Measuring the Difficulty Watching Video With Hemianopia and an Initial Test of a Rehabilitation Approach

Francisco M. Costela¹², Daniel R. Saunders¹², Sidika Kajtezovic¹, Dylan J. Rose¹, and Russell L. Woods¹²

¹ Schepens Eye Research Institute, Massachusetts Eye and Ear, Boston, MA, USA
² Department of Ophthalmology, Harvard Medical School, Boston, MA, USA

Correspondence: Francisco M. Costela, Schepens Eye Research Institute, 20 Staniford St, Boston, MA 02114, USA. e-mail: Francisco_costela@meei.harvard.edu
Received: 22 August 2017
Accepted: 4 June 2018
Published: 16 August 2018

Keywords: hemianopia; rehabilitation; stroke; media; video

Citation: Costela FM, Saunders DR, Kajtezovic S, Rose DJ, Woods RL. Measuring the difficulty watching video with hemianopia and an initial test of a rehabilitation approach. Trans Vis Sci Tech. 2018;7(4):13, https://doi.org/10.1167/tvst.7.4.13
Copyright 2018 The Authors

Purpose: If you cannot follow the story when watching a video, then the viewing experience is degraded. We measured the difficulty of following the story, defined as the ability to acquire visual information, which is experienced by people with homonymous hemianopia (HH). Further, we proposed and tested a novel rehabilitation aid.

Methods: Participants watched 30-second directed video clips. Following each video clip, subjects described the visual content of the clip. An objective score of information acquisition (IA) was derived by comparing each new response to a control database of descriptions of the same clip using natural language processing. Study 1 compared 60 participants with normal vision (NV) to 24 participants with HH to test the hypothesis that participants with HH would score lower than NV participants, consistent with reports from people with HH that describe difficulties in video watching. In the second study, 21 participants with HH viewed clips with or without a superimposed dynamic cue that we called a content guide. We hypothesized that IA scores would increase using this content guide.

Results: The HH group had a significantly lower IA score, with an average of 2.8, compared with 4.3 shared words of the NV group (mixed-effects regression, \( P = 0.001 \)). Presence of the content guide significantly increased the IA score by 0.5 shared words \( (P = 0.03) \).

Conclusions: Participants with HH had more difficulty acquiring information from a video, which was objectively demonstrated (reduced IA score). The content guide improved information acquisition, but not to the level of people with NV.

Translational Relevance: The value as a possible rehabilitation aid of the content guide warrants further study that involves an extended period of content-guide use and a randomized controlled trial.

Introduction

Homonymous hemianopia (HH) is a cortical blindness that eliminates vision on one side in both eyes. The prevalence of homonymous visual field defects has been reported as 0.8% in the general population over 49 years of age.¹ As there were about 100 million people aged 50 years and older in the United States in 2010,² this suggests that there are over 800,000 people in the United States with a homonymous visual field defect. In an acute stroke care setting, the proportion of visual field deficits can be as high as 49%,³ though many visual field deficits resolve in 1 to 6 months.⁴ Of stroke survivors, 8%¹ to 16%⁴ have long-term hemianopia, most of which are partial or complete HH. As there were 6.2 million stroke survivors in the United States in 2010,⁶ this suggests that between half and 1 million stroke survivors in the United States have a permanent hemianopic visual field deficit.

The vision impairment from HH often reduces the quality of life, impacting daily life activities, such as reading, writing, and driving. In addition, individuals with HH may experience difficulties watching videos, which can limit their ability to engage in social activities and stay connected with family and friends. The purpose of this study was to measure the difficulty of following stories in video content and to test the hypothesis that participants with HH would score lower than NV participants, consistent with reports from people with HH that describe difficulties in video watching.
as noticing other persons,7 reading,8,9 walking,10–13 driving,8,15,14–17 and watching television (TV) and other forms of video.18,19 A survey of 46 people with hemianopia and quadranopia, which included one question about TV, found that 30% reported some difficulty with watching TV.19 In a more extensive study of viewing habits,18 the experience of 91 people with hemianopia with various forms of video access (e.g., TV, theater) was compared with that of 192 people with normal vision (NV). Participants with hemianopia were more likely to report difficulty watching TV, movies on a computer, and movies at the cinema, and were less likely to attend the cinema.

With age, the risk of hemianopia increases, as does time watching TV; approximately 46 hours a week on average for people 65 years and older in the United States.20 Despite being more likely to report difficulty watching TV, people with hemianopia watch similar hours of television per day compared with people with NV.18 Many strategies have been suggested for the rehabilitation of people with HH, including compensation (e.g., scanning training),10,21,22 vision restitution (or restoration),23–26 and substitution (e.g., prisms).27–30 Clear evidence of benefit for the patients by randomized studies could be presented only for the compensation method, whereas no benefit could be demonstrated for the restitution approach.25,26,28

We are not aware of any reports of rehabilitation aids to assist TV watching, even though this is a commonly reported difficulty.18,19

Here, we report two studies. In the first study, we objectively measured the information acquisition (IA)31,32 of participants with HH (N = 24) and with NV (N = 60). As used here, IA measures the ability to follow the story (comprehension), which is a primary requirement of watching videos, even when done for pleasure. We hypothesized that participants with HH would score lower than participants with NV, based on the difficulties that they report while watching video.18,19

The second study was an initial test of a novel rehabilitation aid, which was termed “content guide.” The content guide is proposed as an assistive technology, not as a treatment. In other words, it may help when available, but would not be expected to alter task performance without its presence even after extended use of the guide by people with HH. We tested whether there was an improvement in the ability to follow the story with this content guide present, as measured by the IA score.

Materials and Methods

Information Acquisition (IA) Method

IA is an objective measure of the ability to perceive and understand a sensory stimulus, using descriptions of the stimulus made by the observer. In the case of video, IA measures the ability to follow or understand the story. We restricted responses to descriptions of the visual content, even though audio content was available. We have found that, with careful instructions, responses can be restricted to the visual content,32 with no difference in IA when the audio content is not available. Participants viewed 30-second video clips wearing their habitual optical correction. An experimenter gave the instructions and was in the room during data collection, and the MATLAB (MathWorks, Natick, MA) program automatically displayed the prompts after viewing each clip, asking the participant to provide verbal responses to the open-ended queries: “Describe this movie clip in a few sentences, as if to someone who has not seen it,” and then, “List several additional visual details that you might not mention in describing the clip to someone who has not seen it.” Participants were instructed to report, without time constraints, on the visual aspects of the clip only. The spoken responses to each prompt were recorded using a headset microphone and later transcribed.

Video Clips

There were 200 video clips, chosen to represent a range of genres and types of depicted activities. The genres included nature documentaries (e.g., BBC’s Deep Blue, The March of the Penguins), cartoons (e.g., Shrek, Mulan), and dramas (e.g., Shakespeare in Love, Pay it Forward). The clips were 30 seconds in duration and were selected from parts of the films that had relatively few scene cuts, which was reflected in the average number of cuts per minute in our clips being nine, as compared with approximately 12 per minute in contemporary films.32,33 Videos with very fast action are expected to be more challenging for gaze tracking and for enhancement methods that make use of the center of interest. The clips included conversations, indoor and outdoor scenes, action sequences, and wordless scenes where the relevant content was primarily the facial expressions and body language of one or more actors. Participants viewed the clips on a 27-in display (aspect ratio 16:9; Apple 27-in iMac or a Barco F50 projector on an AeroView 70 rear projection screen from Stewart FilmScreen.
Corporation, Torrance, CA) from a 100-cm distance, so the videos were 33° of visual angle wide. The clips were displayed by a MATLAB program using the Psychophysics Toolbox34 and Video Toolbox.35 Participants with HH viewed one set of 20 video clips without the content guide and, if they participated in Study 2, they watched a different set of 20 clips with the content guide. The data from the set of 20 clips without-guide was used in the analyses for both studies. As described previously,31 each participant with NV viewed a different set of 40 clips drawn from our set of 200 video clips that included the 40 video clips viewed by the participants with HH. By doing this, each clip was watched by at least 12 of 60 participants with NV, providing gaze data to obtain the center of interest (see section “Center of Interest [COI] Determination” below) and descriptions of the clip (responses) for the control (“crowd”) response database to which the new response was compared (see section “scoring of description of the video clips” below).31 The analysis in Study 1 for the NV group included all 40 clips watched by each participant. A further 99 participants with NV (crowdsourcing group) viewed the 200 video clips online providing 20 additional responses per video clip.31 for a total of at least 32 responses per video clip in the crowd response database.

Processing of Audio Files
The spoken responses to each prompt were recorded using a headset microphone and later transcribed using MacSpeech Scribe Pro (Nuance Communications, Burlington, MA) to produce a preliminary transcription. Afterward, a group of online participants verified and corrected the preliminary transcript via Amazon Mechanical Turk, a crowdsourcing Internet marketplace enabling individuals and businesses to coordinate the use of human intelligence to perform tasks that computers are currently unable to do.

Scoring of Descriptions of the Video Clips
These natural language, open-ended responses were automatically scored for their relevant content31,32 using a “wisdom of the crowd” approach (i.e., collective opinion of a group of individuals rather than that of a single expert)37 to determine the IA score. We processed the text of responses with the Text to Matrix Generator toolbox for MATLAB (as described31,32). Each response was compared with each of the responses to the same video clip in a database of responses from 159 participants with NV (including responses from both the crowdsourcing participants and the 60 NV participants in Study 1). There were at least 32 responses per video clip in the response database.31 The number of words that two responses shared (after removing stopwords and disregarding repeated instances of the word in either response) produced a shared-word count for each pair of responses. The IA score for each video clip for each study participant was the average of the shared-word counts from the paired comparisons with each of the responses from the response database for the same clip. For participants within the NV group, we removed their own response to a given clip from the response database when calculating the IA scores (“leave one out” approach).

Study 1: Effect of Hemianopia on Information Acquisition From Video
To investigate the impact of HH on IA, we compared the IA scores of 24 participants with HH to those of 60 participants with NV.

Participants
Participants were recruited from the community in and around Boston, Massachusetts, using contact lists, physician referrals, and former study participant lists. As shown in Table 1, there were 24 participants with HH, 13 had their blind hemifield on the left side, five were female, and the median age was 61 years (range, 19–81 years). For all participants with HH, there was no projection of sight of more than 10° into the blind hemifield within the central 30° (Goldmann perimeter). Six participants showed evidence of hemispatial neglect (bells38 and line bisection39 tests) and a further three had a previous history of neglect. All six participants with measured neglect showed mild neglect. The cause of the hemianopia was due to a stroke not during surgery for 15 participants (62%; 5 ischemic, 2 hemorrhagic, and 8 unknown), a stroke during surgery for two participants (8%), traumatic brain injury for four participants (17%), and a brain tumor for three participants (13%). The cause of HH in this sample was similar to previous studies.1,4,40 One of the participants with hemianopia also had glaucoma, but there was no absolute scotoma associated with glaucoma (also participated in Study 2). All 24 participants with HH were able to hold a conversation and name objects, and there was no evidence of expressive aphasia, agrammatism, anoma, or articulation difficulties.

Sixty people with no ocular conditions in self-reported ophthalmic history, no visual field defects (Goldmann perimeter), normal appearance of the
retina (Nidek MP-1; Nidek Technologies, Vigonza, Italy), and binocular visual acuity better than 20/35 constituted the NV group (Table 1). The median age of this group was 66 (range, 23–87) years and 30 were male. There were no known neurologic disorders or evidence of any speech or memory problems among the NV group, apart from one participant who reported having had a stroke 9 years before. Her speech was clear and fluent, albeit a little slower than average. Analyses were conducted with and without this participant, and the interpretation of the outcomes was the same, so we report analyses with this participant included.

Visual acuity for all participants was measured while wearing their habitual optical correction at a distance of 6 m. Though the video viewing distance was 1 m, it is unlikely that the focal difference would make a substantive difference. A reduction of visual acuity to 20/50 did not reduce IA scores in a group of 20 NV subjects.41

All participants, NV and HH, had a Montreal Cognitive Assessment score above 20, indicating no evidence of substantial cognitive deficits that would...
interfere with information processing or language production. Compared to the NV group, the HH group had a higher proportion of males ($\chi^2, P = 0.014$), a lower maximum education (Mann-Whitney $U$ test, $z = 3.7, P < 0.001$), and there were no differences in age (Kolmogorov-Smirnov, $P = 0.35$), race ($\chi^2, P = 0.17$), or visual acuity (Kolmogorov-Smirnov, $P = 0.45$). For MoCA scores there was a tendency for a difference between the two groups (shape of distribution: Kolmogorov-Smirnov $P = 0.09$), but the medians were not different (Mann-Whitney $U$ test, $z = 1.58, P = 0.12$). Informed consent was obtained from each participant prior to data collection. Participants were shown the video clips wearing habitual, not necessarily optimal, optical correction.

Gaze was tracked while watching the video clips. Gaze data will be reported separately.

Study 2: Effect of the Content Guide on Information Acquisition From Video by People With Hemianopia

People with hemianopia miss, completely or partially, information on one side of the object to which they are attending. So, if an object of interest appears on that blind side, they are not aware of its presence unless they look in that direction. To reduce the risk of missing important information (e.g., an object with which they might collide), people with HH often use compensatory scanning, eye movements toward their blind hemifield, to provide information from that side. This strategy is taught in rehabilitation programs and can provide benefit to people with HH during certain tasks.10,21,22,29,43,44 Some people with HH develop this compensatory strategy without training, and many do it in an apparently automatic manner.

However, these scans may result in less time attending to the most informative areas in videos found with most commercial films and TV,12,45 which could lower their ability to understand what the video is about and therefore decrease their information acquisition score. The content guide highlights the objects of interest, which are highly informative as determined by the gaze patterns of NV viewers. The guide eliminates the need to make compensatory scanning eye movements toward the blind side (which are made by people with HH to check whether there are objects of interest), unless participants are trying to find the content guide, which might have drifted into the blind side.

Figure 1. The content guide dynamically directs attention to areas that were fixated on by the majority of healthy sighted viewers ($N = 15$ in this particular clip). (A) Kernel density estimate of the gaze points for this particular frame. (B) Illustration of the content guide as it appeared in the frame. Background image from ‘Julie & Julia.’

Center of Interest (COI) Determination

To determine the center of interest (COI) in each video clip, we tracked the gaze of viewers with normal vision using an EyeLink 1000 system (SR Research Ltd., Mississauga, Ontario, Canada). At least 12 of 60 participants with NV watched each video clip. We removed saccades and blinks from the data and defined the smoothed median gaze location for each video frame as the COI. The median gaze location was calculated from all the gazes for a given frame. Then a smooth function was run through the median gaze location to reduce visual jitter across frames. A kernel density estimate$^{46,47}$ of the fixation points for one frame is shown in Figure 1A.

Content Guide

For this early-stage study, we presented a thin yellow ring centered over the COI (Fig. 1B) in the first or second of two blocks (20 clips per block, with clips and blocks randomly and evenly distributed). Participants were told that the highlighted region in the content-guide condition would identify the objects of most interest, but were not told to look there. An example video clip with the content guide is included.
in the supplemental material. The yellow ring was an arbitrary choice and may not be the best method of presenting the content guide. IA scores were compared between the two viewing conditions (presence or absence of the content guide).

Participants
All participants with HH from Study 1 also participated in Study 2, except for three participants that could not finish the study due to personal reasons, which totaled 21 participants. Ten participants had the blind hemifield on the left, five participants were female, and the median age was 61 years (range, 20–81) years. Five participants showed evidence of hemispatial neglect and two others had a previous history of neglect. The cause of HH was a stroke not during surgery for 14 participants (5 ischemic, 2 hemorrhagic, and 7 unknown), a stroke during surgery for two participants, traumatic brain injury for three participants, and a brain tumor for two participants.

Statistical Analyses
We compared demographics between the groups by applying the two-sample Kolmogorov-Smirnov test for equality of ordered distributions and the $\chi^2$ test for categories (Table 1). Where a difference between groups was noted for an ordered demographic variable, we used the Mann-Whitney $U$ test, which compares the central tendency of the two distributions. To compare IA scores between groups (with and without HH; Study 1), and the effect of the content guide (IA scores within subjects using the content guide or not, Study 2), we used mixed-effects regression analyses$^{48}$ with crossed-random factors of participant and video clip, covariates of sex, age, race, education level, visual acuity, cause of HH, side of HH, and MoCA score. The random effects account for differences between responses to clips (some clips have higher shared-word scores than others) and differences between responses from participants (some subjects are more articulate than others), removing those sources of variance. Weak effects of sex, age, and education level have been found in some groups of participants with NV.$^{31}$ Race and cognitive status (MoCA) could affect the form of responses, and visual acuity might affect the ability to see details. Non-significant terms ($P > 0.10$) were removed from models in a stepwise manner. Covariates are only reported when the covariate was significant. For example, if MoCA score was not a significant covariate in an analysis, then it was not included in the final model. Block order and neglect (measured or history) were included as fixed factors, and an interaction between neglect and content guide was included in the analysis of the content guide. To address the possibility that the number of scene cuts might have an impact on the IA score, we used a mixed-effects model with the number of scene cuts in each clip as a fixed factor. As the sample sizes were small, we accepted $P \leq 0.01$ as significant, and report terms with $0.10 \leq P < 0.01$ as trends. Analyses were conducted using Stata (version 14, StataCorp LP, College Station, TX).

Results

Study 1: Information Acquisition
The HH group (2.8 shared words, 95% confidence interval 2.3–3.3) had an IA score that was 1.5 shared words lower than the NV participants (4.3 shared words, 95% confidence interval 4.0–4.6), when corrected for age (mixed-effects regression, $z = 5.69, P < 0.001$; Fig. 2).

Age was significantly related to the IA score ($z = 3.49, P < 0.001$), with IA score decreasing 0.24 shared-words per decade. Other co-variates, such as sex, maximum education level, MoCA score, race, duration of hemianopia and visual acuity, were not significant ($P > 0.35$), and therefore were not included in the final model. Among the participants with HH, measured neglect or history of neglect ($z = 0.42, P = 0.67$) and cause of HH ($z < 0.30, P > 0.76$) were not related to IA score in our sample. There was a trend for subjects with right HH to have a lower IA score ($-0.80$ shared words, $z = 1.62, P = 0.10$).
Study 2: Content Guide

For the 21 participants with HH, the content guide improved the IA scores by 0.5 shared words on average (mixed-effects regression, $z = 2.18$, $P = 0.03$) from 2.5 (95% confidence interval 1.8–3.0) shared words without guide to 3.0 (95% confidence interval 2.3–3.5) shared words with guide, when corrected for sex. In this sample, males had 1.0 less shared words than females ($z = 1.85$, $P = 0.065$). Except for sex, no covariates were significant ($P > 0.46$), including presence or history of neglect ($P = 0.97$) and block order (i.e., which condition was seen first; $P = 0.84$). There was no difference in the effect of the content guide between the participants with and without neglect ($z = 1.09$, $P = 0.27$). The improvement of 0.5 shared words is a Cohen’s $d$ (effect size) of 0.29, which is considered small to medium.49 As shown in Figure 3, mean IA scores were higher when the content guide was present for 17 of the 21 participants.

When there was a scene cut, the content guide had to relocate to a new COI (that changed instantaneously with the scene cut). When this occurs, a viewer with normal vision will quickly make a saccade to the new COI. Because the content guide followed the COI as defined by the group of viewers with NV, the content guide would move following a scene cut. That move was not as fast as a saccade, since the NV viewers would make their relocation saccades at slightly different times and we restricted how quickly the content guide could move. Even so, it is possible that having more scene cuts might affect the ability to follow the story and the value of the content guide. When viewing original clips, there was no effect of the number of scene cuts on the IA score in the NV group ($z = 0.08$, $P = 0.93$). For the HH group, there was no effect of the number of scene cuts on the IA score when viewing original clips ($z = 0.16$, $P = 0.87$) and this did not change when viewing the content guide ($z = 0.45$, $P = 0.66$).

Discussion

Self-reported difficulties with watching TV, an activity of daily living, have been previously reported for hemianopia.18,19 As the prevalence of hemianopia in the United States is estimated to affect 0.5 to 1 million people,1,2,6 and difficulty watching television is reported by 30% to 56% of the people with hemianopia,18,19 this suggests that between 150,000 and 560,000 people with hemianopia in the United States feel that they have difficulty watching TV because of hemianopia. Study 1 is the first demonstration that watching a video is measurably more difficult for people with hemianopia.50 We are aware of no electronic method to assist people with hemianopia to watch TV. Study 2 investigated our innovative method of presenting a content guide around the COI. Our approach guided the viewer’s gaze toward the most visually informative area of the movie, and was found to improve their information acquisition (IA score). While we found no effect of block order, the improvement in IA could be due to a placebo effect, with the participants noting the intervention (yellow ring at the COI) and trying harder in those trials. We did not tell the participants that this would improve their viewing experience, nor did we ask that they look at the ring; instead, we told the participants that things of interest were highlighted by the ring and allowed them to decide how to make use of it. To fully assess the content guide, we propose a randomized controlled trial in which, for one condition, the content guide will follow a path that is not associated with the clip (i.e., different from the democratic COI for the clip) as a control for a placebo effect from content-guide presence, in addition to the existing control condition with no content guide.

In our study, participants were told that the highlighted region in the content-guide condition
would identify the objects of most interest, but were not told to look there. If the content guide was not visible, they understood that it could be located by making an eye movement (scan) into the blind side. The viewer knows that the content guide contains the region with the most important information. Therefore, making compensatory scans into the blind side to check for important content that might be missed due to the visual field loss is no longer required. This compensatory scanning is a widely used form of rehabilitation treatment that has been shown to improve performance on certain tasks.\(^{10,21,22,29,43,44}\)

The intention of the content guide is not to train users not to make compensatory scans, nor is it expected that using the content guide will cause a reduction in compensatory scans in other settings.

The rationale for the value of the content guide is that compensatory scanning may take attention away from the most informative region (the COI) in a video, and thus these scans may impair the ability to follow a video storyline. Thus, the benefit from the content guide may come both by drawing attention to the most informative region and by the reduction of scans into the blind side. We are examining this hypothesis separately by measuring whether there was a decrease in compensatory scans when viewing video clips with the content guide.

To put the average 0.52 shared words increase in IA scores with the content guide into context, we evaluated the variability of IA scores by comparing the score obtained with the first 20 video clips to that obtained with the second 20 video clips, of the 40 clips seen by the 60 participants in the NV group. While not an ideal evaluation of test–retest repeatability, it provides some estimate of variability in the metric. The repeatability coefficient (95\% confidence interval\(^{51}\)) was ±1 shared word. Thus, the improvement was less than the within-session variability that might be expected. Further, the average IA score of 3.0 shared words with the content guide was still substantially less than the NV group (average 4.3 shared words). Thus, while this preliminary study showed that the content guide provided some improvement, it did not bring IA ability of people with HH up to the level of people with NV.

After the completion of the study we did not ask participants whether they preferred using the content guide or if they noticed an improvement in their ability to interpret videos. It is possible that a lack of experience with the content guide limited the value from the content guide in our study. For future evaluations of the content guide, we intend to give participants a practice period to acclimatize to the content guide and to learn to associate it with areas that are particularly informative, while de-emphasizing areas of the display that are not highlighted by the content guide. Future studies could use alternative methods of presenting the content guide, as the yellow ring was a simple application over the video that may not be the most optimal variant.

Half of the participants with hemianopia in a survey of video viewing habits\(^{18}\) expressed strong interest in assistive technology. One of the concerns with an intervention that modifies the displayed image is that it may be unacceptable to other viewers watching the same display. However, it seems that there would be plenty of opportunity to use such a guide, as watching television or video on a computer while alone was common,\(^{18}\) and, in a previous related study, the median number of TV sets per home was found to be two.\(^{52}\)

Though we found no difference between the participants with and without spatial neglect in this study, this does not mean that neglect may not affect the response to the content guide. All of the participants with neglect had mild neglect. So, it seems that people with mild neglect can benefit from the content guide. However, viewers with more severe neglect may have more difficulty with the IA task and may respond differently to the content guide.

In conclusion, we objectively demonstrated that participants with HH experienced difficulties acquiring information from video (reduced scores in our novel IA metric), and that the content guide provided benefit (increased IA scores), which shows promise as a rehabilitation intervention for people with HH.

**Acknowledgments**

This study was supported by National Institutes of Health Grant R01EY019100 and Core Grant P30EY003790.

Disclosure: F. M. Costela, None; D.R. Saunders, commercial relationship described below; S. Kajtezovic, None; D.J. Rose, None; R.L. Woods, commercial relationship described below.

The information acquisition method used here for scoring natural language responses is described in a patent application submitted by the Schepens Eye Research Institute. The content guide rehabilitation
method is described in a patent application submitted by the Schepens Eye Research Institute.

References

25. Reinhard J, Schreiber A, Schiefer U, et al. Does visual restitution training change absolute hom-