

Objective Assessment of Activity Limitation in Glaucoma with Smartphone Virtual Reality Goggles: A Pilot Study

Rachel L. Z. Goh¹, Yu Xiang George Kong^{1,2}, Colm McAlinden^{3,4}, John Liu⁵, Jonathan G. Crowston^{1,2}, and Simon E. Skalicky^{1,6,7}

¹ The Royal Victorian Eye and Ear Hospital, Melbourne, Australia

² Centre for Eye Research Australia, The University of Melbourne, Melbourne, Australia

³ Department of Ophthalmology, Glangwili Hospital, Hywel Dda University Hospital Board, Carmarthen, UK

⁴ School of Ophthalmology and Optometry, Wenzhou Medical University, Wenzhou, Zhejiang, China

⁵ Monash Medical Centre, Melbourne, Australia

⁶ Discipline of Ophthalmology, The University of Sydney, Sydney, Australia

⁷ Royal Melbourne Hospital, Department of Surgery, University of Melbourne, Melbourne, Australia

Correspondence: Yu Xiang George Kong, The Royal Victorian Eye & Ear Hospital, 32 Gisborne St, East Melbourne, Victoria 3002, Australia. e-mail: gkong008@hotmail.com

Received: 29 May 2017

Accepted: 13 October 2017

Published: 23 January 2018

Keywords: glaucoma; virtual reality; technology; activity limitation

Citation: Goh RLZ, Kong YXG, McAlinden C, Liu J, Crowston JG, Skalicky SE. Objective assessment of activity limitation in glaucoma with smartphone virtual reality goggles: a pilot study. *Trans Vis Sci Tech.* 2018; 7(1):10, <https://doi.org/10.1167/tvst.7.1.10>

Copyright 2018 The Authors

Purpose: To evaluate the use of smartphone-based virtual reality to objectively assess activity limitation in glaucoma.

Methods: Cross-sectional study of 93 patients (54 mild, 22 moderate, 17 severe glaucoma). Sociodemographics, visual parameters, Glaucoma Activity Limitation-9 and Visual Function Questionnaire – Utility Index (VFQ-UI) were collected. Mean age was 67.4 ± 13.2 years; 52.7% were male; 65.6% were driving. A smartphone placed inside virtual reality goggles was used to administer the Virtual Reality Glaucoma Visual Function Test (VR-GVFT) to participants, consisting of three parts: stationary, moving ball, driving. Rasch analysis and classical validity tests were conducted to assess performance of VR-GVFT.

Results: Twenty-four of 28 stationary test items showed acceptable fit to the Rasch model (person separation 3.02, targeting 0). Eleven of 12 moving ball test items showed acceptable fit (person separation 3.05, targeting 0). No driving test items showed acceptable fit. Stationary test person scores showed good criterion validity, differentiating between glaucoma severity groups ($P = 0.014$); modest convergence validity, with mild to moderate correlation with VFQ-UI, better eye (BE) mean deviation, BE pattern deviation, BE central scotoma, worse eye (WE) visual acuity, and contrast sensitivity (CS) in both eyes ($R = 0.243\text{--}0.381$); and suboptimal divergent validity. Multivariate analysis showed that lower WE CS ($P = 0.044$) and greater age ($P = 0.009$) were associated with worse stationary test person scores.

Conclusions: Smartphone-based virtual reality may be a portable objective simulation test of activity limitation related to glaucomatous visual loss.

Translational Relevance: The use of simulated virtual environments could help better understand the activity limitations that affect patients with glaucoma.

Introduction

Glaucoma causes an increasing visual burden for patients around the world. By 2020, it is predicted to affect 79.6 million people.¹

Patients with advancing glaucoma suffer from increasing difficulty with tasks requiring contrast discrimination, light/dark adaptation, and peripheral

vision.^{2–4} These impairments can impact driving, motion perception, adjusting to different levels of illumination, and judging distances.^{5–8} Motor vehicle accidents and fall-related injuries may occur.⁹ Motion perception in particular is anomalous in patients with glaucoma. A recent review of literature identified that patients with glaucoma have elevated motion thresholds compared to normal subjects, and that motion

perimetry discriminated better than conventional perimetry between the two.⁸

Contrast sensitivity (CS), visual acuity (VA), and visual fields are clinical parameters commonly used to quantify functional visual ability.^{10–13} However, they may not reflect real-world visual function. For instance, in everyday life patients are able to move their head and eyes to compensate for loss of peripheral vision. Questionnaires—also known as patient reported outcomes (PROs)—can subjectively evaluate patients' abilities to perform visually related tasks.⁷ However, psychological factors and personality may influence patients' responses, and patients may under- or overestimate their functional impairment. Two patients with the same degree of clinically measured vision loss may rate their disability differently on a questionnaire.^{14,15}

Many of these limitations could potentially be overcome by objective simulation of functional visual ability. However, previous models have significant limitations. The Assessment of Visual Disability Related to Vision (ADREV), a performance-based assessment of functional visual ability, is limited in its applicability to glaucoma as it is predominantly influenced by central visual function.^{2,16,17} It also requires participants to navigate a large space; outcomes may thus be influenced by neurological or musculoskeletal morbidities.^{14,18,19} Recently, our group reported on a Rasch-analyzed objective visual function test called the Cambridge Glaucoma Visual Function Test (CGVFT), which is comprised of tests projected on a screen. The study showed good correlation between CGVFT score and visual field parameters, as well as quality of life (QoL).²⁰ The CGVFT is easier to administer than the ADREV, but is still a two-dimensional test and may thus not reflect a patient's experience in real life.

Recent developments in portable technology have seen a rapid growth in readily accessible virtual reality (VR) headsets that allow users to experience artificially replicated three-dimensional environments that can potentially be useful for vision research.^{21,22} These VR headsets generally rely on the display and sensors from a smartphone placed inside head-mounted goggles with optics that correct for the viewing distance (for example, the Samsung Note 3 smartphone in the original Oculus Rift VR headset). The user is able to “look around” the VR world, as the phone detects any head movement and alters the displayed field of view accordingly. VR goggles have been used to examine balance control in patients with glaucoma.²³ In the study, dynamic visual stimuli simulating falling were presented to patients with

glaucoma and controls; patients with glaucoma were found to have a larger variation in the corrective movement to maintain postural stability. VR goggles have also been used for visual field testing.^{24,25} However, the use of VR goggles to understand patients' perceptions of their environment has not been examined. We thus set out to create a three-dimensional test of visual function using VR technology that would be reflective of real life and correlate with QoL. Given that the two-dimensional CGVFT has been well-validated in patients with glaucoma, we based our test on a similar design.

Methods

Subjects

Participants were recruited from glaucoma clinics in Melbourne, Australia in 2015 at a tertiary referral teaching hospital. During regular follow-up visits, eligible subjects were approached consecutively and invited to participate in the study after providing informed consent. The study adhered to the tenets of the Declaration of Helsinki. Ethics approval was provided by the local hospital network Human Research and Ethics Committee (HREC Number 15/1219H).

Eligibility for this study included age over 18 years and the ability to speak, read, and comprehend English fluently. To be eligible, participants required a diagnosis of chronic open-angle glaucoma (OAG) in one or both eyes. OAG was diagnosed based on an open anterior chamber angle on gonioscopy, characteristic glaucomatous optic disc changes (including rim loss, notching, and/or significant nerve-fiber layer bundle loss), and/or glaucomatous visual field loss demonstrated on the Humphrey Visual Field Analyzer (HFA) (Humphrey Instruments Inc., Zeiss Humphrey, San Leandro, CA). Glaucomatous visual field loss was defined based on the Anderson criteria of a cluster of three or more nonedge points having sensitivity with $P < 5\%$, with at least one point having sensitivity with $P < 1\%$ reproducible on at least two consecutive visual field tests.²⁶ All patients had binocular VA better than or equal to 6/12.

As this was a pilot study; normal patients without glaucoma were not evaluated.

Patients with any nonglaucomatous condition that might influence visual function, such as visually significant cataract (Lens Opacities Classification System III > Grade 2), nonglaucomatous optic neuropathy or other neuro-ophthalmic condition, retinal or macular pathology, or ocular laser or surgery

in the previous 3 months, were also excluded from the study, as were patients without reliable visual field test indices (i.e., more than 15% false-positive errors, false-negative errors, or fixation losses).²⁷

Assessment of Clinical Parameters

Achromatic perimetry was performed using the HFA Swedish Interactive Threshold Algorithm standard 24-2 test. For all measured visual parameters, the better eye (BE) was determined based on the visual field index (VFI); when equivalent between eyes, the mean deviation (MD) was used. A cluster of three or more <1% points within the central 9° was considered as having a central scotoma and a cluster of three or more <1% points outside the central 9° was considered having a peripheral scotoma. Snellen VA was recorded and converted to the logarithm of the minimum angle of resolution (logMAR). CS was recorded using the Pelli-Robson chart both monocularly and binocularly.

Clinical markers of visual function from the better and worse eye (WE) may each influence vision-related QoL; however, BE markers are typically more influential.^{12,28–32} For this reason, clinical markers of visual function (CS, VA, MD, pattern standard deviation [PSD], and VFI) from the BE and WE were used in regression modelling.

The Nelson glaucoma staging system (GSS) was used to stratify glaucoma severity.⁶ This staging system was chosen specifically because it is based on binocular changes, and thus may reflect real world function more closely than monocular staging systems. It was also the staging system used for the CVGFT. The Nelson GSS involves three groups of patients: “mild” (unilateral deficit of < half of the visual field), “moderate” (unilateral deficit of > half of the visual field, or deficit of < half of the visual field in each eye), or “severe” (loss of > half of the visual field in each eye). This GSS strongly correlates with perimetric MD and PSD calculated from binocularly integrated data.⁶

Assessment of Other Risk Factors

Sociodemographic details were obtained by self-report. Covariates included age, gender, education level, employment status, marital status, ethnicity, and driving status.

Subjective Assessment of Vision-Related Activity Limitation

Vision-related activity limitation was assessed subjectively by the Glaucoma Activity Limitation-9

(GAL-9) questionnaire and the Visual Function Questionnaire–Utility Index (VFQ-UI). The GAL-9 comprises nine items that correlate with severity of visual field loss.³³ The VFQ-UI comprises six items that correlate with severity of visual field loss on QoL. It has been validated in patients with glaucoma.³⁴

Rasch Analysis of the GAL-9 and VFQ-UI

Rasch analysis was used to assess the psychometric properties of the GAL-9 and VFQ-UI using the Andrich rating scale model using bespoke Rasch analysis software (Winsteps [version 3.81], Chicago, IL).³⁵ During Rasch analysis, ordinal questionnaire responses are estimated on an interval scale (expressed in logits).³⁶ Increasing person score (in logits) indicates greater activity limitation.³⁷ A scoring algorithm was developed to convert VFQ-UI scores into a utility value using ordinary least squares regression analysis based on general population time trade-off utility scores.³⁸ VFQ-UI utility values range from 0.16 (worst health state) to 1.0 (full health).

The Virtual Reality Visual Function Test (VR-GVFT)

We set out to develop a test that would reflect real world challenges described by patients with glaucoma, such as searching for objects, motion detection, and driving.^{3,8} The VR-GVFT was thus based on objective tests well-validated in patients with glaucoma; the CGVFT and ADREV ([Supplementary Material S1](#)).⁷ Similar to the CGVFT, we utilized indoor and outdoor scenes to recreate the real-life experience of patients with glaucoma ([Fig. 1](#)).

Two major iterative processes were used to develop the study measures. Firstly, we based the design of the CGVFT on the ADREV, and then further refined the CGVFT to the VR-GVFT. Secondly, Rasch analysis is an iterative process that was used to evaluate the VR-GVFT. Rasch analysis involves evaluating test items and to identify items that do not fit the Rasch model. This process is repeated until all remaining test items pass preset Rasch metrics.

The VR-GVFT consists of 38 tests that are related to glaucoma and reflective of daily life. The initial 14 tests are stationary tests, where the time required to identify stationary objects in a VR 180° Photo Sphere (Samsung, Seoul, South Korea) image environment is recorded. The subsequent 24 items are video tests where reaction time to key events is recorded, and consists of two parts: 12 motion ball tests and 12 driving tests. High resolution (8 megapixel) 180°

Photo Sphere images and videos were recorded using the native camera application on the Samsung Galaxy Note 3 smartphone (Samsung). Driving videos of 5 to 12 seconds' duration were taken with the smartphone on a dashboard stabilization attachment.

The 14 stationary tests simulated 10 indoor and four outdoor scenes, with objects ranging in position within the field of view (and therefore the amount of eye and head movement required to be seen):

1. Identifying a fast food store sign at an outdoor intersection (STN 01)
2. Identifying a road work construction sign at an outdoor intersection (STN 02)
3. Identifying a street sign at an outdoor intersection (STN 03)
4. Identifying a general store sign at an outdoor intersection (STN 04)
5. Identifying a microwave in a tea room (STN 05)
6. Identifying a sink in a tea room (STN 06)
7. Identifying a sandwich maker in a tea room (STN 07)
8. Identifying a clock in a tea room (STN 08)
9. Identifying a refrigerator in a tea room (STN 09)
10. Identifying a poster on the wall in a cluttered study (STN 10)
11. Identifying a guitar in a cluttered study (STN 11)
12. Identifying a laptop computer in a cluttered study (STN 12)
13. Identifying a printer in a cluttered study (STN 13)
14. Identifying a row of books on the shelf in a cluttered study (STN 14)

Each stationary test outcome was a binary item of whether the object was seen in the allocated time of 60 seconds (Yes or No). If the object was seen, we also recorded a timing item of the seconds taken to identify the object. The stationary test thus has 28 items prior to Rasch analysis.

The 12 motion ball tests simulated a binocular confrontational visual field, with a single white ball moving from various peripheral positions on the periphery of the screen toward the screen center against a grass background. Eight peripheral positions were tested, with four additional balls being repeats. Participants were asked to immediately indicate when they saw the white ball in their vision. They were encouraged to use their peripheral vision

and combined head/eye movement to locate the moving ball.

The 12 driving tests simulated road hazard perception under various driving conditions. The first four asked the participant to verbally identify when it was safe to start driving, and the latter eight asked the participant to verbally indicate when they would brake to avoid a potential hazard. The scenes ranged in time of day, to reflect real life driving.

Scenes that required participants to verbally identify when it was safe to start driving:

1. At an intersection with traffic light change from red to green (MOV1-1)
2. Behind stationary car (MOV1-2)
3. Behind stationary car at night (MOV1-3)
4. Behind stationary car and cyclist at an intersection without traffic lights (MOV1-4)

Scenes that required participants to verbally identify when to brake to avoid a potential hazard:

1. Identify that a pickup truck ahead slows down and makes a U-turn (MOV2-1)
2. Identify an intersection with a stop sign (MOV2-2)
3. Identify that a car ahead is stationary and has its hazard lights on (MOV2-3)
4. Identify that a roundabout ahead has two cyclists in it (MOV2-4)
5. Identify a zebra crossing ahead in a parking lot in the day time, with pedestrians crossing (MOV2-5)
6. Identify a zebra crossing ahead at a parking lot in the evening, with pedestrians crossing (MOV2-6)
7. Identify a cyclist on the side of the road (MOV2-7)
8. Identify a cyclist on the side of the road after passing through a traffic intersection (MOV2-8)

The VR-GVFT test was preceded by two “practice” stationary scenes (one indoor lounge room scene and one outdoor street scene), which allowed participants to get used to the VR environment and headset. In each scene, participants were asked to describe what they could see and encouraged to move their head and eyes to best simulate real life. For the driving tests, patients were given a four-second countdown prior to each video commencing in order to orient themselves to the task.

For each timed item, timing was recorded with an electronic timer, commencing from the moment the administrator finished reading the task instructions to

the time when the participant verbally indicated the correct identification of the object or completed the task. The view seen by each participant was monitored and recorded throughout by the administrator on a laptop computer. For stationary test items, the item was recorded as a miss if more than 60 seconds were required to identify an object. For driving test items, a miss was recorded if the participant failed to respond within the duration of the video. For moving ball test items, failure to identify a ball within the duration that the ball appeared was recorded as a miss.

The VR-GVFT was administered to participants by one of the test administrators (XYGK, RLZG, or JL). All administrators conferred before, during, and after testing to ensure strict and consistent adherence testing protocols as maintained by the principal investigator (SES). Images and videos were delivered using a Samsung Note 3 smartphone (Samsung) inserted into a low-cost, head-mounted Google Cardboard Project Virtual Reality Adaptor (Google Inc., Mountain View, CA). A commercially available VR display software application, VR Player (version 1.8.2, VimersiV, Montreal, Quebec, Canada) was used to display images and videos on the smartphone. The smartphone was linked to a laptop computer to enable the test administrator to control the flow of images, and for live recording of patient performance using Mobizen screen synchronization software (Mobizen Inc., Seoul, South Korea). If a participant had a refractive error requiring distance correction spectacles, they wore their spectacles under the goggles throughout the test. Optical properties of Google Cardboard have been previously documented,³⁹ with a total field of view of 80° and a nominal virtual image distance of -667 mm. Validation of field of view was performed by taking a 180° Photo Sphere image of calibrated markers placed at 5° intervals from a central location and showing the image to normal volunteers ($n = 3$) (nasal field 38.3 ± 2.9 , temporal field 36.7 ± 2.9).

Rasch Analysis of the VR-GVFT

Rasch analysis was used to assess the psychometric properties of the VR-GVFT. The Rasch model is based on a probabilistic relationship between patient ability and item difficulty, with the difference signifying the functional reserve.⁴⁰ This reserve conveys the probability of any patient being successful on any item. Raw scores are transformed into odds of success to failure, with the natural log of this ratio estimating the difference between patient ability and

item difficulty. Both patient ability and item difficulty are expressed on the same logit scale (with a mean logit of zero).^{33,41} For the VR-GVFT, increasing person score (in logits) indicates poorer visual function.

The pilot VR-GVFT contained 52 items in total. Items 1 to 28 were stationary items with two parts: the first part was a “Yes/No” category and the second part was timed in seconds, which was only applicable if the patient answered “Yes” to part one. Items 29 to 40 were timed moving ball items and items 41 to 52 were timed driving items. Three separate Rasch analyses were performed for the three key item types: stationary, moving ball, and driving.

Infit and outfit statistics may be reported as a mean square (MNSQ) with an ideal fit of 1.⁴² An acceptable range for clinical observations is 0.50 to 1.70, which was set as the criterion in this study.^{43,44} The person separation statistic was used to illustrate how many strata of person ability an instrument can discriminate.³³ A low person separation statistic (< 2) implies the instrument may not be sensitive enough to distinguish effectively between differing levels of patient functioning, and frequently more items are required. Therefore, a minimum acceptable person separation for this study was set at 2.⁴⁵

Targeting reflects the matching of patient ability to item difficulty and can be graphically visualized with a person-item map. Poor targeting occurs when many patients have a higher or lower ability than the most or least difficult item. Traditionally, targeting can be assessed by comparing the mean patient and item values, ideally with a mean difference of zero. Significant mistargeting can be classified with differences > 1 .⁴⁶

Category collapsing is typically used in questionnaire development or validation.⁴⁷⁻⁴⁹ However, category collapsing violates the Rasch model, assuming uncollapsed data strictly follow the Rasch model. As the VR-GVFT is not a questionnaire but a test of competency at a specific task with a timed result, categories were not considered for collapse as each time in seconds is entirely possible.

Statistical Analysis of the VR-GVFT

Statistical Package for Social Sciences (SPSS Inc., Chicago, IL) for Windows (version 16.0, Microsoft Corp., Redmond, WA) was used for statistical analyses. The sample size calculation was based on the modelled standard error of the item calibration. The modelled standard errors are in the range: $2/[\sqrt{\text{sample size}}] < \text{standard error} < 3/[\sqrt{\text{sample size}}]$.

Rearranging this, one can calculate the sample size based on the standard error: $4/[(\text{standard error})^2] < \text{sample size} < 9/[(\text{standard error})^2]$. A sample size was calculated to achieve 95% confidence that any item calibration was within 0.5 logits from its modeled standard error, which equates to a required sample size in the range of 64 to 144.⁵⁰ Hence, the minimum acceptable sample size set for this study was 64. To account for subject drop-out, we aimed to recruit slightly greater numbers.

Demographic variables, BE VA and CS, BE visual field parameters (VFI, MD, and PSD), and VR-GVFT and GAL-9 (logit) scores were compared among mild, moderate, and severe glaucoma patient groups. Intergroup significance was assessed using analysis of variance (ANOVA) for parametric data and the χ^2 test for dichotomous variables.

Validity Evaluation

The following tests were used to validate the VR-GVFT: criterion, convergent, and divergent validity.

Criterion Validity. Criterion validity of the VR-GVFT descriptive system was assessed by evaluating the ability of the VR-GVFT person scores to distinguish between the following groups: mild, moderate, and severe glaucoma. Intergroup significance was assessed using ANOVA, with two-tailed *P*-value considered significant at <0.05 .

Convergent Validity. Convergent validity of the VR-GVFT descriptive system was assessed by exploring the correlation with the activity limitation indices and vision-related indices using the Pearson correlation coefficient.

Divergent Validity. Divergent validity of the VR-GVFT was assessed by evaluating the correlation between VR-GVFT person scores and factors (gender and sociodemographic data: marital status, employment status, and education level) that we hypothesized to have no correlation with VR-GVFT person scores.

Regression Analysis

A univariable regression analysis was performed to examine the association between VR-GVFT person scores (logit) and the following variables: age, gender, demographic data, driving status, employment status, BE MD, PSD, VFI, VA, CS, glaucoma severity, and GAL-9 (logit) scores. To evaluate for multicollinearity, the correlation between all variables was assessed using Spearman and Kendal tau b correlation tests for nonparametric data and Pearson's for parametric

data. All variables with *P*-value <0.05 were then included in a multivariate analysis. All variables were assessed for normality and linearity and transformed as required, and analysis of residuals was performed. Independence of residuals was assessed using the Durbin-Watson statistic. All tests were two-tailed with *P*-value <0.05 considered significant.

Results

The cohort consisted of 93 patients (54 mild, 22 moderate, and 17 severe glaucoma). Mean age was 67.4 (SD 13.2) years; 52.7% were male. The relative distribution of gender and demographic data did not differ significantly among groups (Table 1). The proportion of patients who always or frequently drove was lower in proportion in patients with severe glaucoma (29.4%) compared to other groups (42.9%–53.7%); however, it did not reach statistical significance. Visual field parameters (MD, PSD, and VFI), CS, and VA of either the BE and WE worsened with increasing glaucoma severity.

Rasch Analysis of the GAL-9 and VFQ-UI

The GAL-9 scores displayed good fit to the Rasch model, with no evidence of multidimensionality, ordered thresholds, and no differential item functioning or item misfit. Initially, targeting and scale precision were suboptimal (difference between person and item means 1.88 and person separation 1.95). To achieve satisfactory scale precision, patients with “perfect” scores were removed from the analysis. This was justifiable as unanchored persons with extreme scores provide no information for estimating item measures, and their removal substantially improves measurement precision.³⁵ After removing persons with “extreme” responses (i.e., those who responded “no difficulty” to all items, $n = 12$) from Rasch analysis of the GAL-9, targeting (1.47) and person separation (2.51) improved to acceptable levels.

Rasch analysis of the VFQ-UI indicated no multidimensionality, ordered thresholds, and no differential item functioning. Targeting was initially poor (the difference between person and item means was 3.0 and person separation was suboptimal at 1.2). After removing persons with “extreme” responses ($n = 33$), targeting (1.93) and precision (1.75) improved, but not to acceptable levels.

GAL-9 and VFQ-UI logit scores increased (i.e.,

Table 1. Participant Demographics and Clinical Characteristics Grouped According to Glaucoma Humphrey Field Analyzer Severity

Variable	Mild, <i>n</i> = 54	Moderate, <i>n</i> = 22	Severe, <i>n</i> = 17	<i>P</i> Value
Mean age, <i>y</i> (SD)	67.3 (11.9)	64.6 (17.5)	68.0 (11.1)	0.674
Female, <i>n</i> (%)	27 (50)	12 (54.5)	5 (29.4)	0.384
Demographic data, <i>n</i> (%)				
Caucasian	39 (72.2)	14 (63.6)	14 (82.4)	0.222
Tertiary education	18 (33.3)	6 (27.3)	5 (29.4)	0.756
Married/stable relationship	29 (53.7)	12 (54.5)	8 (47.1)	0.956
Retired/unemployed	35 (64.8)	23 (59.1)	11 (64.7)	0.889
Driving status, <i>n</i> (%)				
Always or frequently drives	29 (53.7)	10 (45.5)	5 (29.4)	0.368
BE logMAR VA mean (SD)	0.02 (0.18)	0.01 (0.13)	0.17 (0.31)	0.018
BE MD mean dB (SD)	-2.10 (2.32)	-5.54 (3.96)	-18.14 (6.03)	<0.001
BE PSD mean dB (SD)	2.53 (1.07)	5.25 (3.17)	9.90 (2.32)	<0.001
BE VFI mean % dB (SD)	96.2 (3.5)	87.7 (10.4)	51.9 (14.3)	<0.001
BE CS, log, mean (SD)	1.41 (0.24)	1.41 (0.31)	1.17 (0.34)	0.008
WE logMAR VA mean (SD)	0.14 (0.57)	0.52 (0.93)	0.95 (1.28)	0.002
WE MD mean dB (SD)	-5.59 (3.17)	-18.00 (5.42)	-22.68 (6.48)	<0.001
WE PSD mean dB (SD)	5.90 (3.04)	9.87 (2.39)	9.06 (2.52)	<0.001
WE VFI mean % dB (SD)	87.7 (8.37)	51.5 (19.1)	31.5 (16.4)	<0.001
WE CS, log, mean (SD)	1.32 (0.30)	1.06 (0.50)	0.74 (0.60)	<0.001
GAL-9 score (logit, SD)	-2.14 (1.43)	-1.33 (1.70)	0.06 (1.49)	<0.001
VFQ-UI utility value (SD)	0.89 (0.10)	0.84 (0.15)	0.80 (0.13)	0.043

dB, decibels.

activity limitation worsened) with increasing glaucoma severity (Table 1).

Rasch Analysis of the VR-GVFT

For the 28 stationary items, the initial person separation was 1.8, with four misfitting items. After removing misfitting items, person separation improved to 3.02 with no misfits, and targeting was 0.

For the 12 moving ball items, the initial person separation was 3.23, with one misfitting item. After removing the misfitting item, the person separation deteriorated to 3.05 with no misfitting items, and targeting was 0.

For the 12 driving items, the initial person separation was 1.41 and one item was found to

misfit. After removing the misfitting item, the person separation deteriorated to a suboptimal 1.39.

Hence, only the stationary and moving ball tests passed Rasch analysis. A Rasch-scaled scoring algorithm for the person scores of these two scales are available on request and provide a linear transformation of ordinal VR-GVFT data.

Statistical Analysis of the VR-GVFT

Criterion Validity

Stationary test person scores increased with greater glaucoma severity, with the average person scores for mild, moderate, and severe glaucoma being -1.45 ± 0.35 , -1.37 ± 0.31 , and -0.96 ± 1.20 , respectively (ANOVA $P = 0.014$) (Table 2).

Table 2. Person Scores (Logit) for Stationary Test and Moving Ball Test According to Glaucoma Humphrey Field Analyzer Severity

Variable	Mild, <i>n</i> = 54	Moderate, <i>n</i> = 22	Severe, <i>n</i> = 17	<i>P</i> Value
Stationary test (logit, SD)	-1.45 (0.35)	-1.37 (0.31)	-0.96 (1.20)	0.014
Moving ball test (logit, SD)	-1.28 (2.71)	-1.87 (2.57)	-0.64 (2.74)	0.364

Table 3. The Association Between Demographic, Clinical, and Patient-Reported Outcome Variables and Person Scores (Stationary Test) in Univariate and Multivariate Regression Models

Variable	B Coefficient	95% CI	R Statistic	F Statistic	P Value*
Univariate analysis					
GAL-9-person score	0.053	−0.027 to 0.133	0.146	1.712	0.195
VFQ-UI utility value	−1.539	−2.978 to −0.100	0.271	4.583	0.037
Age, older	0.010	0.001 to 0.020	0.222	4.726	0.032
Female	−0.112	−0.365 to 0.141	0.092	0.776	0.381
Tertiary education	−0.495	−0.905 to −0.085	0.203	5.746	0.043
Married/stable relationship	0.145	−0.107 to 0.397	0.119	1.303	0.257
Retired/unemployed	0.219	−0.041 to 0.478	0.173	2.805	0.097
BE MD	−0.021	−0.039 to −0.004	0.244	5.784	0.018
WE MD	−0.008	−0.017 to 0.001	0.190	3.074	0.083
BE PSD	0.045	0.009 to 0.081	0.250	6.046	0.016
WE PSD	−0.001	−0.023 to 0.022	0.005	0.002	0.967
BE central scotoma	0.336	0.074 to 0.598	0.258	6.487	0.013
WE central scotoma	0.062	−0.087 to 0.212	0.091	0.690	0.408
BE peripheral scotoma	0.241	−0.019 to 0.501	0.189	3.377	0.069
WE peripheral scotoma	0.149	−0.106 to 0.404	0.127	1.356	0.248
BE VA	0.497	−0.107 to 1.102	0.169	2.669	0.106
WE VA	0.224	0.086 to 0.362	0.321	10.46	0.002
BE CS	−0.546	−0.974 to −0.118	0.257	6.424	0.013
WE CS	−0.495	−0.745 to −0.244	0.381	15.43	<0.001
Multivariate analysis					
WE CS	−0.360	−0.710 to −0.010	0.553	3.627	0.044
Age	0.007	0.002 to 0.013			0.009

* $P < 0.05$ considered significant.

Moving ball test person scores were not significantly different between glaucoma severity groups.

Convergent Validity

For stationary test person scores, the Pearson correlation coefficient (R) with VFQ-UI utility values was 0.271 ($P = 0.037$), indicating a weak correlation with vision-related functional limitation. There was also weak to moderate correlation of person scores with BE MD ($R = 0.244$; $P = 0.018$), BE PSD ($R = 0.250$; $P = 0.016$), BE having central scotoma ($R = 0.258$; $P = 0.013$), poorer WE VA ($R = 0.321$; $P = 0.002$), BE CS ($R = 0.257$; $P = 0.013$), and WE CS ($R = 0.381$; $P < 0.001$) (Table 3).

For moving ball test person scores, the Pearson correlation coefficients with GAL-9 and VFQ-UI were poor ($R = 0.152$, $P = 0.173$ and $R = 0.062$, $P = 0.638$, respectively). There was a weak correlation with having a peripheral scotoma in the WE, but this did not reach statistical significance ($R = 0.199$, $P = 0.068$) (Table 4).

Divergent Validity

A weak correlation was detected between higher stationary test person scores and age ($R = 0.222$; $P = 0.032$) and having tertiary education ($R = 0.203$; $P = 0.043$). No significant correlation was found for gender ($R = 0.092$, $P = 0.381$) or marital status ($R = 0.119$, $P = 0.257$) (Table 3).

There was no significant correlation between moving ball test person scores and age, gender, education level, or marital status (Table 4).

Multivariate Analysis

On multivariate regression modelling, higher stationary test person scores were associated with having worse WE CS (regression coefficient [b], -0.360 ; 95% confidence interval [CI], -0.710 to -0.010 ; $P = 0.044$) and older age (b, 0.007 ; 95% CI, 0.002 to 0.013 ; $P = 0.009$). The variance (R^2) explained by the multivariate model was 0.306.

As none of the univariate correlations between moving ball test person scores and covariates had P -

Table 4. The Association Between Demographic, Clinical, and Patient-Reported Outcome Variables and Person Scores (Moving Ball) in a Univariate Regression Model

Variable	B Coefficient	95% CI	R Statistic	F Statistic	P Value*
Univariate analysis					
GAL-9-person score	0.250	−0.112 to 0.611	0.153	1.890	0.173
VFQ-UI utility value	−1.220	−6.379 to 3.938	0.062	0.224	0.638
Age, older	−0.020	−0.062 to 0.022	0.099	0.909	0.343
Female	−0.018	−1.134 to 1.098	0.003	0.001	0.975
Tertiary education	−0.292	−1.493 to 0.909	0.051	0.233	0.631
Married/stable relationship	−0.584	−1.694 to 0.525	0.109	1.095	0.298
Retired/unemployed	−0.644	−1.793 to 0.505	0.116	1.239	0.269
BE MD	−0.031	−0.111 to 0.048	0.082	0.614	0.435
WE MD	−0.039	−0.110 to 0.032	0.121	1.217	0.273
BE PSD	0.015	−0.149 to 0.179	0.019	0.033	0.857
WE PSD	−0.118	−0.294 to 0.057	0.146	1.791	0.189
BE central scotoma	−0.339	−1.528 to 0.851	0.059	0.320	0.573
WE central scotoma	−0.110	−1.272 to 1.052	0.021	0.036	0.851
BE peripheral scotoma	−0.980	−2.127 to 0.166	0.175	2.886	0.093
WE peripheral scotoma	−1.812	−3.761 to 0.138	0.199	3.416	0.068
BE VA	0.627	−2.067 to 3.320	0.048	0.213	0.645
WE VA	−0.407	−1.040 to 0.226	0.133	1.633	0.205
BE CS	−0.780	−2.719 to 1.159	0.083	0.639	0.426
WE CS	−0.167	−1.355 to 1.021	0.029	0.078	0.781

* $P < 0.05$ considered significant.

value <0.05 , multivariate regression modelling was not performed for the moving ball test person scores.

Discussion

To our knowledge, this proof of concept study is the first VR test designed to simulate visual function limitation related to glaucoma in both stationary and driving tasks. Of the 52 starting items, only the stationary and moving ball tests showed reasonable measurement.

Rasch-analyzed person scores for the stationary test showed good criterion validity; that is, the ability to differentiate between glaucoma severity groups. They also demonstrated reasonable convergence validity with mild to moderate correlation with VFQ-UI, BE MD, BE PSD, BE having central scotoma, WE poorer VA, and either eyes having reduced CS. Multivariate analysis showed that poorer CS in the WE is an independent factor for worse stationary test person scores. Divergence validity, however, was suboptimal, with worse stationary test person scores associated with increasing age; this finding is consistent with our previous study of the

CVGFT.²⁰ Slower response and reduced comprehension in older patients may explain their worse performance.

Rasch-analyzed person scores for the moving ball test showed that it is a valid measuring scale. However, there was poor criterion validity. A weak correlation with having a peripheral scotoma in the WE also did not reach statistical significance. The driving test failed Rasch analysis and hence is unsuitable in its current form as a measure for patient ability.

The lack of measurement validity for the moving ball and driving tests may be because it is unclear what the tests actually measure; we hypothesize that they may have been related to factors that are not measured, such as patterns of head and eye movements, or attributes unrelated to vision, such as familiarity with technology. Unfamiliarity with technology may have been compounded by the short length of videos (5–12 seconds).

The driving test may have involved tasks and commands that were too difficult for the patient to comprehend or perform to allow reliable and consistent measurements. For example, it required that the participant verbally indicate when they would stop/start driving. This may not correlate precisely

with reflex-type motor behaviors, such as braking when a driver sees a cyclist. Future testing may address these issues by using a pedal instead of verbal cues, varying the time for each task, and the use of preceding “practice” videos. As this is a pilot study, the results of these tests form the basis of an ongoing iterative process to optimize study measures for future research design.

Compared to the GAL-9 and VFQ-UI, stationary test person scores had a lower correlation with BE MD (correlation coefficient; -0.546 , -0.400 , and -0.244 , respectively). The strength of the GAL-9 and VFQ-UI result in part from serial refinement. Both were derived from extensively validated questionnaires whose psychometric qualities were refined through Rasch analysis.^{7,33} In comparison, this is the first proof of concept version of the VR-GVFT, which may be improved with future modifications. The VR-GVFT also has advantages over PROs. First, unlike traditional clinical parameters, it visually simulates real life scenarios. This is further enhanced by its integration of the eye and head movement that most activities require as part of the visual search to complete a task. The very realistic nature of the test may allow patients, clinicians, and policy makers to gain greater insights into the potential impact of glaucoma on daily visual function. Second, the VR-GVFT utilizes widely available smartphone technology and low-cost VR head-mounted goggles (current estimated average cost of US\$20). Both are easily portable, allowing for potential use in low-resource, rural, and remote clinical environments. These qualities also give it an advantage over physical simulations of visual function; unlike the ADREV test, the VR-GVFT is useable for patients with manoeuvrability issues, and requires less effort, equipment, and space to set up. Third, it is timed, allowing fine gradations of visual function to be detected. Lastly, the stationary test component correlates well with subjective measures of QoL, such as the VFQ-UI.

By simulating real-world tasks, VR is more likely to reflect the difficulties with day-to-day tasks experienced by each patient. Far more than a visual field test printout, this can heighten a patient’s understanding of their own visual disability and be used by caregivers, clinicians, and policymakers to identify personalized means of optimizing QoL, such as modifications to the home environment catered to a patient’s own vision-related challenges. It may also aid understanding in the doctor–patient relationship, which can influence treatment adherence—one of the key problems in glaucoma management today.⁵¹

This study has potential limitations. First, as this

was a prototype test designed specifically to evaluate visual dysfunction related to glaucoma, patients with coexisting ocular disease were excluded from this study. While this ensured we were measuring mainly glaucomatous visual dysfunction, extrapolation to patients with other ocular and nonocular co-morbidities that can affect vision is limited. Further studies evaluating the influence of these on VR-GVFT performance would be worthwhile.^{52,53}

Second, due to its nature as a pilot study, normal individuals without glaucoma were not included. Using such patients as controls would benefit future evaluation of this test.

Third, we had a larger proportion of mild glaucomatous cases. This is unfortunately a limitation when patients are recruited from clinical patient encounters, in that very advanced cases are generally rarer than moderate or mild cases of glaucoma. However, we feel this may better reflect the proportion of patients with glaucoma in real life.

Fourth, 24-2 Humphrey visual fields were primarily used in this study and this could miss field defects in the far periphery, which may have influenced patients’ performance on the VR-GVFT. However, routine clinical testing was prioritized during recruitment, and the 24-2 is a commonly used standard visual field assessment in the clinical setting and it is sufficient for the purpose of grading mild, moderate, and severe glaucoma in our study.⁵⁴

Fifth, smartphone-based goggles, while attempting to simulate real life, do not capture the precise daily visual challenges experienced by individual patients with glaucoma. Such tasks are impossible to precisely recreate and measure under the conditions of scientific study; like all clinical tests and PROs, the VR-GVFT is at best a potential sample of visual difficulties that might be experienced by patients with glaucoma. For instance, patients with glaucoma may develop compensatory responses in terms of eye or head movements to cope with limitations imposed by their scotoma; such compensation may mean their visual limitation may not be detected by visual challenges.⁵⁵

Lastly, gaze-tracking is not possible with Google Cardboard VR goggles. To minimize this, we monitored the visual experience of the patient from the administrator’s laptop. Gaze-tracking capability on future smartphone technology may lead to improvements in testing,⁵⁶ to both monitor gaze fixation and measure ability to navigate a simulated three-dimensional environment.

We have demonstrated that using readily accessible VR goggles and a structured objective test can



Figure 1. (A–C) Stationary scenes used for stationary test items. (D) Example driving test scene showing potential driving hazard (cyclist).

provide near real-world assessment of how glaucoma affects activities of daily living. Such testing may also help both clinicians and patients communicate how glaucoma may interfere in the practical, day-to-day world of patients. It could thus be an important intervention in patient and community education and a source of information for health policy. Further development is required to improve the precision of this test in glaucoma.

Acknowledgments

Centre for Eye Research Australia receives Operational Infrastructure Support from the Victorian Government (Australia).

Disclosure: **R.L.Z. Goh**, None; **Y.X.G. Kong**, None; **C. McAlinden**, None; **J. Liu**, None; **J.G. Crowston**, None; **S.E. Skalicky**, None

References

1. Quigley H, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol.* 2006;90:262–267.
2. Lorenzana L, Lankaranian D, Dugar J, et al. A new method of assessing ability to perform activities of daily living: design, methods and baseline data. *Ophthalmic Epidemiol.* 2009;16: 107–114.

3. Ramulu P. Glaucoma and disability: which tasks are affected, and at what stage of disease? *Curr Opin Ophthalmol*. 2009;20:92–98.
4. Lahav K, Levkovitch-Verbin H, Belkin M, Glovinsky Y, Polat U. Reduced mesopic and photopic foveal contrast sensitivity in glaucoma. *Arch Ophthalmol*. 2011;129:16–22.
5. Ramulu P, Maul E, Hochberg C, et al. Real-world assessment of physical activity in glaucoma using an accelerometer. *Ophthalmology*. 2012;119:1159–1166.
6. Nelson P, Aspinall P, Papasouliotis O, Worton B, O'Brien C. Quality of life in glaucoma and its relationship with visual function. *J Glaucoma*. 2003;12:139–150.
7. Goldberg I, Clement CI, Chiang TH, et al. Assessing quality of life in patients with glaucoma using the Glaucoma Quality of Life-15 (GQL-15) questionnaire. *J Glaucoma*. 2009;18:6–12.
8. Shabana N, Cornilleau Peres V, Carkeet A, Chew PT. Motion perception in glaucoma patients: a review. *Surv Ophthalmol*. 2003;48:92–106.
9. Haymes SA, Leblanc RP, Nicolela MT, Chiasson LA, Chauhan BC. Risk of falls and motor vehicle collisions in glaucoma. *Invest Ophthalmol Vis Sci*. 2007;48:1149–1155.
10. Burton R, Crabb DP, Smith ND, Glen FC, Garway-Heath DF. Glaucoma and reading: exploring the effects of contrast lowering of text. *Optom Vis Sci*. 2012;89:1282–1287.
11. Chan EW, Chiang PP, Liao J, et al. Glaucoma and associated visual acuity and field loss significantly affect glaucoma-specific psychosocial functioning. *Ophthalmology*. 2015;122:494–501.
12. McKean-Cowdin R, Wang Y, Wu J, et al. Impact of visual field loss on health-related quality of life in glaucoma: the Los Angeles Latino Eye Study. *Ophthalmology*. 2008;115:941–948.
13. Medeiros FA, Gracitelli CP, Boer ER, Weinreb RN, Zangwill LM, Rosen PN. Longitudinal changes in quality of life and rates of progressive visual field loss in glaucoma patients. *Ophthalmology*. 2015;122:293–301.
14. Skalicky S, Goldberg I. Depression and quality of life in patients with glaucoma: a cross-sectional analysis using the Geriatric Depression Scale-15, assessment of function related to vision, and the Glaucoma Quality of Life-15. *J Glaucoma*. 2008;17:546–551.
15. Mabuchi F, Yoshimura K, Kashiwagi K, et al. High prevalence of anxiety and depression in patients with primary open-angle glaucoma. *J Glaucoma*. 2008;17:552–557.
16. Altangerel U, Spaeth GL, Steinmann WC. Assessment of function related to vision (AFREV). *Ophthalmic Epidemiol*. 2006;13:67–80.
17. Richman J, Lorenzana LL, Lankaranian D, et al. Importance of visual acuity and contrast sensitivity in patients with glaucoma. *Arch Ophthalmol*. 2010;128:1576–1582.
18. Warrian KJ, Katz LJ, Myers JS, et al. A comparison of methods used to evaluate mobility performance in the visually impaired. *Br J Ophthalmol*. 2015;99:113–118.
19. Kowalski JW RA, Walt JG, Lloyd A, et al. Rasch analysis in the development of a simplified version of the national eye institute visual-function questionnaire-25 for utility estimation. *Qual Life Res*. 2012;21:323–334.
20. Skalicky SE, McAlinden C, Khatib T, et al. Activity limitation in glaucoma: objective assessment by the Cambridge Glaucoma Visual Function Test. *Invest Ophthalmol Vis Sci*. 2016;57:6158–6166.
21. Riva G, Wiederhold BK. The new dawn of virtual reality in health care: medical simulation and experiential interface. *Stud Health Technol Inform*. 2015;219:3–6.
22. Foerster RM, Poth CH, Behler C, Botsch M, Schneider WX. Using the virtual reality device Oculus Rift for neuropsychological assessment of visual processing capabilities. *Sci Rep*. 2016;6:37016.
23. Diniz-Filho A, Boer ER, Gracitelli CP, et al. Evaluation of postural control in patients with glaucoma using a virtual reality environment. *Ophthalmology*. 2015;122:1131–1138.
24. Wroblewski D, Francis BA, Sadun A, Vakili G, Chopra V. Testing of visual field with virtual reality goggles in manual and visual grasp modes. *Biomed Res Int*. 2014;2014:206082.
25. Matsumoto C, Yamao S, Nomoto H, et al. Visual field testing with head-mounted perimeter 'imo'. *PLoS ONE*. 2016;11:e0161974.
26. Anderson DR. *Automated Static Perimetry*. St Louis, MO: Mosby; 1992:123–133.
27. Keltner JL, Johnson CA, Cello KE, et al. Visual field quality control in the Ocular Hypertension Treatment Study (OHTS). *J Glaucoma*. 2007;16:665–669.
28. Sawada H, Yoshino T, Fukuchi T, Abe H. Assessment of the vision-specific quality of life using clustered visual field in glaucoma patients. *J Glaucoma*. 2014;23:81–87.

29. Dev MK, Paudel N, Joshi ND, Shah DN, Subba S. Impact of visual impairment on vision-specific quality of life among older adults living in nursing home. *Curr Eye Res.* 2014;39:232–238.
30. Jampel H, Friedman DS, Quigley H, Miller R. Correlation of the binocular visual field with patient assessment of vision. *Invest Ophthalmol Vis Sci.* 2002;43:1059–1067.
31. Zhou C, Qian S, Wu P, Qiu C. Quality of life of glaucoma patients in China: sociodemographic, clinical, and psychological correlates—a cross-sectional study. *Qual Life Res.* 2014;23:999–1008.
32. Arora KS, Boland MV, Friedman DS, Jefferys JL, West SK, Ramulu PY. The relationship between better-eye and integrated visual field mean deviation and visual disability. *Ophthalmology.* 2013;120:2476–2484.
33. Khadka J, Pesudovs K, McAlinden C, Vogel M, Kernt M, Hirneiss C. Reengineering the glaucoma quality of life-15 questionnaire with rasch analysis. *Invest Ophthalmol Vis Sci.* 2011;52:6971–6977.
34. Goh RL, Fenwick E, Skalicky SE. The visual function questionnaire: utility index: does it measure glaucoma-related preference-based status? *J Glaucoma.* 2016;25:822–829.
35. Linacre JM. *A User's Guide to Winsteps/Ministeps Rasch-Model Programs.* Chicago, IL: MESA Press; 2005.
36. Mallinson T. Why measurement matters for measuring patient vision outcomes. *Optom Vis Sci.* 2007;84:675–682.
37. Lamoureux E, Pesudovs K. Vision-specific quality-of-life research: a need to improve the quality. *Am J Ophthalmol.* 2011;151:195–197.
38. Rentz AM, Kowalski JW, Walt JG, et al. Development of a preference-based index from the National Eye Institute Visual Function Questionnaire-25. *JAMA Ophthalmol.* 2014;132:310–318.
39. Google, Inc. Google Cardboard Technical Specification. Google, Inc.; 2015.
40. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen, Denmark: Danish Institute for Educational Research; 1960.
41. Khadka J, Gothwal VK, McAlinden C, Lamoureux EL, Pesudovs K. The importance of rating scales in measuring patient-reported outcomes. *Health Qual Life Outcomes.* 2012;10:80.
42. McAlinden C, Khadka J, de Freitas Santos Paranhos J, Schor P, Pesudovs K. Psychometric properties of the NEI-RQL-42 questionnaire in keratoconus. *Invest Ophthalmol Vis Sci.* 2012;53:7370–7374.
43. Wright BD, Linacre JM, Gustafsson JE, Martin-Loff P. Reasonable mean-square fit values. *Rasch Meas Trans.* 1994;8:370.
44. McAlinden C, Jonsson M, Kugelberg M, Lundstrom M, Khadka J, Pesudovs K. Establishing levels of indications for cataract surgery: combining clinical and questionnaire data into a measure of cataract impact. *Invest Ophthalmol Vis Sci.* 2012;53:1095–1101.
45. Gothwal VK, Wright TA, Lamoureux EL, Khadka J, McAlinden C, Pesudovs K. Improvements in visual ability with first-eye, second-eye, and bilateral cataract surgery measured with the visual symptoms and quality of life questionnaire. *J Cataract Refract Surg.* 2011;37:1208–1216.
46. Mallinson T, Stelmack J, Velozo C. A comparison of the separation ratio and coefficient alpha in the creation of minimum item sets. *Med Care.* 2004;42:I17–24.
47. Khadka J, McAlinden C, Pesudovs K. Quality assessment of ophthalmic questionnaires: review and recommendations. *Optom Vis Sci.* 2013;90:720–744.
48. McAlinden C, Pesudovs K, Moore JE. The development of an instrument to measure quality of vision: the Quality of Vision (QoV) questionnaire. *Invest Ophthalmol Vis Sci.* 2010;51:5537–5545.
49. McAlinden C, Skiadaresi E, Gatinel D, Cabot F, Huang J, Pesudovs K. The Quality of Vision questionnaire: subscale interchangeability. *Optom Vis Sci.* 2013;90:760–764.
50. Linacre JM. Sample size and item calibration stability. *Rasch Measurement Transactions.* 1994;7:328.
51. Friedman DS, Hahn SR, Gelb L, et al. Doctor-patient communication, health-related beliefs, and adherence in glaucoma results from the Glaucoma Adherence and Persistency Study. *Ophthalmology.* 2008;115:1320–1327.
52. Musch DC, Gillespie BW, Niziol LM, et al. Cataract extraction in the collaborative initial glaucoma treatment study: incidence, risk factors, and the effect of cataract progression and extraction on clinical and quality-of-life outcomes. *Arch Ophthalmol.* 2006;124:1694–1700.
53. Skalicky SE, Martin KR, Fenwick E, Crowston JG, Goldberg I, McCluskey P. Cataract and quality of life in patients with glaucoma. *Clin Exp Ophthalmol.* 2015;43:335–341.

54. Nelson P, Aspinall P, Papasouliotis O, Worton B, O'Brien C. Quality of life in glaucoma and its relationship with visual function. *J Glaucoma*. 2003;12:139–150.
55. Smith ND, Crabb DP, Garway-Heath DF. An exploratory study of visual search performance in glaucoma. *Ophthalmic Physiol Opt*. 2011;31:225–232.
56. Smith ND, Glen FC, Crabb DP. Eye movements during visual search in patients with glaucoma. *BMC Ophthalmol*. 2012;12:45.