The Graded Incomplete Letters Test (GILT): a rapid test to detect cortical visual loss, implemented in the UK Biobank

Running head: GILT: a short test to detect cortical visual loss

KXX Yong[1], A Petzold[1,2,3,4], P Foster[2,3], A Young[5], S Bell[6], AP Leff[1], S Crutch[1], JA Greenwood[7]

[1]Queen Square Institute of Neurology, University College London, London, United Kingdom

[2]Institute of Ophthalmology, University College London, London, London, United Kingdom

[3]Moorfields Eye Hospital NHS Foundation Trust, London, London, United Kingdom

[4]The National Hospital for Neurology and Neurosurgery, Queen Square, WC1N 3BG, UK & Neuro-ophthalmology Expertise Centre, Amsterdam UMC, NL.

[5]Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, University of Oxford, Oxford, UK

[6]UK Biobank, Stockport, UK

[7]Experimental Psychology, University College London, London, United Kingdom

Corresponding authors

Dr. Keir Yong

Dementia Research Centre, UCL Institute of Neurology Email: keir.yong@ucl.ac.uk

Dr. John Greenwood

Experimental Psychology, UCL: john.greenwood@ucl.ac.uk

Abstract

Higher order impairments of object recognition are core features of neurodegenerative syndromes, in particular posterior cortical atrophy (PCA), the 'visual-variant of Alzheimer's disease'. These impairments arise from damage to cortical visual regions and are often missed or misattributed to one of the more common ophthalmological conditions. Consequently, the diagnosis can be delayed for years with considerable implications for patients. Here we report a new test, the Graded Incomplete Letters Test (GILT), for rapid measurement of cortical visual loss and differentiation of ocular and cortical causes of visual loss. The GILT is an optimised psychophysical variation of a test used to diagnose cortical visual impairment which measures thresholds for recognising letters under levels of increasing visual degradation (decreasing completeness). Consistent with other ophthalmic tests like acuity or contrast sensitivity charts, thresholds are based on performance scored "by line" (cut-off thresholds) and "by letter" (letter-based thresholds). The GILT was administered to UK Biobank volunteers (total n= 2,359) and participants with neurodegenerative conditions characterised by predominant cortical visual (PCA n=10) or memory loss (typical Alzheimer's disease n=9). Findings suggest UK Biobank participants, including those with documented ophthalmological conditions, were able to recognise letters under low levels of completeness. However, participants with PCA exhibited a particular tendency to make errors with modest decreases in completeness, while performance in participants with mostly mild, typical Alzheimer's disease was comparable to that of the UK Biobank sample. Using letter-based thresholds for participants reaching <80% accuracy cut-off, GILT sensitivity and specificity values were 90% and 91% respectively when comparing PCA to UK Biobank participants without or with documented visual conditions (total n without=1757; n with=194). These first release UK Biobank and pilot patient data suggest the GILT is specific to cortical visual loss and PCA. The GILT may have utility in screening for visual loss owing to posterior cortical damage and differentiating cortical visual loss from common eye-related conditions.

Introduction

In routine clinical practice, assessments of visual functions are geared towards evaluating the eye. However, you 'see' things with your brain. Visual agnosias (from Greek: 'not knowing') are the brain-based disorders of perception. Agnosias arise from damage to cortical visual regions following stroke, anoxia or other aetiologies, and are core features of focal neurodegenerative conditions[1–5], in particular posterior cortical atrophy (PCA; the 'visual-variant of Alzheimer's')[6]. Distinctions between agnosias, particularly regarding profiles of cortical visual deficits, have been used extensively across research and clinical contexts for over a century[7], corroborated by the emergence of subsequent neuroimaging and pathologic investigations[1,5,8]. Accurate measurement of these cortical visual deficits is important given their clinical implications[9], from the need to establish their cause to informing treatment and managing associated disability. It is therefore concerning that these deficits are often overlooked or missed for many years. For example, individuals with PCA typically have multiple appointments with optometrists and ophthalmologists before being referred to specialised neurology or dementia specialist services[10,11]. The average time between symptom onset and formal diagnosis in PCA has been reported as 3-4 years compared to 2 years in late-onset dementia[12], precluding opportunities for timely management and treatment.

The various forms of agnosia can be distinguished based on their impairment of distinct levels of the visual processing hierarchy that subserves object recognition. Agnosias cannot be attributed to deficits at the lowest levels of visual processing, such as diminished visual acuity or field defects. Instead, visual agnosias have been attributed to failures of object recognition at two higher levels – (ap)perceptive and associative[13]. Apperceptive agnosia broadly reflects visual degradation, with impaired integration of form and feature information precluding the conscious perception of objects and scenes. Associative agnosia instead reflects impaired access to or degradation of semantic knowledge, with a loss of meaning associated with the object representation despite the preserved ability to consciously perceive objects. For example, someone with apperceptive agnosia may have difficulty matching objects presented from different angles and misperceive visual features (e.g. responding 'wheel'/'wire' to a photograph of a whisk), yet recognise the object based on tactile or verbal cues. Someone with associative agnosia may be able to match such objects but mislabel them as semantically related but visually dissimilar objects (e.g. reporting 'spoon' instead of whisk).Within this distinction, core features of PCA fall into the apperceptive category, while associative agnosia may arise in semantic dementia despite otherwise intact perception.

PCA is the most common atypical AD clinical phenotype characterised by predominant cortical visual relative to memory loss, comprising ~10% of AD patients at specialist centres[14]. While most commonly underpinned by AD pathology, the PCA syndrome may arise from Lewy body pathology, and rarely frontotemporal lobar degeneration with tau or TDP-43 inclusions and other neurodegenerative disease[15]. Visual deficits associated with PCA include an elevation in visual crowding[16] (the disruptive effect of clutter on object recognition[17]), deficits in figure-ground segmentation and shape discrimination, simultanagnosia (an inability to perceive multiple objects at once), and "partonomic" errors (where local features are identified at the expense of the global

object/form), amongst others[18]. The age of onset of PCA is typically earlier than other forms of AD, around 50-65 years[19]. As the disease progresses, symptoms in the various subtypes of AD increasingly converge, those with PCA acquire memory and linguistic difficulties, while those with typical AD can similarly acquire visual deficits in addition to their earlier memory, executive and language deficits.

Various tests exist to evaluate, screen and diagnose cortical visual deficits[20–22] arising from traumatic brain injury, structural damage to the brain (cancer, metastatic disease), stroke[23] or neurodegenerative conditions such as PCA. These tests often use visually degraded conditions like unconventional orientations, silhouettes, and overlapping or 'fragmented' formats to evaluate deficits in object perception corresponding to the apperceptive level described above[20–22,24,25]. Amongst these, a frequently used diagnostic test of object recognition under visually degraded conditions is the recognition of incomplete letters[26]. Impairments in incomplete letter recognition have been well-documented following brain lesions and neurodegenerative disease. Accordingly, incomplete letter stimuli are frequently used in standard measures within dementia clinical and research settings[20–22,24], including in the diagnosis of PCA[19,27,28]. Incomplete letter recognition dissociates with various lower-level deficits, including visual field defects and diminished figure ground and shape discrimination[18,29,30], but has been associated with the increases in visual crowding[31] associated with PCA. Common clinical manifestations of difficulty with degraded letter forms includes struggling reading digital signs or clocks and recognising fragmented visual test stimuli (e.g. failing dotted Ishihara test plate unrelated to unaffected colour vision)[15]. In the context of the dual theory of visual streams, while such deficits are often considered 'ventral' in nature (the 'what?' stream), impaired performance on such tasks has also been documented in patients with right parietal lesions[26,29,30,32] overlapping with 'dorsal' functions (the 'where/how?' stream)[7] (for a critical review of the two systems theory, see[33]) . Incomplete letter recognition is estimated to become abnormal early on in PCA and at intermediate stages of typical AD[8], consistent with the general convergence of symptoms over time in AD, as above. Impairments on incomplete letters in typical AD are associated with a younger age at onset, likely owing to parieto-occipital atrophy[34,35]. Notably, the Incomplete Letters Test has been recommended by eye and neurology professionals to distinguish ocular/optic deficits from cortical visual deficits[27].

Current versions of incomplete letter tasks are however susceptible to particular limitations, especially in dementia clinical and research settings. Firstly, a limitation of the current version of the test is that letters are presented only with a single level of degradation (e.g. 30% complete[20]), with performance measured as percent-correct recognition. This creates a susceptibility to ceiling effects[8] – a general limitation of many routine visual measures[36] – and limits the sensitivity of the test in tracking disease progression. Secondly, their use in clinical practice is essentially restricted to highly specialised professionals in neurology and neuro-ophthalmology, while people with PCA are most commonly initially seen by eye health professionals. In research settings, their use is often restricted to specialised test batteries, with generic batteries featuring few, if any visual measures[36]. Thirdly, in the context of neurodegenerative disease there is mixed evidence regarding the disease-specificity of the impairments in this test. Mixed findings of impaired incomplete letter recognition reflecting AD pathology[37] or mixed pathology[38] have prompted recommendations to better differentiate neurodegenerative conditions, particularly PCA, by evaluating both intact and incomplete letter recognition. Finally, outside professional recommendations[27], there is limited empirical evidence on whether these tests can differentiate cortical from ocular visual deficits (like glaucoma), which are also prevalent in older adults.

Given these gaps in test sensitivity, specificity and utility, and following patient and professional consultation to improve the diagnosis of cortical vision loss[27], we present a novel variation of this test, the rapid Graded Incomplete Letters Test (GILT), to detect cortical vision loss in agnosia and neurodegenerative conditions. We present preliminary normative data from the UK Biobank re-imaging study[39] and compare patients with predominant cortical visual (PCA) or memory loss (typical AD) arising from neurodegenerative disease. The test optimises assessment through psychophysical techniques to measure thresholds for the identification of letters affected by visual degradation (or low completeness). Rather than tests measuring thresholds using letters under varying contrast or brightness, the GILT uses letters which become progressively less complete on a digital interface. The test is designed to be short (<3 minutes), to minimise ceiling/floor, order and letter effects and to enable the sensitive detection of cortical visual abnormalities.

Methods

Participants

Participants from UK Biobank (n=2,359) and the UCL Dementia Research Centre (n=19; 10 with PCA and 9 with typical AD) were administered the GILT. UK Biobank volunteers were administered a version using touchscreen at biobank visits (GILT-UKB), while UCL patient participants were administered the test using a portable laptop for home testing. A number of UK Biobank volunteers had documented conditions which may affect vision (cataract n=109; amblyopia n=88; glaucoma n=60; stroke n=21; low vision [<6/12 acuity] n=4). UCL patients had varying degrees of cortical visual loss which could not be attributed to ophthalmological conditions, stroke or tumour, consistent with clinical diagnoses. See Table 1 for participant demographic and clinical information.

UK Biobank: UK Biobank is a population-based prospective cohort study of >500,000 volunteers aged 40-69 years recruited between 2006-2010 (https://www.ukbiobank.ac.uk). Participants completed a touchscreen questionnaire, cognitive testing, verbal interview, and physical examination and provided biological samples. Ethics Committee approval for UK Biobank was obtained from the North West Multi-Centre Research Ethics Committee (Research Ethics Committee reference: 16/NW/0274). The GILT-UKB was administered to volunteers within the UK Biobank Imaging sub-study; the first release data are presented here.

UCL Dementia Research Centre: pilot testing participants had a diagnosis of PCA or typical AD, and fulfilled consensus criteria for PCA (9 PCA-pure, 1 PCA-plus) and research criteria for probable AD respectively[6,40,41]. PCA and typical AD participants were of comparable age and disease severity; all available molecular pathology was consistent with AD pathology (Table 1). Prior ethical approval for the study was provided by the National Research Ethics Service Committee London Queen Square and informed consent obtained from all participants according to the Declaration of Helsinki.

Stimuli and procedures

The Graded Incomplete Letters Test was developed from the Incomplete Letters subtest from the Visual Object and Space Perception Battery (VOSP)[20]. The VOSP subtest involves the identification of 20 black letters on a white background, visually degraded via random blocks of fragmentation (white sections removed from the black letter) with a fixed black:white ratio of 30:70. The GILT optimises the sensitivity of this test to detect cortical visual abnormalities by adding a range of completeness levels, presented using a modified method of limits procedure with forced-choice responses. For each trial, participants were asked to select the response letter which matches a target letter which progressively decreases in completeness. See Figure 1 for example instructions and trial from the version featured within the UK Biobank study (GILT-UKB).

In detail, all stimuli were presented from a typical viewing distance (approximately 50cm). On each trial, a single uppercase target letter was presented. This item-by-item presentation was selected over a chart format (as in standard visual acuity or contrast sensitivity assessment) to reduce disruption caused by adjacent letters[42], a particular problem for individuals with cortical visual deficits

such as excessive crowding[15]. Target letters were presented in the Sloan font[43]. This font is a standard for acuity testing due to the fixed proportions of the letters – the stroke width is one fifth the letter diameter and matched for all letter features. In this way, the effect of degradation on the features within letters was equated across the various letterforms. Letters were presented with a diameter of 275 pixels, approximately 8.3 degrees of visual angle (100 minutes of arc per letter stroke). Given this corresponds to logMAR 2.0 (Snellen equivalent: 20/2000), acuity limitations on performance would be unlikely. Letters were black-on-white at 100% Weber contrast, further ensuring that any contrast sensitivity losses were unlikely to limit performance.

On each trial, the target was one of 12 distinct uppercase letters (C, D, E, F, H, K, N, P, R, U, V, Z). These letters were selected to be the same as in the UK Biobank visual acuity testing, based on a standard logMAR chart. Note that this set is expanded from the original Sloan letters, similar to other expanded sets[44]. Participants were asked to indicate which of these 12 uppercase letters is presented on each trial using a set of lowercase response items (Figure 1). This forced-choice response has a sufficient number of options to minimise the impact of correct guesses on the threshold estimates[45], whilst also allowing for response options to be presented to participants simply (compared with the use of the full alphabet, for instance). Response options were presented to participants as two rows of six lower case letters, presented at the bottom of a touchscreen (ELO 1715L 17”, 1280x1024). Target and response letters used differing case and font to preclude strategies relying on letter matching rather than recognition. UCL Dementia Research Centre participants were administered the GILT-UKB with the following adjustments to allow for patient in-home assessments: the task was administered using a portable laptop (Dell Latitude 5500 16”, 1280x1024; presenting stimuli at comparable diameter of approximately 7.8 degrees of visual angle) on which the tester used the touchscreen to register patient verbal responses.

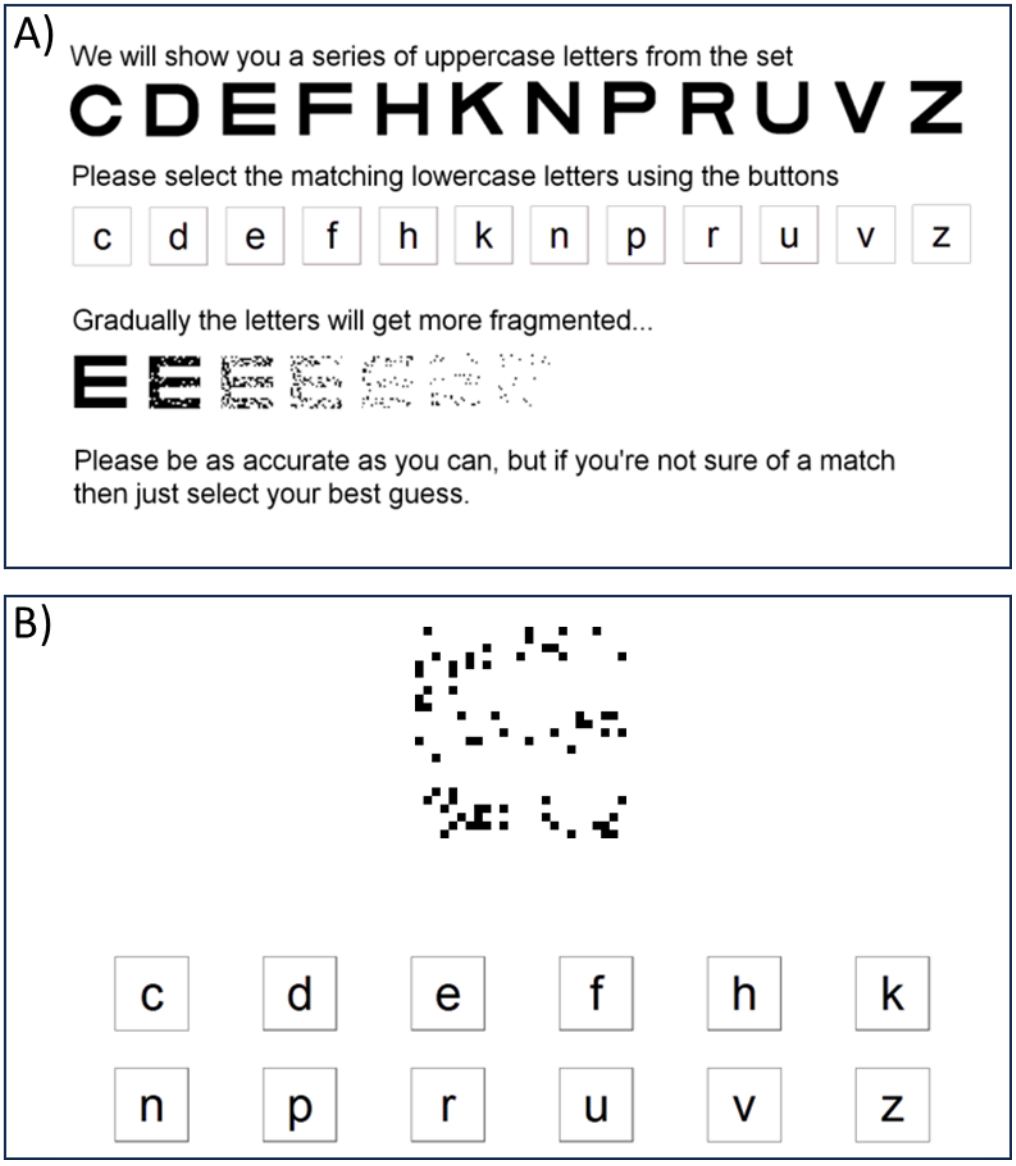

Figure 1 Examples of GILT A) instructions from UK Biobank, B) target uppercase and response lowercase items (target letter: E).

Targets were presented under a total of nine completeness levels, starting with 100% complete and decreasing with 8 subsequent log-spaced levels (each separated by 0.25 log units to give 89.13, 50.12, 28.18, 15.85, 8.91, 5.01, 2.82, and 1.58% completeness). In this way, six of nine levels were of lower completeness than the standard VOSP items (fixed at 30% complete). The incremental fragmentation (Figure 1A) allowed for efficient measurement of letter-identification thresholds, whereby each participant quickly approached the level of completeness sufficient to induce errors, consistent with the design of other tests of visual function (e.g. acuity and Pelli-Robson contrast charts[46,47]). The percentage completeness corresponded to the number of black pixels within each letter divided by the number of black pixels within the complete letter image. Letters of varying completeness were generated using the following steps. Firstly, random noise images were

generated with the same size as the letter image (275×275 pixels) with a Gaussian distribution of grey levels. Greyscale noise images were binarized to be black or white 'checks', and scaled up such that each check within the noise image was one fifth the size of the letter stroke (i.e. each stroke was the width of 5 checks). The mask for each letter was then applied, such that checks outside the letter boundary were removed. By shifting the mean luminance of the greyscale noise image prior to binarization, different levels of completeness could be achieved. This process was repeated iteratively until the final image reached the desired level of completeness.

Trials were presented with a modified method of limits procedure, beginning at 100% completeness and decreasing, again following a similar design to other visual tests[46,47]. Each completeness level was presented in blocks of 5 trials (similar to the 5-letter lines in acuity charts), for a maximum of 45 trials per participant (5 targets under up to nine completeness levels). The task was discontinued when participants reach a pre-specified accuracy level within the current completeness level. For the GILT-UKB version, the accuracy cut-off was taken at 60% correct (i.e. 3/5 correct within the completeness level) or below. The task was discontinued either at this completeness level, or after a maximum time of 180 seconds from task onset. For UCL Dementia Research Centre participants there was no time limit for the task. To control for stimulus order and letter effects, four testing sets with distinct check patterns and pseudorandomised letter order were randomly assigned to each participant. For each testing set, target letters were randomly assigned to each of the 5-letter blocks (i.e. the letters at each completeness level). Both letter order (within block) and the letters selected within the block (from the 12 possible options) were arranged pseudorandomly so that each letter never appeared twice in consecutive trials, and always appeared in every 20 consecutive trials. Each of the four testing sets was generated as a different image set comprising the above completeness levels, each generated with a distinct distribution of checks.

GILT measures

The primary GILT outcome measures were completeness thresholds for letter identification – the lowest completeness level at which letters can be identified. To determine the best practice for obtaining these thresholds, we compared several measurement approaches.

Cut-off thresholds: A common approach in ophthalmic testing is to score performance "by line" (i.e. each difficulty/completeness level), and to take the threshold as the line at which a desired accuracy level is reached[43]. As above, the GILT-UKB was run with a minimum accuracy cut-off of 60%, meaning that thresholds can be taken as the highest completeness level at which at least 3/5 letters are correctly identified. Because higher thresholds are typically taken in ophthalmic practice[46], we also calculated thresholds with accuracy cut-offs of 100% and 80% - the highest completeness level at which at least 5/5 and 4/5 letters are correctly identified. We refer to these as cut-off thresholds, bounded at 100% and 1.58% (highest and lowest completeness level).

Letter-based thresholds: One issue with the above cut-off, or "line-based" measurements is that the resolution of the resulting thresholds is limited to the specific difficulties tested. A common approach to increase this resolution in visual acuity and contrast sensitivity testing is to use "by letter"

scoring[48]. Here, responses to each letter contribute to the threshold estimate. Because our completeness levels are log spaced with 0.25 log units of completeness between them, the correct response to each letter can be considered to contribute 0.05 log units of *sensitivity* (the inverse of threshold, since high sensitivity yields a low threshold). Participants start with a sensitivity of 0 and add 0.05 log units with each correct response, excluding the first 5 trials with 100% complete letters (i.e. beginning from 89.13% complete). Scoring is terminated when the accuracy cut-off is reached (which we calculated with 100%, 80% and 60% cut-offs, as above). To compute the threshold *t* from this value of sensitivity *s*, we convert back from the logarithm and take the inverse, with $t = \left(\frac{1}{10^s}\right) \times 100$. We refer to these as letter-based thresholds, bounded at 100% (sensitivity of 0) and 1% (40 trials x 0.05 log units= sensitivity of 2).

GILT-UKB acquired the above measures with a number of variables relevant to task performance (total number of correct responses; selected response letter; total duration of test; duration between stimulus presentation and selected response letter; duration between stimulus presentation and trial end; reason test concluded; history of actions).

ROC curves were used to investigate the ability of GILT response thresholds to differentiate PCA from UKB participants without documented visual conditions (low vision [<6/12 acuity], cataract, glaucoma, amblyopia) or stroke. Discriminatory ability was assessed using logistic regression models relating GILT response thresholds to odds of PCA (PCA vs UKB) fitted with Firth's penalized likelihood method to reduce small-sample bias.

Results

GILT performance: accuracy

See Figure 2 for distribution of percentage accuracy in UK Biobank (UKB) participants regardless of diagnosis. The majority of participants in UKB reached accuracy cut-offs within the GILT-UKB time limit of 120s (<100% accuracy: n=2,300 (97%); <80% accuracy: n=1,967 (83%); <60% accuracy: n=1,421 (60%)). All PCA and typical AD participants reached accuracy cut-offs as there was no time limit for administration. For all analyses, between-group comparisons were only made between participants who reached each threshold; in this way, we avoided comparing participants who simply ran out of time at a certain completeness level (i.e. accurate, but slow) with those who reached accuracy cut-offs (i.e <100%; <80%; <60%).

GILT mean percentage accuracy was 87.3% SD=8.1 for UK Biobank participants regardless of diagnosis (n=2,359) and 69.7% SD=9.5 in PCA and 83.4% SD=3.5 in typical AD participants. Mean percentage accuracy was above 77.5% in 95% of UKB participants, including those with visual conditions (low vision [visual acuity <6/12] and/or presence of cataract, glaucoma, amblyopia) and stroke. Only two PCA participants were just within this normal range for percentage accuracy based on the total UKB sample (5^{th}%ile: 77.5%), achieving 80% accuracy (PCA range: 50.0-80.0%). In contrast, all but one typical AD participant was within this normal range (typical AD range: 76.0-86.7%). See Supplementary Figure 1 for further details on percentage accuracy in UKB, PCA and typical AD participants.

Table 1 Total UK Biobank and UCL sample demographic and clinical information.

	GILT-UKB (n=2,359)	GILT (n=19)	
Sample	UK Biobank	UCL	
Diagnoses	Cataract n=109; Amblyopia n=88; Glaucoma n=60; Stroke n=21*	PCA n=10	Typical AD n=9
Age (years)	66.8 ± 7.0	69.9 ± 7.3	65.0 ± 8.5
Sex (male:female)	1190:1169	8:2	3:6
β-Amyloid PET/ CSF consistent with AD	-	4/4	7/7
MMSE	-	21.2 ± 4.5	24.0 ± 6.0

CSF: Cerebrospinal fluid, MMSE: Mini-mental state examination *Unique n without diagnoses or low vision =2,094

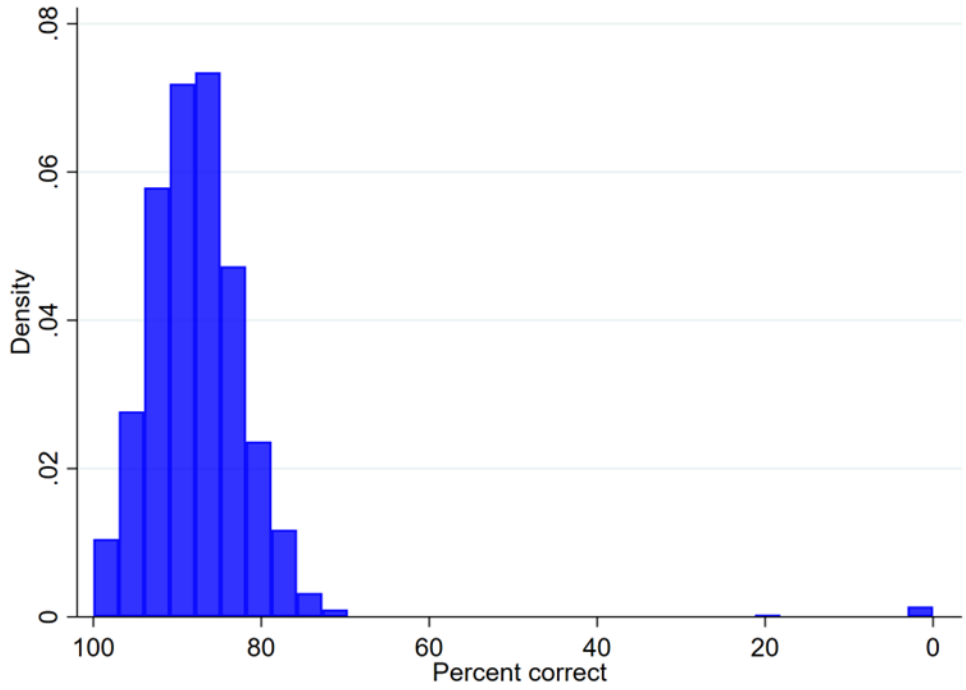

Figure 2 GILT accuracy (percent correct) in the total UK Biobank sample.

GILT Primary outcome: completeness thresholds

Primary GILT outcome measures are completeness thresholds – cut-off or letter-based – for letter recognition[48].

Here we report letter-based thresholds at <80% accuracy. We report letter-based thresholds owing to their improved resolution compared to cut-off thresholds. Cut-off thresholds were restricted to the nine pre-specified completeness levels, while letter-based thresholds were more granular, taking up to 28 values in the UKB sample. We report thresholds at <80% (i.e. making at least two errors on a completeness level) rather than <100% or <60% accuracy. Beyond consistency with 80% thresholds commonly used in ophthalmic practice, the higher 100% threshold was considered unsuitable as these return an infinite number of values at 100%. The 80% rather than 60% threshold was preferrable as only n=1,421 (60%) of the total UKB sample reached <60% accuracy (making at least 3 errors on a completeness level) within the UKB time limit of 120s. See Supplementary Figure 2 for comparisons of cut-off and letter-based thresholds in UKB, PCA and typical AD participants.

See Figure 3 for letter-based thresholds at <80% accuracy. Letter-based thresholds were defined at the level of response - incrementally increasing accuracy per correct response until reaching cut-off thresholds of <80% accuracy. The same UKB participants whose accuracy was at floor had letter-based thresholds at floor and no participant reached ceiling. In the UCL sample, these letter-based thresholds ranged from 6.3-100% in PCA and 4.5-23.4% in typical AD participants.

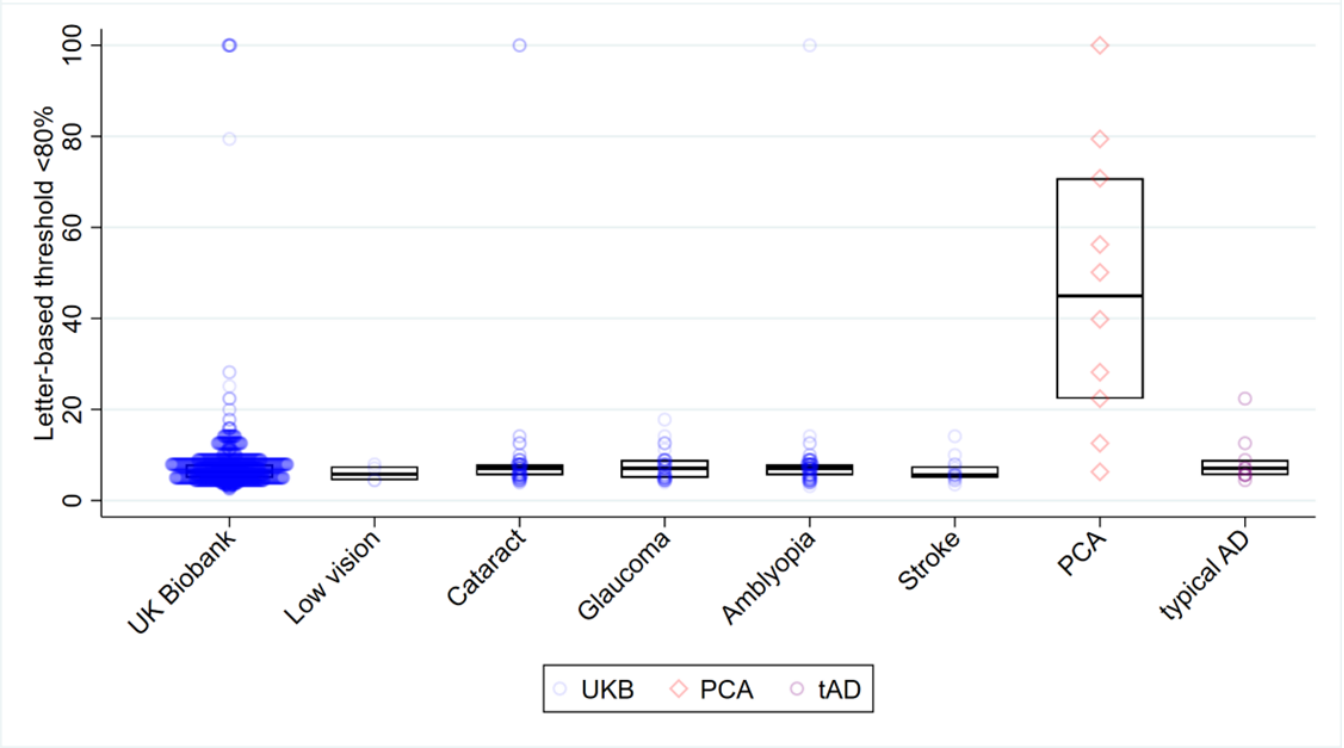

Figure 3 GILT letter-based thresholds, medians and interquartile ranges for UK Biobank (UKB) participants (blue), PCA (red) and typical AD participants (purple). Letter-based thresholds are presented for all UKB participants without or with visual conditions or stroke who reached accuracy cut-offs of <80% within the UKB time limit of 120 seconds (total UKB n=1,967).

See Table 2 for associations between GILT impairment (<5th%ile) and PCA diagnosis compared to UKB participants without or with documented visual conditions and sensitivity, specificity, positive and negative predictive values. Based on patient pilot data, the GILT appears to have good sensitivity and specificity for PCA (90.0% and ≥91.2% respectively) compared to UKB participants without or with common visual conditions (e.g. cataract, glaucoma, amblyopia). Using a penalized likelihood method for small sample sizes, the area under the ROC curve value differentiating PCA from UKB participants without documented visual conditions or stroke was 0.938 using letter-based thresholds at <80% accuracy.

GILT accuracy thresholds compared to VOSP incomplete letter performance

See Figure 4 for PCA and typical AD letter-based thresholds at <80% accuracy shown in relation to number of errors on the VOSP incomplete letters subtest. The VOSP subtest is used in diagnostic settings but features letters at a fixed completeness level (30% complete) rather than varying as in the GILT (100, 89.13, 50.12, 28.18, 15.85, 8.91, 5.01, 2.82, 1.58% complete). Granularity is apparent on letter-based thresholds across the range of GILT performance. Variation in GILT letter-based thresholds is observed towards both upper and lower ends of performance on the VOSP, including for typical AD participants performing at- or near-ceiling, and PCA participants performing near floor.

Most PCA and typical AD participants had concordant impairment on GILT and VOSP incomplete letter recognition, defined as performance <5th%ile of UKB participants without documented visual conditions or stroke (GILT) or published normative data (VOSP[20]). Using this 5th%ile cut-off, 9/10 PCA patients were considered impaired using letter-based thresholds of <80% and <100%; the remaining participant performed at threshold (either PCA participant 10 or 8 using <80% or <100% thresholds respectively; see Figure 4). Using letter-based thresholds, 9/10 PCA and 1/9 typical AD participants exhibited impairment on the GILT compared to 10/10 PCA and 2/9 typical AD participants who exhibited impairment on the VOSP.

Table 2. Association between PCA diagnosis vs UKB (without (left) or with documented visual conditions (right)) and GILT impairment. GILT impairment is defined using a standard cut-off (<5th%ile in UKB without visual conditions) using letter-based thresholds at <80% accuracy. Sensitivity, specificity, positive predictive and negative predictive values (PPV; NPV) are presented. Visual conditions in participants reaching <80% accuracy are cataract (n=69), amblyopia (n=75), glaucoma (n=46) and low vision (n=4). UKB participants with stroke reaching <80% accuracy (n=16) have been excluded.

	Letter-based threshold <80%				
	PCA diagnosis vs UKB (no visual dx)			PCA diagnosis vs UKB (visual dx)	
	Positive	Negative		Positive	Negative
Positive	9	132	Positive	9	17
Negative	1	1625	Negative	1	177
Sensitivity	90.0		Sensitivity	90.0	
Specificity	92.5		Specificity	91.2	
PPV	6.4		PPV	34.6	
NPV	99.9		NPV	99.4	

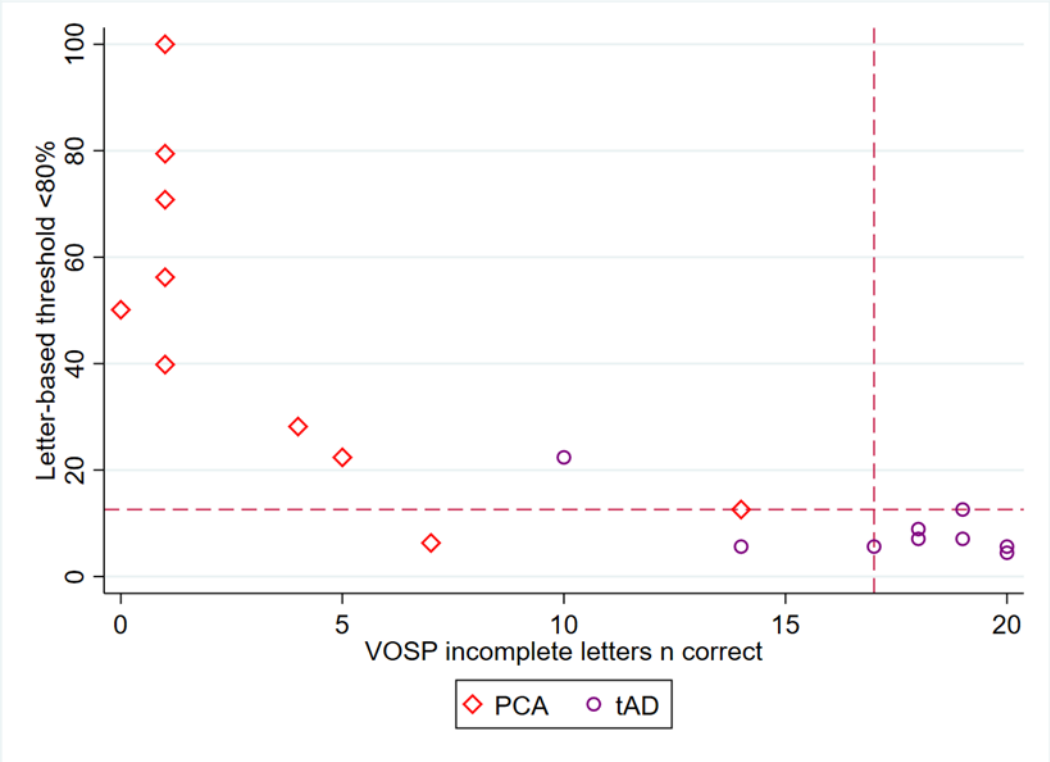

Figure 4 Concordance of GILT letter-based thresholds at <80% accuracy with VOSP incomplete letter accuracy in PCA and typical AD participants. Dashed vertical lines represent the VOSP cut-off indicating impairment (5th%ile: 17; max score: 20[20]). Dashed horizontal lines represent the GILT 5th%ile cut-off based on UK Biobank participants without documented visual conditions or stroke (5th%ile letter-based threshold:12.6; max score: 0).

Discussion

The GILT is a rapid digital test optimised to detect subtle cortical visual abnormalities in the form of a difficulty recognising letters which become progressively visually degraded. Based on normative data from the UK Biobank, we demonstrate that the test can be performed with a high level of degradation (low 'completeness') in typical adults. PCA participants consistently performed below the normal range of the UK Biobank sample. Despite correctly identifying complete letters, PCA participants exhibited a particular tendency to make errors with decreasing letter completeness, in some cases with very subtle decreases in completeness. This is consistent with deficits being specific to stimulus degradation, or low completeness, rather than general letter recognition deficits or other visual issues (e.g. diminished acuity). The same deficits do not manifest in participants with mostly mild, typical AD, whose symptoms primarily concern memory dysfunction. Furthermore, in the UK Biobank, GILT performance was high in both healthy participants and those with visual conditions like glaucoma. In other words, the GILT appears to exhibit a specificity for cortical visual loss arising from PCA-related neurodegeneration. Our pilot data suggest the GILT may have utility for differentiating participants with common visual conditions from those with visual loss owing to posterior cortical damage.

UK Biobank and pilot data from clinically diagnosed PCA and typical AD participants suggest particular advantages of the GILT in research settings. GILT letter-based thresholds exhibited a number of desirable characteristics. Even with brief test administration (<3 minutes), the GILT provides automatic data capture of measures which are granular while limiting ceiling effects – a particular limitation of existing visual tasks in standard batteries[36]. The increased granularity also increases the sensitivity to abnormalities. Beyond current accuracy and primary outcome measures, analyses of GILT error type (e.g. the extent to which errors relate to the target[16,49]) may aid in the differentiation and understanding of abnormal performance. Given that visual symptoms are characteristic initial features of PCA[8], GILT thresholds and error analysis may have particular promise to detect these early deficits and track the progression of the condition.

Limitations of the current study include the relatively small sample of UCL participants diagnosed with PCA or typical AD. While we used a penalized likelihood logistic regression given the small PCA sample, estimates should be interpreted with caution and larger studies are required to validate the GILT across settings. While GILT uses the same stimuli sets and presentation, test administration was adjusted to allow for cognitive impairments so that PCA and typical AD participants made verbal responses rather than using the touchscreen, and only discontinued when reaching cut-off thresholds rather than a timed limit as in UK Biobank. Normative dataset accuracy may therefore have been underestimated owing to time constraints. However, as a proportion of people within this age range may be amyloid positive and relatedly exhibiting subtle cognitive deficits[50,51], it is also possible that normative performance has been overestimated. UK Biobank is not representative of the UK population as a whole. Further investigations are required to determine thresholds which minimize false positives in typical adults as well as false negatives- the current pilot findings suggest low false negatives using only one letter-based threshold. Further testing is required to determine GILT set and item (letter) effects, although these first release data in UK Biobank do not suggest material set

effects. Informing a GILT version for clinical use requires further testing, ideally involving participants with visual conditions across ophthalmic and neurologic settings. Amendments for clinical testing might include adjusting administration (e.g. GILT-UKB responses are restricted to 12 items) and evaluating acceptability. Further pilot testing might determine reasons for participants who make errors with complete letters (0.4% of current UK Biobank sample), including clinical (e.g. visual disorders) and socio-demographic factors (first language, illiteracy).

Further testing might address limitations of the GILT and extend understanding of GILT performance outside the current UK Biobank, PCA and typical AD samples. Neuroimaging investigations are ongoing to assess relationships between GILT measures and the integrity of visual cortical regions. Behavioural investigations adjusting stimulus presentation (e.g. stimulus/check size, blur, eccentricity) are ongoing to evaluate particular deficits (e.g. visual crowding[31]) and mechanisms underpinning abnormal performance.

Ongoing work in the UK Biobank imaging sub-study and clinic-based research cohort studies is investigating brain-behaviour relationships in the context of neurodegenerative disease. Further validation of the GILT is required incorporating clinico-radiological and biomarker measures. In research contexts, salient opportunities afforded by the GILT include detecting subtle cortical visual abnormality at scale in relation to candidate associated clinical, developmental and genetic factors. Ultimate clinical goals include addressing knowledge gaps noted by eye and neurology professionals[27] and diagnostic delays and misdiagnosis faced by individuals with dementia-related visual loss[14].

Acknowledgements

We thank all patient participants for their generous time and efforts supporting research. The UKBB re-imaging study is supported by the Chan Zuckerberg Foundation. The Dementia Research Centre is an Alzheimer's Research UK Co-ordinating Centre and is supported by Alzheimer's Research UK, Brain Research Trust, and The Wolfson Foundation. KY is an Etherington PCA Senior Research Fellow and is funded by the Alzheimer's Society, grant number 453 (AS-JF-18-003). This work was also supported by the NIHR Queen Square Dementia Biomedical Research Unit and by an Alzheimer's Research UK Senior Research Fellowship [ART-SRF2010-3] and ESRC/NIHR [ES/L001810/1] and EPSRC [EP/M006093/1] grants to SC. We are grateful to the UKBB Eye and Vision Consortium and Mr Stephen Bell and Dr Alan Young from the UKBB team for implementation of the test to the UKBB IT platform.

Data availability statement

The data used in this work are available upon application to the UK BioBank (https://www.ukbiobank.ac.uk). The authors do not have permission to share data directly.

Conflicts of interest

Nothing to report.

References

1. Devinsky O, Farah MJ, Barr WB. Chapter 21 Visual agnosia. *Handb Clin Neurol*. 2008;88:417-427. doi:10.1016/S0072-9752(07)88021-3

2. Koedam ELGE, Lauffer V, van der Vlies AE, van der Flier WM, Scheltens P, Pijnenburg YAL. Early-Versus Late-Onset Alzheimer's Disease: More than Age Alone. *J Alzheimer's Dis*. 2010;19(4):1401-1408. doi:10.3233/JAD-2010-1337

3. Snowden JS, Stopford CL, Julien CL, et al. Cognitive phenotypes in Alzheimer's disease and genetic risk. *Cortex*. 2007;43(7):835-845. doi:10.1016/S0010-9452(08)70683-X

4. Neary D, Snowden ; J S, Gustafson ; L, et al. Frontotemporal lobar degeneration A consensus on clinical diagnostic criteria. Published online 1998.

5. Hodges JR, Patterson K. Semantic dementia: a unique clinicopathological syndrome. *Lancet Neurol*. 2007;6(11):1004-1014. doi:10.1016/S1474-4422(07)70266-1

6. Crutch SJ, Schott JM, Rabinovici GD, et al. Consensus classification of posterior cortical atrophy on behalf of the Alzheimer's Association ISTAART Atypical Alzheimer's Disease and Associated Syndromes Professional Interest Area. *Alzheimer's Dement*. 2017;13:870-884. doi:10.1016/j.jalz.2017.01.014

7. Milner AD, Cavina-Pratesi C. Perceptual deficits of object identification: apperceptive agnosia. *Handb Clin Neurol*. 2018;151:269-286. doi:10.1016/B978-0-444-63622-5.00013-9

8. Firth NC, Primativo S, Marinescu RV, et al. Longitudinal neuroanatomical and cognitive progression of posterior cortical atrophy. *Brain*. 2019;142(7):2082-2095. doi:10.1093/brain/awz136

9. Coslett HB. Apraxia, Neglect, and Agnosia. *Contin Lifelong Learn Neurol*. 2018;24(3, BEHAVIORAL NEUROLOGY AND PSYCHIATRY):768-782. doi:10.1212/CON.0000000000000606

10. Yong KXX, Graff-Radford J, Ahmed S, et al. Diagnosis and Management of Posterior Cortical Atrophy. *Curr Treat Options Neurol*. 2023;25(2):23-43. doi:10.1007/S11940-022-00745-0/TABLES/2

11. Harding E, Sullivan MP, Woodbridge R, et al. "Because my brain isn't as active as it should be, my eyes don't always see": a qualitative exploration of the stress process for those living with posterior cortical atrophy. *BMJ Open*. 2018;8(2):e018663. doi:10.1136/bmjopen-2017-018663

12. O'malley M, Parkes J, Stamou V, Lafontaine J, Oyebode J, Carter J. Young-onset dementia: scoping review of key pointers to diagnostic accuracy. Published online 2019. doi:10.1192/bjo.2019.36

13. Shallice T, Jackson M. Cognitive Neuropsychology Lissauer on Agnosia. *Cogn Neuropsychol*.

1988;5(2):153-156. doi:10.1080/02643298808252931

14. Graff-Radford J, Yong KXX, Apostolova LG, et al. New insights into atypical Alzheimer's disease in the era of biomarkers. *Lancet Neurol*. 2021;20(3):222-234. doi:10.1016/S1474-4422(20)30440-3

15. Yong KXX, Graff-Radford J, Ahmed S, et al. Diagnosis and Management of Posterior Cortical Atrophy. *Curr Treat Options Neurol*. 2023;25(2):23-43. doi:10.1007/S11940-022-00745-0/TABLES/2

16. Yong, Shakespeare TJ, Cash D, et al. Prominent effects and neural correlates of visual crowding in a neurodegenerative disease population. *Brain*. 2014;137(12):3284-3299. doi:10.1093/brain/awu293

17. Whitney D, Levi DM. Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn Sci*. 2011;15(4):160-168. doi:10.1016/J.TICS.2011.02.005

18. Lehmann M, Barnes J, Ridgway GR, et al. Basic visual function and cortical thickness patterns in posterior cortical atrophy. *Cereb Cortex*. 2011;21(9):2122-2132. doi:10.1093/cercor/bhq287

19. Schott J, Crutch S. Posterior cortical atrophy. *Contin Lifelong Learn Neurol*. 2019;25(1):52-75. doi:10.1212/CON.0000000000000696

20. Warrington EK, James M, Thames Valley Test Company. *The Visual Object and Space Perception Battery*. Thames Valley Test Company; 1991. Accessed April 24, 2017. http://sfx.ucl.ac.uk/sfx_local?sid=google&auinit=EK&aulast=Warrington&title=The+visual+object+and+space+perception+battery&genre=book&isbn=0951432273&date=1991

21. James M, Plant GT (Gordon T, Warrington EK, Thames Valley Test Company. *Corvist : Cortical Vision Screening Test : Manual and Test Materials*. Thames Valley Test Co; 2001.

22. Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR. The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *Int J Geriatr Psychiatry*. 2006;21(11):1078-1085. doi:10.1002/gps.1610

23. Lopes R, Bournonville C, Kuchcinski G, et al. Prediction of Long-term Cognitive Function After Minor Stroke Using Functional Connectivity. Published online 2021. doi:10.1212/WNL.0000000000011452

24. Riddoch MJ, Humphreys GW, Lincoln N. Birmingham object recognition battery. Disability and Rehabilitation. Published 1995. Accessed August 31, 2023. https://psycnet.apa.org/record/9999-13731-000?doi=1

25. Torfs K, Vancleef K, Lafosse C, Wagemans J, de-Wit L. The Leuven Perceptual Organization Screening Test (L-POST), an online test to assess mid-level visual perception. *Behav Res Methods*. 2014;46(2):472-487. doi:10.3758/s13428-013-0382-6

26. Warrington EK. Neuropsychological studies of object recognition. *Philos Trans R Soc London B, Biol Sci*. 1982;298(1089):15-33. doi:10.1098/RSTB.1982.0069

27. Bowen M, Zutshi H, Cordiner M, Crutch S, Shakespeare T. Qualitative, exploratory pilot study to investigate how people living with posterior cortical atrophy, their carers and clinicians experience tests used to assess vision. *BMJ Open*. 2019;9(3):e020905. doi:10.1136/bmjopen-2017-020905

28. Yong KXX, Hardy CJD, Petzold A, Crutch SJ. Posterior cortical atrophy: an overview for optometrists - College of Optometrists. Optometry In Practice. Accessed July 14, 2022. https://www.college-optometrists.org/professional-development/college-journals/optometry-in-practice/all-oip-articles/volume-23,-issue-2/posterior-cortical-atrophy-an-overview-for-optomet

29. Warrington EK, James M. Visual Apperceptive Agnosia: A Clinico-Anatomical Study of Three Cases. *Cortex*. 1988;24(1):13-32. doi:10.1016/S0010-9452(88)80014-5

30. McCarthy RA, Warrington EK. Cognitive neuropsychology: A clinical introduction. Published online 1990.

31. Strappini F, Pelli DG, Di Pace E, Martelli M. Agnosic vision is like peripheral vision, which is limited by crowding. *Cortex*. 2017;89:135-155. doi:10.1016/J.CORTEX.2017.01.012

32. Warrington EK, James M. Disorders of visual perception in patients with localised cerebral lesions. *Neuropsychologia*. 1967;5(3):253-266. doi:10.1016/0028-3932(67)90040-1

33. Rossetti Y, Pisella L, McIntosh RD. Rise and fall of the two visual systems theory. *Ann Phys Rehabil Med*. 2017;60(3):130-140. doi:10.1016/J.REHAB.2017.02.002

34. Smits LL, Pijnenburg YAL, Koedam ELGE, et al. Early Onset Alzheimer's Disease is Associated with a Distinct Neuropsychological Profile. *J Alzheimer's Dis*. 2012;30(1):101-108. doi:10.3233/JAD-2012-111934

35. Pavisic IM, Firth NC, Parsons S, et al. Eyetracking Metrics in Young Onset Alzheimer's Disease: A Window into Cognitive Visual Functions. *Front Neurol*. 2017;8:377. doi:10.3389/fneur.2017.00377

36. Bellio M, Oxtoby NP, Walker Z, et al. Analyzing large Alzheimer's disease cognitive datasets: Considerations and challenges. *Alzheimer's Dement Diagnosis, Assess Dis Monit*. 2020;12(1). doi:10.1002/dad2.12135

37. Boyd CD, Tierney M, Wassermann EM, et al. Visuoperception test predicts pathologic diagnosis of Alzheimer disease in corticobasal syndrome. *Neurology*. 2014;83(6):510-519. doi:10.1212/WNL.0000000000000667

38. Salmon DP, Smirnov DS, Coughlin DG, et al. Perception of Fragmented Letters by Patients With Pathologically Confirmed Dementia With Lewy Bodies or Alzheimer Disease. *Neurology*. 2022;99(18):e2034-e2043. doi:10.1212/WNL.0000000000201068

39. Foster PJ, Atan D, Khawaja A, et al. Cohort profile: rationale and methods of UK Biobank repeat imaging study eye measures to study dementia. *BMJ Open*. 2023;13(6):e069258. doi:10.1136/BMJOPEN-2022-069258

40. McKhann GM, Knopman DS, Chertkow H, et al. *The Diagnosis of Dementia Due to Alzheimer's Disease: Recommendations from the National Institute on Aging-Alzheimer's Association Workgroups on Diagnostic Guidelines for Alzheimer's Disease*. Alzheimer's & Dementia 7, 263-269 (2011). doi:10.1016/j.jalz.2011.03.005

41. Dubois B, Feldman HH, Jacova C, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol*. 2014;13(6):614-629. doi:10.1016/S1474-4422(14)70090-0

42. Song S, Levi DM, Pelli DG. A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *J Vis*. 2014;14(5):3-3. doi:10.1167/14.5.3

43. Sloan LL. New test Charts for the Measurement of Visual Acuity at far and Near Distances. *Am J Ophthalmol*. 1959;48(6):807-813. doi:10.1016/0002-9394(59)90626-9

44. Shah N, Dakin SC, Redmond T, Anderson RS. Vanishing Optotype acuity: repeatability and effect of the number of alternatives. *Ophthalmic Physiol Opt*. 2011;31(1):17-22. doi:10.1111/J.1475-1313.2010.00806.X

45. Carkeet A. Modeling logMAR visual acuity scores: Effects of termination rules and alternative forced-choice options. *Optom Vis Sci*. 2001;78(7):529-538. doi:10.1097/00006324-200107000-00017

46. Bailey IL, Lovie JE. New Design Principles for Visual Acuity Letter Charts*. *Optom Vis Sci*. 1976;53(11). https://journals.lww.com/optvissci/Fulltext/1976/11000/New_Design_Principles_for_Visual_Acuity_Letter.6.aspx

47. Pelli DG, Robson JG, Wilkins AJ. The design of a new letter chart for measuring contrast sensitivity. *Clin Vis Sci*. 1988;2(3):187-199. doi:10.11432/JPNJVISSCI.29.67

48. ELLIOTT DB, BULLIMORE MA, BAILEY IL. Improving the reliability of the Pelli-Robson contrast sensitivity test. *Clin Vis Sci*. 1991;6(6):471-475.

49. Nandy AS, Tjan BS. The nature of letter crowding as revealed by first- and second-order classification images. *J Vis*. 2007;7(2):5-5. doi:10.1167/7.2.5

50. Lu K, Nicholas JM, Pertzov Y, et al. Dissociable effects of APOE ε4 and β-amyloid pathology on visual working memory. *Nat Aging 2021 111*. 2021;1(11):1002-1009. doi:10.1038/s43587-021-00117-4

51. Lane CA, Barnes J, Nicholas JM, et al. Associations between blood pressure across adulthood

and late-life brain structure and pathology in the neuroscience substudy of the 1946 British birth cohort (Insight 46): an epidemiological study. *Lancet Neurol*. 2019;18(10):942-952. doi:10.1016/S1474-4422(19)30228-5