Abstract Word Count: 289

Word Count: 5568

Title:

Cardiff ONline Cognitive Assessment (CONCA): Results from a web-based national population cohort

Running Title: Cardiff ONline Cognitive Assessment (CONCA)

Authors:

Amy J Lynham, PhD, Ian R Jones, MRCPsych, PhD, James TR Walters, MRCPsych, PhD

Affiliations:

Division of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, United Kingdom

Corresponding Author:

James T R Walters

MRC Centre for Neuropsychiatric Genetics and Genomics

Division of Psychological Medicine and Clinical Neurosciences, School of Medicine,

Cardiff University

Maindy Road

Cardiff

CF24 4HQ

Email: waltersjt@cardiff.ac.uk

Abstract

Background: Psychiatric disorders are associated with cognitive impairment. We have developed a web-based, 9-task cognitive battery to measure the core domains affected in people with psychiatric disorders. To date, this assessment has been used to collect data on a clinical sample of participants with psychiatric disorders.

Objectives: The aims of this study were: (1) to establish a briefer version of the battery (called the Cardiff ONline Cognitive Assessment, CONCA) that can give a valid measure of cognitive ability (‘g’), and (2) to collect normative data and demonstrate CONCA’s application in a health population sample.

Methods: Based on six criteria and data from our previous study, we selected 5 out of the original 9 tasks to include in CONCA. These included 3 core tasks that were sufficient to derive a measure of ‘g’ and 2 optional tasks. Participants from a web-based national cohort study (HealthWise Wales) were invited to complete CONCA. Completion rates, sample characteristics, performance distributions and associations between cognitive performance and demographic characteristics and mental health measures were examined.

Results: A total of 3679 completed at least one CONCA task, of which, 3135 participants completed all three core CONCA tasks. Performance on CONCA was associated with age (B=-0.05, SE=0.002, $P<.001$), device (tablet computer: B=-0.26, SE=0.05, $P<.001$; smartphone: B=-0.46, SE=0.05, $P<.001$), education (degree: B=1.68, SE=0.14, $P<.001$), depression symptoms (B=-0.04, SE=0.01, $P<.001$) and anxiety symptoms (B=-0.04, SE=0.01, $P<.001$).

Conclusions: CONCA provides a valid measure of ‘g’, which can be derived using as few as 3 tasks that take no more than 15 minutes. Performance on CONCA showed associations with demographic characteristics in the expected direction and was associated with current depression and anxiety symptoms. The effect of device on cognitive performance is an important consideration for research utilising web-based assessments.

Keywords: Cognition, mental health, online, digital assessment, normative data, mobile phone

Introduction

Background

Cognitive function has been shown to be associated with health, such that those who perform better on cognitive assessments have better health outcomes including decreased mortality risk, on average, than those with lower cognitive function [1-4]. A number of mental and physical conditions are associated with cognitive impairments, including common conditions such as depression [5], anxiety [6], hypertension [7] and diabetes [8]. More pronounced cognitive impairments are seen in those with a diagnosis of a severe mental disorder, such as schizophrenia [9] or bipolar disorder [10]. The severity of these impairments is an important predictor of occupational and social functioning in participants diagnosed with these disorders [11, 12].

Existing cognitive research is limited by sample size, as collecting cognitive data traditionally involves a face-to-face assessment and can be labour intensive. However, the rise in internet use over the past few decades and the development of digital assessments has presented researchers with new opportunities to collect large datasets [13]. At the MRC Centre for Neuropsychiatric Genetics and Genomics, we have developed and used a web-based cognitive assessment to collect data on over 1000 participants diagnosed with a range of psychiatric disorders [14]. To date, we have: (1) established validity against a gold standard measure of cognition in psychiatric research (MATRICS Consensus Cognitive Battery, MCCB); (2) reported an association between performance on the battery and functioning in a cohort of participants with psychiatric disorders; and (3) demonstrated that performance on the battery discriminates between controls and participants with schizophrenia, bipolar disorder, and major depressive disorder. However, we have not reported normative data for the battery, measured the distribution of scores or examined associations between performance on the battery and demographic factors in a population sample. Although normative data for some of the individual tasks already exist, it is crucial that normative data for web-based tasks are collected on the web using the same platform [13]. In addition, the original battery consisted of nine tasks with an administration time of up to 50 minutes. However, given that some of the correlations between the web-based tasks and the MCCB were small and there were concerns about the length of the battery, we have developed a briefer version of the battery with an improved user-friendly interface (Cardiff ONline Cognitive Assessment, CONCA). This new version of the battery was specifically designed to provide a brief, valid measure of general cognitive function ('g'). A measure of general cognition ('g') was considered appropriate given the literature showing that cognitive impairment in psychiatric disorders (particularly schizophrenia) is characterised by widespread, global impairment rather than specific localised dysfunction and this global impairment is predictive of poor community functioning [15].

HealthWise Wales

In addition to cognitive assessments, web-based technologies have provided the opportunity to recruit population cohorts for epidemiological research. One such cohort is HealthWise Wales, a Welsh Government-funded digital health project that has recruited a web-based cohort of people living or receiving healthcare in Wales [16]. The aim of HealthWise Wales is to understand factors that impact health and wellbeing, including social inequalities, environment, and health behaviours, through web-based data collection and linkage to routine healthcare records. This cohort provides an opportunity to examine cognitive performance in the general population.

Study Aims

This study had two aims. First, we established a core battery (CONCA) that can provide a valid measure of ‘g’ in less than 15 minutes. To do this, we used data from our previous study [14] to evaluate the original nine cognitive tasks against set criteria. Second, we aimed to derive normative data for CONCA and demonstrate its application in a health population sample by collecting cognitive data from HealthWise Wales. This study is presented in two parts to reflect these aims.

Methods

Part 1: Establishing the Cardiff ONline Cognitive Assessment (CONCA)

Participants

Full details of the original study have been previously published [14]. Briefly, participants were recruited from the databases of two existing studies of psychiatric disorders within the MRC Centre for Neuropsychiatric Genetics and Genomics: Cognition in Mood, Psychosis and Schizophrenia Study (CoMPaSS, [17]) and National Centre for Mental Health (NCMH, [18]). For the purposes of this study, we included only participants with data on the full nine tasks (N=841).

Measures

The Cardiff ONline Cognitive Assessment (CONCA) was developed to assess cognitive function in individuals with a history of mental illness. All tasks (including source code) were developed by The Many Brains Project, a not-for-profit organisation that develops open-source web-based tools to assess cognitive function [19, 20]. We selected 9 tasks to assess, as closely as possible, the domains outlined by the National Institute for Mental Health's Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) initiative [21]. To improve the battery, we aimed to reduce the length to 5 tasks with a maximum administration time of 30 minutes that would provide a brief, valid measure of 'g'.

We selected the MATRICS Consensus Cognitive Battery (MCCB) as our comparison measure to validate CONCA due to the rigorous selection procedure employed in its development and its widespread adoption in mental health research. The MCCB consists of 10 pen and paper tasks assessing the 7 domains outlined by the MATRICS initiative [21]. It was developed using expert panels, consultations with scientists, evaluations of psychometric properties and assessments of tolerability and practicality, with the explicit aim of creating a gold standard battery for use in schizophrenia research [22].

Participants also completed the 12-item version of the World Health Organisation Disability Assessment Schedule (WHODAS, [23]), which assesses six domains of functional impairment: understanding and communicating, mobility, self-care, social interactions, life activities and participation in the community.

Study Design

The study design was cross-sectional. Selection of tasks for the new CONCA battery was guided by the findings in our previous study [14] and we additionally conducted some new analyses. This study design has been previously described [14] but briefly, participants who had consented to be contacted about follow-up studies were invited via email or letter to complete the original 9-task battery. A subset of participants (N=65) additionally completed the MCCB as a gold standard comparison measure.

Following discussions within our research team and consultation with our health professional and patient representatives, we outlined 6 criteria to be used to guide task selection. To be considered for inclusion, we sought to demonstrate that each task was: (1) correlated with its equivalent task in the MCCB, (2) correlated with general cognitive function 'g' derived using the MCCB, (3) associated with functioning as measured by the WHODAS [23], (4) loaded onto a measure of 'g' derived from the 9-task battery using factor analysis, (5) considered acceptable based on participant feedback with no insurmountable technical issues reported, and (6) translatable into other languages to support our international collaborations. Tasks were considered "translatable" if it would be possible to translate the instructions and materials without fundamentally changing the measurement properties of the task (e.g., tasks with non-verbal stimuli). Correlations between the CONCA tasks and the MCCB (criteria 1 and 2), associations with functional outcomes (criterion 3) and technical issues and participant feedback (criterion 5) have been previously published in Lynham et al. [14] (a summary of these results can be found in Table S1 in Multimedia Appendix 1). We conducted further analyses (see next section) to determine whether tasks met criterion 4 and to evaluate the validity of the new battery. As far as possible, we selected tasks that were representative of different domains as opposed to similar tasks to ensure CONCA was a well-rounded measure of global cognitive function.

Analysis

The structure of the 9-task web-based cognitive battery was examined using exploratory factor analysis. The number of factors was identified using scree plots and parallel analysis. Principal axis factoring with oblique rotation (direct oblimin) was conducted to identify the factors.

To evaluate the validity of ‘g’ derived using the new CONCA battery, we examined correlations between ‘g’ derived using the MCCB, and ‘g’ derived using the new CONCA battery. This analysis was conducted on the subset of participants with MCCB data available (n=65). ‘g’ was derived using multidimensional scaling [24], which is an approach analogous to principal component analysis, with the first component extracted as ‘g’.

Part 2: Assessing Cognition in HealthWise Wales

Participants

Participants were recruited from HealthWise Wales, a web-based national population cohort [16]. Adults aged 16 and above who live or receive their healthcare in Wales are eligible for inclusion in HealthWise Wales. Participants consent to be contacted for follow-up data collection with new questionnaires added to the website and advertised via email invitations every six months. Ethical approval for HealthWise Wales was obtained from Wales Research Ethics Committee 3 (reference: 15/WA/0076). HealthWise Wales data is collected and stored in the Secure Access Portal and Protected HWW Information Repository (SAPPHIRe), which is powered by the UK Secure e-Research Platform (UKSeRP) [25]. The Cardiff ONline Cognitive Assessment (CONCA) was added as a module on the HealthWise Wales website in January 2020 and email invitations sent to all participants in the cohort

(N=29,492). Ethical approval for CONCA was granted by Cardiff University's School of Medicine Research Ethics Committee (reference: 15/64).

Measures

Participants completed CONCA, the WHODAS, the Hospital Anxiety and Depression Scale (HADS, [26]), as well as providing basic demographic information (age, gender, education, and device used). The data collected was also linked with existing data from HealthWise Wales to determine whether participants had ever been diagnosed or treated for a mental health problem [16].

Study Design

The study design was cross-sectional. Participants completed the study by either clicking on the link in their email invitation or clicking on the module on the HealthWise Wales home screen. This took participants to the CONCA webpage, where they could read the information sheet, provide informed consent, and complete all the measures.

Analysis

All analyses were conducted using R version 3.6.1. For each task, z scores were derived using the mean and standard deviation of the sample. Two measures of 'g' were derived using multidimensional scaling (MDS) [24]: (1) using the scores on three core CONCA tasks only (Core 'g'), (2) using scores on the complete (Full 'g'). These two measures of 'g' were highly correlated (r=.93).

Completion rates for each task were calculated. To examine predictors of completing the optional tasks, we performed a logistic regression to test the association between completion of at least one optional task and the following variables:

cognitive performance on the core tasks (Core ‘g’), age, gender, education, device and ever received diagnosis and/or treatment for a mental health problem.

We performed multiple linear regression to test the association between cognitive performance (‘g’) and the following demographic variables in a single model: age, gender, education, and device. We repeated this analysis for each cognitive task. *P* values were corrected using the false discovery rate (FDR) method.

As CONCA was developed as a tool for mental health research, we evaluated whether performance on CONCA was associated with two measures of mental health: (1) whether participants had ever been diagnosed or treated for a mental health problem, (2) scores on the HADS subscales, depression and anxiety. Each mental health variable (ever diagnosed, HADS depression, HADS anxiety) were entered as predictors into separate linear regressions with ‘g’ as the outcome and age, gender, education, and device as covariates.

Statement of Ethical Approval

Ethical approval for HealthWise Wales was obtained from Wales Research Ethics Committee 3 (reference: 15/WA/0076). Ethical approval for CONCA was granted by Cardiff University’s School of Medicine Research Ethics Committee (reference: 15/64). All participants indicated their informed consent by selecting “yes” in response to the statement, “I agree to take part in this study and know that I am free to leave the study at any point” at the start of the study. No personal identifiers were collected as part of the study, as all data was linked to an ID number. Participants did not receive compensation for their time.

Results

Part 1: Establishing CONCA

Factor Loadings

Examination of the scree plot and parallel analysis indicated 2 factors with eigenvalues above 1. All the measures except Vocabulary and Balloon Analogue Risk Task loaded onto the first factor (Table 1). Only Vocabulary had a high loading on the second factor.

Table 1 Factor loadings of the web-based tasks

Task	Factor 1	Factor 2
Matrix Reasoning	.56	.29
Multiple Object Tracking	.7	.04
Balloon Analogue Risk Task	.18	.27
Backward Digit Span	.43	.29
Verbal Paired Associates Test	.4	.24
Digit Symbol Coding	.81	-.11
Morphed Emotion Identification	.56	.07
Vocabulary	-.07	.66
Hartshorne Visual Working Memory	.66	-.16
Proportion of variance explained	.76	.24

Selection of the final CONCA battery

The final battery consisted of 3 core tasks with an administration time of 15 minutes and 2 optional tasks (total administration time of 30 minutes). Once the final tasks were selected, we consulted with patient representatives to design a new user-friendly website for CONCA [27].

Task 1: Digit Symbol Coding

This task is an adapted web-based version of the well-validated measure of processing speed [28]. Performance on the task was correlated with its MCCB equivalent (r=.73) and 'g' (r=.74), had the strongest association with functional outcome, a high factor loading (.81) and is easily translatable.

Task 2: Backward Digit Span

This task is a web-based version of the well-validated measure of working memory [29]. Performance on the task was correlated with its MCCB equivalent (r=.34), was strongly associated with functional outcome, and had a short administration time (3 minutes).

Task 3: Vocabulary

Participants are shown a target word and asked to select which of four words is closest in meaning to the target word [28]. This task was included as a measure of crystallised intelligence based on its correlation with the National Adult Reading Test (r=.64) [30]. Performance on the task did not load onto the web-based 'g' in the 9-task factor analysis but was correlated with MCCB 'g' (r=.36), associated with functioning and was the only well-tolerated verbal task.

Task 4: Morphed Emotion Identification (Optional Task)

Participants are presented with a face and must decide whether the face looks angry, fearful, happy, or disgusted [31, 32]. Faces are morphed between a neutral face and each emotion at varying intensities. The correlation between this task and its MCCB equivalent was low (r=.26), likely reflecting the different methodologies of the tasks. However, the task was correlated with 'g' (r=.58), strongly associated with functional outcome, and captured social cognition.

Task 5: Matrix Reasoning (Optional Task)

This task is based on the well-validated Matrix Reasoning test used in the Wechsler Abbreviated Scale of Intelligence II [28, 33]. This task was correlated with both its MCCB equivalent (*r*=.53) and ‘g’ (*r*=.59), was associated with functional outcome and had a high factor loading (.56). However, it was included as an optional task due to its long administration time (up to 15 minutes if all trials are completed).

Excluded Tasks

Hartshorne Visual Working Memory and Balloon Analogue Risk Task were excluded due to low correlations with ‘g’ (0.3 and 0.11 respectively). Verbal Paired Associates was poorly tolerated by participants who voted it “worst task” in their feedback and could not be easily translated. Multiple Object Tracking met all inclusion criteria, but participants reported difficulties completing it on smaller touchscreen devices, which could not be easily resolved.

Validity of CONCA-derived ‘g’

We calculated correlations to compare MCCB ‘g’ with three measures of ‘g’ from the web-based batteries: (1) original 9-task battery, (2) CONCA 5-task battery and (3) CONCA 3-task battery. Correlations were similar between MCCB ‘g’ and ‘g’ from all three versions (original 9-task battery: *r*=.78, 95% CIs: .66-.86; CONCA 5-task battery: *r*=.78, 95% CIs: .67-.86; CONCA 3-task core battery: *r*=.71, 95% CIs: .57-.81). Finally, the factor analysis was repeated including only the final selection of CONCA tasks and indicated that all tasks contributed to ‘g’ with factor loadings between 0.51 and 0.66 (see Supplementary Table S2, Multimedia Appendix 1 for full results).

Part 2: Assessing Cognition in HealthWise Wales

Completion rates

A total of 3889 participants from HealthWise Wales consented to the study (response rate = 3889/29,492, 13.19%). Of these, 3679 participants completed at least one cognitive task (3679/3889, 94.6%). Completion of the core battery was high (3135/3889, 80.61%), including 2048 who completed the core battery and both optional tasks (2048/3889, 52.66%, Table 2). After FDR correction, participants with higher scores on the core tasks were more likely to complete at least one optional task (OR=1.4, 95% CIs: 1.26-1.55, *P*<.001). None of the other variables significantly predicted completion of the optional tasks (see Table 3).

Table 2 Task completion rates and summary statistics

Task	Scoring	N	Mean	SD	Median	IQR
Digit Symbol Coding	Correct responses in 90 seconds	3679	41.71	10.72	41	15
Backwards Digit Span	Longest correctly recalled digit span	3199	4.44	1.62	4	2
Vocabulary	Correct responses (Max.=20)	3135	16.77	3.17	17	4
Emotion Identification	Correct responses (Max.=60)	2319	34.92	6.54	35	10
Matrix Reasoning	Correct responses (Max.=35)	2444	24.08	5.74	25	7

SD: Standard Deviation; IQR: Interquartile Range

Table 3 Predictors of optional task completion

	OR	95% CIs	*P*
Core ‘g’	1.4	1.26-1.55	<.001
Age	1.01	1-1.02	.14
Gender (reference: women)	0.95	0.74-1.23	.79
Education (reference: none)			
GCSE / O-levels	0.75	0.45-1.23	.44
A-levels	0.76	0.46-1.22	.44
Degree	0.79	0.48-1.27	.49
Post-graduate degree	0.74	0.44-1.21	.44
Device (reference: desktop/laptop)			
Smartphone	0.95	0.71-1.27	.79
Tablet	1.03	0.77-1.38	.84

Ever diagnosed with or treated for a mental health problem (reference: none)	1.38	1.07-1.76	.07

Results of a logistic regression where outcome is completion of at least one optional task (1 – completed, 0 – not completed). OR: Odds Ratios; CIs: Confidence Intervals

Sample characteristics

Sample characteristics were examined including all participants who had completed at least one cognitive task (N=3679, see Table 4). Most participants were women (69.55%) and had a mean age of 55.86 years (SD=15.05, range=16-93). Participants reported high levels of education; 1095/3557 (30.78%) reported an undergraduate degree as their highest level of education and 732/3557 (20.58%) reported a post-graduate degree as their highest level of education. Just under half of participants used a laptop or desktop computer to complete the study (1781/3672, 48.5%), whilst 803/3672 (21.87%) used a tablet device and 1088/3672 (29.63%) used a smartphone. The number of participants who reported a previous diagnosis of or treatment for a mental health condition was 1212 out of 3309 (36.63%).

Table 4 Sample characteristics

Sample Characteristics	**N**	**% of available data**	**Data available (N)**	**HealthWise Wales: Whole Sample (%)**[1]	**Population Data for Wales (%)**
Gender (Women)	2551	69.55	3668	72	50.69[2]
Highest education level			3557		
No GCSEs	259	7.28			7.3[3]
GCSE or equivalent	524	14.73			30.3[3]
A-level or equivalent	947	26.62			21.3[3]
Undergraduate degree	1095	30.78			29.2[3]
Postgraduate degree	732	20.58		N/A	11.9[3]
Device used			3672	N/A	N/A
Laptop / desktop	1781	48.5			

Tablet	803	21.87			
Smartphone	1088	29.63			
Ever diagnosed with or treated for a mental health problem	1212	36.63	3309	32	11[4]
45 years or older	2802	76.16	3679	60	47.25[2]
	Median	**IQR**	**Data available (N)**		
Age	59	21	3679	N/A	42.4[5]
WHODAS Total	5	11	1033	N/A	N/A
HADS Anxiety	6	7	1034	N/A	N/A
HADS Depression	5	7	1034	N/A	N/A

Information on population data was obtained from sources dated as close to the point of CONCA data collection as possible (January 2020). N/A – Data not available or not applicable. [1]Published data from HealthWise Wales [16]; [2]Office for National Statistics' national level population estimates for Wales in 2020 (note: sex not gender was recorded) [34]; [3]Office for National Statistics' highest qualification data in 2020 [35] (note: these education categories have been mapped as closely as possible to the study data); [4]National Survey for Wales 2019-2020 [36]; [5]Office for National Statistics population estimates for the UK and its constituent countries in 2020 [34]

Cognitive performance and demographic variables

There was evidence of a ceiling effect on Vocabulary amongst those aged 60 years and older, as 13.3% (251/1887 participants) achieved the maximum score (see Figure 1). Summary statistics for each of the tasks are presented by gender and age group in Supplementary Table S3 and by educational attainment in Supplementary Table S4 (Multimedia Appendix 1). These summary statistics can be used to generate age- and gender-adjusted z-scores using the formula:

$$Z = \frac{X_{ti} - M_{tga}}{SD_{tga}}$$

where X_{ti} is the score for individual *i* on test *t* and M_{tga} and SD_{tga} represent the mean and standard deviation for test *t* for that individual's corresponding age group *a* and gender *g*.

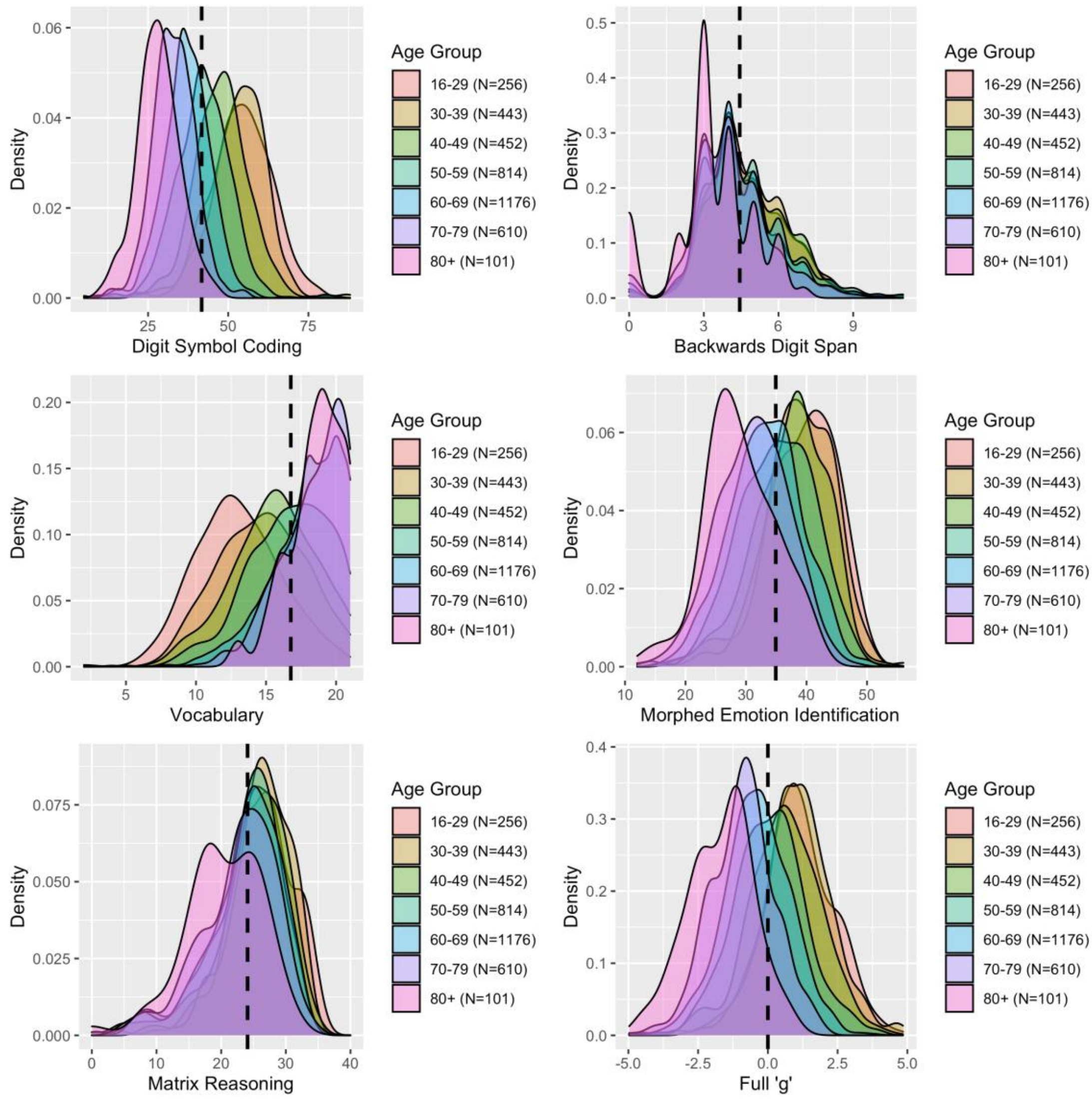

Figure 1 Distributions of performance on tasks by age group

From top left to bottom right, density plots stratified by age group for Digit Symbol Coding, Backwards Digit Span, Vocabulary, Morphed Emotion Identification, Matrix Reasoning, and ‘g’. Dashed line indicates mean performance for each task.

Cognitive performance (Full ‘g’) was associated with age (B=-0.05, SE=0.002, *P*<.001), device (tablet computer: B=-0.27, SE=0.06, *P*<.001; smartphone: B=-0.45, SE=0.05, *P*<.001) and education (degree: B=1.68, SE=0.14, *P*<.001; see Table 5), such that older age, use of a tablet computer or smartphone rather than a laptop or desktop and lower educational attainment were associated with lower cognitive performance (results for individual tasks can be found in Supplementary Table S5 in

Multimedia Appendix 1). Gender was not associated with ‘g’ (B=-0.002, SE=0.05, *P*=0.97) but was associated with performance on three tasks; men performed better on Vocabulary (B=0.1, SE=0.03, *P*=0.004) and Matrix Reasoning (B=0.2, SE=0.04, *P*<.001), whilst women performed better on Morphed Emotion Identification (B=-0.24, SE=0.05, *P*<.001). The proportion of variance in full ’g’ and core ‘g’ explained by demographic variables were .34 and .36 respectively (adjusted R^2).

Table 5 Associations between demographic variables and cognitive performance

	B	**SE**	***P***
Full ‘g’			
Age	-0.05	0.002	<.001
Gender (reference: women)	-0.002	0.05	.97
Education (reference: no qualifications)			
GCSE or equivalent	1.15	0.14	<.001
A-levels or equivalent	1.39	0.14	<.001
Undergraduate degree	1.68	0.14	<.001
Post-graduate degree	1.87	0.14	<.001
Device (reference: desktop/laptop)			
Smartphone	-0.45	0.05	<.001
Tablet	-0.27	0.06	<.001
Core ‘g’			
Age	-0.04	0.001	<.001
Gender (reference: women)	-0.02	0.04	.69
Education (reference: no qualifications)			
GCSE or equivalent	0.64	0.11	<.001
A-levels or equivalent	0.72	0.11	<.001
Undergraduate degree	0.88	0.11	<.001
Post-graduate degree	1.03	0.11	<.001
Device (reference: desktop/laptop)			
Smartphone	-0.18	0.04	<.001
Tablet	-0.13	0.04	.003

B: Linear regression coefficients; SE: Standard Error

Cognitive performance and mental health

Lower scores on the HADS depression subscale were associated with higher general cognitive ability ‘g’ (Full ‘g’: B=-0.04, SE=0.01, *P*<.001; Core ‘g’: B=-0.03, SE=0.01, *P*<.001). Lower scores on the HADS anxiety subscale were also associated with higher ‘g’ scores (Full ‘g’: B=-0.04, SE=0.01, P<.001; Core ‘g’: B=-0.03, SE=0.01, *P*<.001). Self-report of any mental health problem was associated with lower performance on the core CONCA tasks (Core 'g': B=-0.11, SE=0.04, *P*=.01) but this association was not found for Full 'g' (B=-0.09, SE=0.05, *P*=.07).

Technical Issues

Technical issues were reported by 52 participants (52/3679, 1.4%) and 17 unique problems were identified. Three of these problems were determined as bugs in the website coding and were resolved. Where the problems were the result of bugs in the assessment and participants were unable to view the stimuli, participants were given the opportunity to complete the task once the issue was resolved. Five were identified as issues that were specific to those users' devices and further technical support was provided by our team to support each participant to complete the tasks if possible. For the remaining 9 issues, insufficient information was provided and attempts to contact the participants for further information were unsuccessful.

Discussion

Principal Findings

The aims of this study were to further develop CONCA to provide a brief measure of 'g', to recruit from a large web-based population study and demonstrate CONCA's application in a health population sample. Results from each aim are outlined in the sections below.

Part 1: Establishing CONCA

The number of tasks in CONCA was reduced from 9 to 3 core tasks and 2 optional tasks. All these tasks loaded onto a single factor, 'g', which supported our decision to reduce the number of tasks in the battery for the purpose of creating a brief assessment that provides a measure of ‘g’. The measure of ‘g’ obtained using the tasks from the core CONCA battery was correlated with ‘g’ derived from the MCCB, which indicates that the three tasks are sufficient to obtain a valid measure of ‘g’. This correlation increased when the two optional tasks were included, suggesting that whilst the optional tasks are not essential to derive a measure of ‘g’, they do have added value.

Part 2: Assessing Cognition in HealthWise Wales

To demonstrate CONCA’s application in a health population sample, we examined completion rates, technical issues, and performance distributions. This enabled us to determine whether the tasks were sufficiently engaging and challenging for a general population sample. Completion rates for the core CONCA tasks were high indicating acceptable levels of tolerability and engagement. These rates were similar to those reported in our previous study [14]. Over half the sample completed both additional optional tasks (2048/3679 participants, 52.66%), which suggests that participants were sufficiently engaged with the core tasks and our research to be motivated to complete additional measures. It should be highlighted that participants with higher scores on the core tasks were more likely to complete the optional tasks. This suggests that those who find the tasks more difficult may be demotivated and choose not to complete the optional tasks leading to a less representative sample for these tasks. The number of technical issues reported was low with only 52 of 3679 participants (1.41%) reporting a problem. Combined with the high completion rates, this suggests that most participants were able to complete the tasks without a problem. The distributions of scores for most of the tasks were relatively normal, except for Vocabulary where there was evidence of a potential ceiling effect, particularly amongst older participants. This ceiling effect among older people has been identified in a previous report examining the psychometric properties of Vocabulary [28].

The relationship between performance on the tasks and age, gender and education were in the expected direction. Older age and lower education levels were associated with lower scores on all tasks and measures of ‘g’, except for Vocabulary where older participants performed better. Men performed better on Vocabulary and Matrix Reasoning than women, whilst women had higher scores on Morphed Emotion Identification. This is consistent with previous studies assessing emotion recognition [37, 38] and matrix reasoning [28]. In contrast, a previous report assessing the psychometric properties of the Vocabulary task showed marginally better performance in women [28].

We found lower performance amongst those using touchscreen devices (tablet computer or smartphone) compared to those using a laptop or desktop computer, which is consistent with two other studies using these tasks [28, 39]. This effect was seen across all the tasks suggesting that it cannot be explained by response times alone, as some tasks such as Vocabulary do not have a timed component. The lower performance may be partly explained by screen size, particularly as lower performance was found amongst participants using smartphones compared to those using tablet computers. This is supported by the findings of Passell et al. [39] who demonstrated that performance on Digit Symbol Coding and Vocabulary was impacted by screen size, input type and the type of internet browser used. Device use has been associated with age, gender, and education [39], all of which were controlled for in the current study but there may be other factors that were not measured in this study. Smartphones and tablet computers may be cheaper and more accessible, as they do not rely on a home broadband connection and have relatively straightforward interfaces compared to traditional computers. Therefore, their use may be influenced by socioeconomic factors or computer literacy, which may also be associated with performance on the tasks. Consistent with this, a report by the UK's communications regulator, Ofcom, found that people in manual occupations, unemployed or considered financially vulnerable were most likely to use a smartphone exclusively to access the internet [40]. The portable nature of touchscreen devices means that participants may be more likely to complete the tasks in locations outside the home or whilst conducting other activities and therefore may be subject to more distractions. These results highlight the importance of controlling for device effects when analysing cognitive data from web-based studies.

CONCA was designed to be a measure of cognition in psychiatric populations. Therefore, we evaluated whether the mental health measures collected were associated with performance on CONCA. We found that higher levels of depression and anxiety symptoms, and self-reported history of diagnosis or treatment for a mental health problem were associated with lower overall performance on the core CONCA tasks. This suggests that CONCA is sufficiently sensitive to the cognitive differences associated with mental health disorders. This is also a novel finding of the study, as to our knowledge, few studies have examined the relationship between depression and anxiety symptoms and cognition in a general population sample.

Sample Representativeness

The response rate of 13.9% raises the issue of participation bias. There was evidence of bias in the demographic distributions of the CONCA sample. Compared to population estimates for Wales, the sample was older, more educated, and predominantly women. We did not stratify the data by ethnicity as 99% of participants reported their ethnicity as White, which was a consequence of recruiting from the wider HealthWise Wales sample (98% White). The bias reported in this study is in part a reflection of the original HealthWise Wales sample, which has a higher proportion of women, older people and White people [16]. However, even amongst the least represented groups (e.g. men aged 16-40), the number of participants in our sample exceeds the amount of normative data collected for other mental health cognitive batteries, such as the Brief Assessment for Cognition in Schizophrenia [41] and the MATRICS Consensus Cognitive Battery [42]. Whilst the sample did contain a higher number of participants with postgraduate degrees than expected, it is important to note that the representation across the education groups was satisfactory with at least 200 participants in each group. The proportion of participants reporting no qualifications was also comparable to estimates for the Welsh population, which alleviates concerns that the sample may be under-represented by those with lower educational attainment. We are currently undertaking targeted recruitment to collect data on younger people with a particular focus on recruiting more men into the sample.

Strengths and Limitations

We have collected a large cognitive dataset on a population sample that spans a wide range of ages and enabled us to derive age-, gender- and education-based norm scores for CONCA. However, results should be interpreted with the consideration of the potential biases in the sample, as detailed below. CONCA has several advantages over existing assessments (such as BACS [43] or CANTAB [44]) including a user-friendly website designed with input from patient representatives and health professionals, a large normative dataset collected online, and it can be completed on the participants' own devices (including touchscreen tablets and smartphones) rather than relying on specific hardware or software that can be required for similar assessments.

Sample representativeness is a clear limitation of this study, as highlighted in the previous section. In addition, participants with high scores in the core tasks were more likely to complete the optional tasks. This needs to be considered when interpreting results using the Matrix Reasoning and Morphed Emotion Identification tasks and is another source of bias. It should also be noted that the response rate for this study was 13.9%. Recruitment for this study commenced in January 2020 and overlapped with the initial months of the COVID-19 pandemic and UK lockdown. There is evidence that the pandemic negatively impacted research participation, with current research participants less able and/or willing to participate in ongoing research [45]. The main limitations of CONCA include a lack of verbal or episodic memory tasks, and a lack of evidence for its use as a longitudinal assessment, although some data on practice effects have been previously published [28].

Conclusions

CONCA provides a valid measure of 'g', which can be derived using as few as 3 tasks that take no more than 15 minutes. We have demonstrated that the battery is sufficiently engaging and challenging for use in a general population sample with the potential exception of Vocabulary in older adults. Based on our findings, we recommend that CONCA is suitable for use in general population samples and may be particularly useful for studies of the relationship between cognition and mental health, but caution is advised for the use of Vocabulary in older adults (60 years and older) given the potential for ceiling effects. Factors that impacted performance on CONCA included age, gender, education, and type of device and these should be controlled for in analyses as appropriate. The primary purpose of this study was to introduce the new CONCA battery, provide normative data and demonstrate the associations between CONCA and demographic variables. The recruitment of a web-based normative sample is an important step forward in the development of CONCA, although more work is needed to ensure the data is representative of the population, particularly in terms of education levels. However, we have also reported some novel findings, namely that symptoms of depression and anxiety are associated with cognitive function in a general population sample, as well as demonstrating the effect of device when measuring cognition. Now that we have established normative performance on CONCA, we intend to investigate the clinical utility of CONCA, including the development of new features to support health professionals in interpreting their patient’s performance on the battery when administered in a clinical setting.

Declarations

Acknowledgements

Authors' contributions: AJL is the lead author and was involved in all aspects of the study, including designing the assessment tools and study methodology, overseeing recruitment of participants, conducting analyses and interpretation of the data, and drafting and redrafting the manuscript. JTRW is the senior author, the principal investigator of CoMPaSS and was involved in all aspects of the paper. IRJ is the chief investigator and Director of the NCMH and advised on the methodology and interpretation of the results. All authors critically revised the paper and approved the final version to be submitted.

Funding: This work was supported by a Medical Research Council (MRC) PhD studentship to A.L and by the National Centre for Mental Health (NCMH). The NCMH is a collaboration between Cardiff, Swansea and Bangor Universities and is funded by Welsh Government through Health and Care Research Wales. The NCMH provided resources and staff to assist with the development of the battery. The further development of the battery has been supported by two MRC grants, MRC Mental Health Data Pathfinder award (MC_PC_17212) and MRC Confidence in Concept award (MC_PC_17170). The MRC Pathfinder award (MC_PC_17212) and NCMH funded A.L's postdoctoral position. The MRC Confidence in Concept award (MC_PC_17170) is a proof-of-concept grant, which funded the further development of CONCA, including creating a user-friendly interface and supporting public involvement. The work at Cardiff University was funded by Medical Research Council (MRC) Centre (MR/L010305/1) and Program Grant (G0500509).

This study was facilitated by HealthWise Wales, the Health and Care Research Wales initiative, which is led by Cardiff University in collaboration with SAIL, Swansea University. Responsibility for the interpretation of the information provided in this manuscript is the authors' alone. We thank all the participants from HealthWise Wales who contributed their data to the CONCA study.

This research was conducted with the support of The Many Brains Project, Inc. The web-based tasks remain the intellectual property of The Many Brains Project, which retains the copyright for these tasks and the source code. The authors thank Dr Laura Germine for her support in developing this battery.

We thank the members of NCMH-PAR and the patients and clinicians from Headroom Cardiff and Vale UHB Youth Psychosis Service for contributing to the development of CONCA.

Conflicts of Interest

All authors were involved in the design of CONCA, the cognitive assessment described in the study methods. Professor Walters has received grant funding from Takeda Pharmaceuticals for research unrelated to this manuscript.

Data Availability

The datasets generated and analysed during the current study are available through HealthWise Wales on reasonable request (https://www.healthwisewales.gov.wales/for-researchers/).

Multimedia Appendix

Multimedia Appendix 1: "Lynham Walters CONCA – Supplementary.docx". This file contains additional information on the task selection process and supplementary results tables.

References

1. Shipley BA, Der G, Taylor MD, and Deary IJ, Cognition and all-cause mortality across the entire adult age range: Health and lifestyle survey. Psychosomatic medicine, 2006. 68(1). PMID:16449407
2. Bosworth HB, Schaie KW, and Willis SL, Cognitive and sociodemographic risk factors for mortality in the seattle longitudinal study. The Journals of Gerontology: Series B, 1999. 54(5): P273-P282. PMID:10542820
3. Wraw C, Deary IJ, Gale CR, and Der G, Intelligence in youth and health at age 50. Intelligence, 2015. 53: 23-32. PMID:26766880
4. Batty GD, Deary IJ, and Gottfredson LS, Premorbid (early life) iq and later mortality risk: Systematic review. Annals of Epidemiology, 2007. 17(4): 278-288. PMID:17174570
5. Bora E, Harrison BJ, Yücel M, and Pantelis C, Cognitive impairment in euthymic major depressive disorder: A meta-analysis. Psychological Medicine, 2012. 43(10): 2017-2026. PMID:23098294
6. Gualtieri CT and Morgan DW, The frequency of cognitive impairment in patients with anxiety, depression, and bipolar disorder: An unaccounted source of variance in clinical trials. The Journal of clinical psychiatry, 2008. 69(7): 1122-1130. PMID:18572982
7. Ou Y-N, Tan C-C, Shen X-N, Xu W, Hou X-H, Dong Q, Tan L, and Yu J-T, Blood pressure and risks of cognitive impairment and dementia. Hypertension, 2020. 76(1): 217-225. PMID:32450739
8. Xue M, Xu W, Ou Y-N, Cao X-P, Tan M-S, Tan L, and Yu J-T, Diabetes mellitus and risks of cognitive impairment and dementia: A systematic review and meta-analysis of 144 prospective studies. Ageing Research Reviews, 2019. 55: 100944. PMID:31430566
9. Heinrichs RW and Zakzanis KK, Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. Neuropsychology, 1998. 12(3): 426-445. PMID:9673998
10. Bora E, Yucel M, and Pantelis C, Cognitive endophenotypes of bipolar disorder: A meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. Journal of Affective Disorders, 2009. 113(1–2): 1-20. PMID:18684514
11. Depp CA, Mausbach BT, Harmell AL, Savla GN, Bowie CR, Harvey PD, and Patterson TL, Meta-analysis of the association between cognitive abilities and everyday functioning in bipolar disorder. Bipolar Disorders, 2012. 14(3): 217-226. PMID:22548895
12. Fett A-KJ, Viechtbauer W, Penn DL, van Os J, and Krabbendam L, The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. Neuroscience & Biobehavioral Reviews, 2011. 35(3): 573-588. PMID:20620163
13. Feenstra HEM, Vermeulen IE, Murre JMJ, and Schagen SB, Online cognition: Factors facilitating reliable online neuropsychological test results. The Clinical Neuropsychologist, 2016: 1-26. PMID:27266677

14. Lynham AJ, Jones IR, and Walters JTR, Web-based cognitive testing in psychiatric research: Validation and usability study. J Med Internet Res, 2022. 24(2): e28233. PMID:35142640
15. Dickinson D and Harvey PD, Systemic hypotheses for generalized cognitive deficits in schizophrenia: A new take on an old problem. Schizophrenia Bulletin, 2009. 35(2): 403-414. PMID:18689868
16. Hurt L, Ashfield-Watt P, Townson J, Heslop L, Copeland L, Atkinson MD, Horton J, and Paranjothy S, Cohort profile: Healthwise wales. A research register and population health data platform with linkage to national health service data sets in wales. BMJ Open, 2019. 9(12): e031705. PMID:31796481
17. Lynham AJ, Hubbard L, Tansey KE, Hamshere ML, Legge SE, Owen MJ, Jones IR, and Walters JT, Examining cognition across the bipolar/schizophrenia diagnostic spectrum. Journal of psychiatry & neuroscience: JPN, 2018. 43(4): 245-253. PMID:29620518
18. Underwood JF, Kendall KM, Berrett J, Lewis C, Anney R, Van den Bree MB, and Hall J, Autism spectrum disorder diagnosis in adults: Phenotype and genotype findings from a clinically derived cohort. The British Journal of Psychiatry, 2019. 215(5): 647-653. PMID:30806336
19. The Many Brains Project. *Testmybrain*. Accessed: 28 October 2020; Available from: https://www.testmybrain.org.
20. Germine L, Nakayama K, Duchaine B, Chabris C, Chatterjee G, and Wilmer J, Is the web as good as the lab? Comparable performance from web and lab in cognitive/perceptual experiments. Psychonomic Bulletin & Review, 2012. 19(5): 847-857. PMID:22829343
21. Marder SR and Fenton W, Measurement and treatment research to improve cognition in schizophrenia: Nimh matrics initiative to support the development of agents for improving cognition in schizophrenia. Schizophrenia Research, 2004. 72(1): 5-9. PMID:15531402
22. Nuechterlein K, Green M, Kern R, Baade L, Barch D, Cohen J, Essock S, Fenton W, Frese F, and Gold J, The matrics consensus cognitive battery, part 1: Test selection, reliability, and validity. American Journal of Psychiatry, 2008. 165(2): 203-213. PMID:18172019
23. Rehm J, Üstün TB, Saxena S, Nelson CB, Chatterji S, Ivis F, and Adlaf E, On the development and psychometric testing of the who screening instrument to assess disablement in the general population. International Journal of Methods in Psychiatric Research, 1999. 8(2): 110-122. DOI: 10.1002/mpr.61
24. Tucker-Drob EM and Salthouse TA, Methods and measures: Confirmatory factor analysis and multidimensional scaling for construct validation of cognitive abilities. International Journal of Behavioral Development, 2009. 33(3): 277-285. PMID:20963182
25. Jones KH, Ford DV, Jones C, Dsilva R, Thompson S, Brooks CJ, Heaven ML, Thayer DS, McNerney CL, and Lyons RA, A case study of the secure anonymous information linkage (sail) gateway: A privacy-protecting remote access system for health-related research and evaluation. Journal of biomedical informatics, 2014. 50(100): 196-204. PMID:24440148
26. Zigmond AS and Snaith RP, The hospital anxiety and depression scale. Acta Psychiatrica Scandinavica, 1983. 67(6): 361-370. PMID:6880820

27. Lynham A, Jones I, and Walters J. *Cardiff online cognitive assessment*. Accessed: 4 May 2022; Available from: https://www.conca.wales/about.html.
28. Passell E, Dillon DG, Baker JT, Vogel SC, Scheuer LS, Mirin NL, Rutter LA, Pizzagalli DA, and Germine L, Digital cognitive assessment: Results from the testmybrain nimh research domain criteria (rdoc) field test battery report. PsyArXiv, 2019. DOI:10.31234/osf.io/dcszr
29. Wechsler D, Wais-iii, wechsler adult intelligence scale: Administration and scoring manual. Psychological Corporation; 1997. ISBN: 0158981030
30. Nelson H and Willison J, Nart: National adult reading test. Windsor: NFER ‐ Nelson; 1991. ISBN: 9780700504756
31. Frigerio E, Burt DM, Montagne B, Murray LK, and Perrett DI, Facial affect perception in alcoholics. Psychiatry Research, 2002. 113(1): 161-171. PMID:12467955
32. Kessels RP, Montagne B, Hendriks AW, Perrett DI, and Haan EH, Assessment of perception of morphed facial expressions using the emotion recognition task: Normative data from healthy participants aged 8–75. Journal of neuropsychology, 2014. 8(1): 75-93. PMID:23409767
33. Wechsler D and Hsiao-pin C, Wasi ii: Wechsler abbreviated scale of intelligence. 2nd. Psychological Corporation, 2011. ISBN:9780158981567
34. Office for National Statistics. *Population estimates for the uk, england and wales, scotland and northern ireland: Mid-2020*. Accessed: 21 October 2022; Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2020#age-structure-of-the-uk-population.
35. Office for National Statistics. *Annual population survey: Levels of highest qualifications held by working age adults by area and gender*. Accessed: 5 April 2022; Available from: https://statswales.gov.wales/Catalogue/Education-and-Skills/Post-16-Education-and-Training/Data-For-Regions-of-Wales/highest-qualification-level-of-working-age-adults-by-region-and-local-authority.
36. Welsh Government. *National survey for wales - adult general health and illness by age and gender (2019-2020)*. Accessed: 5 April 2022; Available from: https://statswales.gov.wales/Catalogue/National-Survey-for-Wales/Population-Health/Adult-general-health-and-illness/genhealthillness-by-age-gender.
37. Hoffmann H, Kessler H, Eppel T, Rukavina S, and Traue HC, Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. Acta Psychologica, 2010. 135(3): 278-283. PMID:20728864
38. Collignon O, Girard S, Gosselin F, Saint-Amour D, Lepore F, and Lassonde M, Women process multisensory emotion expressions more efficiently than men. Neuropsychologia, 2010. 48(1): 220-225. PMID:19761782
39. Passell E, Strong RW, Rutter LA, Kim H, Scheuer L, Martini P, Grinspoon L, and Germine L, Cognitive test scores vary with choice of personal digital device. Behavior Research Methods, 2021. 53(6): 2544-2557. PMID:33954913

40. Ofcom. *Adults' media use and attitudes report*. Accessed: 22 February 2023; Available from: https://www.ofcom.org.uk/__data/assets/pdf_file/0020/234362/adults-media-use-and-attitudes-report-2022.pdf.
41. Keefe RSE, Harvey PD, Goldberg TE, Gold JM, Walker TM, Kennel C, and Hawkins K, Norms and standardization of the brief assessment of cognition in schizophrenia (bacs). Schizophrenia Research, 2008. 102(1): 108-115. PMID:18495435
42. Kern RS, Nuechterlein KH, Green MF, Baade LE, Fenton WS, Gold JM, Keefe RSE, Mesholam-Gately R, Mintz J, Seidman LJ, Stover E, and Marder SR, The matrics consensus cognitive battery, part 2: Co-norming and standardization. American Journal of Psychiatry, 2008. 165(2): 214-220. PMID:18172018
43. Atkins AS, Tseng T, Vaughan A, Twamley EW, Harvey P, Patterson T, Narasimhan M, and Keefe RSE, Validation of the tablet-administered brief assessment of cognition (bac app). Schizophrenia Research, 2017. 181: 100-106. PMID:27771201
44. Backx R, Skirrow C, Dente P, Barnett JH, and Cormack FK, Comparing web-based and lab-based cognitive assessment using the cambridge neuropsychological test automated battery: A within-subjects counterbalanced study. J Med Internet Res, 2020. 22(8): e16792. PMID:32749999
45. Cardel MI, Manasse S, Krukowski RA, Ross K, Shakour R, Miller DR, Lemas DJ, and Hong Y-R, Covid-19 impacts mental health outcomes and ability/desire to participate in research among current research participants. Obesity, 2020. 28(12): 2272-2281. PMID:32845582