

# What Visual Targets Are Viewed by Users With a Handheld Mobile Magnifier App

Gang Luo<sup>1</sup>

<sup>1</sup> Schepens Eye Research Institute, Massachusetts Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, MA, USA

**Correspondence:** Gang Luo, 20 Staniford Street, Boston, MA 02114, USA. email: [gang\\_luo@meei.harvard.edu](mailto:gang_luo@meei.harvard.edu)

**Received:** December 22, 2020

**Accepted:** February 4, 2021

**Published:** March 15, 2021

**Keywords:** smartphone; low vision; visual demands

**Citation:** Luo G. What visual targets are viewed by users with a handheld mobile magnifier app. *Trans Vis Sci Tech.* 2021;10(3):16, <https://doi.org/10.1167/tvst.10.3.16>

**Purpose:** Mobile video magnifier apps are used by many visually impaired people for seeing details that are beyond their visual capacity. Understanding the common types of visual targets will be importantly informative for low-vision research and assistive technology development. This study addressed this question through analysis of images captured by magnifier app users pursuing their daily activities.

**Methods:** An iOS magnifier app, free to the public, was used to capture and upload images to the Azure Computer Vision cloud service for object recognition. Returned object tag results for each image were uploaded to the Umeng analytics server for aggregated tallies. Consolidated data from 24,295 users across 1 month were analyzed. More than 1300 types of object tags found in 152,819 images were grouped into 11 categories. The data collection and analyses were conducted separately for users who toggled on or off iOS vision-accessibility features.

**Results:** For accessibility and nonaccessibility user groups, 60% to 70% of objects were nontextual, such as an indoor scene, human, or art. More than 40% of the images contained more than one object category. Accessibility users viewed textual objects more frequently than nonaccessibility users (41.1% vs. 29.8%), but overall, the probability ranking of categories was not significantly different between the two groups.

**Conclusions:** Nontextual objects make up a major portion of visual needs of magnifier users across a wide range of vision loss.

**Translational Relevance:** Low-vision research and vision assistance technology development should address the need for nontextual object viewing.

## Introduction

People with vision loss have difficulty with a wide range of daily visual tasks. Some need multiple visual aids for different situations, such as a handheld magnifier for near distances and a telescope for far distances. To help visually impaired people better cope with the difficulties they experience in a fast-changing world, frontline development must be based on a comprehensive understanding of their visual needs.

Conventionally, visual needs can be probed using short questionnaires, such as the National Institute of Health, Visual Function Questionnaire (NIH VFQ), or large questionnaires that include hundreds of visual task items, such as the Activity Inventory.<sup>1</sup> Through the questionnaires, the visual tasks that a person has diffi-

culty with may be identified. As modern life evolves, however, some of the questionnaire items may become less relevant. For instance, searching in phone books is no longer common, as online searches now prevail. Therefore, the questionnaires, or even the survey methods, may need to be updated to suit changing lifestyles. While questionnaire methods are useful and can be validated, they lack open-ended questions. The predefined questionnaires designed by researchers can miss important aspects of the real-world experience.

Recently, Starke et al.,<sup>4</sup> without presetting any visual tasks, classified the visual needs of 32 people with low vision, based on captured images, over a 1-week period, of daily-life scenes and corresponding participant narration. They identified some visual demands that were not included in previous survey studies.<sup>2</sup>

Some of the visual needs were related to use of mobile devices, which were not widely available just 10 years ago. The study by Starke et al. highlights the value of a naturalistic survey method, which records the actual daily activities of participants rather than subjective recollections.

The study presented in this article adopted a novel naturalistic survey method, based on a visual aid used in daily activities. As mobile devices become ubiquitous, many vision-assistive mobile apps have been developed. Most are video magnifier apps that use the built-in cameras to turn the mobile devices into handheld electronic magnifiers. A search using the keywords “magnifying glass” or “magnifier” returned more than 100 apps at the Apple App Store (iOS) as well as the Google Play Store (Android). In this study, a video magnifier app developed by us, SuperVision+ Magnifier, was used to peek into the visual tasks for which people used help from the app. The strength of this study is that the investigation was conducted based on big data collected from tens of thousands of users worldwide, thanks to the app’s popularity in the low-vision community.

## Methods

### Data Collection

Data presented here were collected for 4 weeks in October 2020 using the free SuperVision Magnifier iOS app (v1.8.1) as the platform, which was released to the public worldwide in September 2020. The app provides vision assistance features commonly avail-

able in other smartphone magnification apps, such as snapshot, contrast enhancement, and color inversion. It can be considered representative of many other similar magnifier apps. The app also provides a unique live-image stabilization feature, which is activated when a user presses the touchscreen. In our previous user-behavior study,<sup>3</sup> it was found that live-image stabilization was used more often than the snapshot feature (which captures a still image). Both techniques mitigate image-shaking problems. In this study, when the stabilization or snapshot was activated within 15 seconds of app launch, an image frame at that moment was uploaded to the Microsoft Azure Computer Vision service for image analysis. If neither image stabilization nor snapshot was activated by the time 15 seconds had elapsed, one image frame was captured for image analysis. No additional images were captured per launch. No images were captured from activations shorter than 15 seconds with no stabilization or snapshot trigger, as such short uses may include some accidental, unintentional app launches. Those omitted app launches accounted for about 25% of all app uses.

Object recognition results returned to the SuperVision app from Azure’s service provide a list of objects and image/object property tags. Tags with a confidence score higher than 80% were sent to an analytics server (Umeng, Beijing, China), where the data were consolidated for the entire app user population. No individual-identifiable data or images are saved by the Azure server, Umeng server, or the SuperVision app.

Figure 1 shows some examples of object recognition results returned from the Azure Computer Vision service. For each analyzed image, one or more tags may be received.

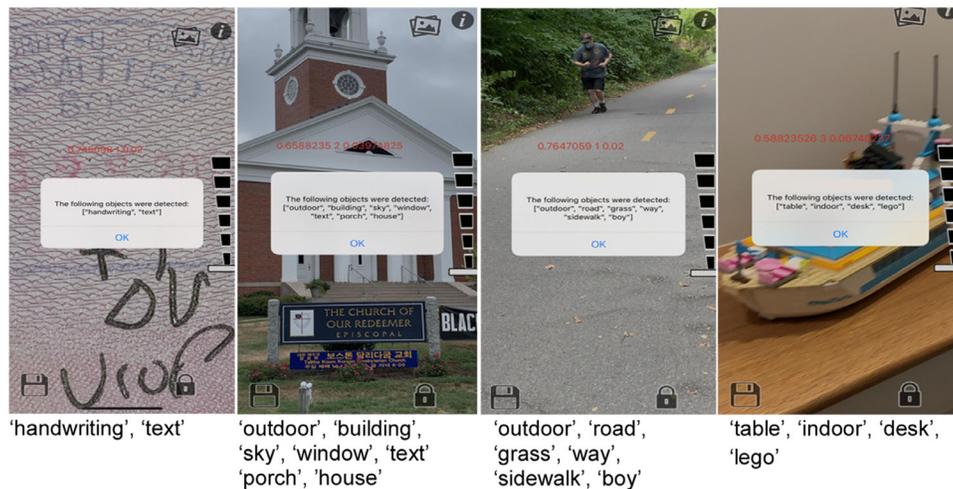


Figure 1. Examples of object recognition results provided by Microsoft Azure’s Computer Vision cloud service. Images taken with the SuperVision Magnifier app are sent to the Azure server via the Internet for processing. The popup message boxes showing the returned object tags were used for code debugging. No such messages were shown to actual users.

## Cohorts

During the data collection period, 38,749 unique users used the SuperVision Magnifier app, according to Umeng analytics report. Most of these users installed the new version (v1.8.1) through update. During the data collection period, about 13% of the users were new according to the weekly tally, and 22% of the users were new according to the monthly tally. When the new version was opened for the first time, a message box popped up to ask the users if they permitted the image to be uploaded to a cloud server for analysis. Some of them did not grant permission, so no image was analyzed for them. Therefore, the number of users who participated in the data collection was less than the total number of active users. The total number of participants in this study during the data collection period was 24,295.

The vision status of the users is unknown, as the users were not asked to provide that information, mainly for three reasons: (1) the sampling would not be biased to users who are willing to disclose the information, (2) the study only aims to investigate the common patterns in the entire cohort, and (3) sample size can thus be maximized. In this study, however, the users were divided into two groups and sampled separately according to their smartphone accessibility settings, which were detected automatically by the app. The first group of users comprised those who toggled on any of the iOS vision accessibility features (e.g., color invert, voiceover), and the second group comprised those who did not toggle on any accessibility features. It is very likely that the accessibility feature users met clinical definitions of visual impairment. Because the accessibility mode is not user-friendly to normally sighted people, users without great visual difficulty usually do not toggle on those features. The nonaccessibility

feature users may also include many with severe visual impairments, but it is reasonable to assume that overall, their vision impairment was less severe than that of the accessibility feature users.

After the object tag data for each image were uploaded to the Umeng analytics server, the data were tallied in an aggregated manner, without any individually identifiable information being saved. For instance, if 100 people used the app for viewing restaurant menus once in a day, the messages we received from the Umeng server for that day would include “menu -100.” The study was approved by the Institutional Review Board of Partners Health Care, which ruled that this survey study is exempt from human subject study regulations because individually identifiable information is not collected.

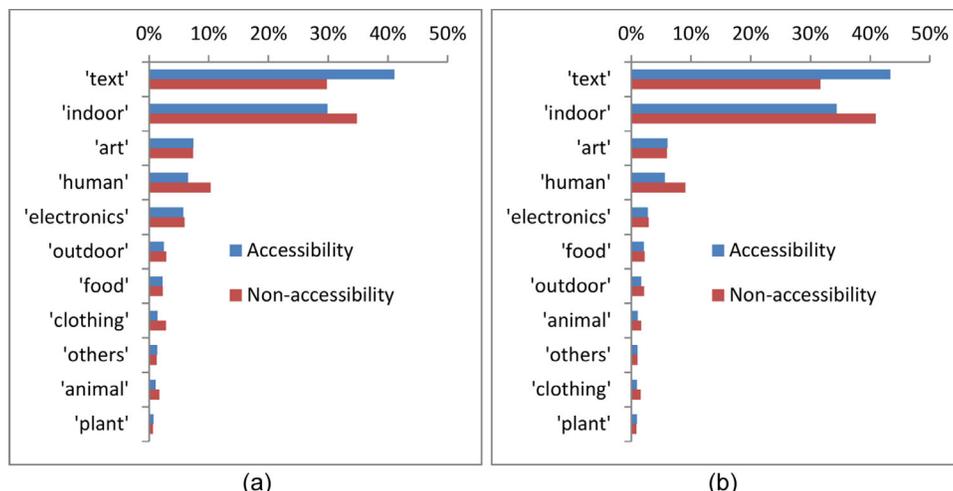
## Statistical Analyses

The statistical analyses were conducted offline, based on the data downloaded from the Umeng analytics server. Prior to data analysis, the data went through preprocessing to remove image property tags, which describe the color (e.g., red, pink), action (e.g., diving, riding), object property (e.g., tall, round), and image quality (e.g., blur, fog). These property tags were not of interest for this study.

As Azure’s Computer Vision can recognize thousands of types of objects, objects needed to be grouped into categories to gain a high-level understanding of the visual targets viewed through the video magnifier. Based on observation of the object tag data, 18 categories were first used to group the objects. After the initial tallying was completed, the category list was condensed to 11 by combining some related and minor groups to make it easier to understand the data. [Table](#) lists the categories used in final analysis

**Table.** Category List of Objects

Category	Example Object Types
Text	Printed text; books, letters, menu, product labels
Indoor living	Indoor scenes, house structure (e.g., wood floor), furniture, appliances, household items
Art	Drawing, sketch, sculpture, musical instruments
Human	Person, face, body parts (e.g., hand)
Electronics	Computers, circuit, mobile phone
Clothing	Clothing, cosmetics, accessories
Outdoor	Outdoor scenes, street, snow, beach, rock
Food	A variety of foods, including drink, bakery, vegetable, fruit
Animal	Pet, wild animal, insects
Plant	Unspecified plant, tree, flower, grass
Others	Sports gear, game, vehicle



**Figure 2.** Percentage of captured images with each object category. (a) The percentage is calculated based on NP counts. (b) The percentage is calculated based on WP counts. The top five categories were the same based on the two measures.

and their descriptions. There may be more than one object in an image belonging to a given category. In our analyses, a category is counted only once per image. The object categories were tallied in two ways: the number of presences (NP) and the weighted presences (WP). If a category is present in the interpretation of an image, it will be tallied once for NP regardless of the other categories in the image. For WP, the weight of a category is defined as 1 divided by the number of categories in the image. Normalizing the NP and WP category sums by the total tally gives the overall percentage of each category.

One object may be associated with multiple tags. For example, a water bottle with a textual label on it may be labeled by two tags, “water bottle” and “text.” Thus, category combinations (e.g., “food\_text”) were also tallied for a separate analysis of the image scenes. Images with one category or a two-category combination were investigated in this study. Images with more than two categories were rare and excluded from this analysis.

## Results

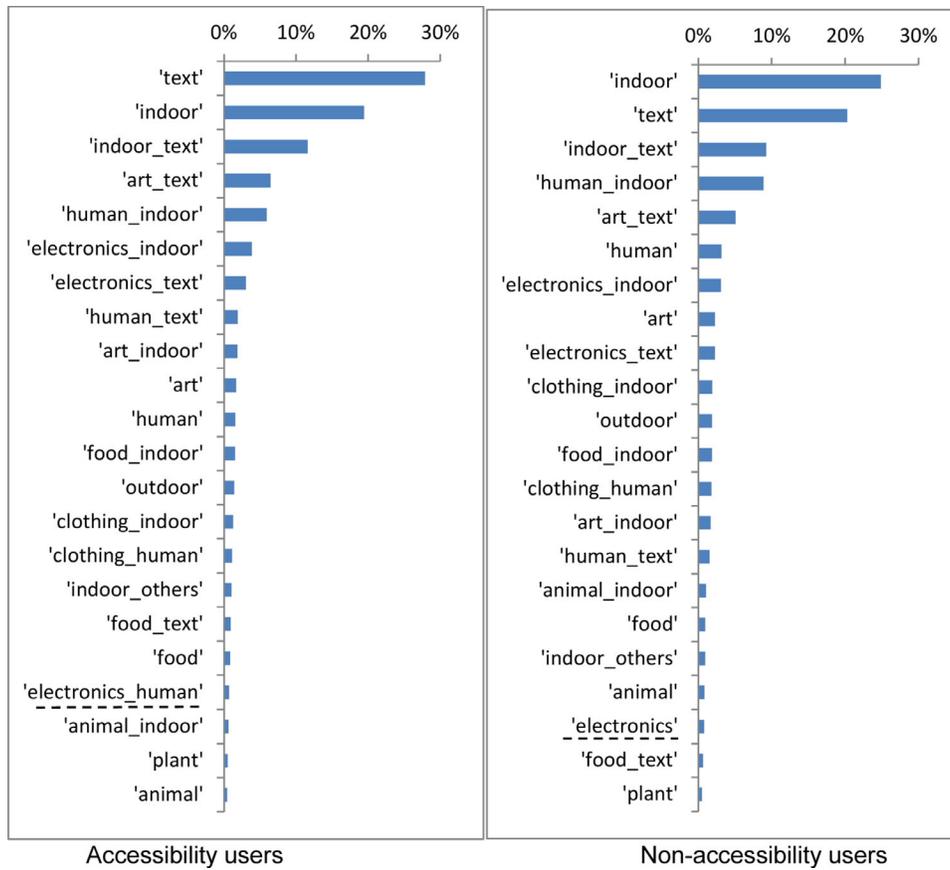
In total, 49,877 images from accessibility users and 102,942 images from nonaccessibility users were successfully processed by Azure Computer Vision. Figure 2a shows the percentage of different object categories presented in those images. For accessibility users, the percentage of textual objects was the highest, 41.1%, and it was the second highest for

nonaccessibility users, 29.8%. Overall, most objects were nontextual for both groups of users.

Figure 2b shows the weighted percentage of object categories, which is calculated based on WP. The overall pattern of weighted percentage is very similar to the percentage of presences because the NP and WP counts are highly correlated ( $P < 0.001$ ,  $R^2 = 0.99$ ). There were some differences in the rank orders for minor categories. The high correlation between NP and WP suggests that each of the categories had about the same probability to coexist with other categories in an image. If a category occurs only in single-category images, its WP would be high because the weight would be always 100%. However, this was not the case.

Qualitatively, the object category percentages in images captured by accessibility and nonaccessibility users seemed to be very similar (Fig. 2). For instance, they all viewed textual and indoor objects much more than outdoor objects and plants. To quantitatively compare the two groups of users, a Wilcoxon signed rank test was conducted based on NP percentage and WP percentage. The difference was not statistically significant ( $P = 0.283$  based on NP,  $P = 0.275$  based on WP).

Figure 3 shows the percentage of category combinations. This measure indicates whether there was more than one category in an image and, if yes, what the category combinations were. Images containing three or more categories are rare, so they are not included in the analysis. The images containing two categories were 45.3% and 43.9% for accessibility and nonaccessibility user groups, respectively. Commonly seen category combinations are “indoor\_text,” “human\_indoor,”



**Figure 3.** Presence percentages of category combinations seen by the accessibility and nonaccessibility groups. There may be one or more object categories in an image. The graphs show the combinations that were present in the analyzed images. If an image contains only one category, there is no combination. The lists are cut off at 95% cumulative percentages. The two rankings include almost the same category combinations, although the specific rank order and percentages are different. Items underlined by *dashed lines* indicate categories that are not in the other group’s list.

and “art\_text.” The ranked lists of category combinations for accessibility and nonaccessibility users include almost the same items, although the specific rank order and presence percentages are different.

## Discussion

People with low vision have a wide range of visual needs in their daily lives. On the basis of 612 scenes captured within a week from the daily activities of 32 low-vision participants, Starke et al.<sup>4</sup> identified their need for assistance with activities such as finding things on a crowded shelf, reading package labels, reading newspapers, watching TV, using devices, and crossing streets. It can be expected that mobile magnifier apps can help with some of the visual tasks, such as reading package labels. Based on usage data collected by our SuperVision Magnifier app over 1 month, from more

than 24,000 users during their daily activities, this study specifically characterized the visual targets that were viewed. This “big data” sample is likely indicative of needs being met by mobile electronic magnifier apps. While it is not surprising that magnifier apps are used to view text, we, as in Starke et al.,<sup>4</sup> found that more than half of the visual targets viewed through our app are nontextual, such as objects indoors. The high frequency of such uses indicates that the users found value in using the magnifier app for nontext visual targets. The utility of magnifier apps for this important application scenario should be considered in future low-vision rehabilitation research.

Text-reading performance evaluation is one of the commonly used assessment paradigms in low-vision research.<sup>5,6</sup> A variety of vision rehabilitation interventions have been studied based on text-reading performance.<sup>7,8</sup> As mobile devices become an increasingly common platform for low-vision rehabilitation, some studies have investigated the use of mobile devices for

text reading. Gill et al.<sup>9</sup> reported that reading speed on an iPad was faster than on paper. Morrice et al.<sup>10</sup> found reading performance with an iPad was not different from that with a CCTV. Walker et al.<sup>11</sup> evaluated a special tablet app that presents text in a scrolling fashion, and they showed reading performance with scrolling text was better in terms reading error rate than static text.

As our study has found that text was a quite common visual target, text-reading studies are certainly a meaningful and relevant part of the daily activities of people with low vision. However, our study also found that, in total, the wide variety of nontextual targets exceeded textual targets. These nontextual visual targets are probably associated with some of the visual demands found in Starke et al.,<sup>4</sup> such as “finding things on a crowd shelf” and “appreciate environment.” The prominence of nontextual visual targets manifested in magnifier use indicates that it is important to include the visually challenging activities in low-vision research, for instance, object search.<sup>12–14</sup>

Starke et al.<sup>4</sup> reported that most visual tasks that people with low vision would need help with were ad hoc or short. Our recent study<sup>3</sup> with the SuperVision Magnifier app also showed that the app was primarily used for short spot viewing; about 51% of the app uses were shorter than 1 minute. The spot-viewing breakdown is not yet clearly known, but given the visual target tally reported here, it is speculated that the spot-viewing tasks were involved with both textual and nontextual targets. Our previous visual search studies on low-vision people<sup>12–14</sup> involved spot viewing of nontextual objects. The visual tasks in our keyword search app evaluation study<sup>15</sup> and the Wittich et al.<sup>16</sup> study with iPad magnifier apps are spot reading of textual objects, such as product labels. Actually, the keyword search app (SuperVision Search), which can localize keywords provided by users in images captured by phone cameras, was developed by us to help facilitate spot reading of text.<sup>15</sup> The finding of the high use for viewing textual objects justifies the significance of the app development work. In addition, this study showed that the scenarios in which people with low vision need vision assistance are quite diversified and often not restricted to a single object category within a snapshot (Fig. 3). This suggests that there is a need to develop vision-aiding apps that are more versatile than those solely designed for reading text.

There have been some low-vision research efforts to develop testing batteries for performance evaluation of visual tasks involved with nontextual targets. It is argued that visual tasks of this type are an important part of daily life and should be included in assessments of functional vision and effectiveness of rehabilitation

interventions. For instance, the Timed Instrumental Activities of Daily Living test developed by Owsley et al.<sup>17</sup> includes finding tools in a crowded drawer. The Melbourne Low-Vision ADL test<sup>18</sup> includes recognizing faces and threading a sewing needle. According to the big data collected from actual daily activities, our study confirms that nontextual visual targets are involved in most visual tasks. Furthermore, we argue that such visual tasks should make up the majority of tasks in the testing batteries used in many low-vision studies. Ideally, the composition of visual tasks should roughly match the spectrum of visual needs in a cohort’s actual daily activities. Otherwise, laboratory studies may miss important needs while overrepresenting others. The methods of our study, of course, can only shed light on those tasks for which people chose to use a handheld magnifier. Complementary observational techniques are needed to ensure that significant needs are not missed.

A lack of detailed demographic information and vision status of the users was a limitation of this study. This prevents investigation of the relationship between behaviors and user characteristics. Only common patterns for the entire set of users can be identified. Nevertheless, by distinguishing those who did and did not use any iOS vision accessibility features, this study infers that the main difference associated with vision loss severity might be in visual demands for reading text. The accessibility group needed to use visual aids to read textual information more frequently than the nonaccessibility group (Fig. 2). This is not a surprise, since critical print size is correlated with visual acuity for people with low vision.<sup>19</sup> Presumably, the accessibility users are more likely to encounter print text smaller than their limits than the people who do not need to use accessibility features. It should be noted that the nonaccessibility group may include some users with severe vision impairment, but they did not toggle on any accessibility for some reason (e.g., lack of skills). If this is true, it may mean that the need for assisted text reading is even larger for people with severe vision loss. An interesting finding, though, is that the overall distribution of visual targets was very similar for both groups. This might imply that, with help from the smartphone magnifier, they all were able to accomplish similar types of visual tasks in their daily lives.

In future studies, the limitation mentioned above will be addressed. For instance, age of users can be obtained via a simple survey. Age factor is expected to have impacts on smartphone use behaviors. Older people make up a significant portion of the visually impaired population, and many older people with visual impairments are using vision assistance apps.<sup>20</sup> It has been predicted that the market share of dedicated,

expensive low-vision devices may shrink substantially due to the increased availability of smartphone-based vision assistance technologies.<sup>21</sup> Following the trend, more elderly people may adopt the technologies. How they use the vision apps and how to improve the usability of vision apps for them can be investigated by using behavioral research approaches similar to that presented in this article.

## Conclusions

Through a novel survey approach, this big-data study revealed that the majority of visual targets people use smartphone magnification apps to help to see are nontextual. Overall, the visual demands of daily activities were very similar whether or not any iOS vision accessibility features were enabled, although vision impairment severity was likely greater for those using accessibility features. It might be likely that users with more severe vision loss needed to use the magnification app for reading text slightly more than users with less vision loss, but the importance of nontextual targets was still surprisingly significant. These insights should be considered in low-vision research and assistive technology development.

## Acknowledgments

The author thank Anurag Shubham for upgrading the SuperVision Magnifier app, which was used in this study. We also thank Henry Apfelbaum for paper editing and insightful discussions.

Disclosure: **G. Luo**, None

## References

1. Massof RW, Ahmadian L, Grover LL, et al. The activity inventory: an adaptive visual function questionnaire. *Optom Vis Sci.* 2007;84(8):763–774.
2. Taylor DJ, Hobby AE, Binns AM, Crabb DP. How does age-related macular degeneration affect real-world visual ability and quality of life? A systematic review. *BMJ Open.* 2016;6(12):e011504.
3. Luo G. How 16,000 people used a smartphone magnifier app in their daily lives. *Clin Exp Optom.* 2020;103(6):847–852.
4. Starke SD, Golubova E, Crossland MD, Wolffsohn JS. Everyday visual demands of people with low vision: a mixed methods real-life recording study. *J Vis.* 2020;20(9):3.
5. Legge GE, Bigelow CA. Does print size matter for reading? A review of findings from vision science and typography. *J Vis.* 2011;11(5):8.
6. Legge GE, Ross JA, Luebker A, Lamay JM. Psychophysics of Reading: VIII. The Minnesota Low-Vision Reading Test. *Optom Vis Sci.* 1989;66(12):843–853.
7. Ortiz A, Chung ST, Legge GE, Jobling JT. Reading with a head-mounted video magnifier. *Optom Vis Sci.* 1999;76(11):755–763.
8. Culham LE, Chabra A, Rubin GS. Clinical performance of electronic, head-mounted, low-vision devices. *Ophthalmic Physiol Optics.* 2004;24(4):281–290.
9. Gill K, Mao A, Powell AM, Sheidow T. Digital reader vs print media: the role of digital technology in reading accuracy in age-related macular degeneration. *Eye.* 2013;27(5):639–643.
10. Morrice E, Johnson AP, Marinier JA, Wittich W. Assessment of the Apple iPad as a low-vision reading aid. *Eye (Lond).* 2017;31(6):865–871.
11. Walker R, Bryan L, Harvey H, Riazi A, Anderson SJ. The value of tablets as reading aids for individuals with central visual field loss: an evaluation of eccentric reading with static and scrolling text. *Ophthalmic Physiol Optics.* 2016;36(4):459–464.
12. Satgunam P, Luo G. Does central vision loss impair visual search performance of adults more than children? *Optom Vis Sci.* 2018;95(5):452–456.
13. Luo G, Satgunam P, Peli E. Visual search performance of patients with vision impairment: effect of JPEG image enhancement. *Ophthalmic Physiol Optics.* 2012;32:421–428.
14. Satgunam P, Woods R, Luo G, Bronstad M, Reynolds Z, Ramachandra C, Mel B, Peli E. Effects of contour enhancement on low-vision preference and visual search. *Optom Vis Sci.* 2012;89(9):1364–1373.
15. Pundlik S, Singh A, Baghel G, Baliutaviciute V, Luo G. A mobile application for keyword search in real-world scenes. *IEEE J Transl Eng Health Med.* 2019;7:1–10.
16. Wittich W, Jarry J, Morrice E, Johnson A. Effectiveness of the Apple iPad as a spot-reading magnifier. *Optom Vis Sci.* 2018;95(9):704–710.
17. Owsley C, McGwin G, Jr, Sloane ME, Stalvey BT, Wells J. Timed instrumental activities of daily living tasks: relationship to visual function in older adults. *Optom Vis Sci.* 2001;78(5):350–359.
18. Haymes SA, Johnston AW, Heyes AD. The development of the Melbourne Low-Vision ADL Index: a measure of vision disability. *Invest Ophthalmol Vis Sci.* 2001;42(6):1215–1225.

19. Massof RW. Relation of reading performance to visual acuity and perceived reading ability in low vision. *Invest Ophthalmol Vis Sci.* 2003; 44(13):1284–1284.
20. Bhakhri R, Chun R, Coalter J, Jay WM. A survey of smartphone usage in low vision patients. *Invest Ophthalmol Vis Sci.* 2012;53(14):4421–4421.
21. Karmel M. Tools for low vision patients: high hopes for high-tech gadgets. *Eyenet Magazine.* 2012:31–33.