Noname manuscript No.
(will be inserted by the editor)

Personalized Presentation Annotations Using Optical HMDs

Wyko Rijnsburger · Sven Kratz

Received: date / Accepted: date

Abstract It is difficult to adjust the content of traditional slide presentations to the knowledge level, interest and role of individuals. This might force presenters to include content that is irrelevant for part of the audience, which negatively affects the knowledge transfer of the presentation. In this work, we present a prototype that is able to eliminate non-pertinent information from slides by presenting annotations for individual attendees on optical head-mounted displays. We first create guidelines for creating optimal annotations by evaluating several types of annotations alongside different types of slides. Then we evaluate the knowledge acquisition of presentation attendees using the prototype versus traditional presentations. Our results show that annotations with a limited amount of information, such as text up to 5 words, can significantly increase the amount of knowledge gained from attending a group presentation. Additionally, presentations where part of the information is moved to annotations are judged more positively on attributes such as clarity and enjoyment.

Keywords head-mounted device (HMD);smart spaces;presentations;annotations;

Sven Kratz
3174 Porter Drive
Palo Alto, CA, 94304
USA
Tel.: +1-650-842-4800
E-mail: kratz@fxpal.com

Wijko Rijnsburger
University of Amsterdam
1012 WX Amsterdam
Netherlands
Tel.: +31-20-525-9111
utwijko@gmail.com

1 Introduction

Nowadays, presentations are a widely used means for transferring knowledge within a company or in an educational setting. A typical setting for a presentation is an individual standing in the front of a room, directed towards a physically present audience of more than two people [40]. To transfer knowledge in the most efficient way, the presenter adjusts his or her content to the knowledge level and interest of the audience. The visual element of a presentation is most often a set of slides created in PowerPoint [24] or similar presentation software.

However, it is difficult to adjust slides to the knowledge level, interests and roles of specific attendees or groups with differing backgrounds. According to Tufte [35], presentations stand or fall on the quality and relevance of the content. Bartsch and Cobern [1] conclude that non-pertinent information included in slides can decrease the amount of knowledge acquired by the attendees from the presentation.

One way of eliminating non-pertinent information from slides is by presenting personalized content for individual attendees on optical head-mounted-displays (OHMDs).

OHMDs can be used to show information personalized for each presentation attendee by showing it on a display directly in the wearer's field of vision. Various big tech companies are introducing OHMDs, such as Google with Glass [9] and Microsoft with HoloLens [23]. These devices differ in many ways, but their goal is the same: to present information to the user in a hands-free manner and to augment the existing reality with relevant data for the user's current information need.

This approach offers several advantages. First, only relevant content is shown to attendees, which reduces the amount of information they have to process. Secondly, privacy-sensitive information can be shared during group meetings: imagine attending a company meeting where employee bonuses are awarded, but the amount of the individual bonuses is not public information. In this case, the bonus amounts could be shown in an annotation on an OHMD for each employee to view their bonus while keeping it hidden from other attendees. Lastly, showing individualized content on OHMDs allows users to process information at their own speed. During parts of a presentation that users perceive as being irrelevant to their interest, they can review previous information that was relevant to them.

In this work, we discuss and evaluate a prototype that presents individuals with personalized annotations on an optical head-mounted display during group presentations. In a first study, we evaluate several ways of presenting OHMD-based annotations alongside several standard (main) slide designs, resulting in a set of design guidelines for creating annotations for OHMDs. Then, in a second study, we evaluate and compare two types of presentations: a traditional slide-based presentation and a slide-based presentation with individualized, OHMD-based annotations created using the guidelines from the previous study. We compare both presentations by looking at qualitative factors and at the knowledge gain from viewing the presentation.

We present the following core contributions in this work: **(1)** a prototype presentation systems that allows authoring presentations with personalized OHMD-based annotations. **(2)** We present insights from an eye-tracking experiment studying how users experience watching OHMD-based annotations. We were able to determine how much time users spend looking at the display and how this changes based on the content that is displayed. **(3)** We present results from an experiment that compares a standard presentation with one enhanced with OHMD-based annotations and show that personalized annotations can significantly enhance the knowledge gain from presentations, and make the presented content (subjectively) easier to understand and less overwhelming.

The remainder of this work is structured as follows: first, we discuss related work and introduce our prototype. Then, we describe the setup and the results of our two studies. Finally, we summarize our findings and discuss possible directions for future research.

2 Related Work

The following section provides an overview of related work regarding presentations, cognitive overload, head-mounted displays and user modeling.

2.1 Presentations

Savoy et al. [31] indicate that the effectiveness of slide presentations in an educational setting depends on the type of knowledge that needs to be transferred. Student's retention of graphic information (i.e. figures and tables) was better when the lecturer used PowerPoint compared to a lecture without a slide presentation. There was no difference in the retention of alphanumeric information (i.e. text and numbers), but the retention of auditory information (i.e. verbalized) was worse when using PowerPoint. Additionally, even when the information retention was worse, students still preferred lectures with slide presentations to traditional lectures.

Lai et al. [16] have created and evaluated a dual-slide PowerPoint Presentation in an educational classroom environment. They conclude that PowerPoint presentations enhanced with extra annotations on a second projector enhances the students comprehension of the learning material because students could select the annotations that best fit their needs and preferences. Their results indicate that personalized annotations may be even more helpful.

A number of recent works have researched ways of improving presentation creation, quality, structuring, linking and audience reception: Karrer et al. [17] analyzed the authoring process as well as audience reception of canvas-based presentations vs. standard slide-based presentations. Pschetz et al. [26] examined presentation authoring from a storytelling perspective. Lastly, Trinh et al. examined the use of a virtual avatar acting as a co-presenter [34]. The main goal of their system was to increase speaker confidence and boost presentation quality through the help of the virtual co-presenter. *Hyperslides* [6] investigated the process of presenting and used these insights to create a markup language for hyperlinked presentations.

We position our current work within this line of previous research by contributing and investigating a technique to improve (personal) content retention for viewers and suggest design guidelines for applying this technique to small OHMDs.

2.2 Cognitive Overload

Mayer and Moreno [22] indicate that humans can only process limited amounts of visual information at a time. When someone's processing capacity is exceeded, cognitive overload occurs. This decreases the person's ability to understand and retain information. They propose several solutions to reduce cognitive overload in multimedia learning situations, one of which is *'weeding'*. Weeding eliminates extraneous content to make a multimedia narration, which in our case is a presentation, as concise and coherent as possible. This approach proved to be effective in increasing understanding of content in an educational setting [21], which suggests this might also be the case for presentations.

2.3 Head-Mounted Displays

There are two types of head-mounted displays [37]. Virtual reality (VR) devices replace the user's entire vision with computer generated graphics. Through the use of external cameras that provide a feed of the real-world, these devices can become video see-through devices that replace the user's vision with a virtual representation of reality. VR devices were not suitable for this project, because we believe that the conversion from reality to a virtual representation of reality would harm the user's interactivity with the presenter and other attendees.

Optical head-mounted displays (OHMDs) are wearable devices that can enhance a user's vision by overlaying computer generated images on the real world [30]. While wearing such a device, a user is immersed in an augmented reality, in which the real world is supplemented with computer data such as audio or imagery [28]. OMHDs aim to increase the accessibility of information by providing it in such a way that does not require the user to use devices such as smartphones or laptops, which would distract the user's attention from the real world. Additionally, they can be used hands-free. This could mean that the user could take notes while viewing both the information on the shared presentation and the personalized information shown on the OHMD.

Various implementations of OHMDs attach the display onto a traditional glasses frame. These devices are called smartglasses [18]. Smartglasses are a good way to integrate OHMDs into traditional working environments, because they are not as intrusive and noticeable as other types of HMDs and thus are more likely to enjoy social acceptance [3].

Hong [12] indicates that there is a negative sentiment towards OHMDs with attached camera units because of their recording capabilities. People have argued that OHMDs may enable the infringement of privacy due to their ability to record data without this being apparent. However, Hong argues that the effect of OHMDs on privacy is yet to be determined because it is not yet clear how the general public will use these devices. Mann [19] argues that as these devices become more widespread, new protocols and social attitudes toward the capture and sharing of visual information have to be adapted, but that these protocols should not discriminate against users of these devices.

This literature suggests that introducing our prototype into a working environment might result in a negative sentiment to the wearer's of the prototype. However, this sentiment is likely to change in the future with wider adoption of head-mounted devices.

2.4 User Modeling

To determine what information needs to be shown to users through the OHMD, we need to build a model of the user. A user model allows the system to adapt a user's specific needs [15]. User modeling can capture the user's interest in a working domain, which forms the basis for providing personalized information [13].

There are different approaches to modeling the user. Stereotype modeling classifies users into common stereotypes and has the advantage of requiring little data. Adaptive user models adapt to the specifics of a single user based on gathered data for that user. For Study 2, we use a stereotype modeling approach to group presentation attendees into certain roles, such as 'developer' or 'sales manager'.

3 Prototype

We created a prototype implementation capable of presenting presentation attendees with personalized annotations in order to answer the following research questions through controlled user experiments:

1. *How can we format annotations on an OHMD in such a way that it enhances the understanding of a presentation?*
2. *Can personalized annotations during group presentations increase the amount of knowledge gained by attendees from the presentation?*

In this section, we describe the individual components of prototype. First, we discuss our choice of OHMD. Then, we show how we created presentations with annotations. Finally, we discuss how these annotations are send to a real-time database and shown on the OHMD. Figure 1 shows an overview of how an annotation gets from the presentation to the user's device.

Fig. 1: Diagram detailing the different parts of the system along with the associated technologies.

3.1 Hardware

The field of OHMDs is rapidly developing, e.g., as Microsoft's recent announcement of its new OHMD *HoloLens* [23] has shown.

However, for this experiment we chose to use *Glass*, an OHMD developed by Google, for two reasons: it is arguably the most mature OHMD on the market and it runs on the Android platform. This offers several development advantages such as the use of existing libraries and a large knowledge and support base.

3.2 Presentations

We created presentations using a modified version of Reveal.js, a JavaScript presentation framework [8]. This framework allowed us to create presentations in HTML and add extra, optional fields where annotations can be added. An annotation consists of up to 200 characters of text, an image, or both. The style in which the annotation is shown, which will be discussed in the next section, can be set directly in HTML. Multiple annotations can be added to an individual slide by duplicating the slide and creating separate annotations. This way, the presenter can control when the next annotation is shown.

3.3 Real-time Database

To enable real-time synchronization of state changes between the presentation client and the HMD, we created and hosted a real-time database using Firebase [7]. A real-time database instantly informs clients when certain data is updated and thus does not require periodic checking for state changes. When a presenter changes slides, the new annotation, consisting of text, an image or both, is send to the database and stored under the unique key for that presentation using the Firebase JavaScript library. The Glass application for displaying annotations listens for changes in that document, fetches the annotation and automatically displays it on the screen.

3.4 Google Glass Application

We created a Google Glass application using the Glass Development Kit [10]. Synchronized via the real-time database, the application shows the annotation for the current slide and also allows users to scroll through previously shown annotations.

3.4.1 Pairing

We tested two methods of pairing the Glass application with a presentation: using a QR-code and using image recognition.

By placing a QR-code at the start of the presentation, users can scan the code using the built-in camera and pair with the presentation. However, attendees joining

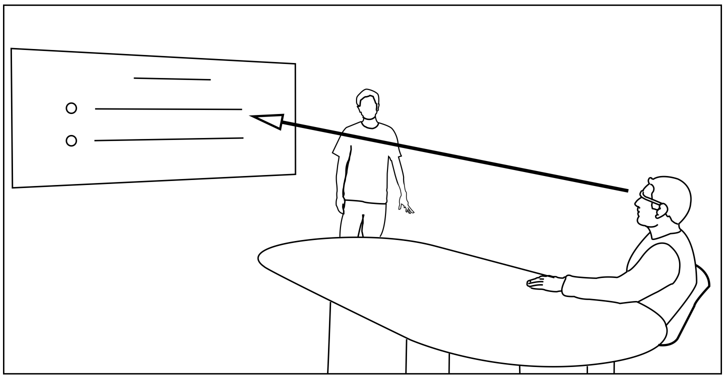
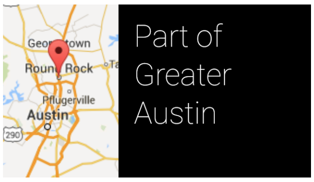
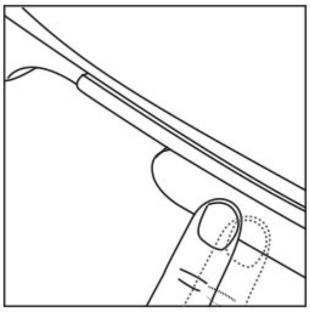


Fig. 2: An attendee pairs with a presentation by scanning the slides with Google Glass's built-in camera.



a: Example of an annotation as displayed on the Glass display.



b: Gesture used to scroll between annotations using the Glass touchpad.

Fig. 3: Example of an annotation and explanation of gesture used to scroll between annotations.

the presentation at a later state are not able to scan the code and thus cannot complete the pairing.

The other approach uses image recognition to recognize the presentation slides. A presentation attendee aims his or her Glass at the presentation and activates the camera, as demonstrated in Figure 2. This was implemented using the Wikitude Augmented Reality SDK [39]. We use screenshots of each slide of a presentation, to extract visual features. These features are then used as targets for image recognition. When an image from a slide is recognized, the unique key associated with that presentation is used by the Glass application to fetch correct annotations from the real-time database.

There are, however, two cases in which this approach does not work correctly. Firstly, slides with little content, such as slides containing only a title, do not have enough unique features to be recognizable. Secondly, slides containing videos or animated images cannot be recognized since their features change over time. These problems could potentially be fixed by putting unique markers in the sides or corners of the presentation.

Several factors influence the reliability of both approaches, such as the distance and the angle at which the user is looking at the slides, the lightning in the room, and the amount of unique features of the displayed slide.

3.4.2 Annotations

After the application has been successfully paired with a presentation, the application shows the annotation for the current slide. Anytime the presenter switches to a new slide, the new annotation is 'inserted' after the previous annotation. This is indicated to the user through an animation where we zoom out on the previous annotation, after which a new annotation appears beside it, on which we then zoom in. When the presenter switches back to a previous slide, a similar animation transfers the user back to the associated annotation.

Users can view annotations from previous slides by sliding back and forth on the Glass touchpad, as shown in Figure 3b. A scrolling bar at the bottom indicates the user's scrolling position within the annotations. When a slide without an annotation is shown, the Glass screen dims. However, users are still able to scroll back to previous annotations. Figure 3a shows an example of an annotation, which is seen by the user on a Google Glass display.

4 Study 1

In a first study, we evaluated several combinations of layout styles for annotations and slides, in order to determine what type of annotation layout design is most suitable for viewers of presentations.

One of the challenges of head-mounted devices is that they generally have a narrow field of vision. The field of vision is the extent of the observable world that is seen at any given moment, which is 150 to 200 degrees for normal humans combining eye and head rotation [36]. While the initial trend was to create OMHDs with a wide display and a wide field of vision, recent devices have opted for devices with a more narrow display, because this increases the pixel density of the display [29]. Google Glass has a display that compares to viewing a normal 24" display from a distance of 2.4 meters. This means that the information must be communicated to the individual in a small part of their vision.

Slides vary widely in their format. Some slides contain images, while others may contain videos, bullet points or just a title. Annotations can also be displayed in a variety of ways. Google Glass offers several ways to present text and images on the display. Table 1 shows the most common slide and annotation styles that we have identified.

In order to present information on an OHMD in the most effective way possible, we conducted a study in which we gathered guidelines for creating visually optimal annotations.

Slide Style	Annotation Style
S1: Title	**A1:** No annotation
S2: Title & Image	**A2:** Text (1-5 words)
S3: Video	**A3:** Text (6-25 words)
S4: Title & Bullet Points	**A4:** Image
S5: Title & Image & Bullet Points	**A5:** Image & Text (side by side)
	A6: Image & Text (1-5 words) (Overlay)
	A7: Image & Text (6-25 words) (Overlay)

Table 1: The most common slide and annotation styles that we have identified.

4.1 Hypotheses and Research Questions

In particular, the study aimed to answer the following questions:

1. *How can annotations on a Head-Mounted Display be presented to optimize the knowledge gained from a slide presentation?*
2. *How long do users look at the annotations and does this differ per annotation format?*

Prior to the study, we formulated the following hypotheses:

– **H1:** Users spend more time looking at an annotation as the amount of content in the annotation increases.
– **H2:** When the distraction level of an annotation increases, the understanding of a presentation decreases.
– **H3:** Annotations with images will be regarded as the most visually pleasant.

4.2 Methods

The experiment tested the effect of two independent variables: annotation style and slide style. To test all possible combinations of the seven annotation styles and the five slide styles, we created a total of 35 presentations.

4.2.1 Choice of Topic

We used 35 different cities in the United States as the topics for the presentations. Cities can be described using standard information available for every city, such as the population, the state and the date of establishment. Additionally, facts about cities should be understandable for any participant regardless of prior knowledge.

The 35 cities were chosen randomly from a list of the 295 most populous cities in the United States [38], and the information about the cities was taken from their respective Wikipedia pages. For the seven video presentations, we used existing videos about cities. For four of the cities, we used the first 30 seconds of the city's "Expedia Travel Guide", e.g., the guide for Dallas, Texas [41]. There was no such guide available for the other three cities, so we used a video made by the city's tourist board.

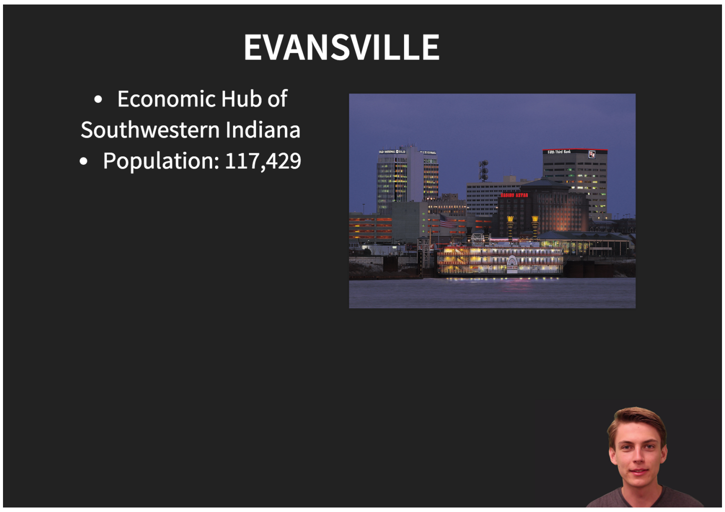


Fig. 4: Presenter inserted into the lower right corner of the presentation.

Since each slide used a standard format and style, we wrote the content of the presentations in JSON, which we then converted into HTML using semantic templates created using the Handlebars.js framework [11].

4.2.2 Presentation Set-up

To allow us to perform gaze tracking, subjects needed to sit in a consistent place in front of a display and an eye tracker. This made it impossible to give a presentation in-person. However, a presentation consisting of just audio and slides, without a visual embodiment, is not representative of a real presentation setting. According to Kelley and Gorham [14], eye contact and physical immediacy account for 20 percent of the overall variance in the recall of information.

Therefore, we chose to record the presentations in front of a green-screen and integrate the recordings into the presentation slides, as shown in Figure 4. This setup was the best approximation for a real-world setup while fitting our constraints. Additionally, this removed any inevitable inconsistencies that would occur when giving the same presentation multiple times to different participants.

The end result is a collection of 35 presentations about cities combining all different annotation layouts with all slide layouts, with a presenter integrated into the slides.

4.2.3 Experiment Design

We used a within-subjects factorial design, with annotation style and slide style as independent factors, which meant that all participants evaluated all 35 presentations. The order of the 35 presentations was randomized for each participant as to compensate for potential carryover effects. The experiment was conducted using an adjusted version of the prototype described in the previous section. Unnecessary features,

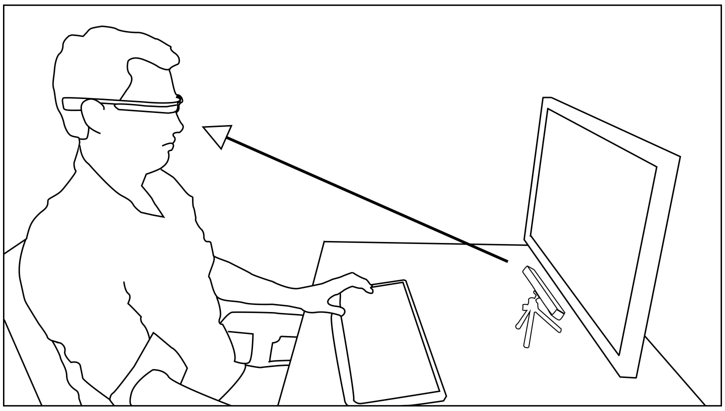


Fig. 5: Study 1 experiment setup and eye-tracking positioning.

such as presentation pairing and annotation scrolling, were removed. After each presentation, we asked the subject to express his or her opinion about the shown presentation and annotation through three statements on a five-point Likert scale, each measuring a single variable. The three statements along with their variables were:

- **Distraction:** *The annotation distracted me from the slides and/or the presenter.*
- **Understanding:** *I followed and understood everything shown in the presentation and the annotation.*
- **Visual Pleasantness:** *The annotation was pleasant to look at.*

The statements for *distraction* and *visual pleasantness* were only applicable when an annotation was shown, thus we added an extra option to the Likert scale for these statements: “No annotation shown”.

Additionally, we tested the participant’s understanding of the content of each presentation with a three-option multiple-choice question. As an example, the multiple choice question for the presentation about Cincinnati was: *“How did Cincinnati rank among US cities by population in 1840?"*. These questions also made sure that the participants remained focused on the presentations and the annotations.

4.2.4 Eye Tracking

During the experiment, we captured the participant’s focus of attention using eye tracking. While attending a presentation using our prototype, an attendee’s focus of attention is divided between the following elements: *Slide Content*, *Presenter*, *Glass Display* and *Miscellaneous*.

For eye-tracking, we initially tried using a mobile eye tracker. However, we ran into radio-frequency interference issues of the eye tracker signal, likely caused by the radios and electronics of Glass. So instead, we used a conventional eye tracker from Eye Tribe [33]. This eye tracker runs at 60 Hz and has an average accuracy of 0.5°. It was placed at the bottom of the display on which the presentations were shown and placed in such a way that the prism of Glass did not block the eye tracker's view of the subject's pupils, as shown in Figure 5.

When a subject sits at the recommended 65 centimeter distance from the eye tracker, the tracking area of the eye tracker is 40 by 30 centimeters. Therefore both the presentation display and the glass screen needed to be in the 30 centimeter vertical range. To maximize the quality of the eye tracking data, subjects were instructed to limit their head movement and keep the glass display between an area, which we marked on the display where the presentations were shown.

We took great care in ensuring the validity of the eye-tracking data we did gather. We discarded the results of several participants after evaluating the eye-tracking data. We used a large (27 inch diagonal) monitor where only the center portion of the screen was occupied by the presentation. If the participant looked at the Glass display, the participant's effective gaze was in a measured area on the top part of the screen. We tested this setup extensively and found that we were able to accurately measure if a participant was looking at either Glass, the presentation or the presenter (or somewhere else entirely).

The presentations were run in *iBrowser*, a web browser that is able to synchronize eye tracking data with browser actions [27]. The eye tracking was calibrated and tested for each subject before the start of the experiment. During the experiment, we monitored the accuracy of the eye tracker and made adjustments to the angle of the eye tracker and the posture of the participant when necessary. Participants viewed all 35 presentations and completed their associated questionnaires in one sitting of about forty minutes.

After all presentations were finished, we asked the participants two open-ended questions, aimed at gathering information to improve the prototype:

- *How did the annotations distract you from the presentation?*
- *How would you improve the annotations?*

In the final part of the experiment, subjects ranked all annotation styles according to their preferences. For each annotation style, two examples were shown. The starting order of the ranking was randomized for each participant to eliminate any influence the order might have on the results.

4.3 Results

Ten participants (two female) from an industrial research lab between the age of 24 and 38 evaluated a total of 350 half-minute presentations. Six of the ten participants had used a Google Glass device before. They did not receive compensation for their participation.

4.3.1 Content Questions

Figure 6 shows the amount of correct answers to the questions about the content of the presentation, grouped by the style in which in the annotation was shown. In total, participants answered 66.3% of the content questions correctly. More correct

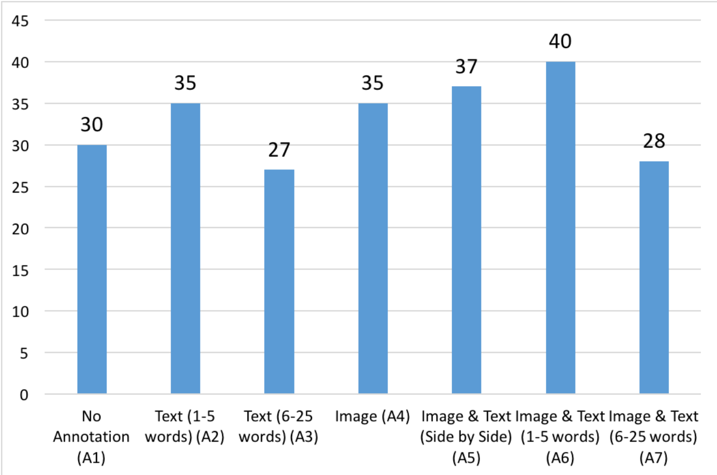


Fig. 6: The amount of correct answers (out of 50) per annotation style.

answers were given when an annotation was shown, with the exception of the two styles displaying more than five words of text (**A3** and **A7**). However, using logistic regression (with $\alpha = 0.05$), we can conclude that only *text annotations with 6-25 words* (**A3**) performed significantly worse and *text up to 5 words with an image background* (**A6**) performed significantly better.

4.3.2 Questionnaire

For the analysis of the presentation questionnaire, we excluded all presentations without an annotation. For the 50 presentations without an annotation, subjects correctly identified that there was no annotation in 49 of the 50 cases.

Although Likert scales nominally provide ordinal data, we treated the data as continuous, as argued by Norman [25], in order to perform an ANOVA.

The blue bars in Figure 8d show that subjects considered the *textual annotation with up to 5 words* (**A2**) to be the least distracting. Using a one-way ANOVA ($\alpha = 0.05$), the style of the annotation significantly determined the perceived distraction ($p < .001$). A Tukey HSD pairwise comparison showed that the perceived distraction for styles *text (6-25 words)* (**A3**) and *image + text (6-25 words)* (**A7**) was significantly higher than for the other styles.

The annotation style also significantly determined the visual pleasantness of the slide ($p < .001$). Styles *'image'* (**A4**) and *'text + image (side by side)'* (**A5**) were regarded to be the most visually pleasant (Figure 8d, gray bars). However, they only significantly differ from styles **A3** and **A7** and not from the other two styles.

The perceived understanding (Figure 8d, orange bars) of the content followed the same trend. It was significantly determined by the annotation style ($p < .001$), but with a lower number of significant differences: **A3** had a significant perceived understanding compared to styles **A2**, **A4** and **A5**.

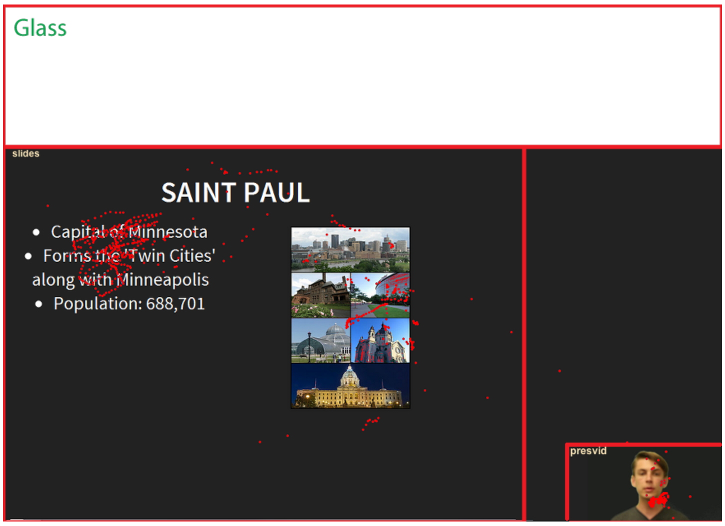


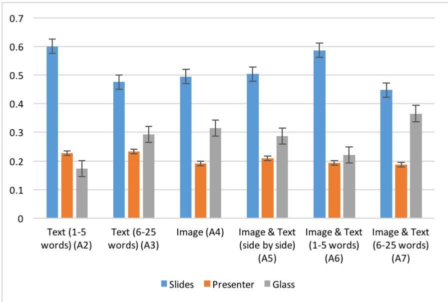
Fig. 7: Example gaze data visualization. The red dots represent user fixations for this session.

The style of the slide does not have a significant impact on the perceived distraction ($p = .0615$) or the visual pleasantness of the annotation ($p = .1972$). However, there is a significant impact on the perceived understanding ($p < .001$). Participants indicated that they understood slides showing a video (**S3**) significantly better than the other styles of slides. This is supported by the amount of correct answers to the content questions for video slides: 77.1% compared to 63.6% for the other types of slides. This could be attributed to the fact that the videos were travel guides which conveyed less information in the same amount of time that the presenter did in the other presentations.

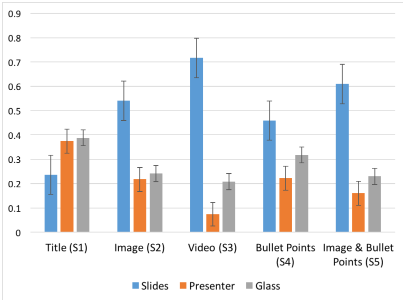
4.3.3 Eye Tracking Analysis

We analyzed the eye tracking data in two ways. First, we visualized the raw gaze data using the iBrowser visualization tool, as shown in Figure 7. This was done to verify the accuracy of the data. For example, we checked if the gaze points matched actual content, such as the presenter, instead of a collection of random points on the screen. Secondly, for each individual presentation we calculated a distribution of the participants gaze by classifying the users fixations as either looking at the *'glass display'*, *'slide'* or *'presenter'* based on the smoothed X and Y coordinates gathered by the eye tracker. The areas used for classification are marked with a red border in Figure 7. Fixations outside these predefined areas, which made up about 15% of the total fixations, were discarded.

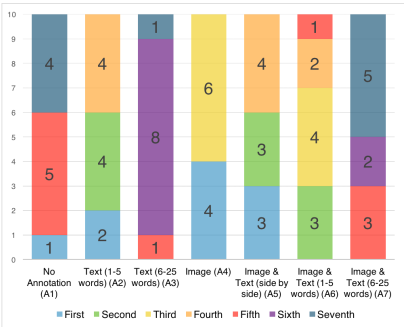
A one-way ANOVA indicated that the annotation style significantly determined the amount of time spent looking at the Glass display ($p < .001$). Figure 8a shows that subjects spent 36.5% of the time looking at *annotations with more than 6 words and an image background* (**A7**) while only 17.2% looking at *textual annotations of*



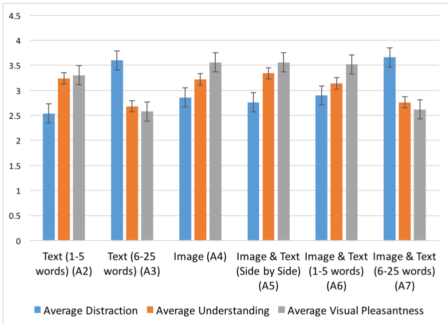
a: Gaze distribution per annotation style.



b: Gaze distribution per slide style.



c: The voting distribution for the ranking of the annotation styles by preference.



d: Average distraction, understanding and visual pleasantness scores based on a five-point Likert scale for each annotation style.

Fig. 8: Results of study 1.

up to 5 words (**A2**). A Tukey HSD pairwise comparison showed that subjects spent significantly less time looking at style **A2**.

The style of the slide significantly impacted the time spent looking at slide, compared to the presenter or the Glass display ($p < .001$). Figure 8b shows that during a video slide (**S5**), participants spent 71.7% of the time looking at the slide, compared to 23.6% during slides containing just a title (**S1**), which is a significant difference. This is also reflected in the percentages for the other parts of the presentation. During slides with style **S1**, participants spent significantly more time (38.9%) looking at annotations than during slides with other designs.

4.4 Ranking

The different rankings of the annotation styles were aggregated according to the Schultz method [32] using the online tool Condorcet Vote [4]. The results are shown in the following list:

1. Image

2. Text + image side by side
3. **Tie:** Text (1-5 words) + image (overlay) / Text (1-5 words)
4. No Annotation
5. Text (6-25 words)
6. Text (6-25 words) + image (overlay)

Figure 8c shows a visualization of the vote distribution.

4.5 Open Questions

Almost all participants indicated that annotations with large amounts of text distracted them. One participant (**P3**) said that *"reading long sentences from a heads-up display required focus and distracted from listening to the audio"*. Another participant (**P6**) thought *"text overlapping on images was hard to read"*.

Subjects indicated that annotations should contain only a small amount of information. **P3** indicated that *"maps and little pieces of information were easier."*. **P8** advised us to *"keep the annotation short and clean"* and to only use key words or pictures.

4.6 Discussion

The eye tracking data shows that the participants spent significantly more time looking at annotations with more content, thus we accept H1. There is also a clear correlation between the perceived understanding and the distraction level. The annotation styles with the highest distraction (**A3 & A7**) have the lowest perceived understanding, which successfully affirms H2. Although the annotation styles *image* and *text + image (side by side)* are regarded to be the most visually pleasant and were ranked as the most preferred styles, there is no significant difference with the textual annotation of 1-5 words, thus we reject H3.

While the results might be influenced by pre-existing knowledge, we can see initial signs that the use of annotations with small amounts of information correlates with an increased amount of correct answers to questions about content conveyed in a presentation. Annotations with 6+ words (**A3 & A7**) that are perceived as distracting resulted in less correct answers.

The annotation style *text (1-5 words) + image* (**A6**) differed significantly in gaze time compared to an image without text, when in fact it has less content. This might mean that subjects saw the image as simply a background and did not process the information from the image.

To summarize our findings, we formulate the following guidelines that a presenter should use when creating OHMD-based annotations:

- Annotations should contain no more than five words.
- If the content of an image is important, do not overlay text on the image
- Image annotations should be simple and convey a single concept.

5 Study 2

To evaluate the effectiveness of our proposed system in combination with the guidelines formed in the previous study, we conducted a further user study. We compared several factors, such as the knowledge gained from a presentation and the perception of a presentation between two groups. One group of subjects used the prototype while viewing a presentation where part of the information had been moved into annotations. A control group viewed a traditional presentation without using the prototype. Both participant groups, however, were exposed to the same amount of information: additional content was displayed on the slides for the control group, whereas participants from the prototype group saw this content on their OHMD.

5.1 Methods

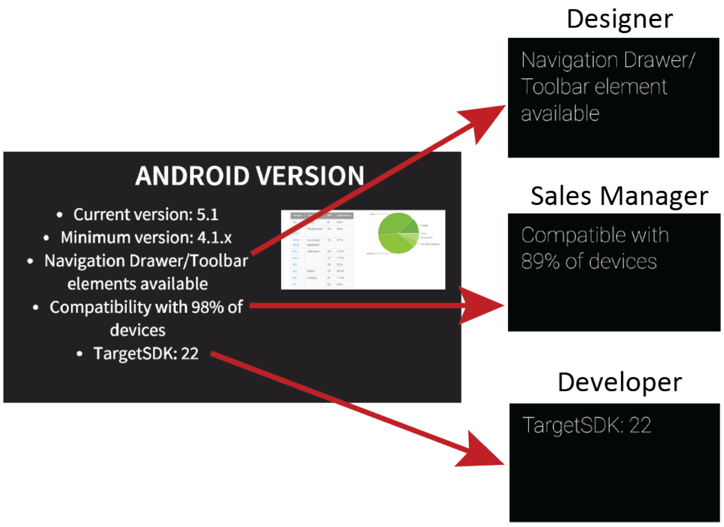
For the purposes of setting up a presentation scenario for the study, we created a fictive presentation. The topic of this presentation was the internal announcement of a new fictional technology product within a fictional company. We used the same setup as in the first study, with participants sitting in front of a display looking at presentation slides with the presenter inserted into the bottom corner. Eye tracking was not needed for this study, thus we ran the presentations in a standard browser (Google Chrome). We chose a fictional subject in order to test the knowledge gained from watching a presentation without the possibility of preexisting knowledge influencing the results. This removed the necessity of a test controlling for preexisting knowledge, as all information in the fictive presentation is new to all participants.

We first created a traditional presentation, containing information that might only be relevant to specific attendees, such as designers, developers or sales managers. For the condition with annotations, we moved part of the information off the original slides to annotations for specific roles, as shown in Figure 9a. This resulted in main slides with less information, as shown by the example slide in Figure 9b.

Participants were randomly assigned to either the control group (no annotations) or the prototype group (annotations shown). Participants were instructed to imagine they were attending a presentation as *Scotty*, an Android developer at the fictional company *InstaCorp*. Participants of the prototype group were then then shown an instructional text and an animation explaining how to scroll through annotations. Afterwards, the presentation was shown.

Directly after the presentation, each subject completed a questionnaire that tested their knowledge of the conveyed information. The questionnaire consisted of 20 four-option multiple choice questions. Afterwards, we asked participants to judge the presentation on the following five aspects on a five-point Likert scale:

- **Understanding:** *I was able to follow and understand the presentation.*
- **Clarity:** *The presentation was coherent and clear.*
- **Overwhelmingness:** *I was overwhelmed by the amount of information in the presentation.*
- **Enjoyment:** *I enjoyed the presentation.*
- **Visual pleasantness:** *The presentation was pleasant to look at.*



a: A traditional slide contains information that is only of interest to specific attendees. This information can be moved to annotations that are only visible to these specific attendees.

b: This results in a slide with less information, where the information is of interest to all attendees.

Fig. 9: Process of converting a traditional slide into a slide where part of the information is moved into annotations for various roles.

Wheres all participant groups The prototype group completed an additional questionnaire measuring the usability of the prototype. We chose to use the System Usability Scale (SUS) [2], which we slightly modified. First, we replaced the word "system" with the word "application" and we indicated that it referred to the application showing the annotations on the Google Glass. Then, we tweaked the wording of questions 1 and 3. At the end of the study, we asked the participants what they liked and what they disliked about the prototype.

5.2 Hypotheses

We formulated the following hypotheses prior to the study:

– **H4:** The prototype group will gain more knowledge from the presentation than the control group.

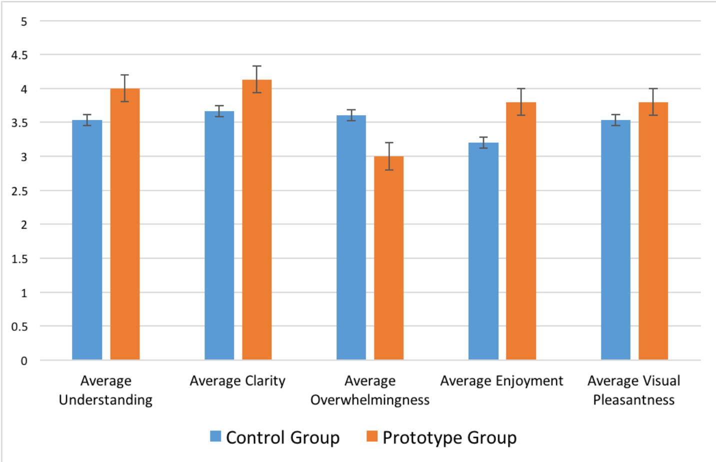


Fig. 10: Chart showing the difference in opinions stated in the Likert scale questions between the traditional presentation and the prototype presentation where information has been partially moved to the annotations.

– **H5:** The presentation using the prototype will be judged more positively by the following criteria: *understanding*, *clarity*, *overwhelmingness*, *enjoyment* and *visual pleasantness*.

5.3 Results

30 subjects (9 female) from an industrial research lab between the age of 17 and 52 participated in the study, divided evenly over the control group and the prototype group. 6 of the 15 participants in the prototype group had previously used Google Glass. One participant had to be reassigned from the prototype group to the control group due to the Google Glass not fitting well over his glasses.

The results for the SUS and the open questions had to be discarded for one participant who interpreted the questions as being about the fictional product instead of the Google Glass application. This left us with 14 valid SUS's.

5.3.1 Content Questions

Nominal logistic regression indicated that the group using the prototype answered significantly more questions correctly than the control group ($p = .007$). The prototype group answered 72% percent of the questions correctly in contrast to the 61% for the control group.

5.3.2 Presentation Questions

The presentation for the prototype group was judged more positively on all measured factors, as shown in Figure 10, although not necessarily significantly. Perceived understanding of the prototype presentation was higher with a small significant effect ($p = .040$), as was the clarity with a stronger significant effect ($p = .012$). The prototype group felt less overwhelmed by the presentation with a small significant effect ($p = .029$). While the prototype group judged the enjoyment and the visual pleasantness of the presentation more positively, there was no significant difference ($p = .115$ and $p = .178$).

5.3.3 SUS

A SUS score was calculated by aggregating all individual scales. Our prototype system got a score of 65.36, which indicates average usability.

5.3.4 Open Questions

Subjects were generally positive about the prototype. **P1** mentioned that it eliminated the need to "*search through slides to find the snippet of information that pertains to me*". **P10** liked the sliding functionality: "*I liked that I can scroll back and see slides with past information*". **P12** felt that it made attending the presentation a "*more interactive*" experience.

However, participants were also critical of several aspects. **P12** indicated that the screen of the Google Glass was hard to see and that looking at it was uncomfortable. Two participants (**P1 & P25**) mentioned that they were distracted from the presentation by the annotations. **P22** indicated that it was confusing that the display was empty during slides with no annotation: "*I didn't know if it was not working or if there was no content*". Finally, the scrolling functionality was not clear to some users. **P26** mentioned that it "*wasn't clear when I could scroll for more information*" and **P27** thought it was "*hard to use*".

5.4 Discussion

Our results indicate that personalized annotations can significantly enhance the knowledge gained from attending a group presentation. Additionally, moving information from the presentation to annotations makes the presentation clearer, easier to understand and less overwhelming. However, the presentation is not significantly more visually pleasant, nor is the enjoyment of the presentation significantly higher. Thus, we can fully accept H4 and partially accept H5.

While these results are encouraging, we think they can be further improved by adjusting the prototype. Comments from the participants and the average usability score both indicate that the prototype needs to be made more usable, specifically regarding the following points.

First, users need to be made more aware when they can scroll through annotations. This could be done by permanently showing the scrolling bar whenever scrolling is possible. Secondly, we need to add extra visual cues to the application to make users more aware of the annotations. While the current prototype uses an animation to indicate that an annotation is available, this is not sufficient for some users. Finally, empty annotations cause confusion, with some users wondering if the system stopped working. We should add a visual cue indicating that the system is still running without distracting from the presentation, which is the point of an empty annotation.

The discomfort that occurs when looking at the display remains a problem, even after limiting the amount of information shown in the annotation based on the results of our first study. This seems to be a specific limitation of the design and positioning of the Google Glass display.

6 Conclusion and Future Work

In this work we have presented a prototype that allows the removal of non-pertinent information from slides by showing individual attendees personalized annotations on optical head-mounted displays. In a first study, we evaluated several ways of formatting annotations. We concluded that annotations on small displays such as Google Glass should contain a limited amount of information, such as a single image representing a single concept or text up to five words, to prevent distraction from the main presentation.

In a second study, we used these guidelines to compare two presentations: a traditional presentation and a presentation where part of the presentation content was moved to annotations shown on OHMDs. Our results indicate that using these guidelines, providing individuals with personalized annotations on OHMDs, significantly increases the knowledge gained from group presentations. Additionally, our results indicate that it positively influences an attendee's opinion about a presentation regarding factors such as enjoyment and clarity.

We think that these benefits will only increase as OHMDs are improved and people become familiar with using them. In this study, we used Google Glass, a device with a narrow display, to show the annotations. Novel devices with a wider display offer ways to present the annotations in ways that are less distracting and more comfortable to the user. As the quality of the devices improves, we think that the benefits of personalized information will only increase. One way that a wider display could be used is by showing the annotations in an augmented reality (AR).

To evaluate the feasibility of this approach, we built an AR version of our prototype using the Wikitude SDK [39]. Slides from the presentation are recognized and image and textual annotations are placed alongside or over the recognized slides, as shown in Figure 11. Annotations are formatted and added to the slides in the same way as the normal version of the application.

While we argue that Google Glass's small display is not suitable for AR, we think that future HMDs with a wide display that approximates our human vision, such as Microsoft's Hololens, can utilize this approach to present annotations in an immersive way that limits distraction from the main presentation.

In the future, we are interested in studying the effects of our prototype on the length of presentations. Certain information could be communicated solely through annotations, which should make presentations shorter and more concise. Additionally, we want to explore the possibility of including automatically generated annotations, using the content on the slide and a profile of the user to query internal knowledge

Fig. 11: Example AR annotation as displayed on the Glass display.

bases within a company or open structured information databases such as DBPedia [5].

A limitation of this research is that our prototype was tested with a tech-savvy demographic, as indicated by the relatively high number of people that had previously used Google Glass. Other demographics might react differently to the annotations and OHMDs in general. In the future, we want to test the prototype in real-world situations with a more varied demographic.

7 Acknowledgements

We would like to thank our study participants, *anonymized* and *anonymized* for their supervision and help with this research.

References

1. Bartsch, R.A., Cobern, K.M.: Effectiveness of powerpoint presentations in lectures. Computers & Education **41**(1), 77–86 (2003)
2. Brooke, J.: Sus-a quick and dirty usability scale. Usability evaluation in industry **189**(194), 4–7 (1996)
3. Cakmakci, O., Rolland, J.: Head-worn displays: a review. Journal of Display Technology **2**(3), 199–216 (2006)
4. Condorcet.Vote: Condorcet Vote. Retrieved on 9/16/2015 from `http://www.condorcet.vote/` (2015)
5. DBpedia: DBpedia. Retrieved on 9/16/2015 from `http://wiki.dbpedia.org/` (2015)
6. Edge, D., Savage, J., Yatani, K.: Hyperslides: dynamic presentation prototyping. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 671–680. ACM (2013)
7. Firebase: Firebase. Retrieved on 9/16/2015 from `https://www.firebase.com/` (2015)
8. GitHub User Hakimel: reveal.js. Retrieved on 9/16/2015 from `https://github.com/hakimel/reveal.js/` (2015)
9. Google: Google Developers: Glass. Retrieved on 9/16/2015 from `https://developers.google.com/glass/` (2015)

10. Google: Google Glass Development Kit. Retrieved on 9/16/2015 from **https://developers.google.com/glass/develop/gdk/** (2015)
11. handlebars.js: Handlebars.js Semantic Web Templates. Retreived on 9/16/2015 from **http://handlebarsjs.com/** (2015)
12. Hong, J.: Considering privacy issues in the context of google glass. Commun. ACM **56**(11), 10–11 (2013)
13. Jiang, X., Tan, A.H.: Learning and inferencing in user ontology for personalized semantic web search. Information sciences **179**(16), 2794–2808 (2009)
14. Kelley, D.H., Gorham, J.: Effects of immediacy on recall of information. Communication Education **37**(3), 198–207 (1988)
15. Kobsa, A.: Generic user modeling systems. In: The adaptive web, pp. 136–154. Springer (2007)
16. Lai, Y.S., Tsai, H.H., Yu, P.T.: Integrating annotations into a dual-slide powerpoint presentation for classroom learning. Educational Technology & Society **14**(2), 43–57 (2011)
17. Lichtschlag, L., Hess, T., Karrer, T., Borchers, J.: Fly: studying recall, macrostructure understanding, and user experience of canvas presentations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1307–1310. ACM (2012)
18. Mann, S.: Wearable computing: A first step toward personal imaging. Computer **30**(2), 25–32 (1997)
19. Mann, S.: Wearable technology as a human right. MIT Technology Review (2015)
20. Mann, S., Nolan, J., Wellman, B.: Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments. Surveillance & Society **1**(3), 331–355 (2002)
21. Mayer, R.E., Heiser, J., Lonn, S.: Cognitive constraints on multimedia learning: When presenting more material results in less understanding. Journal of educational psychology **93**(1), 187 (2001)
22. Mayer, R.E., Moreno, R.: Nine ways to reduce cognitive load in multimedia learning. Educational psychologist **38**(1), 43–52 (2003)
23. Microsoft: Microsoft Hololens. Retrieved on 9/16/2015 from **http://www.microsoft.com/microsoft-hololens/en-us** (2015)
24. Microsoft: Powerpoint. Retrieved on 9/16/2015 from **http://products.office.com/en-us/powerpoint** (2015)
25. Norman, G.: Likert scales, levels of measurement and the "laws" of statistics. Advances in health sciences education **15**(5), 625–632 (2010)
26. Pschetz, L., Yatani, K., Edge, D.: Turningpoint: narrative-driven presentation planning. In: Proceedings of the 32nd annual ACM conference on Human factors in computing systems, pp. 1591–1594. ACM (2014)
27. Qvarfordt, P., Golovchinsky, G., Dunnigan, T., Agapie, E.: Looking ahead: Query preview in exploratory search. In: Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, pp. 243–252. ACM (2013)
28. Rolland, J., Cakmakci, O.: Head-worn displays: the future through new eyes. Opt. Photon. News **20**(4), 20–27 (2009)
29. Rolland, J., Hua, H.: Head-mounted display systems. Encyclopedia of optical engineering pp. 1–13 (2005)
30. Rolland, J.P., Fuchs, H.: Optical versus video see-through head-mounted displays in medical visualization. Presence: Teleoperators and Virtual Environments **9**(3), 287–309 (2000)
31. Savoy, A., Proctor, R.W., Salvendy, G.: Information retention from powerpoint and traditional lectures. Computers & Education **52**(4), 858–867 (2009)
32. Schulze, M.: A new monotonic and clone-independent single-winner election method. Voting matters **17**(1), 9–19 (2003)
33. Tribe, T.E.: The Eye Tribe Tracker. Retrieved on 9/16/2015 from **https://theeyetribe.com/products/** (2015)
34. Trinh, H., Ring, L., Bickmore, T.: Dynamicduo: Co-presenting with virtual agents. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, pp. 1739–1748. ACM, New York, NY, USA (2015). DOI 10.1145/2702123.2702584. URL **http://doi.acm.org/10.1145/2702123.2702584**
35. Tufte, E.R.: The cognitive style of PowerPoint, vol. 2006. Graphics Press Cheshire, CT (2003)
36. U.S. Department of Defense: Design Criteria For Military Systems, Equipment, And Facilities, MIL-STD-1472F. Department of Defense Design Criteria Standard, Human Engineering (1998)
37. Van Krevelen, D., Poelman, R.: A survey of augmented reality technologies, applications and limitations. International Journal of Virtual Reality **9**(2), 1 (2010)
38. Wikipedia: List of United States Cities by Population. Retrieved on 9/16/2015 from **http://en.wikipedia.org/wiki/List_of_United_States_cities_by_population** (2015)

39. Wikitude: Wikitude Augmented Reality SDK. Retrieved on 9/16/2015 from http://www.wikitude.com/products/wikitude-sdk/ (2015)
40. Yates, J., Orlikowski, W.: The powerpoint presentation and its corollaries: how genres shape communicative action in organizations. Communicative practices in workplaces and the professions: Cultural perspectives on the regulation of discourse and organizations pp. 67–91 (2007)
41. Youtube Account of Expedia.com: Expedia Travel Guide for Dallas. Retrieved on 9/16/2015 from https://www.youtube.com/watch?v=Kt7lpqdbNgA (2015)