Looking From a Different Angle: Placing Head-Worn Displays Near the Nose

Yukun Song*
Georgia Institute of Technology
Atlanta, Georgia, United States
sparklight@gatech.edu

Parth Arora*
Georgia Institute of Technology
Atlanta, Georgia, United States
partharora@gatech.edu

Srikanth T. Varadharajan
Georgia Institute of Technology
Atlanta, Georgia, United States
svaradharajan6@gatech.edu

Rajandeep Singh
Georgia Institute of Technology
Atlanta, Georgia, United States
rajandeep.singh@gatech.edu

Malcolm Haynes
United States Military Academy
West Point, New York, United States
malcolm.haynes@westpoint.edu

Thad Starner
Georgia Institute of Technology
Atlanta, Georgia, United States
thad@gatech.edu

ABSTRACT

Head-worn augmented reality displays such as Engo Eyewear avoid placing the virtual image in the user's primary position of gaze (PPOG) to allow a clear view of the user's primary task. Several studies suggest that horizontally offsetting the virtual image toward the ear provides good performance and comfort during different types of tasks. Less research focuses on offsetting the image toward the nose. Extending a previous study with displays positioned at 0°, +10°, +20°, and +30° (defining toward the ear as the positive direction), we run two studies each with four conditions and 12 participants (24 participants total) comparing user comfort at -30°, -20°, -10°, and 0° and -15°, 0°, +15°, and +25°. We follow the previous study's procedures, using a 30-minute reading task and a video display terminal as an emulated right-eye monocular display with a smartphone-sized field of view (FOV). Comparing the results from all three studies suggests that reading on displays with pixels between -24.6° and +19.6° may be comfortable, with users tolerating negative offsets better than positive.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; *User studies*; *Mixed / augmented reality*.

KEYWORDS

Wearable Computers, Virtual/Augmented Reality, Empirical study that tells us about people, Usability Study

ACM Reference Format:
Yukun Song, Parth Arora, Srikanth T. Varadharajan, Rajandeep Singh, Malcolm Haynes, and Thad Starner. 2024. Looking From a Different Angle: Placing Head-Worn Displays Near the Nose. In *The Augmented Humans International Conference (AHs 2024), April 04–06, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3652920.3652946

*Both authors contributed equally to this research.

AHs 2024, April 04–06, 2024, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-0980-7/24/04
https://doi.org/10.1145/3652920.3652946

1 INTRODUCTION

In the $5.5 billion estimated head-worn display (HWD) market [31], manufacturers have made products with various combinations of different technologies and design components, seeking a way to reach better user experiences. Balancing HWD parameters under the constraint of manufacturing resources is not an easy task, as it involves treatment of focal planes, controlling the luminance and chromaticity, sizing the perceptual eyebox, as well as considerations in dozens of other aspects [37]. One of the most impactful factors is the position of the display.

Past HWD products exhibit different answers from the manufacturers. While the Google Glass [17] family of displays as well as the Nreal/Xreal [46] product series places the image centered above the user's primary position of gaze (PPOG), the Epson [11] series of smart glasses places the image in front of users' PPOG. Tooz [42] provides slices of optical combiners in the center of users' view with a slight horizontal offset towards the ear, while the Engo 1 [16] positions the image above the PPOG and horizontally offsets it toward the nose.

Extensive research has been done to explore human preference for positions of display under different types of tasks, and for many tasks placing the display in PPOG is unfavorable. Chua et al. and Lin et al. showed that participants find virtual images placed in the PPOG obtrusive [6, 29]. Moreover, Foyle et al. found that placing the head-up display (HUD) on PPOG can lead to cognitive capture as participants struggled to maintain altitude in a flight simulator when the HUD was in the PPOG compared to when it was offset diagonally at 8.14° and 16.28° above and towards the left of PPOG [14]. Hershberger and Guerin compared the PPOG position to positions below the PPOG and found that the latter significantly reduced binocular rivalry [22]. Furthermore, Dowell et al. showed that cognitive capture can be avoided by displacing the HUD 8° horizontally from PPOG [9].

However, it is more complex to derive an agreement on where to offset the display and how much the offset should be. The ideal placement for the imaging display may be highly dependent on the task, which can be classified into four categories: single tasks, alternating tasks (i.e., tasks that require users to switch attention between the HWD and the physical world), background tasks, and dual tasks [29]. Similar to Chua et al., other studies have also looked at and compared different positions of head-up displays (HUDs) while driving (a dual task) [4, 15, 23, 30, 32, 45]. Even earlier studies focused on lateral and horizontal displacements for HWDs and their effects on pilot performances while flying [14, 22, 43]. In a study on a single task of 30-minute reading from an emulated HWD, Haynes et al. concluded that displays centered at 0°, 10°, and 20° from the PPOG towards the ear are significantly more comfortable than those at 30° [21]. Lin et al. further tested these offsets for alternating tasks and conclude that 0°, 10°, and 20° are visually more comfortable than 30° towards the ear [29]. Our study, focusing on the same single reading task in Haynes et al.'s study, expands the investigation over a wider range of conditions.

Despite the diversity of current research, the results often cannot provide a guideline to manufacturers for positioning the image, as the scope of angle displacements is often limited due to the focus of the experiment, or biased due to the choice of HWD device used in the experiments. For instance, Chua et al. studied lateral displacements both towards the ear and the nose while performing a dual task (driving) using Google Glass [6]. They concluded that a 17.28° horizontal displacement towards the ear was significantly more comfortable than the same displacement towards the nose. This result may be specific to the dual task of driving and could also be highly influenced by the design of Google Glass used in the study, which occludes the vision to the right of the display. In addition, due to the hardware design of Google Glass, participants' comfort may also be impacted by moving the Google Glass toward the nose direction, which further makes this conclusion less generalizable to general HWD experiences.

Among the experiments that examine enough conditions in a wide range of display displacements in personal HWDs, most focus on dual or alternating tasks [25, 39, 47, 48], which often have specific definitions varying in different experiments, making it difficult for manufacturers to generalize and utilize the results. One exemplary study conducted by Katsuama et al. analyzed both vertical (+15°, 0°, -15°, upward defined as positive) and horizontal (0°, +20°, +35°, +45°, right defined as positive) offsets of the display using two cathode ray tubes (CRT) under a dual Unstable Event-Monitoring (UEM) task [25]. While they found that the error rate of task performance increased along with the increment of display offset, the task itself restricted drawing general conclusions for tasks performed on HWDs.

Few current studies focus on human preference for displays that offset towards the nose. Therefore, to complement the existing results investigating the positioning of HWDs, we perform studies that give insights into a wide range of horizontal angle offsets that include both directions starting from users' PPOG. We hope to provide manufacturers with information on where to put the display on the horizontal axis so as to be comfortable for the task of reading text, which we identify as one of the most visually intensive and common tasks for computer displays.

1.1 Contributions

Specifically, the contributions of this paper are

- Based on Haynes et al.'s experiment process [21] on people's eye comfort in single-task HWD reading experiences from different horizontal angle offsets, we performed two new studies (each with 12 participants), expanding the studied angle range from [0°, +30°] to [-30°, +30°], with 10 different angle offsets in total (we herein define the positive direction as horizontally offsetting towards the ear from the eye's PPOG, and the negative direction as offsetting towards the nose).
- Using the results from our two studies, we provide two additional direct comparisons of reading comfort among groups of horizontal angle offsets (-30°, -20°, -10°, 0°) and (-15°, 0°, +15°, +25°). Normalizing our results with those from Haynes et al.'s previous study [21], we provide the first 5°-granular, bidirectional, and wide-range [-30°, +30°] comparison of people's eye comfort during HWD reading experiences.
- With the three studies and prior literature, we provide the guideline of limiting display pixels between the angle range of [-24.6°, +19.6°]. In addition, we illustrate our observation of the implications and potential of centering displays at a negative angle offset.

2 CONSIDERATIONS

Multiple factors affect the positioning of HWDs, including 1) monocularity of the display, 2) number of display segments in each eye, 3) field of view (FOV) of each display, 4) shape of the display, 5) primary task of the display, and 6) the display's angular offset from users' PPOG.

While binocular displays help avoid problems with eye dominance and eyebox and can improve perceived brightness, most displays on the market intended for extended everyday use are monocular (due to price, head weight, nose weight, and power consumption). Therefore, we choose monocular displays for our study. We assume the monocular display will be placed in front of the right eye as most users are right-eyed and almost all in-glasses displays on the market are right-eyed.

For factors 2) and 4), most HWDs in the market only display a single, rectangular image in each eye. Even the Tooz, which consists of several segments of optical combiners, appears to the user as one rectangular image [42]. As a result, we choose to investigate the use case of a single rectangular display. For factor 3), even though HWD manufacturers often target making displays with larger FOV in both vertical and horizontal directions, we should still choose a relatively small but acceptable FOV for the display so that the experiment can be sensitive while still being practical. Currently, smartphones are the dominant platform for consuming content while having a relatively small FOV (approximately 18° diagonal). Clearly, smartphone FOVs are sufficiently large to be usable yet are similar to monocular FOV HWDs on the market (the Vuxiz Blade is 19° diagonal; Google Glass XE is 14.6° diagonal)

For factor 5), we need to select a common, single task that is preferably a sub-task for most daily tasks on HWDs, as some dual/alternating task experiments may not translate to other use cases than the one tested, and emulating dual/alternating tasks would also involve more complex visual phenomenons such as rivalry and interference [27].

With the requirements of controlling factors described above, we are inspired by the study conducted by Haynes et al. [21], where the participants perform a reading task (*a pervasive task*) with only their right eyes (*right-eye monocular*) viewing a (*single, rectangular*) video display terminal (VDT) having a 9.2° horizontal by 16.3° vertical

FOV (*smartphone-sized FOV*). Notice that in this study, a VDT is used instead of existing HWDs, which eliminates the concern of biased results due to novelty effects. Using a familiar flat-screen VDT also eliminates complications due to optics or the discomfort of the head and nose-weight of wearing a HWD for an extended time. However, the studies conducted by Haynes et al. [21] only cover a few cases (0°, +10°, +20°, +30°) in angular offsets towards the ear. Therefore, we decide to replicate Haynes et al.'s study and cover more cases in both the negative and positive directions. In this way, we control factors 1) to 5) and focus our study on 6) the display's angular offset from users' PPOG. For comparison, Haynes's data from his dissertation are re-graphed for this paper with permission.

2.1 Implications and Potential of Negative Angular Offset

Covering more cases in both negative and positive angular offsets may have more theoretical value than simply comparing users' preferences on a specific direction of angular-position offsets. In addition to traditional functionalities achieved by displays that have a positive angular offset, displays that have a negative angular offset can potentially support more system features with unique layouts, especially when binocular displays are manufactured with displays on both eyes positioned near the nose. As shown in Figure 1, which represents a model of symmetric binocular displays near the nose, the displays could form an area (Area 1) that supports stereoscopic vision by combining images from both eyes, similar to the video see-through solutions that form stereoscopic displays [24]. In addition, due to the special structure of the display and the optical see-through (OST) property, the model can also support two separate areas (Area 2 and 3) for displaying regular content by single (2D) displays. This technique forms an OST system structure that allows the user to view 3D virtual content in the center while viewing 2D virtual content on the sides of their view, in addition to the information from the real world.

Notice that due to the minimal near point of human eyes, which averages 8.7cm [10] for the global average median age of 30 [38], the angle θ between the near-nose end of the displays and the horizontal line of sight has a lower bound. Taking 63mm as the average inter-pupillary distance (IPD) of adults [12], we can derive that $\theta \geq \arctan(\frac{87}{31.5}) \approx 70.1°$. Therefore, if θ is smaller than 70.1°, parts of Area 1 become practically ineffective. Moreover, if $\theta+\phi$ is smaller than 70.1°, then the entire Area 1 becomes practically ineffective and the system degenerates to two independent biocular displays. According to Song et al.'s research [41], which examines the disruption in users' vision caused by the transparency difference between the optical combiner for a see-through head-worn display and the surrounding glass, the right edge of a right-eye monocular HWD is recommended to have a horizontal angle offset farther than -15°, confining $\theta + \phi$ below 75°. However, this limitation can be mitigated with improved technology which makes the optical combiner effectively disappear for the user, as shown in the latest released model of Vuzix Ultralite [44], which has an indistinguishable border between the optical combiner and the surrounding glass. Therefore, the results of this study can help determine the lower bound for θ and hence the angular scope of the display, which affects the effectiveness of the potential stereoscopic optical see-through system as discussed above.

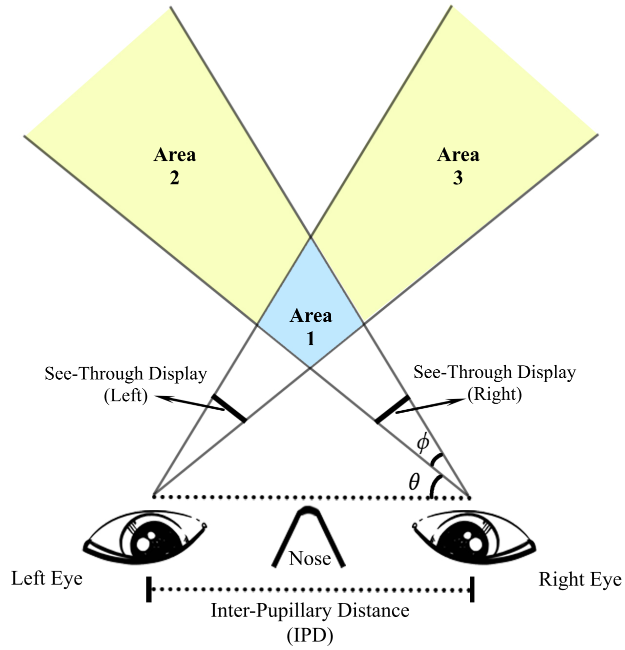
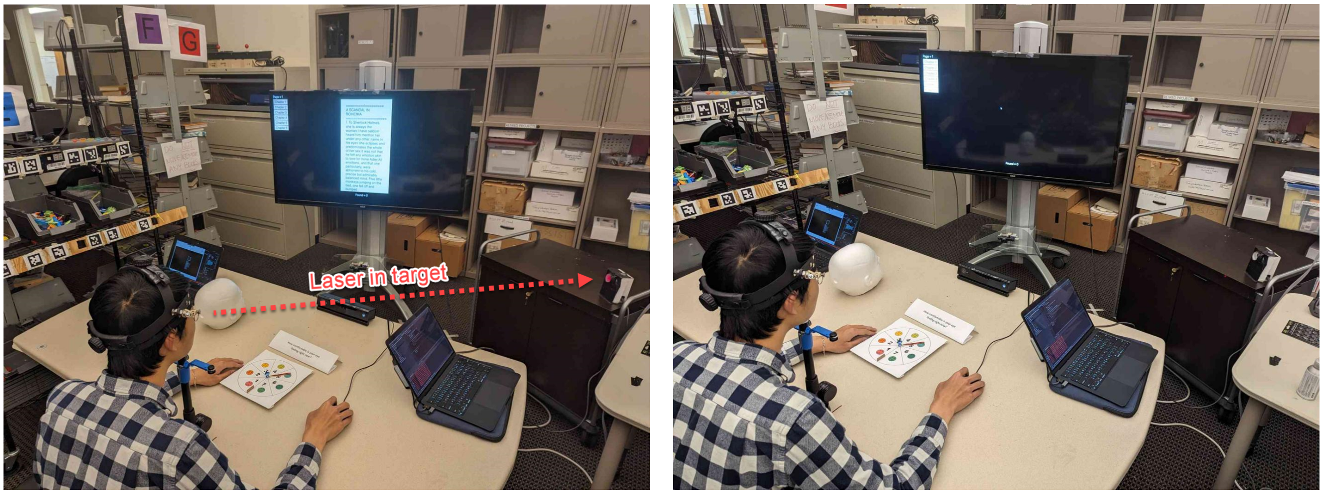


Figure 1: Conceptual model of a symmetric binocular display with both displays positioned near the nose.

3 EXPERIMENT

Per the considerations described, we follow the experimental setup and procedures published by Haynes et al. [20, 21] in our studies. At the beginning of the experiments, the participants were tested for visual acuity and IPDs. They were then seated in a chair with a table in front and a chin rest anchored to keep their head movement to a minimum. After that, the setup is calibrated to that user and the 30-minute reading procedure begins. The participants were asked to read the novel, *The Adventures of Sherlock Holmes* by Sir Arthur Conan Doyle, presented on a video display terminal (VDT) centered with different angle offsets. The novel presented has an FOV of 9.2° horizontal by 16.3° vertical (the same as typical for a hand-held iPhone 6 [2]) from the participants' perspective. The participant's left eye is covered using an eye patch throughout the reading experience. During the study, participants' head movements are strictly controlled using the apparatus described in the following sections. In Haynes' study, angle offsets of 0°, +10°, +20°, and +30° were covered. In our first study, we test the offsets -30°, -20°, -10° and 0°. In our second study, offsets -15°, 0°, +15°, and +25° are tested. To ensure and measure participants' focus on the text content and their reading accuracy, random nursery rhymes are placed in between the stories sentences, and the participants need to record the number of nursery rhymes they find with mouse clicks. We measure the comfort levels of participants' eyes by asking how comfortable their eyes feel, explicitly instructing them to ignore potential discomfort caused by supplemental experiment instruments (i.e. the chin rest), and letting them express their comfort using a rotary knob (Figure 3) corresponding to a 7-point Likert scale to keep verbal interference to a minimum. We ask for their comfort before the



(a) Text is displayed when the laser is pointed to the photo-resistor apparatus **(b) Text is hidden when the laser is not pointed to the photo-resistor apparatus**

Figure 2: The setting environment of the experiment, including all relevant apparatus.

Figure 3: Seven-position rotary knob to elicit comfort ratings

30-minute reading period to get an individual baseline of comfort ratings. Every five minutes during the 30-minute period, a timer will ring and participants give their comfort ratings again using the knob. Participants are prevented from making changes to the knob except when we ask them or the timer rings. In total, seven comfort readings were taken including the one at the beginning. Once the participants finish a reading session, they are presented with a NASA Task Load Index (NASA-TLX) survey, a comprehension quiz with 10 multiple-choice questions, and an asthenopia survey with one preference question and nine representative questions asked in positive and negative wordings [19–21, 33, 34] to reflect on their comfort during the study. Participants are repeatedly tested for different conditions, with each reading session being at least six hours apart. Figure 2 shows the exact experiment setting.

3.1 Emulating an HWD Experience

One essential factor to emulate an experience similar to reading in an HWD is simulating the exact FOV, focus distance, angle offset, and size of the text presented to the participants. To make our results comparable to Haynes et al.'s study [21], we replicated their experiment settings. The VDT we use is an RCA 55" LED HDTV, and to emulate the exact FOV of 9.2° by 16.3° as used in Haynes et al.'s study [21], we customize the computer program to only display texts within the area defined by the FOV. We emulate the focus distance of HWDs as we put the VDT 2 meters away from the participant, which is a common focus distance for HWDs [21]. We also format the page and paginate the text similar to the reading experience in popular reading applications like Kindle in order to maximize familiarity of the reading experience, utilizing Helvetica as the font with a 20 pt font size [21], with 462 characters on a single page on average.

To position the VDT at the exact position with correct angle offsets, a set of laser pointers attached to the apparatus is used together with markers fixed on the floor. The markers on the ground denote angles of interest on the circumference of a circle with a radius of 2 meters and the vertical projection of the participants's right eye as the center. The center of the circle is 3.15 cm to the right of the vertical projection of the chinrest, compensating the IPD for participants assuming an average 63mm adult IPD [12]. One laser is used to determine that the normal vector of the screen is always pointing perpendicular to the vertical center of the circle. A second laser is used to determine that the bottom center of the VDT is vertically aligned with the markers denoting the angle of interest for the specific study session. In addition, the height of the VDT screen is fixed to be 85.5 cm from the ground, and participants adjust the chair and chinrest before the beginning of each session to compensate for the height difference among participants.

3.1.1 Controlling Head Movement. Another crucial factor to emulate HWD experience is to accurately control the participant's head movements while they are reading. The relative position of HWD display to the user's head position should always be fixed, as they are in the real situation. Therefore, we need to keep the participant's head straightforward and in a fixed location throughout the session as they are reading the text from the VDT. We opted to implement an apparatus similar to the one used by Haynes et al. [21].

To keep the participant's head static in position, we utilized a chinrest with adjustable height. The participants adjust the height to their comfort in order to compensate for their height difference against the VDT before each 30-minute reading session. In order to keep their head static in rotation, we use a headband equipped with a laser pointer that points toward a 1-inch circle positioned 1.75 meters away. We build an apparatus with a photo-resistor within the circle, which senses whether the laser is pointing at the circle and hides the text on the display (Figure 2) if the laser strays away from it. This procedure restricts the participant's head movement within a range of approximately ±0.416°. In addition, to avoid participants giving lower comfort ratings due to the discomfort introduced by these apparatus, we explicitly instruct the participants prior to every session to focus on their eye comfort when they give ratings using the rotary knob.

In addition, to make sure the participants' head movement is straight (0° yaw and 0° roll, facing 0° in our angle offset coordinates) at the beginning of each session, we perform a calibration before we turn on the laser and adjust the photo-resistor apparatus. The calibration is done via readings (with less than ±0.5° error) from the Microsoft Kinect V2 depth camera (giving less than 2mm error within 3m [26]), which is anchored to the center of the table and aligned to the 0° angle offset. The participant is instructed to move their head by mimicking our movements on a head model. After the participant moves their head to reach 0° yaw and roll, we turn on the laser and adjust the photo-resistor apparatus, finishing the calibration process. During the 30-minute reading session, the readings from Kinect V2 are not tracked, as we use the laser with photo-resistor apparatus instead. After each reading session, we check the reading of Kinect V2 again with the participant pointing the laser at the photo-resistor apparatus to make sure that the laser did not shift during the process. On average, the laser shifts within ±0.5° during the experiment.

3.1.2 Ensuring Reading Accuracy. It is important to encourage participants to actively participate in the reading as they would during regular reading experiences. The participants are tasked with reading "The Adventures of Sherlock Holmes" by Sir Arthur Conan Doyle, which is at a Scholastic reading level grade 7. To navigate among pages, they were handed a normal mouse that had left, right, and scroll-wheel buttons. The left button navigates to the previous page, and the right button goes to the next page. To ensure that participants read the text carefully, random nursery rhymes were included in the text (1 rhyme sentence per page on average). Participants were instructed to click the center button of the mouse every time they encountered a rhyme, which increments the counter displayed on the screen. They were asked not to double count them if lines spanned from the end of one page to the beginning of the next page. The reading accuracy is then defined by the percentage of the target sentences found by the participants. On average, each page contains 462 characters, including the rhyme sentences. Moreover, to ensure that the participants are not just simply scanning the passage for the rhyme sentences, they are given comprehension quizzes after each reading session with 10 multiple-choice questions on each quiz.

3.2 Choice of Conditions - Two Studies

In Haynes et al.'s study [21], four angle offsets were tested on 12 participants: 0°, +10°, +20°, and +30°. Notice again that we define positive direction as horizontally offset from the user's PPOG toward the ear. Per the goal of our experiment, which is to provide a wider-ranging curve showing trends of users' comfort over different angle offsets for an HWD, we performed two studies sequentially to evaluate the exact opposite conditions, fill in the gap between the conditions, and compare the conditions with offsets in both directions. In both studies, we involved four conditions to keep consistent with Haynes et al.'s study and facilitate comparison across studies. We keep the 0° condition across the studies, which serves as a de facto comfort baseline to compare the data across studies. In pilot testing, we find that no participant can tolerate offsets greater than -50° or +50° for even five minutes, which provides an upper threshold for discomfort for all the experiments.

3.2.1 First Study. In our first study, we tested the exact opposite conditions as Haynes et al. did on 12 participants: -30°, -20°, -10°, 0°. Observing the results of this study (discussed in following sections), we discovered a less significant internal contrast among conditions, which was distinct from the trend exhibited in Haynes et al.'s study. As this inconsistency might indicate a higher tolerance of eye discomfort when users are reading in HWDs that have angular offsets towards the nose, we decided to perform the second study to verify our findings and analyze the difference between the two directions.

3.2.2 Second Study. Our second study focuses on filling the gaps between tested conditions while directly comparing different directions of angular offsets. We tested the angles -15°, 0°, +15°, and +25° on another 12 participants, following the same experiment procedure. Specifically, we wished to compare -15° and +15° as these locations seem potentially desirable for HWD manufacturers, as all pixels would be more than 8° from PPOG as recommended by Dowell [9] while still being near enough to PPOG for fast reference. Note that our emulated display centered at -15° has pixels ranging from -19.6° to -11.4°, and +15° has pixels between +11.4° and +19.6°. Furthermore, to confirm the sensitivity of our testing, we include the +25° condition, which is close to the extreme +30° condition and should be significantly more uncomfortable than the 0° condition based on the previous study done by Haynes et al.

3.3 Participants

We required the participants to (1) not wear glasses (including any kind of contact lenses and orthokeratology lenses), (2) have Snellen visual acuity of 20/40 or better on both eyes, and (3) have not read the novel "The Adventures of Sherlock Holmes." The reason why participants cannot wear glasses is that most eyeglasses and contact lenses bend light to a different extent depending on the entrance angle of the light regarding users' PPOG, which results in a disproportional distortion of the image from different angles.

The prohibition of orthokeratology lenses (lenses people wear only during sleep to temporarily correct vision) in this study is based on the feedback of an early participant whose visual acuity varies largely during the day (who also was excluded from the experiment results due to the same reason). The second requirement of visual acuity guarantees that the participants can read the text content effortlessly in normal conditions. The third requirement encourages the active participation of the participants during reading sessions.

Table 1 shows the demographic composition of participants in our first (n=12) and second (n=12) studies, respectively. For the computer usage time question, the participants are given the ranged time lengths such as “0-1 hours” and “1-3 hours”. The average time computer usage per day is calculated by taking the weighted average of the lowest value and the highest value within the range of participants’ selections. The eye dominance of each participant is determined using their hands and a small target, as described by Peli [34].

3.4 Data Collection

3.4.1 Bias Mitigation. Multiple potential sources of bias are involved in the experiment process, which we try to address through improvements to the experiment process utilized by Haynes et al.’s study.

1) *Avoiding carryover effect.* Each of the participants involved in this study completes the four conditions separately and in a specific order. In order to offset the bias caused by the carryover effect (the experience of former conditions affects participants in latter conditions), we utilize a 4x4 balanced Latin Square to define the order of conditions for each group of four participants [18]. Participants are assigned to groups based on their registration sequence. This way, given any two conditions A and B, the number of participants experience A before B is equal to the number of participants experience B before A.

2) *Adding training sessions.* To avoid potential learning effects [21], participants receive a 30-minute training session before they begin with the four testing conditions using the Latin Square sequence. Therefore, the participants are always familiarized with the equipments before participating in any testing conditions. In addition to the 30-minute training session as used in the study of Haynes’ et al. [20], a 5-minute introduction period is given to the participants before the training session so that they familiarize with the operation methods (e.g. navigating between the pages, recording the occurrence of nursery rhyme sentences, etc.). The angle offsets used in training sessions and the chapter read by participants during the training sessions are different from what is used in the testing sessions. For instance, in the first study when we are testing —30°, -20°, -10°, and 0°, we use +15° as the training condition. For each study we perform, we use the same angle offset and training chapter for all training sessions.

3) *Controlling subjective bias.* Since we are investigating the comfort of participants between conditions, it is important to design the experiment so minimal subjective bias is caused by participants’ prior knowledge or misunderstanding of the question. This involves preparation before the participant enters the room, phrasing of the question during the experiment, and the design of the survey questions. First, the location of the VDT is always set to the angle offset condition for the upcoming session before the participant enters the room. This technique avoids giving participants a subjective impression that the VDT “should” be in some place before the start of the experiment. Second, note that the apparatus to control participants’ head movement, including the chin rest, the headband with laser, and the eye patch to cover the participant’s left eye, brings extra discomfort to the participants on their general body. To compensate for this effect, we phrase the question to the participants as “*how comfortable is your eye feeling right now?*” such that they give ratings only accounting for the comfort of their eyes. We also print and put the question in front of the rotary knob to alert the participants when they need to express their feelings using the knob. Furthermore, in the asthenopia survey after the experiment, we asked the 9 questions in both positive and negative ways (hence 18 questions in total) about participants’ comfort feelings to derive their asthenopia score [20, 21]. For instance, “reading was comfortable” and “reading was uncomfortable” both exist in the survey, where participants need to indicate their extent of agreement to the statements using a 7-point Likert Scale. This way, the participants’ bias caused by the question wording is counterbalanced.

4) *Increasing rating sensitivity.* Notice that the comfort knob used in Haynes et al.’s study only corresponds to a 5-point Likert Scale, which gives participants fewer options to choose from and hence diminishes their accuracy of response. As suggested by psychometric literature, the 7-point Likert Scale is more sensitive to participants’ feedback while not increasing the stress for participants during the choice [13, 36, 40]. In addition, the 7-point Likert Scale is found better than the 5-point Likert Scale in online surveys [28], which corresponds to our use of online surveys after each session. To investigate a more accurate comfort range to put the displays in the view of HWD users, we prioritize the sensitivity of the metric in our studies. As a result, we utilize the 7-point Likert Scale during and after the experiment, and we convert the Likert Scale back to 5-point following conversion rules suggested by previous literature[8, 36] to facilitate data comparison with Haynes et al.’s study in the data analysis process.

5) *Reducing human interference.* Since our experiment is a timed and controlled process, we need to avoid participants being interrupted unexpectedly or receiving different experiences. We also need to ensure the accuracy of data collected by us. Therefore, to minimize the latency in data collection, we automated this process entirely with code. The computer program records when the participants move their heads away from the desired location, the time they spend on each page, as well as the number of sentences they find, and the computer will automatically stop the experiment after 30 minutes. For the comfort level expressed by the participants using the rotary knob during the session, we are also notifying the participants using alarms and noting down their comfort level without verbal communication when possible. Furthermore, to reduce human error in data collection, we utilized the official NASA-TLX mobile application [1] as well as Google Form to distribute the surveys to the participants after each 30-minute reading session.

3.4.2 Measurements. During the study, multiple data are collected from each participant aside from the demographic information in each angle offset condition with which they are tested.

Table 1: Demographic composition of participants in the first study (top half of the chart, n=12) and the second study (bottom half of the chart, n=12). Categories provided but not selected by any participants (e.g., "Other" in the question about gender) are not shown. Categories are ranked in alphabetical order.

Age		Ethnicity		Gender		Eye Dominance		Computer Usage
Range	**Count**	**Type**	**Count**	**Type**	**Count**	**Type**	**Count**	**Estimated Avg.**
18-21	4	Asian	11	Female	4	Left	4	7.96 (hrs/day)
22-25	7	White	1	Male	8	Right	8	
26-29	1							
Range	**Count**	**Type**	**Count**	**Type**	**Count**	**Type**	**Count**	**Estimated Avg.**
18-21	6	African American	1	Female	3	Left	2	7.83 (hrs/day)
22-25	5	Asian	9	Male	9	Right	10	
26-29	1	Hispanic/Latino	1					
		White	1					

1) *Comfort levels.* As stated above, we collect comfort levels from the participants using the rotary knob, which corresponds to a 7-point Likert Scale. We collect the comfort levels by asking participants the question (printed as a question card) "how comfortable is your eye feeling right now?" The collections are made before the 30-minute reading session begins (0 minutes) and after each 5-minute interval during the reading session. In total, seven comfort levels are collected during each session.

We utilized 7-point Likert Scale scores throughout our study, whereas Haynes et al. utilized 5-point Likert Scale scores [21]. Therefore, before we do the analysis and compare our results with results from Haynes et al.'s study, all the Likert Scale scores we have are converted to a 5-point scale using the following equation suggested by the conversion rule in previous literature [8, 36].

$$X_{5-point} = (X_{7-point} - 1) * 4/6 + 1$$

The conversion to 5-point Likert Scale instead of 7-point Likert Scale helps readers to compare the graphs established in this study with the graphs in Haynes et al.'s previous study.

2) *Reading task performance.* We collect the reading speed (seconds spent per page), the rate of nursery rhymes found during the reading session, and the comprehension score (number of questions answered correctly in the comprehension quiz) from each session.

3) *Subjective preference, Asthenopia score, and NASA-TLX.* In the post-session survey about asthenopia symptoms, the question "I prefer a display at this position" is asked in addition to the 9 questions (asked twice in positive and negative wordings) related to potential asthenopia symptoms [33, 34]. The asthenopia score is calculated by the average of all the 7-point Likert Scale ratings in the negatively worded questions and (7 - rating) of all the ratings in the positively worded questions. The higher the asthenopia score is, the worse the participants feel during the experiment. The NASA-TLX score is calculated using the weighted average of the task load indices in each section [19].

4) *Head Stability.* We also collect the number of seconds when the participant doesn't keep their head straight (fail to align the laser with the photo-resistor apparatus).

4 RESULTS

We first define the primary metric for comfort analysis. In Haynes et al.'s study, the summed comfort was used as a metric, which was defined as the sum of the seven comfort levels collected from participants in each condition [21]. However, the validity of such a metric depends on all participants giving seven (the highest comfort level) as their initial comfort level during the experiment (i.e., a participant giving sequential ratings 5,4,4,4,3,3,3 should be interpreted as having the same level of discomfort as another participant giving ratings 4,3,3,3,2,2,2). Among all reading sessions in our experiment, participants started their ratings from a lower level (mostly a 5 or 6 instead of a 7 rating) 43% of the time. However, few participants started an experimental condition (3 out of 96 total testing sessions conducted) at lower than 7 and ended the condition at a 1. In other words, the floor effect is negligible, and using the change in comfort over time seems a viable metric across conditions. Specifically, we use summed delta comfort (SDC) as our major metric for comfort analysis, which is defined below (where C_x represents the comfort level given at the x-th minute in the 30-minute reading session). A higher SDC indicates a higher total discomfort experienced during the reading session.

$$\text{Summed Delta Comfort (SDC)} = \sum_{i=0}^{6} (C_0 - C_{5i})$$

4.1 Analysis

For statistical tests, we first utilize the Friedman Test [35] to test our hypothesis that the measurement results are significantly different between conditions. If affirmative, we proceed with the Conover-Iman method [7] adjusted by the Benjamini-Hochberg false discovery rate procedure [3] to explore which specific pairs of conditions are different. This procedure analyzes the variance among different samples from the same participant group and limits type I errors [5, 21]. Notice that the use of either a 7-point Likert Scale or a 5-point Likert Scale does not affect the results of this statistical test procedure, as it primarily investigates results based on the ranking difference among the conditions.

In the first study, we test the conditions -30°, -20°, -10°, and 0°. Surprisingly, the Friedman test does not yield a statistically significant result ($\chi^2 \approx 4.37, p \approx 0.22$) among the conditions, indicating

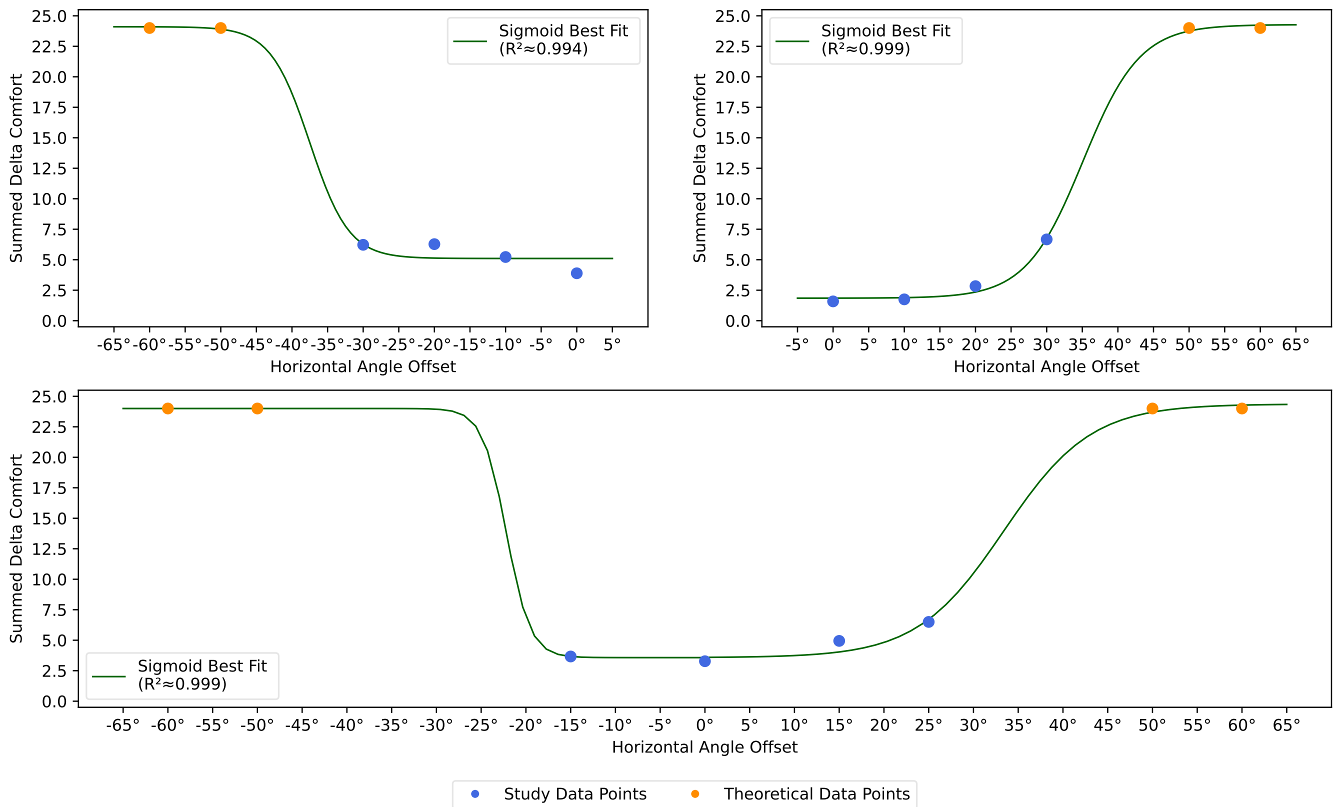


Figure 4: Separate sigmoid best-fit curves using data from study 1 (top left), Haynes et al.'s study [21] (top right), and study 2 (bottom). Note that the steep slope at -15 for the bottom graph is misleading, as we suspect, given the results from study 1 (top left), that the increase is much more shallow. Deltas are first converted to 5-point scale.

no significant difference among these conditions, challenging the assumption based on Haynes et al.'s study that there should at least be a significant difference between the -30° and 0° conditions, as 0° and +30° are significantly different. With this counter-intuitive result, we perform our second study, which tests the conditions -15°, 0°, +15°, and +25°. As explained above, the conditions in the second study fill the gaps between the data points and put more emphasis on the positive direction to test the sensitivity of our results based on the prior knowledge derived from Haynes et al.'s study [21]. Surprisingly again, the Friedman test gets a statistically significant result ($\chi^2 \approx 11.32, p \approx 0.01$) comparing the conditions -15°, 0°, +15°, and +25°. Using the Conover-Iman method [7] adjusted by the Benjamnyi-Hochberg false discovery rate procedure [3] to further test the results, we found statistically significant differences between -15° and +25° ($p \approx 0.0084$), as well as between 0° and +25° ($p \approx 0.0007$). However, we did not find a statistically significant difference between the +15° and +25° conditions. Therefore, the results in the second study effectively indicate an asymmetric trend in the negative direction (horizontally offsetting towards the nose from the right-eye PPOG) and the positive direction (horizontally offsetting towards the ear from the right-eye PPOG).

Figure 4 shows psychometric sigmoids fit to the data of the three experiments. Visually, the comfort scores at +25° and +30° positions suggest increasing discomfort, which is supported by the statistical analysis. However, the trend between -30° and 0° is less clear.

4.2 Trends

Comparing data across experiments is fraught with possible confounds (e.g., different participants, different times of year, etc.), which is why we use 0° as a baseline, "least discomfort" condition from which we offset the other conditions in a given experiment. Fortunately, +/-50° provides an upper limit on discomfort, providing scaling. Thus, we can attempt to unify the curves to provide the reader with a better intuition of the results. We recommend, however, that new experiments should be run for directly comparing if one offset is better than another instead of relying on these curves.

First, we define offsetted summed delta comfort (OSDC) as follows, where $SDC_{p,a}$ represents the SDC of participant p during the reading session where the VDT is centered at the angle offset a, and $OSDC_{p,a}$ represents the OSDC of such participant at the corresponding reading session:

$$OSDC_{p,a} = SDC_{p,a} - SDC_{p,0}$$

In this way, we align all data points with the baseline of zero. We then plot the participants' OSDC relative to different angle offsets tested, including the data gathered from Haynes et al.'s published study [20, 21]. In the following graphs, we utilize opacity to denote the number of data points overlapped.

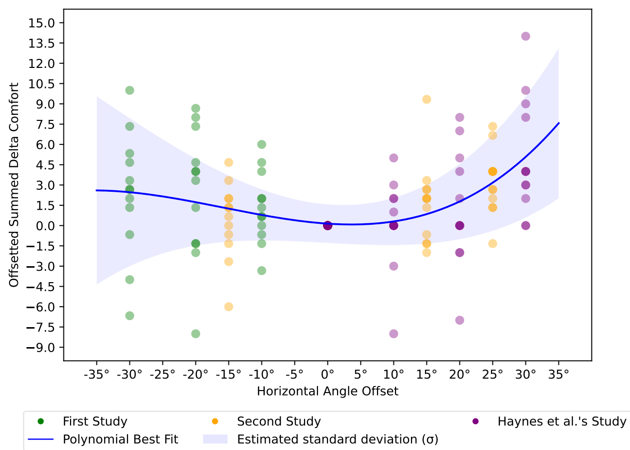


Figure 5: 3-degree polynomial best fit of Offsetted Summed Delta Comfort (OSDC) (deltas first converted to 5-point scale)

1) *Polynomial Fit.* We first utilize a 3-degree polynomial line of best fit (Figure 5) to create a polynomial trend across the SDC data derived from all the studies. As exhibited by the polynomial line, the angle offsets in the negative direction are more tolerable than the angle offsets in the opposite direction. Furthermore, as the absolute value of the horizontal angle offset increases, the estimated standard deviation (using sample standard deviation) also increases, indicating a more unstable preference level across the population. The polynomial fit presents a trend within the [-30°,+30°] range of cases tested.

2) *Sigmoid Fit.* As tested in the study of Haynes et al.[21], the end-to-end change in comfort ($C_0 - C_{30}$) shows a psychometric sigmoid trend in the positive direction of horizontal angle offsets, which is also true for the SDC metric we use, as they are highly related metrics. Therefore, it is reasonable to assume that the SDC metric we use also shows a psychometric sigmoid trend in the negative direction of horizontal angle offsets. That is, the SDC is expected to follow an S curve of increment, increasing slowly under a certain threshold of negative angle offset and starting to increase dramatically until it reaches close to the theoretical upper bound of SDC, which is $6 * (5 - 1) = 24$ on a 5-point scale on known extreme conditions of -50°, -60°, +50°, and +60°. It is simple to test the assumption of the higher bound as a quick self-study shows that the -40° condition is around the edge of making people give the lowest comfort level after the first 5 minutes, and -50° or -60° is physiologically unfeasible for people to read.

Therefore, following this knowledge, we can fit all data to sigmoidal functions, with theoretical data points having the highest OSDC possible at the -50°, -60°, +50°, and +60° conditions and appropriate weight (each of the points should represent data from 12 participants), as shown in Figure 4 (using SDC data from three studies to create separate sigmoid fits) and Figure 6 (using OSDC data from three studies for a single bidirectional sigmoid fit). It is worth noting that the bidirectional sigmoidal curve fitted ($R^2 \approx 0.962$) using all data from the three studies highly resembles the fitted sigmoidal function in the positive direction [21]. More detailed data and trends derived from each participant are included in Appendix A and B.

4.3 Other Metrics

In addition to the primary SDC metric measuring participants' comfort levels during the reading period, additional data collected, such as the subjective preference, asthenopia score, and NASA-TLX, are also analyzed post-hoc using the same testing process as we use for testing the SDC metric. Detailed data for these metrics can be viewed in Table 2. Viewing the results of statistical tests, it is worth noting that the Friedman test in subjective preference among conditions in the second study also shows significant results ($\chi^2 \approx 21.03$, $p \approx 0.0001$). Further analysis using the Conover-Iman method [7] with the Benjamini-Hochberg false discovery rate procedure [3] demonstrates a significant preference for -15° ($p \approx 2 * 10^{-7}$), 0° ($p \approx 3 * 10^{-5}$), and +15° ($p \approx 0.0059$) against the +25° condition, and the 0° condition is also significantly more preferred ($p \approx 0.0012$) than the +15° condition. On the other hand, the Friedman test on subject preference in our first study yielded a significant but relatively weaker result ($\chi^2 \approx 9.97, p \approx 0.019$) and only one pair of significant preference for 0° ($p \approx 0.0021$) against -30°.

5 DISCUSSION

The asymmetric trend derived from the study metrics as well as the polynomial and sigmoidal curves in Figure 6 suggest that participants have a higher tolerance for displays offset toward the negative direction (horizontal offsets toward the nose from the PPOG) compared to their tolerance in the positive direction (toward the ear). The curve suggests a range between [-20°, +15°] where horizontal offsets have minimal effect on participants' increase in discomfort compared to facing straightforward over a 30-minute reading period. Being more aggressive, the curve also indicates a rough tolerance range of approximately [-30°, +25°], where participants' OSDC is predicted to be lower than 3, indicating a relatively acceptable discomfort. However, the standard deviation of participants' OSDC is expected to increase to approximately 4.58 at -30° and 4.22 at +25°, which increases the possibility of making at least some users of the display more uncomfortable. Therefore, we suggest between [-20°, +15°] for HWD manufacturers to center their displays. Note that our "display" had a horizontal FOV of 9.2°, suggesting that pixels of the display should not exceed [-24.6°, +19.6°].

This increased tolerance for positions closer to the nose matches intuition. As an object approaches a user's nose, the eyes naturally converge to track the object. However, the eyes rarely diverge past the stage when viewing an object at a distance (beyond about 10 meters, changes in eye focus and divergence are minimal). Thus, eye muscles are exercised more frequently towards the nose than towards the ears.

Combining this result with the theoretical system (Figure 1) proposed in the Considerations section, we have shown that starting the display with θ = 70.1° (corresponding to a -19.9° horizontal angle offset of the display's edge) is feasible from the point of considering participants' comfort of reading angles.

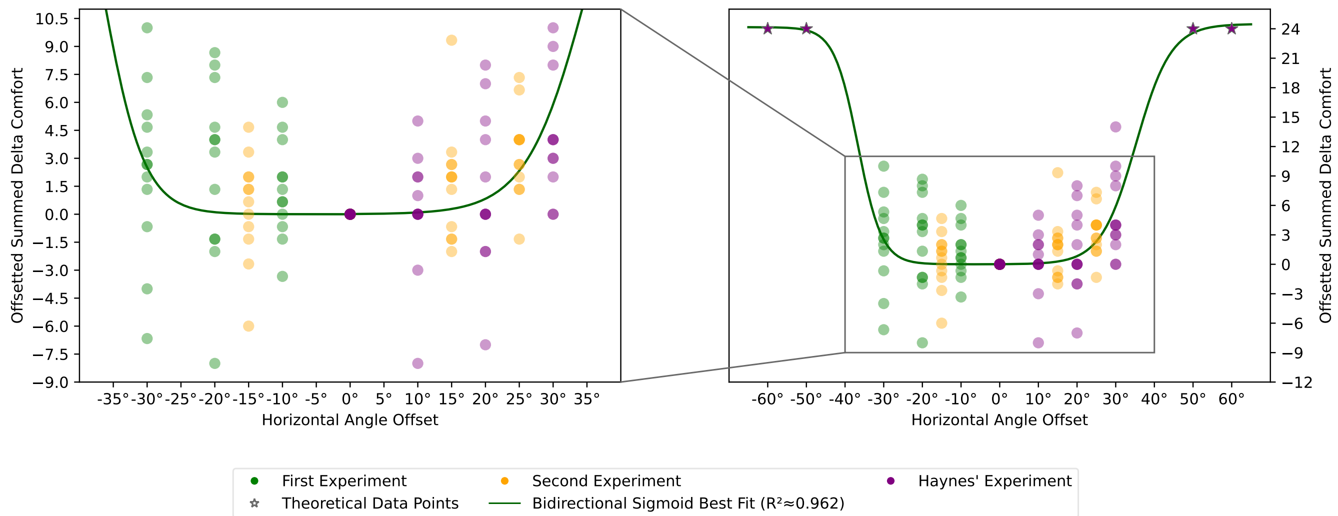


Figure 6: OSDC with bidirectional sigmoid best fit (deltas first converted to 5-point scale). The left image magnifies the region of interest from the right image. Theoretical data points encompass the corresponding weights of all three studies during the sigmoid fit process.

Table 2: Median ± median absolute deviation for summed delta comfort, subjective preference, and asthenopia. Mean ± standard deviation for other measurements. Bold texts indicate best conditions w.r.t measurements. The upper part represents data and comparisons from our first study, and the lower part represents the second study. Data surrounded by parentheses denotes statistically significant advantages over the cases included in the exponents marked on the top right of the parentheses.

Measurement	-30°	-20°	-10°	0°
Summed Delta Comfort	6.33 ± 2.67	7.33 ± 2.67	5.0 ± 2.33	**3.67 ± 1.0**
Subjective Preference	1.33 ± 0.33	3.0 ± 0.67	3.33 ± 1.0	$(\mathbf{4.33 \pm 0.0})^{\{-30^\circ\}}$
Asthenopia Score	32.75 ± 4.0	28.75 ± 7.75	29.25 ± 5.0	**25.75 ± 5.75**
NASA-TLX	57.78 ± 14.86	**48.33 ± 19.37**	51.05 ± 15.42	48.86 ± 18.01
Reading Speed (pgs)	76.25 ± 29.22	**83.83 ± 44.24**	69.17 ± 32.71	72.83 ± 31.65
Reading Accuracy (%)	88.38 ± 16.13	86.06 ± 13.0	84.92 ± 14.33	**90.7 ± 10.7**
Reading Comp. (# correct)	6.58 ± 1.55	6.75 ± 2.2	6.08 ± 2.66	**7.58 ± 1.61**
Head Stability (%)	93.21 ± 2.77	93.65 ± 3.97	94.11 ± 4.37	**95.12 ± 2.96**

Measurement	-15°	0°	+15°	+25°
Summed Delta Comfort	$(\mathbf{3.33 \pm 2.67})^{\{+25^\circ\}}$	$(3.67 \pm 2.0)^{\{+25^\circ\}}$	4.0 ± 2.0	7.0 ± 1.67
Subjective Preference	$(3.33 \pm 1.0)^{\{+25^\circ\}}$	$(\mathbf{4.33 \pm 0.67})^{\{+15^\circ,+25^\circ\}}$	$(2.0 \pm 0.33)^{\{+25^\circ\}}$	1.67 ± 0.67
Asthenopia Score	25.75 ± 8.25	**23.75 ± 8.25**	38.0 ± 7.0	38.25 ± 5.5
NASA-TLX	46.03 ± 17.91	**41.72 ± 14.43**	50.72 ± 15.96	57.89 ± 18.98
Reading Speed (pgs)	76.17 ± 13.64	72.5 ± 16.33	75.17 ± 22.99	**77.33 ± 17.06**
Reading Accuracy (%)	**92.81 ± 11.68**	91.66 ± 8.39	87.16 ± 14.37	89.61 ± 13.99
Reading Comp. (# correct)	6.58 ± 2.1	**7.08 ± 2.66**	6.25 ± 2.59	6.33 ± 2.32
Head Stability (%)	**94.04 ± 2.93**	93.49 ± 3.32	90.98 ± 4.69	89.55 ± 4.82

For many tasks, users may tolerate even more extreme angles, as the users are alternating their visual attention between the off-center HWD and a task in PPOG, providing the eyes with small breaks. That possibility was noted by both Lin [29] and Haynes [20]. On the other hand, a reading task implies that the user is scanning their eyes across the page. A task that requires the user to spend more of their visual attention time on the more extreme angles may result in more discomfort. Thus, we suggest that interface components that are used most often be placed at the edge of the display closest to PPOG.

6 FUTURE WORK AND LIMITATIONS

In our second study, we only directly compared the difference between the -15° and +15° conditions. Deriving from the trend we predicted, future work can select and directly compare conditions farther from users' PPOG, such as -25° and +25° to further test the prediction. Furthermore, the feasibility of the proposed system with both a stereoscopic view and a separate view is also dependent on users' eye strain caused by different focal distances, especially when the users' eyes are focusing on the stereoscopic view located at a position close to the user. Also, our study only focuses on a single task (i.e. reading). As Lin et al.'s study [29] further validates the findings in Haynes et al.'s prior study [21] (which we base our approach on) in an alternating task with HWD (i.e. order-picking), future studies focusing on the same topic but with different tasks are also essential in expanding the scope of our study results. While our research extensively studies the effect of horizontal offsets, vertical offsetting the display could yield other positions for the display that users may find comfortable. In addition, testing the prediction of our study by crafting HWD prototypes or applying the result to positioning larger FOV displays are also recommended directions for future research.

7 CONCLUSION

Based on the procedure of Haynes et al.'s study testing participants' reading comfort with displays centered at the horizontal angle offsets 0°, +10°, +20°, and +30° [20, 21] from users' PPOG, we further tested the conditions -30°, -20°, -10°, and 0° in the first study and -15°, 0°, +15°, and +25° in the second study. Our second study shows that -15° and 0° conditions are statistically significantly more comfortable for reading than +25°, while +15° is not significantly better than +25°. Here, the positive direction is defined as horizontally offsetting the HWD towards the ear from the users' PPOG. Our first study yields small differences between conditions, suggesting an asymmetric trend of increasing discomfort with respect to increasing absolute horizontal angle offsets. The bidirectional psychometric sigmoidal curve derived from the combination of our and Haynes et al.'s data aligns with the predicted curve in Haynes et al.'s study. Furthermore, the sigmoidal curve suggests a general range of [-20°, +15°] for manufacturers to center displays, which limits the angular position of lighted pixels in the range [-24.6°, +19.6°] for a HWD having a 9.2° horizontal FOV. From the aspect of users' reading comfort, the data suggests the feasibility of a two-eyed AR system, which can potentially support one stereoscopic view and two 2D views.

ACKNOWLEDGMENTS

To accurately replicate and improve the study performed by Haynes et al., we contacted Haynes et al. and inquired about specific details of the study [20, 21]. We hereby express our appreciation for their help during the preparation phase of our experiment.

REFERENCES

[1] National Aeronautics and Space Administration. 2017. Human Systems Integration Division NASA Ames - Outreach & Publications. https://hsi.arc.nasa.gov/awards_pubs/news_view.php?news_id=117
[2] Yuliya Bababekova, Mark Rosenfield, Jennifer E. Hue, and Rae R. Huang. 2011. Font Size and Viewing Distance of Handheld Smart Phones. *Optometry and Vision Science* 88, 7 (July 2011), 795–797. https://doi.org/10.1097/OPX.0b013e3182198792
[3] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57, 1 (Jan. 1995), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
[4] Chun-Wen Chao, Cheng Hung Huang, and Ji-Liang Doong. 2009. See-through head-up display of in-vehicle information systems for vision enhancement. In *The International Association of Societies of Design Research Conference*. Korea Design Center, Coex, Seoul, Korea, 877–886.
[5] Shi-Yi Chen, Zhe Feng, and Xiaolian Yi. 2017. A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease* 9, 6 (June 2017), 1725–1729. https://doi.org/10.21037/jtd.2017.05.34
[6] Soon Hau Chua, Simon T. Perrault, Denys J. C. Matthies, and Shengdong Zhao. 2016. Positioning Glass: Investigating Display Positions of Monocular Optical See-Through Head-Mounted Display. In *Proceedings of the Fourth International Symposium on Chinese CHI*. ACM, San Jose USA, 1–6. https://doi.org/10.1145/2948708.2948713
[7] W Conover and R Iman. 1979. *Multiple-comparisons procedures. Informal report.* Technical Report LA-7677-MS, 6057803. Los Alamos National Laboratory. 18 pages. https://doi.org/10.2172/6057803
[8] John Dawes. 2008. Do Data Characteristics Change According to the Number of Scale Points Used? An Experiment Using 5-Point, 7-Point and 10-Point Scales. *International Journal of Market Research* 50, 1 (Jan. 2008), 61–104. https://doi.org/10.1177/147078530805000106
[9] Susan R. Dowell, David C. Foyle, Becky L. Hooey, and Jennifer L. Williams. 2002. The Effect of Visual Location on Cognitive Tunneling with Superimposed Hud Symbology. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 46, 1 (2002), 121–125. https://doi.org/10.1177/154193120204600125
[10] Alexander Duane. 1922. Studies in Monocular and Binocular Accommodation with their Clinical Applications. *American Journal of Ophthalmology* 5, 11 (Nov. 1922), 865–877. https://doi.org/10.1016/S0002-9394(22)90793-7
[11] Epson. 2023. Smart Glasses | Augmented Reality Headsets | Epson.com | Epson US. https://epson.com/moverio-smart-glasses-augmented-reality-devices-headsets
[12] Hamid Fesharaki, Leila Rezaei, Fereidoun Farrahi, Taghi Banihashem, and Ahmad Jahanbakhshi. 2012. Normal Interpupillary Distance Values in an Iranian Population. *Journal of Ophthalmic & Vision Research* 7, 3 (July 2012), 231–234. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3520592/
[13] Kraig Finstad. 2010. Response Interpolation and Scale Sensitivity: Evidence Against 5-Point Scales. *Journal of Usability Studies* 5, 3 (May 2010), 104–110.
[14] David C. Foyle, Robert S. McCann, Beverly D. Sanford, and Martin F.J. Schwirzke. 1993. Attentional Effects with Superimposed Symbology: Implications for Head-up Displays (HUD). *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 37, 19 (1993), 1340–1344. https://doi.org/10.1518/107118193784162227
[15] J. Fukano, S. Okabayashi, M. Sakata, and T. Hatada. 1995. Automotive head-up displays for navigation use. *Proceedings of International Technical Conference on the Enhanced Safety of Vehicles* 1995 (1995), 306–314.
[16] Engo Glasses. 2023. Engo Eyewear, Real time sports data displayed where you need it most. https://engoeyewear.com
[17] Google. 2013. Google Glass. https://www.google.com/glass/start/ the product has been discontinued.
[18] D. A. Grant. 1948. The latin square principle in the design and analysis of psychological experiments. *Psychological Bulletin* 45, 5 (1948), 427–442. https://doi.org/10.1037/h0053912
[19] S. G. Hart and L. E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183.
[20] Malcolm Haynes. 2017. *Lateral Positioning of a Monocular Head-worn Display*. Ph. D. Dissertation. Georgia Institute of Technology, Atlanta, GA, USA. Advisor(s) Starner, Thad.
[21] Malcolm Haynes and Thad Starner. 2018. Effects of Lateral Eye Displacement on Comfort While Reading from a Video Display Terminal. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 138 (January 2018), 1–17. Issue 4. https://doi.org/10.1145/3161177
[22] M. L. Hershberger and D. F. Guerin. 1975. *Binocular Rivalry in Helmet-Mounted Display Applications:*. Technical Report. Defense Technical Information Center, Fort Belvoir, VA. https://doi.org/10.21236/ADA013838
[23] Yasuhiro Inuzuka, Yoshimasa Osumi, and Hiroaki Shinkai. 1991. Visibility of Head up Display (HUD) for Automobiles. *Proceedings of the Human Factors Society Annual Meeting* 35, 20 (1991), 1574–1578. https://doi.org/10.1177/154193129103502033 arXiv:https://doi.org/10.1177/154193129103502033
[24] M. Kanbara, T. Okuma, H. Takemura, and N. Yokoya. 2000. A stereoscopic video see-through augmented reality system based on real-time vision-based registration. In *Proceedings IEEE Virtual Reality 2000 (Cat. No.00CB37048)*. IEEE Comput. Soc, New Brunswick, NJ, USA, 255–262. https://doi.org/10.1109/VR.2000.840506
[25] R.M. Katsuyama, D.L. Monk, and E.P. Rolek. 1989. Effects of visual display separation upon primary and secondary task performances. In *Proceedings of the IEEE National Aerospace and Electronics Conference*, Vol. 2. IEEE, IEEE, Dayton, OH, USA, 758–764. https://doi.org/10.1109/NAECON.1989.40297
[26] Gregorij Kurillo, Evan Hemingway, Mu-Lin Cheng, and Louis Cheng. 2022. Evaluating the Accuracy of the Azure Kinect and Kinect v2. *Sensors (Basel, Switzerland)* 22, 7 (March 2022), 2469. https://doi.org/10.3390/s22072469

[27] Robert S. Laramee and Colin Ware. 2002. Rivalry and Interference with a Head-Mounted Display. *ACM Transactions on Computer-Human Interaction* 9, 3 (sep 2002), 238–251. https://doi.org/10.1145/568513.568516

[28] James R. Lewis. 1993. Multipoint scales: Mean and median differences and observed significance levels. *International Journal of Human-Computer Interaction* 5, 4 (Oct. 1993), 383–392. https://doi.org/10.1080/10447319309526075

[29] Georgianna Lin, Malcolm Haynes, Sarthak Srinivas, Pramod Kotipalli, and Thad Starner. 2021. Towards Finding the Optimum Position in the Visual Field for a Head Worn Display Used for Task Guidance with Non-registered Graphics. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (March 2021), 1–26. https://doi.org/10.1145/3448091

[30] Tadashi Lino, Tatsumi Otsuka, and Yoshiyuki Suzuki. 1988. Development of Heads-Up Display for a Motor Vehicle. In *SAE International Congress and Exposition*. SAE International, Warrendale, PA, USA, 12 pages. https://doi.org/10.4271/880217

[31] MarketsAndMarkets. 2021. HMD (head mounted display) market size, share, industry trends, companies, Growth Analysis 2030. https://www.marketsandmarkets.com/Market-Reports/head-mounted-display-hmd-market-729.html

[32] Kibum Park and Youngjae Im. 2020. Ergonomic Guidelines of Head-Up Display User Interface during Semi-Automated Driving. *Electronics* 9, 4 (April 2020), 611. https://doi.org/10.3390/electronics9040611

[33] Eli Peli. 1998. The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display. *Vision Research* 38, 13 (June 1998), 2053–2066. https://doi.org/10.1016/S0042-6989(97)00397-0

[34] Eli Peli. 1999. *Optometric and perceptual issues with head-mounted displays*. McGraw-Hill Professional, New York, NY, USA, 205–276.

[35] Dulce G. Pereira, Anabela Afonso, and Fátima Melo Medeiros. 2015. Overview of Friedman's Test and Post-hoc Analysis. *Communications in Statistics - Simulation and Computation* 44, 10 (Nov. 2015), 2636–2653. https://doi.org/10.1080/03610918.2014.931971

[36] Carolyn C Preston and Andrew M Colman. 2000. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica* 104, 1 (March 2000), 1–15. https://doi.org/10.1016/S0001-6918(99)00050-5

[37] C. E. Rash, M. B. Russo, T. R. Letowski, and E. T. Schmeisser. 2009. *Helmet-Mounted Displays: Sensation, Perception and Cognition Issues*. U.S. Army Aeromedical Research Laboratory, Fort Novosel, AL, USA.

[38] Hannah Ritchie and Max Roser. 2019. Age Structure. https://ourworldindata.org/age-structure

[39] Rufat Rzayev, Paweł W. Woźniak, Tilman Dingler, and Niels Henze. 2018. Reading on Smart Glasses: The Effect of Text Position, Presentation Type and Walking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3173619

[40] Jeff Sauro Sauro. 2010. Should You Use 5 Or 7 Point Scales? – MeasuringU. https://measuringu.com/scale-points/

[41] Yukun Song, Parth Arora, Rajandeep Singh, Srikanth T. Varadharajan, Malcolm Haynes, and Thad Starner. 2023. Going Blank Comfortably: Positioning Monocular Head-Worn Displays When They are Inactive. In *Proceedings of the 2023 ACM International Symposium on Wearable Computers* (Cancun, Quintana Roo, Mexico) *(ISWC '23)*. Association for Computing Machinery, New York, NY, USA, 114–118. https://doi.org/10.1145/3594738.3611375

[42] Tooz Technologies. 2023. tooz Smart Glasses — Stylish, Lightweight, Unobtrusive | tooz.com. https://tooz.com/

[43] Patricia M. Ververs and Christopher D. Wickens. 1998. Head-Up Displays: Effect of Clutter, Display Intensity, and Display Location on Pilot Performance. *The International Journal of Aviation Psychology* 8, 4 (1998), 377–403. https://doi.org/10.1207/s15327108ijap0804_4 arXiv:https://doi.org/10.1207/s15327108ijap0804_4 PMID: 11542276.

[44] Vuzix Corporation. 2023. Vuzix VX Ultralite Video. https://www.youtube.com/watch?v=RENbuIX1TM0

[45] Hiroshi Watanabe, Herbert Yoo, Omer Tsimhoni, and Paul A. Green. 1999. The Effect of HUD Warning Location on Driver Response. In *Proceedings of 6th World Congress on Intelligent Transport Systems (ITS)*. ITS World Congress, Toronto, Canada, 10 pages.

[46] Xreal. 2021. Xreal Product page. https://www.xreal.com/air Formarly Know as Nreal.

[47] Tao Yang and Young Mi Choi. 2015. Study on the Design Characteristics of Head Mounted Displays (HMD) for Use in Guided Repair and Maintenance. In *Virtual, Augmented and Mixed Reality*, Randall Shumaker and Stephanie Lackey (Eds.). Springer International Publishing, Cham, 535–543.

[48] Xianjun Sam Zheng, Cedric Foucault, Patrik Matos da Silva, Siddharth Dasari, Tao Yang, and Stuart Goose. 2015. Eye-Wearable Technology for Machine Maintenance: Effects of Display Position and Hands-Free Operation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2125–2134. https://doi.org/10.1145/2702123.2702305

A LINE GRAPH FOR EACH PARTICIPANT

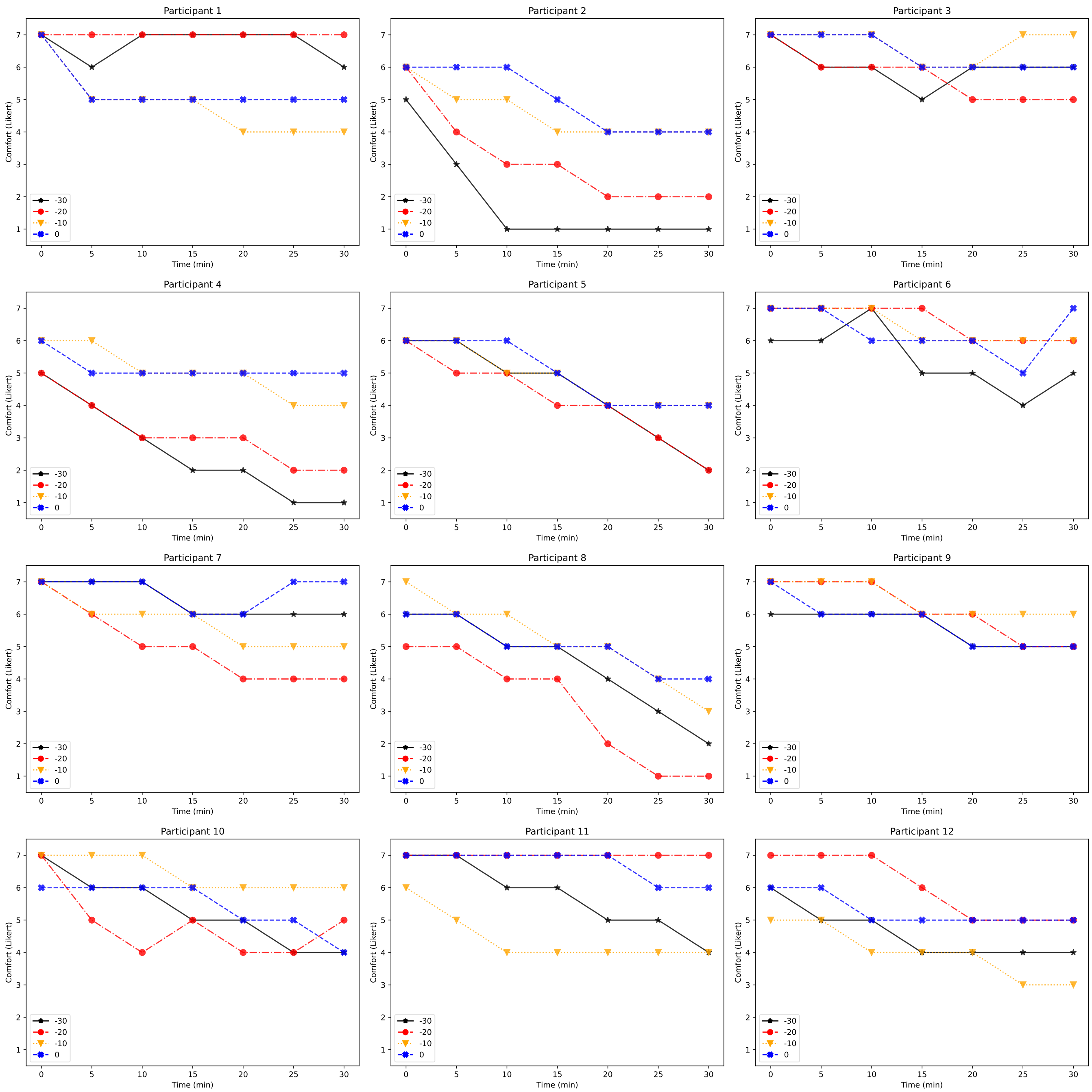


Figure 7: Line Graphs of Comfort Data Collected From Participants in Study 1 (7-Point Scale)

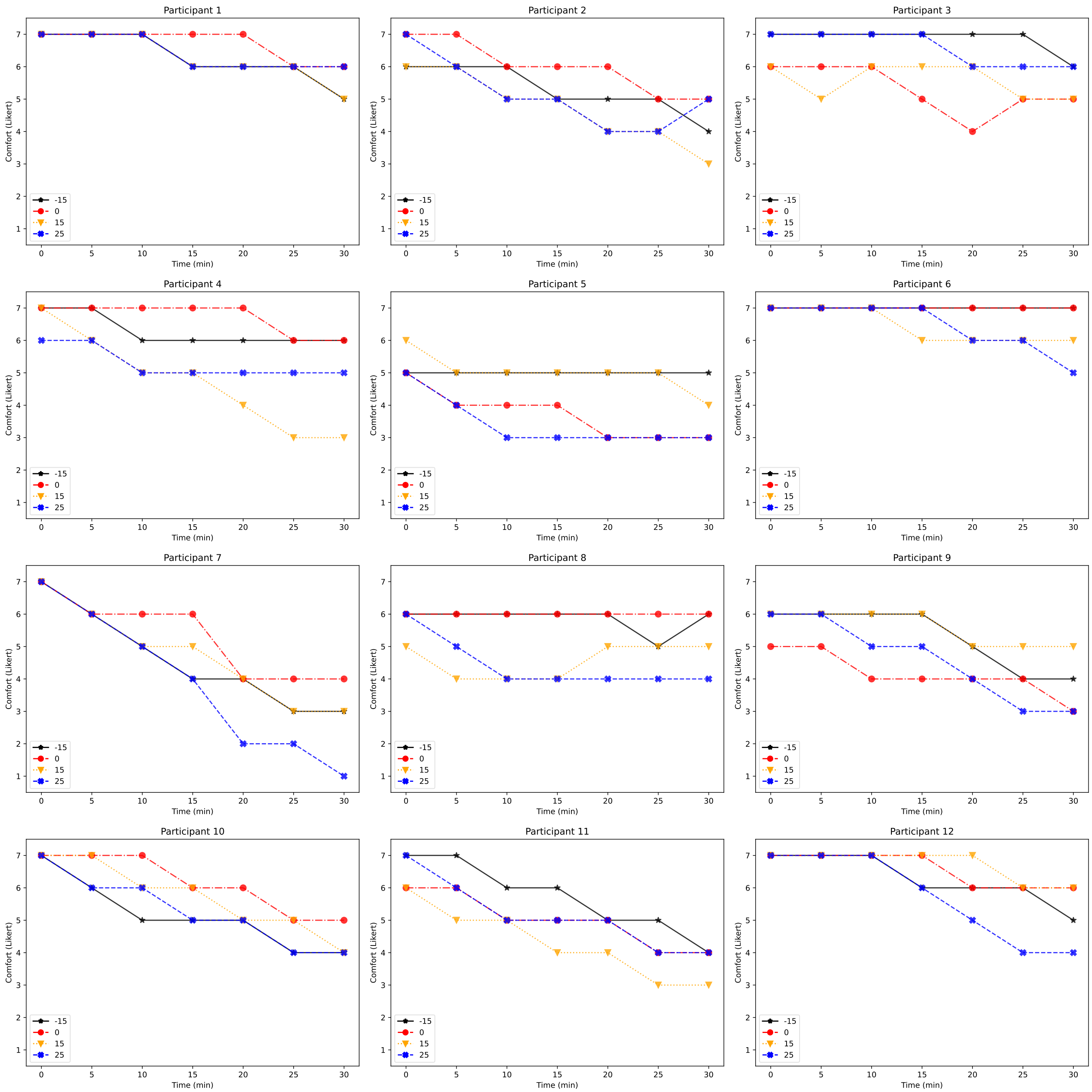


Figure 8: Line Graphs of Comfort Data Collected From Participants in Study 2 (7-Point Scale)

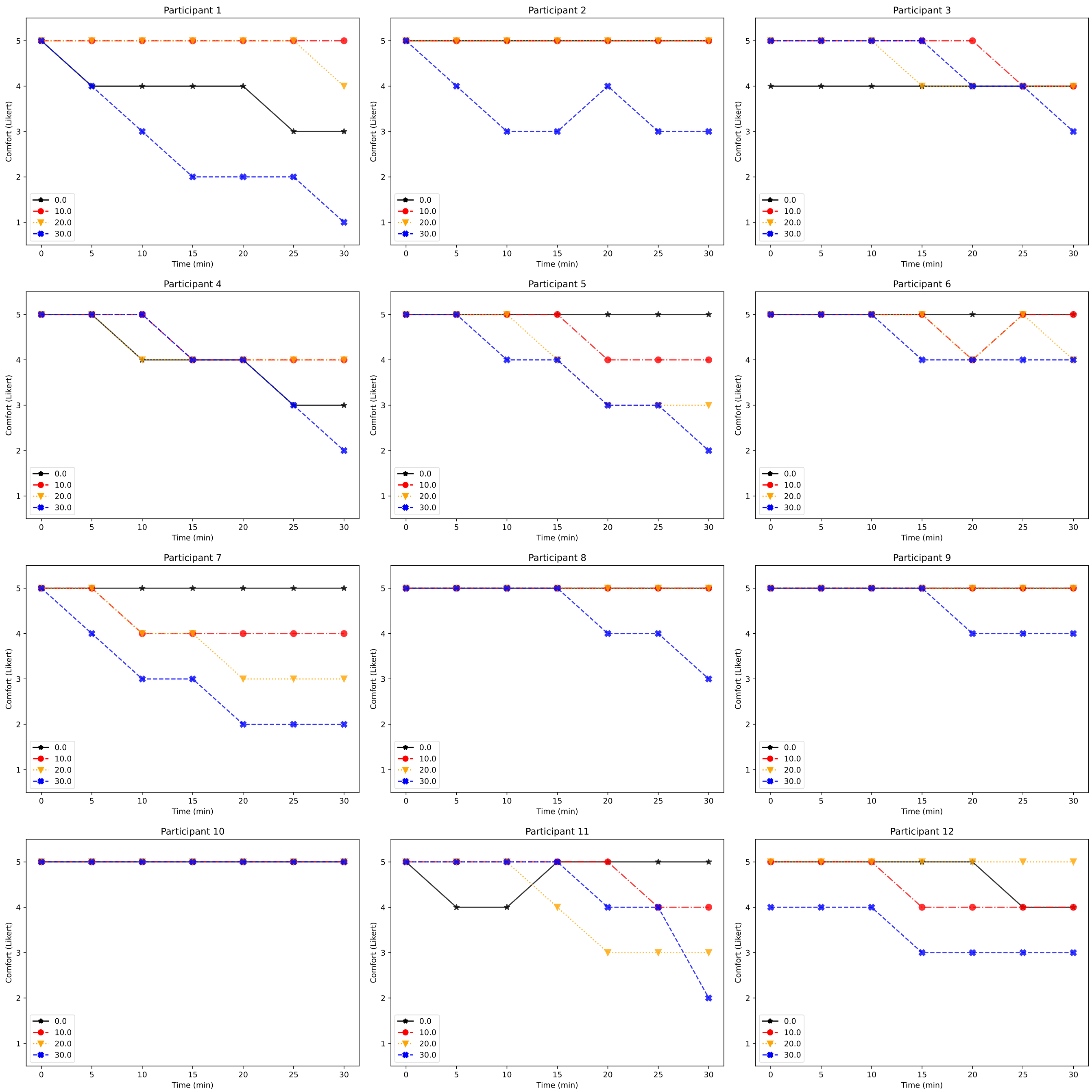


Figure 9: Line Graphs of Comfort Data From Haynes et al.'s study [21] (5-Point Scale)

B SIGMOID BEST FIT FOR EACH PARTICIPANT

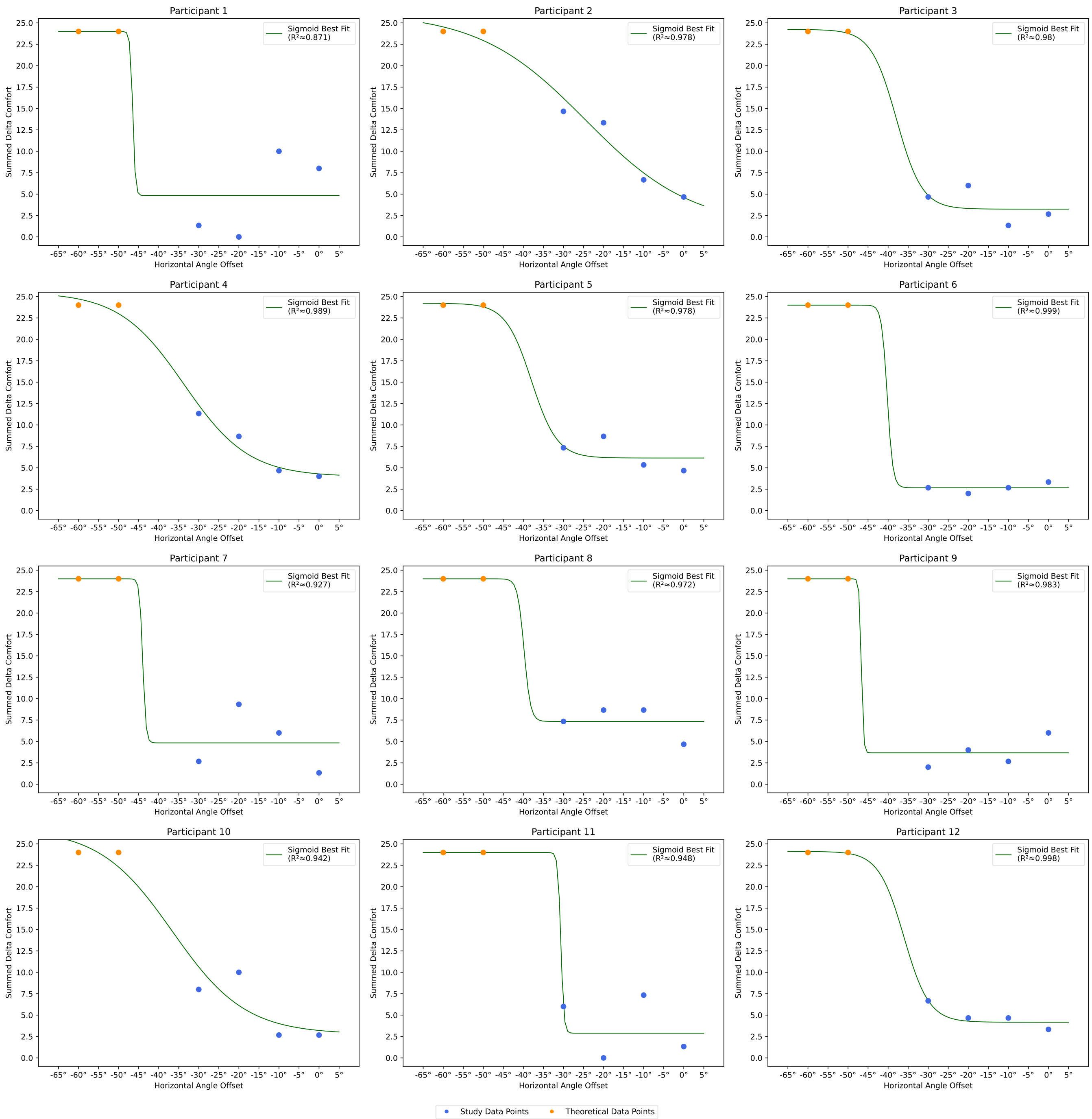


Figure 10: Sigmoid Best Fits for Each Participant in Study 1 (Converted to 5-Point Scale)

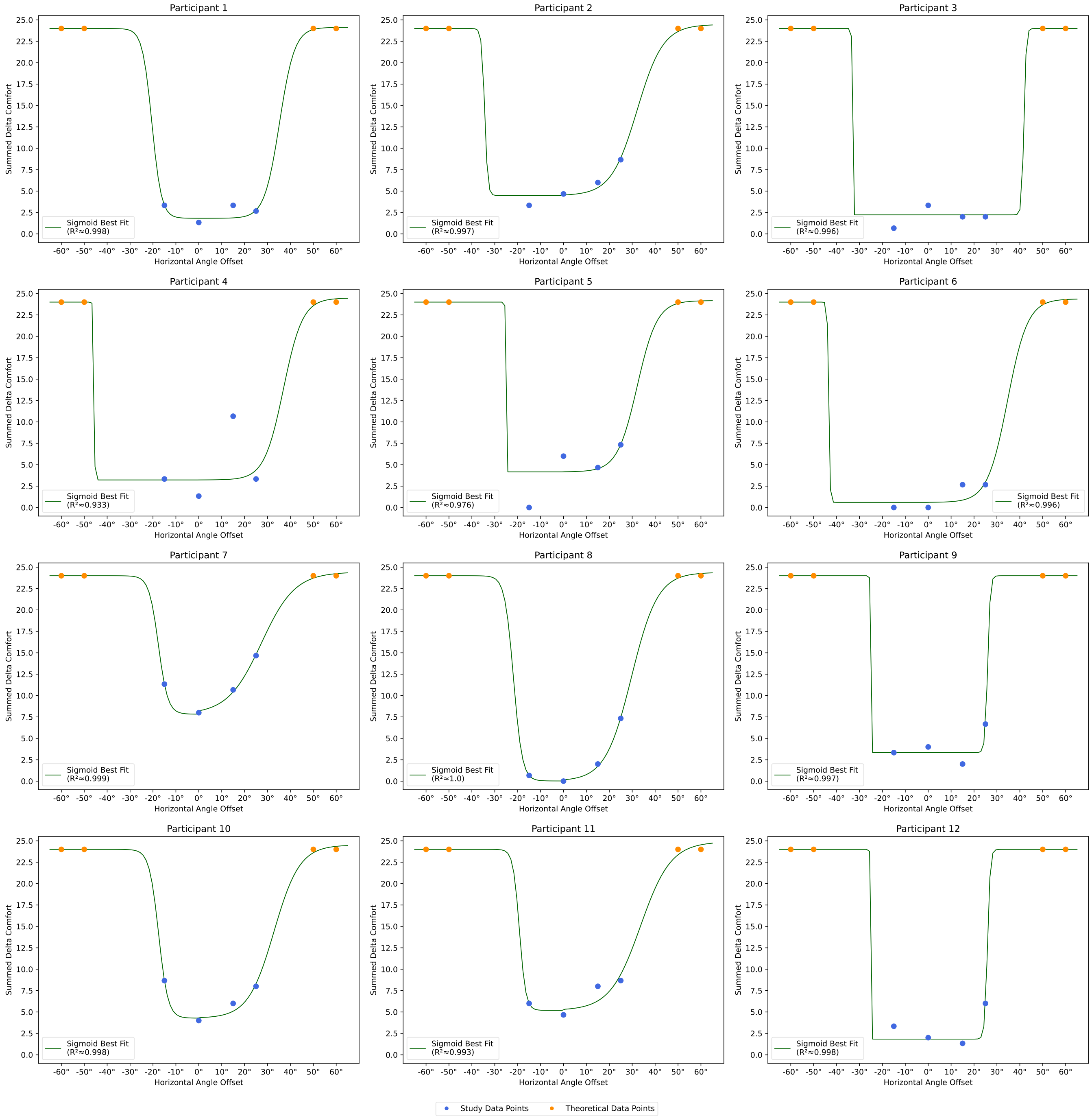


Figure 11: Bidirectional Sigmoid Best Fits for Each Participant in Study 2 (Converted to 5-Point Scale)

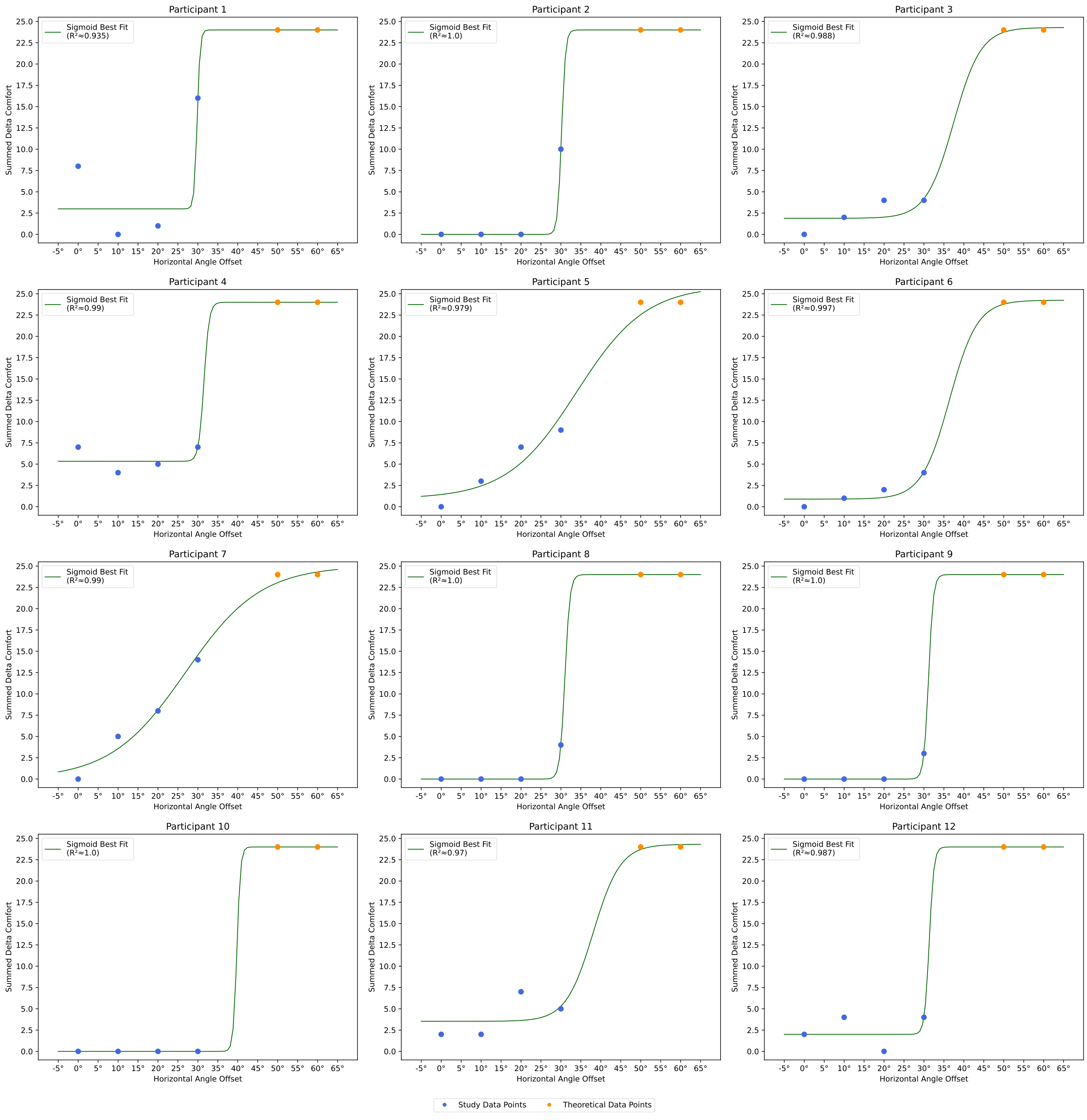


Figure 12: Sigmoid Best Fits for Each Participant in Haynes et al.'s Study[21] (5-Point Scale)