studies and discuss their results individually before presenting overall implications for the design of tiled displays based on their collective results.

## 3 EXPERIMENT 1: INTERACTION AT A DISTANCE

To begin our investigation, we aimed to understand if bezels impact human perception of elementary visual variables on tiled displays from a distance. This approach had the advantage of establishing a baseline for interaction, free from known confounds such as viewing angle [6, 37]. We also believed that interaction at a distance would be the most familiar setting for participants, resembling interactions with displays such as those in meeting spaces or theatres. Our second experiment investigates the complimentary usage case where interaction occurs at arm's length via a touch screen. We wanted to specifically test whether or not the introduction of bezels would increase error in judgements made by participants, if those judgements would take longer to make, and if different types of stimuli would be affected more significantly in the presence of bezels. Our hypotheses were:

H1 As bezel width increases, relative magnitude judgement accuracy decreases

H2 As bezel width increases, relative magnitude judgement time increases

H3 Judgement accuracy differs for different visual variables

H4 Judgement time differs for different visual variables

## 3.1 Experimental Design and Task

Our experiment utilized a 3 (Visual Element Type) × 5 (Bezel Width) within-subjects design that drew on that of Wigdor et al. [37], Bezerianos et al. [6], and Cleveland and McGill [10]. Rather than providing magnitude estimates, as in [37], each participant in our study was tasked with resizing a stimulus to match a corresponding modulus in each trial, thus performing a magnitude *production* task [6]. As noted by Bezerianos et al. [6], our pilot testing revealed that some participants experienced difficulty translating their perceived judgements into numerical estimates; thus, we used a magnitude production task where participants manipulated the on-screen stimulus using a keyboard to reduce the influence of estimation error. Our experiment included two independent variables: visual element type and bezel width.

### 3.1.1 Visual Elements

We chose a subset of three of the visual elements described by Cleveland and McGill [10] consisting of length, angle, and area
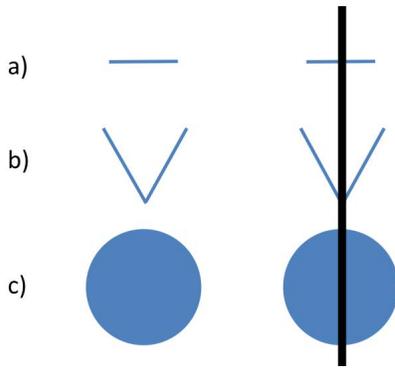
Figure 2: Participants resized three different visual element types (a) length, (b) angle, and (c) area. In conditions where a bezel was present, the stimulus was bisected (right).



Figure 3: Participants sat in a chair in front of the display and interacted via keyboard. In the second experiment, participants stood in front of the display and interacted via touch.

(Figure 2.). These three elements were also used by Bezerianos et al. [6], and have previously been found to represent shapes of varying degrees of difficulty for the judgement task through empirical studies [37, 6, 35]. For each visual variable, we asked participants to resize a stimulus element to match a presented modulus for 6 different modulus magnitudes, that were selected as multiples of 10 ranging in value from 10% to 70% of a base modulus value, depending on the type of visual element: 200 pixels for length judgements, $80°$ for angle judgements, and a radius of 200 pixels for area judgements.

### 3.1.2 Bezel Widths

We investigated 5 bezel widths, reflecting both commercially available hardware, and existing literature that investigated the effect of bezels on human perception: 0, 0.5, 1, 2, and 4cm. We surveyed available LCD and plasma displays suitable for use in video walls and found that they typically were available at different price points ranging from 0.5cm to 2cm, with smaller bezels being sold at a premium relative to larger ones. The 4cm level was included to facilitate comparisons to larger bezels studied in the literature (e.g. [7]), and is representative of panels intended for use with desktop computers, typically available at a much lower price than those intended for use in video wall installations. The width of each condition was measured to accurately reflect on-screen dimensions. The experimental design is summarized as:

> 15 Participants $\times$
>
> 5 Bezel Widths (0cm, 0.5cm, 1cm, 2cm, 4cm ) $\times$
>
> 3 Visual Element Types (length, angle, area) $\times$
>
> 6 Magnitude judgements (10% to 70% of modulus)

For a total of 1350 comparisons.

Participants completed one block of trials for each of the bezel width conditions, the order of which was balanced using a latin-square design. Within each block, the ordering of visual element type and stimuli magnitudes were randomized.

### 3.2 Participants, Procedure, and Apparatus

15 participants (9 men and 6 women) between the ages of 22 and 35 (average age 25) were recruited to participate in this study. Participants were all Science, Technology, Engineering, and Mathematics students who were enrolled at the University of Waterloo, and received a $5 gift card for their participation.

Upon arriving, participants were seated in an adjustable chair at a conference table with a wireless keyboard and a large, projected display approximately 3m in front of them, and instructed to adjust the chair and position themselves so that they were comfortable. The display measured 2m wide by 1.5m tall, and was projected at a resolution of $1024 \times 768$ pixels (Figure 3). Participants were then briefed on the task, and asked to complete an informed consent form and brief demographic questionnaire. To ensure that participants understood the task and the simulated bezel, they were then presented with a series of 20 modulus/stimulus pairs in a practice period. During the practice period, participants were able to pose any clarifying questions, and were asked to confirm that they felt comfortable with the task before proceeding.

Finally, participants were instructed to complete the trials as accurately and as quickly as possible, and completed a block of trials for each bezel condition. As stimulus/modulus pairs were presented during each trial, stimuli were randomly sized within +/-20% of their corresponding modulus. Participants then resized the stimulus using the up and down arrow keys until they felt that it accurately matched the modulus, and pressed return to submit their final judgement. Both the modulus and stimulus remained visible to participants for the duration of the trial. After submitting their magnitude judgement, participants were presented with an opportunity to rest before making their next judgement, and were asked to press the space bar to proceed to the next trial. Experimental sessions lasted approximately 30 minutes in total.

### 3.3 Data Collection and Analysis

Participant interaction data were logged to computer files. In order to evaluate our hypotheses, our primary measures were *judgement error* and *task time*. We measured the accuracy of participant judgements based on how closely their manipulated stimulus matched the presented modulus, calculated as a percentage of the size of the modulus. As in [37, 6], judgement error was defined by:

$$\text{error} = |\text{ judged percent} - \text{true percent }|$$

We also investigated the time required to make each judgement, measured as the time starting from when the modulus and stimuli first appeared, until the participant pressed the key to submit their answer. Thus, this measure included the time required to view and adjust the stimuli. Repeated measures analysis of variance (RM-ANOVA) statistical tests were conducted to examine differences between bezel conditions, with post-hoc pairwise comparisons made using the Bonferroni adjustment. All statistical tests used an alpha-value of 0.05.
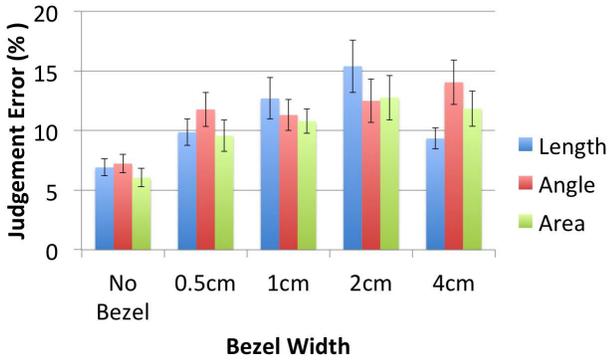
Figure 4: Judgement error results (as a percent of the modulus) for each visual element type (length, angle, and area) and bezel condition (0, 0.5, 1, 2, and 4cm). (Error bars are standard error)
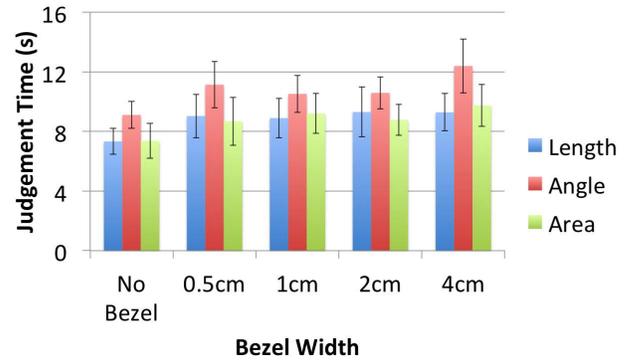


Figure 5: Trial time results (in seconds) for each visual element type (length, angle, and area) and bezel condition (0, 0.5, 1, 2, and 4cm). (Error bars are standard error)

## 3.4 Results

Overall, participants submitted solutions with an average error of 10.8% ($\sigma = 3.41$) and took an average of 9.4 seconds ($\sigma = 4.2$) to make each judgement. Below, we discuss results relevant to our hypotheses.

### 3.4.1 Judgement Error

A main effect of bezel width on accuracy ($F_{4,56} = 10.704, p < 0.001, \eta_p^2 = .433$) revealed judgements made in the No Bezel condition had an average error of 6.7% and were more accurate than any condition where bezels were present: 0.5cm ($\bar{x} = 10.4\%, p = 0.004$), 1cm ($\bar{x} = 11.6\%, p = 0.001$), 2cm ($\bar{x} = 13.6\%, p = 0.005$), 4cm ($\bar{x} = 11.7\%, p = 0.006$). Our analyses did not identify any differences between conditions where interior bezels were present.

We found no effect of visual element type on accuracy ($F_{2,28} = .721, p = 0.495, \eta_p^2 = .049$). Participants' average error for length judgements was 10.8% ($\sigma = 3.9$), for angle judgements was 11.4% ($\sigma = 4.5$), and for area judgements was 10.2% ($\sigma = 3.7$). An interaction effect between visual element type and bezel was found ($F_{8,112} = 2.778, p = .008, \eta_p^2 = .166$), where accuracy error increased with bezel width, except for length judgements made with a 4cm bezel, where they decreased. Post hoc comparisons showed that length judgements made with 4cm bezels were more accurate ($\bar{x} = 9.3\%, \sigma = .884$) than angle ($\bar{x} = 14.1\%, \sigma = 1.85$) or area ($\bar{x} = 11.8\%, \sigma = 1.47$) judgements ($p < .0001$).

### 3.4.2 Trial Time

We found that bezel width had a significant effect on trial time ($F_{4,36} = 10.704, p = 0.026, \eta_p^2 = .433$), where magnitude judgements made in the No Bezel condition ($\bar{x} = 7.9s, \sigma = 1.79$) were faster than those made in the 4cm Bezel condition ($\bar{x} = 10.4s, \sigma = 5.4, p = 0.006$). However, we did not find differences between any of the other conditions.

Finally, our analyses revealed that shape had a significant effect on trial time ($F_{2,18} = 11.2255, p < 0.001, \eta_p^2 = .950$). Length ($\bar{x} = 8.7s, \sigma = 4.52$) and area ($\bar{x} = 8.7s, \sigma = 4.57$) judgements took less time than angle judgements ($\bar{x} = 10.7s, \sigma = 4.59, p = 0.003, p = 0.446$, respectively). No interaction effect between visual element type and bezel width was found for time ($F_{8,112} = .451, p = 0.888, \eta_p^2 = .031$)

## 3.5 Discussion

Our results show that the presence of interior bezels negatively affects magnitude judgements made from a distance on a large dis-

play. With significant differences between the No Bezel condition (0cm) and those where interior bezels were present (0.5, 1, 2, 4cm), we partially confirm **H1**. That is, the presence of any bezel impacted magnitude judgements, but we did not find clear evidence that increased bezel width further impacted accuracy. The average accuracy without any bezels is between 3.7% and 6.9% more accurate, or 5.1% greater than the 11.8% average accuracy across conditions with interior bezels.

Our findings partially support **H2**. Magnitude judgements in the 4cm bezel condition took 2.5s longer than the No Bezel condition, an increase of 25%. However, our analysis did not reveal any time differences between the other bezel widths. When considered together, our findings for accuracy and time suggest a small, but potentially critical effect of bezel width.

Unlike other studies [35, 6], we did not find differences in magnitude judgement accuracy for different visual element types for most of our conditions. However, for judgements made in the 4cm bezel conditions, participants' length judgements were approximately 2.5% more accurate than area judgements, and 5% more accurate than angle judgements. While these findings partially confirm **H3**, our analysis of effect size suggests that the type of visual element accounted for less than 5% ($\eta_p^2 = .049$) of the variance in accuracy. We also found that angle judgements took more time than other visual element types, partially confirming **H4**. We interpret these findings as indicative of a trade-off between judgement time and accuracy, as reported in the literature [35].

## 4 EXPERIMENT 2: INTERACTION AT ARM'S LENGTH

Having investigated magnitude judgements at a distance, we wanted to explore the use of tiled displays in conditions that more closely resembled those we envisioned for typical use cases on a touch-interactive tiled display. In particular, we wanted to capture interaction at arm's length to the large display, and to explore whether a user's proximity to the display may change the role that bezels play in determining the effectiveness at magnitude judgement tasks. We also wanted to investigate whether the use of bezel compensation techniques would mitigate the performance differences observed in Experiment 1. To investigate these questions, we conducted a second study that replicated the conditions of the first. Our hypotheses for this follow-up experiment are:

H5 As bezel width increases, the accuracy of relative magnitude judgements decreases

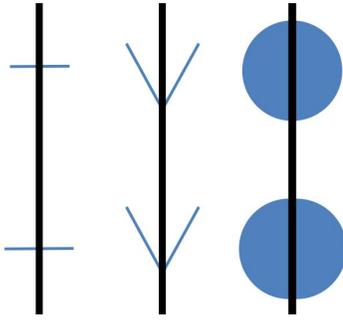H6 As bezel width increases, the time taken to make relative magnitude judgements increases

Figure 6: Participants completed trials under two bezel compensation conditions (top) bezel compensation, in which visual elements' size remained constant regardless of bezel width, and (bottom) no bezel compensation, as in Experiment 1.

H7  Judgement accuracy differs for different visual variables

H8  Judgement time differs for different visual variables

H9  Bezel compensation will improve judgement accuracy

H10  Bezel compensation will reduce judgement time

### 4.1  Experimental Design and Task

Our experiment utilized a 3 (Visual Element Type) $\times$ 5 (Bezel Width) $\times$ 2 (Bezel Compensation) design, where bezel compensation was a between-subjects factor and visual element type and bezel width were within-subjects factors.

#### 4.1.1  Bezel Widths

As in the Experiment 1, we investigated 5 bezel widths: 0, 0.5, 1, 2, and 4cm.

#### 4.1.2  Visual Elements

As in Experiment 1, visual elements consisted of length, angle, and area (Figure 6.). For each visual variable, participants made judgements for 6 different magnitudes, that were selected as multiples of 10 ranging in value from 10% to 70% of a base modulus value, depending on the type of visual element: 200 pixels for length, 80° for angle, and a radius of 200 pixels for area judgements.

#### 4.1.3  Bezel Compensation

Participants completed all of their trials in one of two bezel compensation conditions: either where no bezel compensation was present, as used in Experiment 1, or in a second condition in which bezel compensation was enabled. Our bezel compensation implementation simulated the 'French Window' [13] appearance, where data is hidden behind bezels, and which has been previously shown to improve performance for physical interactions also at arm's length. With compensation enabled, their magnitude in physical space is preserved instead of their magnitude in display space. The experimental design is summarized as:

> 10 Participants $\times$
>
> 2 Bezel Compensation (Present, Absent) $\times$
>
> 5 Bezel Widths (0cm, 0.5cm, 1cm, 2cm, 4cm ) $\times$
>
> 3 Visual Element Types (length, angle, area) $\times$
>
> 6 Magnitude judgements (10% to 70% of modulus)

For a total of 1800 comparisons.

Participants completed one block of trials for each of the bezel width conditions, the order of which was balanced using a latin-square design. Within each block, the ordering of visual element type and stimuli magnitudes were randomized.

### 4.2  Participants, Procedure, and Apparatus

20 participants (8 men and 12 women) between the ages of 22 and 28 ($\bar{x} = 23$) were recruited to participate in this study. Participants were all Science, Technology, Engineering, and Mathematics students who were enrolled at the University of Waterloo, and received $5 for their participation. The experimental procedure for this experiment was identical to that of Experiment 1, except for two factors: bezel compensation and touch interaction.

Unlike Experiment 1, a 'French Window' bezel compensation technique was introduced as a between subjects factor in this experiment. Unlike previous work [13], our 'French Window' implementation did not utilize head tracking technology, and replicated functionality available on commercially available graphics cards that compensates for bezels by translating the projected image by a constant amount. To ensure that participants understood the effect of bezel compensation on the stimulus/modulus pairs, we explained the bezel compensation technique to participants in lay terms before participants completed their practice session (e.g. the bezel 'covers up' part of the shape, or that they would need to account for a 'gap in the shape'). To ensure that participants understood the task and the simulated bezel, they were then presented with a series of 20 modulus/stimulus pairs in a practice period, and given the opportunity to ask questions and confirm their understanding of the display.

For each trial, participants stood in front of a large, projected display at a comfortable distance for touch interaction (determined by the participants' arm length and personal preference). Touch interaction was enabled via a PQ Labs 85" touch frame, supporting up to 12 simultaneous touch points, allowing participants to interact with a projected image at a resolution of 1920 $\times$ 1080 pixels (Figure 3). During each trial, participants resized stimuli by touching and dragging at any point within a 400 pixel radius of its projected image, with drags towards its centre reducing the stimulus' magnitude, and drags away from centre increasing magnitude. Finally, participants were instructed to complete trials as accurately and as quickly as possible, and completed a block for each bezel width condition.

### 4.3  Data Collection and Analysis

As in Experiment 1, our primary measures were *judgement error* and *task time*, calculated from data collected to computer log files. We measured the accuracy of participant judgements based on how closely their manipulated stimulus matched the presented modulus, calculated as a percentage of the size of the modulus. We also investigated the time required to make each judgement, measured as the time starting from when the modulus and stimuli first appeared, until the participant pressed the key to submit their answer. Repeated measures analysis of variance (RM-ANOVA) statistical tests were conducted to examine differences between bezel conditions, with post-hoc pairwise comparisons made using the Bonferroni adjustment. All statistical tests used an alpha-value of 0.05.

### 4.4  Results

Overall, participants submitted magnitude judgements with an average error of 11.54% ($\sigma = 7.45$), and took an average of 10.4 seconds ($\sigma = 7.08$) to make each judgement. Below, we discuss results relevant to our hypotheses.

#### 4.4.1  Judgement Accuracy

A main effect exists for bezel width on accuracy ($F_{4,72} = 2.567, p = 0.045, \eta_p^2 = .125$). Tests revealed that No Bezel ($\bar{x} = 9.981\%, \sigma = 1.2$) and 0.5cm Bezel ($\bar{x} = 9.863\%, \sigma = 0.961$) conditions are more accurate than the 4cm bezel condition ($\bar{x} = 12.246\%, \sigma = 1.3$,
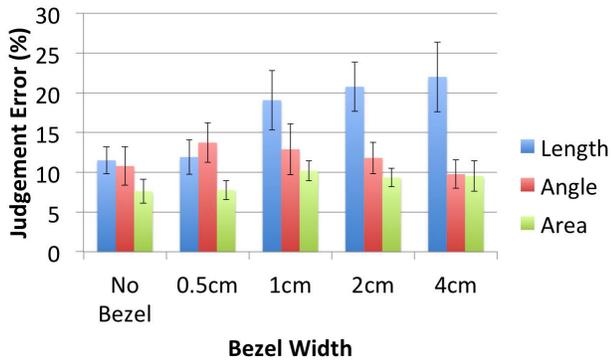
Figure 7: Judgement error results for trials without bezel compensation (as a percent of the modulus) for each visual element type (length, angle, and area) and bezel condition (0, 0.5, 1, 2, and 4cm) in the no bezel compensation conditions.
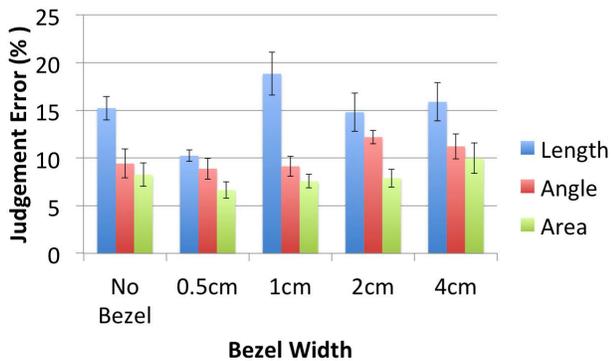


Figure 8: Judgement error results for trials with bezel compensation (as a percent of the modulus) for each visual element type (length, angle, and area) and bezel condition (0, 0.5, 1, 2, and 4cm) in the bezel compensation conditions.

$p = 0.013, 0.28$ respectively). Average judgement errors in the 1cm ($\bar{x} = 12.797\%, p = 1.25$) and 2cm ($\bar{x} = 12.808\%, p = 0.888$) bezel conditions are not significantly different from other bezel widths.

A main effect exists for visual element type on accuracy ($F_{2,36} = 14.433, p < 0.001, \eta_p^2 = 0.445$), and tests revealed that errors in length judgements ($\bar{x} = 15.148\%, \sigma = 1.2$) are less accurate then both Angle ($\bar{x} = 11.0\%, \sigma = 1.1$) and Area ($\bar{x} = 8.50\%, \sigma = 0.74$) judgements ($p = .029, .0001$, respectively). No difference was found between Angle and Area judgements ($p = .088$). An interaction effect between visual element type and bezel width is most relevant ($F_{8,144} = 3.128, p = 0.003, \eta_p^2 = .148$), and our tests reveal that length judgements are approximately 53% less accurate in 1cm, 2cm, and 4cm conditions ($\bar{x} = 17.6$) compared to the No Bezel and 0.5cm conditions ($\bar{x} = 11.5, p < .0001$).

No effect for bezel compensation was found ($F_{1,18} = 0.728, p = 0.405, \eta_p^2 = .039$). Judgements with no compensation ($\bar{x} = 12.2\%, \sigma = 1.01$) appear similar to those in the bezel compensation conditions ($\bar{x} = 10.9\%, \sigma = 1.01$). Figures 7 and 8 illustrate judgement results with and without bezel compensation data.
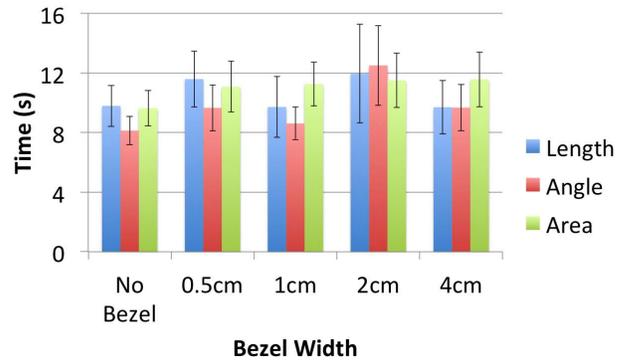
Figure 9: Trial time results in seconds for each visual element type (length, angle, and area) and bezel condition (0, 0.5, 1, 2, and 4cm).

### 4.4.2 Trial Time

There are no significant effects of bezel width on trial time ($F_{4,72} = 1.747, p = 0.149, \eta_p^2 = .088$), visual element type on trial time ($F_{2,36} = 1.961, p < 0.155, \eta_p^2 = .098$), bezel compensation technique on trial time ($F_{1,18} = 0.862, p = 0.365, \eta_p^2 = .046$).

### 4.5 Discussion

Our results show that large interior bezels have a negative effect on magnitude judgement accuracy when made at arm's length. These results are more nuanced than the clear differences between bezel and no bezel for judgements made from a distance. At arm's length, judgements with no bezel and the very small 0.5cm bezel are 2.8% more accurate than judgements made with the large 4cm bezel. Thus, we have some evidence to support **H5**, but unlike Experiment 1, we must reject **H6** given no significant differences for judgement time. An analysis of effect size suggests that bezel width plays a minor role in the accuracy of magnitude judgements ($\eta_p^2 = 0.125$) – visual element type accounted for close to half of the observed variance in our model ($\eta_p^2 = .445$).

Unlike judgements made from a distance, when at arm's length visual element types are a significant factor in accuracy, confirming **H7**. Length judgements in particular are less accurate than area and angle judgements for larger bezels (1cm, 2cm, 4cm). However, there are no differences in judgement time across visual element type, and we reject **H8**. These findings contrast with Experiment 1, where at a distance judgement accuracy is not affected by visual element type but angle judgements take more time. This difference may be attributed to differences in keyboard and touch input: fine-tuning a choice using the keyboard's precise up and down arrows in Experiment 1 could be easier than using the touch screen in this experiment.

Finally, we found no evidence that bezel compensation techniques impacted magnitude judgement accuracy or time. These results can be explained by the small effect size observed during our study, accounting for less than 3.9% and 4.6% of variance in our model. Thus, we reject hypotheses **H9** and **H10**.

## 5 IMPLICATIONS FOR DESIGN

Our results translate directly into practical advice for building very large display systems. If size and resolution can be achieved with a non-tiled single display, is the significant increase in price justified for the intended application? For example, at the time of writing, Panasonic manufactures a 3.9m (152 inch), 8 megapixel display – but it retails for more than USD$500,000, compared to approximately USD$50,000 for a 12 megapixel tiled display of comparable

size. When seeking to build displays that offer even larger physical size or resolution, tiling individual displays offers even more substantial savings, and more flexibility in display size and aspect ratio. Assuming a tiled display is the only economically viable way forward, should premium thin-bezel displays be used? Are the extra expense, space, and calibration difficulties that arise when using an array of LCD projectors justified to reduce inner bezels to more subtle display seams?

Based on our results regarding the fundamental perceptual operation of magnitude judgement, we offer three design implications that inform these decisions: 1) designers should expect a 5% increase in judgement error when interior bezels are present; 2) length judgements made at arm's length are particularly prone to reduced accuracy with wider bezels; and 3) bezel compensation techniques provide a limited benefit.

### 5.1 When are Bezels a Design Consideration?

When viewing data from a distance, the presence of any interior bezel increases magnitude judgement error by 5.1%. When at arm's length, judgement error increases by 2.8% for very wide 4cm inner bezels compared to no bezel or thin 0.5cm bezels. Since a large display is often used both from a distance and at arm's length, the most prudent guideline is to adopt a simple rule-of-thumb to avoid bezels if a 5% increase in magnitude error is a concern. For example, maps used in military situation rooms and emergency response centres require accurate (and fast) magnitude judgements. Similarly, accurate critique of renderings for architecture and auto design may be compromised by magnitude judgement error caused by bezels. If 5% error could be detrimental, then a higher budget should be considered for displays with thin bezels or no bezel at all.

For many tasks, users, and environments, a 5% increase in accuracy error may be less critical. For example, a $1°C$ difference in temperature is likely less sensitive to misinterpretation in the home than in a nuclear reactor. This is especially true if judgements will be made at arm's length where the effect on error is less pronounced, or data can be positioned away from interior bezels [1]. Assuming that the effect of interior bezels on mouse interaction [7], visual search [30], and target acquisition [27] is also small for the usage context, lower cost displays with wider bezels may be used. In these contexts, designers may benefit from focusing on other factors, such as viewing angle [6], that have been identified as having a larger effect on magnitude judgement.

### 5.2 Length Judgements made at Arm's Length

Our analyses revealed that length judgements made at arm's length were particularly prone to increased error; in these settings, accuracy error increased to 18.9% in the 4cm bezel condition, marking a 5% increase over the 13.5% error we observed in the no bezel condition. This increased error may arise due to differences in viewing angle when working close to the display [37, 6]. We anticipate that these issues will most significantly impact the interpretation of geographic maps, scientific graphs, and charts. Designers may want to provide simple tools to assist people when interpreting data close to the display. For example, allow data to be easily translated so that it does not span a bezel, or augment the visualization with numerical values to eliminate length judgements.

Alternatively, it may be beneficial for people to select data from the tiled display and then view that data on a personal device, eliminating magnitude judgements spanning inner bezels. Perhaps counter intuitively, our results suggest that interaction on personal devices may provide a more significant benefit when at arm's length from a large displays than at a distance – a consideration that is not often discussed in the literature. For example, many projects have investigated the use of personal devices to enable interaction at a distance (e.g. [5, 20]), however there may be advantages to exploring methods of displaying content [36] or relocating applications [8] to personal devices when at arm's length.

### 5.3 Utility of Bezel Compensation Techniques

We did not find any benefit to using a 'French Window' [13] bezel compensation technique at arm's length. We observed only a 1.3% improvement in participant accuracy when bezel compensation was present; a difference that was not found to be statistically significant, nor is likely to have significance in practice. Further, our analysis revealed that bezel compensation technique accounted for less than 4% of the variance in our model, typically interpreted as a minimal effect [12]. Thus, we suggest implementing bezel compensation only when it is not a significant investment of resources. For example, many commercial graphics cards provide compensation as a built-in feature and can support small and medium sized tiled displays (2 – 24 displays), however, for deployments of more than 24 displays, there is currently no hardware support for bezel compensation. In this case, our results suggest that it is advisable to avoid implementing bezel compensation given its' lack of impact on magnitude judgement. However, there may be settings in which aesthetic design necessarily outweighs perceptual performance considerations, such as for artistic installations or advertising, where the representational quality of imagery is more important than its accurate lower level perception.

## 6 Limitations and Future Work

By designing two experiments that carefully controlled the circumstances under which magnitude judgements were made, we reduced previously identified confounds such as viewing angle [37, 6], and focused on the effect of interior bezels on magnitude judgement. This work is only a first step in developing guidelines for tiled display design across a variety of tasks, settings, and users, and care needs to be taken in generalizing these results to a broader usage context. For example, our results differ from those of Grüninger and Krüger [18], who suggest a threshold of 1.2cm for bezels to minimally impede depth perception. Similarly, care must be taken in generalizing these results across users of different ages, heights, or with varying abilities such as visual acuity or spatial perception. An open question is how to synthesize results from these independent studies into guidelines that can inform designers as to what degree interior bezels impact higher order tasks such as reading, writing, sketching, or navigation. Given the variety of tasks, environments, and users potentially addressed by tiled displays, no single study can adequately address all questions. We reassert Bi et al.'s [7] suggestion that further studies examining the effects of interior bezels are warranted, particularly where control is leveraged to understand how physical and visual variables impact human perception.

## 7 Conclusion

This research addresses the need [7] for more focused studies that explore the impact of interior bezels on user performance with large, tiled displays. We conducted two empirical experiments that examined the impact of the presence and width of interior bezels on magnitude judgement at a distance and at arm's length, as well as the utility of bezel compensation techniques. Our results show that the presence of interior bezels impacts user performance for magnitude judgement tasks from a distance, but when at arm's length the effect is less pronounced except for length judgements. Based on our findings, we provide three practical implications for the design of large, tiled displays: (1) a 5% decreased accuracy guideline when considering budgets which minimize interior bezels in large displays; (2) length judgements made at arm's length are particularly susceptible to wider bezels; and (3) bezel compensation techniques provide only a minimal benefit for magnitude judgement tasks.

This work contributes towards an understanding of interaction and visualization on tiled displays that is grounded in the theo-

ries of psychophysics, cognition, and human-computer interaction. Our results demonstrate the importance of understanding the effect of interior bezels on human perception, and that practical design guidelines can be formed from such an understanding. As display technologies continue to become more affordable, and tiled displays become available to a variety of usage contexts, such guidelines will serve to guide designers towards appropriate hardware and software support for their target users, settings, and tasks.

## REFERENCES

[1] Andrews, C., Endert, A., Yost, B., and North, C. Information visualization on large, high-resolution displays: Issues, challenges, and opportunities. *Information Visualization 10*, 4 (2011), 341–355.

[2] Anslow, C., Marshall, S., Noble, J., Tempero, E., and Biddle, R. User evaluation of polymetric views using a large visualization wall. In *Proc. SOFTVIS 2010*, ACM (2010), 25–34.

[3] Ball, R., and North, C. Effects of tiled high-resolution display on basic visualization and navigation tasks. In *Proc. CHI 2005*, CHI EA '05, ACM (2005), 1196–1199.

[4] Ball, R., Varghese, M., Sabri, A., Cox, E., Fierer, C., Peterson, M., Carstensen, B., and North, C. Evaluating the benefits of tiled displays for navigating maps. In *Proc. IASTED-HCI 2005* (2005), 66–71.

[5] Baudisch, P., Cutrell, E., Hinckley, K., and Gruen, R. Mouse ether: accelerating the acquisition of targets across multi-monitor displays. In *CHI EA 2004*, ACM (2004), 1379–1382.

[6] Bezerianos, A., Isenberg, P., et al. Perception of visual variables on tiled wall-sized displays for information visualization applications. *IEEE Transactions on Visualization and Computer Graphics 18*, 12 (2012).

[7] Bi, X., Bae, S.-H., and Balakrishnan, R. Effects of interior bezels of tiled-monitor large displays on visual search, tunnel steering, and target selection. In *Proc. CHI 2010*, CHI '10, ACM (2010), 65–74.

[8] Biehl, J. T., and Bailey, B. P. Aris: an interface for application relocation in an interactive space. In *Proc. GI 2004*, Canadian Human-Computer Communications Society (2004), 107–116.

[9] Campbell, C. S., and Maglio, P. P. Segmentation of display space interferes with multitasking. In *Proc INTERACT 2003* (2003), 575–582.

[10] Cleveland, W. S., and McGill, R. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association 79*, 387 (1984), pp. 531–554.

[11] Cleveland, W. S., and McGill, R. Graphical perception and graphical methods for analyzing scientific data. *Science 229*, 4716 (1985), pp. 828–833.

[12] Cohen, J. A power primer. *Psychological bulletin 112*, 1 (1992), 155.

[13] de Almeida, R. A., Pillias, C., Pietriga, E., and Cubaud, P. Looking behind bezels: french windows for wall displays. In *Proc AVI 2012*, ACM (2012), 124–131.

[14] Forlines, C., Shen, C., Wigdor, D., and Balakrishnan, R. Exploring the effects of group size and display configuration on visual search. In *Proc CSCW 2006*, ACM (2006), 11–20.

[15] Gescheider, G. *Psychophysics: the fundamentals*. Lawrence Erlbaum Associates, Incorporated, 1997.

[16] Gescheider, G. A. Psychophysical scaling. *Annual review of psychology 39*, 1 (1988), 169–200.

[17] Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., and Wang, M. Proxemic interactions: the new ubicomp? *interactions 18*, 1 (Jan. 2011), 42–50.

[18] Grüninger, J., and Krüger, J. The impact of display bezels on stereoscopic vision for tiled displays. In *Proc. VRST '13*, ACM (2013), 241–250.

[19] Hall, E. *The Hidden Dimension*. Anchor Books, 1966.

[20] Haller, M., Leitner, J., Seifried, T., Wallace, J. R., Scott, S. D., Richter, C., Brandl, P., Gokcezade, A., and Hunter, S. The nice discussion room: Integrating paper and digital media to support co-located group meetings. In *Proc. CHI 2010*, ACM (2010), 609–618.

[21] Hawkey, K., Kellar, M., Reilly, D., Whalen, T., and Inkpen, K. M. The proximity factor: impact of distance on co-located collaboration. In *Proc. GROUP 2005*, ACM (2005), 31–40.

[22] Hutchings, D. An investigation of fitts' law in a multiple-display environment. In *Proc. CHI 2012*, ACM (2012), 3181–3184.

[23] Isenberg, P., Isenberg, T., Hesselmann, T., Lee, B., von Zadow, U., and Tang, A. Data visualization on interactive surfaces: A research agenda. IEEE Computer Society (2013).

[24] Jiang, X., Hong, J. I., Takayama, L. A., and Landay, J. A. Ubiquitous computing for firefighters: field studies and prototypes of large displays for incident command. In *Proc. CHI 2004*, ACM (2004), 679–686.

[25] Lee, S., Kim, H., Lee, Y.-k., Sim, M., and Lee, K.-p. Designing of an effective monitor partitioning system with adjustable virtual bezel. In *Human Centered Design*, M. Kurosu, Ed., vol. 6776. Springer Berlin Heidelberg, 2011, 537–546.

[26] Mayer, T. The 4k format implications for visualization, vr, command & control and special venue application. In *Proc. EDT 2007*, ACM (2007).

[27] McNamara, A. M., Parke, F., and Sanford, M. Evaluating performance in tiled displays: navigation and wayfinding. In *Proc. VRCAI 2011*, ACM (2011), 483–490.

[28] Moreland, K. Redirecting research in large-format displays for visualization. In *Proc. LDAV 2012* (oct. 2012), 91 –95.

[29] Renambot, L., Jeong, B., Jagodic, R., Johnson, A., Leigh, J., and Aguilera, J. Collaborative visualization using high-resolution tiled displays. In *Proc. CHI Workshop on Information Visualization Interaction Techniques for Collaboration Across Multiple Displays*, ACM (2006).

[30] Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., and Tan, D. The large-display user experience. *Computer Graphics and Applications, IEEE 25*, 4 (2005), 44–51.

[31] Shupp, L., Andrews, C., Dickey-Kurdziolek, M., Yost, B., and North, C. Shaping the display of the future: The effects of display size and curvature on user performance and insights. *Human–Computer Interaction 24*, 1-2 (2009), 230–272.

[32] Stevens, S. S. On the psychophysical law. *Psychological review 64*, 3 (1957), 153.

[33] Su, R. E., and Bailey, B. P. Put them where? towards guidelines for positioning large displays in interactive workspaces. In *Proc. INTERACT 2005*. Springer, 2005, 337–349.

[34] Tan, D. S., and Czerwinski, M. Effects of visual separation and physical discontinuities when distributing information across multiple displays. In *Proc. INTERACT 2003* (2003).

[35] Wagner, M. *The geometries of visual space*. Lawrence Erlbaum, 2006.

[36] Wallace, J. R., Mandryk, R. L., and Inkpen, K. M. Comparing content and input redirection in mdes. In *Proc. CSCW 2008*, ACM (2008), 157–166.

[37] Wigdor, D., Shen, C., Forlines, C., and Balakrishnan, R. Perception of elementary graphical elements in tabletop and multi-surface environments. In *Proc. CHI 2007*, ACM (2007), 473–482.

[38] Yamaoka, S., Doerr, K.-U., and Kuester, F. Visualization of high-resolution image collections on large tiled display walls. *Future Gener. Comput. Syst. 27*, 5 (May 2011), 498–505.

[39] Yang, X.-D., Mak, E., McCallum, D., Irani, P., Cao, X., and Izadi, S. Lensmouse: augmenting the mouse with an interactive touch display. In *Proc. CHI 2010*, ACM (2010), 2431–2440.

[40] Yost, B., Haciahmetoglu, Y., and North, C. Beyond visual acuity: the perceptual scalability of information visualizations for large displays. In *Proc. CHI 2007*, ACM (2007), 101–110.