PLAGIARISM DECLARATION

PLAGIARISM

This means that you present substantial portions or elements of another's work, ideas or data as your own, even if the original author is cited occasionally. A signed photocopy or other copy of the Declaration below must accompany every piece of work that you hand in.

DECLARATION

1. We know that Plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. We have used the American Psychological Association formatting for citation and referencing. Each significant contribution to, and quotation in, this essay/report/project from the work or works, of other people has been attributed, cited, and referenced.
3. This essay/report/project is our own work.
4. We have not allowed, and will not allow, anyone to copy our work with the intention of passing it off as their own work.

SIGNATURE: Harry Stanley, Danielle Mostert

DATE: 27 10 2022

Supervisor: Bjorn Christ

Department: Psychology

Exploring Accuracy-Based Measures of Intraindividual Variability in Alzheimer's Disease

Harry Stanley & Danielle Mostert

Supervisor: Dr Bjorn Christ

Department of Psychology

University of Cape Town

Course Convenor: Dr Leigh Schrieff-Brown

Thesis Word Count: 7988

Abstract word count: 232

Abstract

Age is one of the biggest risk factors for the onset of neurodegenerative diseases, such as Alzheimer's disease (AD), and with life expectancy on the rise around the world the prevalence of age-related diseases such as AD is likely to increase. With the scarcity of research into this disease as well as the socioeconomic burden of AD in South Africa, there is an urgent need for local AD research. One promising line of inquiry uses within-person fluctuation of cognitive performance, or intraindividual variability (IIV), to try and capture the earliest stages of age-related cognitive change. The present study used the Paired Associates Learning (PAL) task and List Learning task as two such measures and the Mini-Mental State Examination (MMSE) and Cambridge Examination for Mental Disorders of the Elderly-Revised (CAMCOG-R) as our outcome measures of general cognitive functioning. The aim of this study was to assess whether these accuracy-based measures of IIV are useful in identifying and capturing AD. We found significant between-group differences in accuracy-based measures of IIV but that they were not able to predict clinical group membership or change in cognitive performance over time. Overall, accuracy-based measures have significant clinical utility in the identification and capture of AD, but accuracy-based measures of IIV in the PAL seem to be less powerful in this regard when compared both to list learning recognition and the overall reaction time measures of IIV.

Keywords: Alzheimer's disease; accuracy; intraindividual variability; list learning; paired-associates learning

Life expectancy is on the rise around the world, and with age being the biggest risk factor in the onset of dementia, the prevalence of these diseases is likely to increase (De Jager et al., 2015; Lopez et al., 2010; Olayinka & Mbuyi, 2014). As of 2015 there were an estimated 46-million people living with dementia worldwide with predictions that this number will increase to 131.5 million by 2050 (Prince et al., 2016). Dementia is a neurological disease caused by the degeneration of brain tissue that impacts various cognitive faculties such as memory, language, orientation, and executive functioning. It is usually progressive and chronic with Alzheimer's Disease (AD) being the most commonly diagnosed form of dementia (De Jager et al., 2015; Rasmussen & Langerman, 2019). There is a distinct poverty of research into dementia and AD in sub-Saharan Africa, especially South Africa, which results in a unsurprising lack of understanding of the condition and misattribution of symptoms to normal aging (De Jager et al., 2015; Olayinka & Mbuyi, 2014). Dementia creates an immense burden on the patient as well as their families due to the loss of independence that dementia poses. With the ever-increasing prevalence of age-related major neurocognitive disorder such as AD there is a need for accurate measures and robust methods that are sensitive to cognitive changes at the earliest stages of the disease process, and that are easily implemented and which may supplement existing diagnostic tools (Christ et al., 2018; Rasmussen & Langerman, 2019). This will have the benefit of facilitating earlier and more informed treatment plans as well as contribute to a deeper understanding of the risk factors that facilitate rapid cognitive decline in AD.

One method, using intraindividual variability (IIV), has been shown to have promise in detecting mild cognitive impairment (MCI) (Jackson et al., 2012; Roalf et al., 2016; Troyer et al., 2016; Vaughan et al., 2013) and predicting cognitive decline in dementia patients (Bangen et al., 2019; Christ et al., 2018; Costa et al., 2019; Holtzer et al., 2008; Vaughan et al., 2013). With the majority of research in this area focusing on latency measures of reaction

time, there is little research on the utility of accuracy based IIV measures of cognitive functioning. Thus, research that helps validate these measures as indicators of IIV in the detection of dementia pathology and in the prediction of longitudinal change has significant clinical utility (Christ et al., 2018).

Intraindividual Variability

IIV can be referred to as the within-person variation in an individual's cognitive or behavioural performance in a given measure (MacDonald et al., 2006). Measures of IIV are superior predictors of neuropathological outcomes, when compared to standard neuropsychological assessments that rely on mean-level measurement of cognitive functioning, as IIV predicts outcome over and above such assessments and provides clinically relevant and unique information over and above mean performance (Costa et al., 2019; Haynes et al., 2017; Jackson et al., 2012). Furthermore, IIV can be operationalised as either inconsistency or dispersion where inconsistency refers to an individual's performance variability within a specific task over several occasions, and dispersion refers to an individual's performance variability across several different tasks on a single occasion (Halliday et al., 2018).

The presence of IIV in cognitive tasks is an early indicator of both healthy and pathological aging. As healthy adults age there is increased IIV in cognitive performance indicating age-related cognitive decline, and this increase in IIV is more distinct in neuropathology such as brain dysfunction, neurodegeneration, and other brain-related disorders (Haynes et al., 2017; Lin & McDonough, 2021; MacDonald & Stawski, 2020; MacDonald et al., 2006). Research has shown that increased IIV predicts cognitive decline (Bielak et al., 2010), functional decline (Bangen, et al., 2019), and the subsequent disease progression from normal aging to MCI/dementia/AD (Anderson et al., 2016; Gorus et al., 2008; Holtzer et al., 2008; Kälin et al., 2014; Roalf et al., 2016; Troyer et al., 2016). With the

urgent need for the early detection of AD, studies show that IIV is a sensitive measure of MCI/ prodromal AD (Kälin et al., 2014; Roalf et al., 2016) and may contribute to the early detection of dementia (Vaughan et al., 2013).

Reaction (RT) Measures

Most of the research in IIV has focused on the use of RT measures (e.g. Bielak et al., 2010; Jackson et al., 2012; MacDonald & Stawski, 2020; Stawski et al., 2019; Tales et al., 2012). Investigations into IIV using RT measures have established its utility as a marker of normal and pathological aging, early cognitive decline, MCI, AD, dementia and the associated severity of impairment (Bielak et al., 2010; Gorus et al., 2008; Jackson et al., 2012; MacDonald & Stawski, 2020; Tales et al., 2012). One of the reasons why research in IIV has primarily been focused on RT measures is that correlations between RT measures and different variables (e.g., age, clinical group status) would remain significant after controlling for mean performance, whereas for accuracy- based measures, correlations would fail to remain significant after controlling for the mean (Christ, et al., 2018). This indicates that accuracy-based measures of IIV do not have useful predictive power outside of what is indicated from an individual's mean-level of performance. (Salthouse et al., 2006). The most likely reason for this is that there is higher temporal resolution of reaction time data compared to accuracy-based data due to the larger number of data points that are captured from the participant (Christ et al., 2018). This higher temporal resolution captures the variance in the participants performance more accurately, and thus facilitates more accurate and more powerful predictive modelling of any relevant data being used in IIV research.

Accuracy-Based Measures

An alternative to RT measures in IIV research is the use of accuracy-based measures. Fewer researchers use accuracy-based measures in IIV due to their lack of predictive power (e.g. Christ et al., 2018; Kälin et al., 2014; Murphy et al., 2007; Tractenberg & Pietrzak,

2011). Accuracy-based measures in IIV research have been reported to detect memory distortions due to aging, cognitive decline and functional decline, and to predict MCI, AD and incident dementia (Holtzer et al., 2008; Kälin et al., 2014; Kliegel & Sliwinski, 2004; Morgan et al., 2012; Murphy et al., 2007; Tractenberg & Pietrzak, 2011).

Accuracy-based IIV measures are most useful as indicators of a specific type of cognitive impairment associated with a distinct neurological or cognitive domain that has been damaged but are less useful as indicators of diffuse neurological or cognitive deterioration (Christ et al., 2018). Such measures are used more widely than RT-based measures in everyday clinical routine as part of the patients' diagnostic and disease management process because accuracy-based IIV measures can be calculated without using additional tasks or adding multiple trials of the same task to the standard test battery, therefore reducing the testing burden on patients (Christ et al., 2018; Kälin et al., 2014). Therefore, validating accuracy-based measures of IIV to detect pathology and predicting longitudinal change has significant potential clinical benefit (Christ et al., 2018).

There is a need for sensitive and accurate cognitive measures that can predict cognitive decline in AD patients in order to supplement existing clinical diagnostic tools in AD's treatment. Accuracy-based measures of cognitive decline are more widely used for diagnostic purposes and in the disease management process, but their usefulness can be extended. Most studies in the literature on this topic of cognitive decline have focused on RT measures with very few investigating the efficacy of accuracy-based measures to predict long-term cognitive change, and this study hoped to bridge that discrepancy in the literature.

Rationale, Aims and Hypotheses

With the prevalence of age-related diseases likely to increase as a result of increased life expectancy, there is a dire need for the increased study and awareness of such diseases (De Jager et al., 2015; Lopez et al., 2010; Olayinka & Mbuyi, 2014). Awareness of AD in

South Africa is exceptionally low which can lead to the appropriate preventative measures not being sought out. Additionally, at time of writing there is no national plan to deal with dementia as a whole in South Africa (De Jager et al., 2015; Prince etal., 2016).

Hence, this study aimed to contribute to the literature on the overall utility of IIV in accuracy-based measures in the identification and capturing of AD. An additional aim was to contribute to a better understanding of predictive methods in AD as a supplement to existing clinical diagnostic tools in its treatment, specifically in the South African context. Furthermore, this study aimed to contribute to the literature on IIV research by expanding our understanding of useful measures of IIV through investigating two accuracy-based measures that are scarcely mentioned in IIV literature.

Finally, the central research question that this study intended to answer was whether accuracy-based measures of IIV are useful in identifying AD, capturing AD, and predicting cognitive decline in AD. Based on this study's aims and research question, we hypothesized that: (**H_1**) there is a significant difference between the clinical and control groups for accuracy-based measures; (**H_2**) There is a significant difference between the clinical and control groups for accuracy-based measures of IIV; (**H_3**) Clinical group membership can be predicted with accuracy-based measures of IIV.; and (**H_4**) Cognitive decline in AD can be predicted by accuracy-based measures of IIV.

Method

Design and Setting

This study was a secondary analysis and was part of a larger ongoing project on the utility of using IIV methods to track the trajectory of cognitive decline in the progression of AD (see Christ et al., 2018). The parent study employed a measurement-burst design (Nesselroade, 1991) which involves every participant undergoing three periods of serial testing throughout the course of 15 months. Each testing period involved three testing

sessions taken over the course of two weeks. This study employed data from the first testing period referred to hereon as interval one.

Participants

The participants in the parent study consisted of a control group (n = 26) that was cognitively healthy as well as a clinical test group (n = 26) which consisted of mild-to-moderate stage possible or probable AD patients, according to NINCDS-ADRDA clinical criteria (McKhann et al., 1984; Appendix A). All clinical participants in the parent study were recruited by clinicians from the Groote Schuur Hospital Memory Clinic and Geriatric Unit following standard international criteria for the diagnosis of AD. The participants recruited for the control group were from the surrounding Cape Town area and were informed of the study through flyers at seniors' clubs, retirement villages, old age homes, and word of mouth.

Four sets of inclusion criteria were established for participation in the study: a) access to accurate medical history of clinical participants; b) those who were 55 years of age or older; c) those with adequate proficiency in English; and d) the availability of someone who could elucidate the authors of the parent study about sudden or recent changes in cognitive functioning. Additionally, six sets of exclusion criteria were established for participation in the study: a) a diagnosis of a condition that would confound results from cognitive measures such as HIV/AIDS, uncontrolled diabetes mellitus, uncontrolled hypertension; b) a recent diagnosis of a psychiatric illness; c) a score of > 15/30 on the Geriatric Depression Scale at any of the testing phases (GDS; Yesavage et al., 1982); d) any major neurological disorder being present such as Huntington's, Parkinson's disease, or a history of cerebrovascular accidents such as a stroke within the last 2 years; e) a smoking (> 20 per day) or alcohol/drug abuse history; and f) a score of <12 on the Mini-Mental State Examination at any of the

testing phases (MMSE; Folstein et al., 1975). The above criteria applied to participants for both the clinical test and control groups.

Participation did not involve any direct physical, psychological, or social risks. However, participants may have been exposed to potential sources of discomfort. For instance, participants were required to travel to GSH to attend three testing sessions in each testing period, that is if they preferred to be tested at GSH instead of their home. Furthermore, the cognitive assessments were burdensome and time-consuming.

Participants received compensation of R70 at each test session for travelling costs to and from GSH. Additionally, participants who displayed potentially treatable illness at the screening session were, with consent of the participant, referred to the hospital for treatment.

Measures

Participants in the parent study were administered a battery of eleven tests. This study used five tests from the original battery: Two tests measured accuracy-based performance in both visual and verbal memory, and three tests measured affect and general cognitive functioning as part of the study's screening procedure.

Accuracy-Based Measures of Memory

Paired Associates Learning (PAL) Task. This test forms a part of the Cambridge Neuropsychological Test Automated Battery (CANTAB; Cambridge Cognition, 2004) and evaluates visual memory and learning. Performance on this test is affiliated with ideal functioning of the medial temporal lobe (Sahakian & Owen, 1992) (see appendix B). The test begins with a blank screen and white boxes on the periphery of the screen, and each box is then randomly opened and subsequently closed. A pattern contained in one or more of the boxes is displayed when a box containing said pattern is opened. Following all the boxes' 'contents' being revealed, the patterns that were previously shown to the participant are then exhibited in the middle of the screen sequentially. Participants are directed to match the

patterns to the specific box from which they originated on the periphery of the screen. If an error is made, the patterns are redisplayed to remind the participant of their original placements. The test itself is comprised of five different stages that became progressively more difficult with the first tweo stages having two boxes, the third stage having three boxes, the fourth stage having sex boxes and the final stage having eight boxes. Each stage consisted of six trials and only a single block of trials was given to participants. Since only a single block of trials was administered, the test took an estimated 10 minutes to complete for each participant.

Whilst the PAL task is used fairly often in AD research such as in the detection of early AD, there is little literature on IIV using the test (Christ et al., 2018). The PAL has been shown to have more than adequate test-retest reliability with one month retest with reliability correlations in the range of .86 to .88 (Fowler et al., 1995; Low et al., 1998). More recent literature suggests that the PAL has moderately acceptable reliability with three-month test-retest reliability with reliability coefficients of .73 (Karlsen et al., 2016).

List Learning Task. The battery that this test forms a part of is the RBANS battery as was originally constructed to screen for dementia in the elderly (Randolph, 1998). This task involved participants being read a list of 10 words who were then told to recall as many of them as possible immediately after presentation. This procedure was repeated again four additional times and after a 25 to 30 minute delay, recall of the list was prompted once again. Following this a recognition task was administered to the participant where they were requested to point out the words that were originally on this list from a new list of 20 words. There were 10 targets and 10 foils, and the test itself took an estimated 15 minutes to complete.

Duff et al. (2005) showed that test-retest reliability at a one year follow up of the RBANS list learning memory task had reliability coefficients between .53 and .67. The IIV

literature reveals that word list-learning and recognition tasks are employed quite frequently and are able to discriminate between older and younger adults (Murphy et al., 2007), as well as between those with early AD and neurologically intact older participants (Hultsch et al., 2000).

Screening Measures

The measures described below are the measures of affect and general cognitive functioning used in this study.

The Cambridge Examination for Mental Disorders of the Elderly-Revised (CAMCOG-R). As a measure of general cognitive functioning, our study employed data collected by the parent study using the *CAMCOG-R* (Huppert et al., 1999). The original CAMCOG was developed by Huppert et al. (1986), revised in 1999, and forms a part of the Cambridge Examination for Mental Disorders in the Elderly (CAMDEX) used as an early detection tool for elderly dementia (Zeltzer et al., 2009). The *CAMCOG-R* contains 67 items which collectively measure eight different functional domains including perception, orientation, memory, language, attention, calculation, praxis, and abstract thinking. The measure demonstrates moderate to high test-retest reliability and inter-rater reliability (Lima & Lorenco, 2010; O' Conner et al., 1989). The *CAMCOG*-R takes an estimated 25-minutes to complete and has been shown to be suitable in low-to-middle-income-countries (LMIC's), and whilst the battery overall is sensitive to levels of education, the recent memory and learning memory subscales are not affected by levels of education (James et al., 2015). Both of these types of subscales are sensitive to picking up MCI and early AD (James et al., 2015). Within the *CAMCOG*-R there are 19 items which form the Mini-Mental State Examination.

Mini-Mental State Examination (MMSE). As a measure of screening for dementia and identifying participants with severe dementia, our study employed data collected by the parent study using the MMSE (Folstein et al., 1975) (see Appendix C). The measure consists

of 19 items with each question scored out of 5 except for the recall question which is scored out of 3 (Galea, 2005). The questions altogether form seven sub-categories: orientation to place, orientation to time, three-word registration, calculation and attention, language, three-word recall, and visual construction. Overall, the test has shown to have good test-retest reliability with correlations between .8 and .95 (Folstein et al., 1975; Tombaugh & McIntyre, 1992). The MMSE is the popular assessment of cognitive functioning, though its use in a South African context is being contested in contemporary literature (Schutte et al., 2021).

Geriatric Depression Scale (GDS). To identify participants with clinical levels of depression, the GDS was employed in the parent study (see Appendix D). A score greater than nine out of 30 is indicative of depression and participants within this range were excluded from the parent, and derivatively, this study. The measure itself was developed for detecting depression in older adults and is a self-report questionnaire with 30 items. A reliability generalization study of the GDS showed that across 338 studies, the average reliability coefficient of the GDS is very high at .85 (Kieffer & Reese, 2002).

Procedure

The parent study was granted ethical clearance by the Human Research Ethics Committee of the Faculty of Health Sciences of the University of Cape Town in 2014 (HREC/REF: 167/2014) (See Appendix E for the parent study's approval letter).

The team involved in the parent study administered the screening measures after participants had read and signed an informed consent form (see Appendix F). This document explained the parent study's purpose, procedures, as well as the risks and benefits of participation. Clinical participants had a relative or friend present to ensure that consent was informed and voluntary. Participants in the control group were assessed individually while clinical participants were assessed with a relative or friend present. Participants were either assessed in a private examination room in the Geriatric Unit at GSH, or at the participant's

home. The scores of each screening measure were calculated by the test administrator and then checked against the eligibility criteria. The participants who did not meet the eligibility criteria were excluded from future participation in the study after receiving a valid explanation. Individuals who were excluded based on their GDS scores received a note stating their score to give to their primary care physician. Individuals who were excluded based on a DSM-IV diagnosis of major neurocognitive impairment received a referral to the GSH Memory Clinic. Those participants who met the eligibility criteria were scheduled to attend their first testing session.

The testing phase commenced within 30 days of the screening phase. As in the screening phase, control group participants were tested individually while clinical participants were tested with a relative or friend present. Participants were either tested in a private examination room in the Geriatric Unit at GSH, or at the participant's home. In each testing period, data was collected from three testing sessions taken over the course of two weeks. Each testing session differed in the order of test administration so as to prevent order effects. In order to maintain confidentiality, participants' medical records were retained at GSH, their identities and medical histories were only known by the team involved in the parent study, and the data itself was anonymized. All hardcopy data were stored in a secure filing cabinet, and all data in digital format were password protected.

Data Preparation And Statistical Analysis

Data Preparation

The collected data from the PAL and List Learning Task measures were prepared by calculating the scores for each of the Accuracy formula A, Accuracy formula B, Reaction Time, Errors and Recognition variables. We extracted intraindividual standard deviations (iSDs) from the scores of each of these variables. The means as well as the iSDs of each of these variables were used as predictor variables in our statistical analyses. The collected data

from the MMSE and CAMCOG-R measures were prepared by calculating the rate of change. The rate of change for both the MMSE and the CAMCOG-R measures as well as their baseline measures were used as outcome variables in our statistical analyses.

Accuracy Scores. We developed two different formulas for calculating accuracy-based data on the PAL measure.

Accuracy A. The formula used to calculate the scores for Accuracy A is as follows:

$$\sum \frac{Si}{\sqrt{Ai}}$$

S- Stage difficulty
A-attempts for that stage
i-Stage number

The formula for Accuracy A sums the number of correct responses in a stage divided by the square root of number of attempts it took to complete that stage.

Accuracy B. The formula used to calculate the scores for Accuracy B is as follows:

$$\sum \frac{Si}{Ai}$$

S- Stage difficulty
A-attempts for that stage
i-Stage number

The formula for Accuracy B sums the number of correct responses divided by the number of attempts it took to complete the stage.

Reaction time. To calculate the reaction time scores on data from the PAL measure, the reaction times for all the correct responses were averaged.

Errors. We used the total errors (adjusted) on the data from the PAL measure generated by CANTAB software. The total number of errors were adjusted for each stage not attempted as a result of failing a previous stage. To calculate the adjustment of the total errors, the software summed the number of patterns that were not attempted, and then from this amount it subtracted the number of patterns divided by the number of boxes. It then multiplied this result by the number of trials allowed per stage, which for this study was six trials (Cambridge Cognition, 2004).

List Learning Recognition. The List Learning Recognition data is from the List Learning Task measure. The recognition scores were obtained from the parent study.

Rate of change. The rate of change was calculated using data from the MMSE and CAMCOG-R measures and were used as output variables. The rate of change is the difference between the test scores of participants' first and last screening sessions, divided by the amount of time (months) between those sessions.

Extraction of iSDs. This study followed the extraction of iSDs method described by Christ et al. (2018). Before calculating the iSDs, which are the IIV scores for each of the Accuracy A, Accuracy B, Reaction Time, Errors and List Learning Recognition predictor variables, it was important to partial out any systematic effects, such as group and time-on-task effects, that may have had an influence on mean performance (Christ et al., 2018). To extract the ISDs, we ran a random intercept model on each of the variables to determine which group and time-on-task effects had a significant influence on the means of each of these variables. We added session and test order to evaluate the influence of time-on-task effects on mean performance, and group, age, sex, years of education, and income to evaluate the influence of group effects on mean performance. (Hultsch et al., 2000, 2002, 2008; see Christ et al., (2018) for detailed explanation). The random intercept models found the following fixed-effects that significantly influenced the mean performance of each predictor variable: for Accuracy A, group and age significantly influenced mean performance; for Accuracy B, group, age and session significantly influenced mean performance; for Reaction Time, group and age significantly influenced mean performance, for Errors, group, age, years of education, and session significantly influenced mean performance; for List Learning Recognition, group, age and session significantly influenced mean performance. Based on these significant main effects, the next step in the extraction of the iSDs involved running random coefficient models with random slopes on sessions for each of the predictor variables

and their significant fixed effects and included session in each model to test whether it was significant or not. These models purify the data of any group and time-on-task effects. We saved the residuals from each of the models, converted the residuals to z-scores, converted these to *t*-scores and then to compute the iSDs we calculated the SDs of the *t*-scores.

Inferential Statistical Analyses

SPSS (version 28) was used for all statistical analyses in this study, with α set at .05. We ensured that all assumptions were upheld before statistical analyses were conducted.

In the first part of our statistical analysis, we generated a series of independent *t*-tests for parametric continuous variables, Mann-Whitney tests for non-parametric variables, and chi-square tests for categorical variables to assess between-group differences in affective, cognitive, and sample demographics. The same analyses were done to assess between-group differences in the predictor variables. In the second part of our statistical analyses, in preparation for regression modelling, we examined the bivariate associations using Pearson's *r* correlation coefficient between each of the predictor variables and rate of change outcome variables. In the final part of our statistical analyses, we generated a logistic stepwise regression model to test the hypothesis that clinical group membership can be predicted by accuracy-based measures of IIV.

Our five predictor variables were derived from the PAL and the List Learning task. From the PAL our predictors were two formula that we created for the purpose of this study (A and B), total adjusted error rate and average reaction time. From the List Learning task our fifth predictor variable was recognition of list-learnt words. Participants that had missing data for thee predictor variables in any three of the testing sessions for interval one were removed as all three sessions were needed for the extraction of IIV. In our calculation for rate of change over time for cognitive performance measured with the MMSE and CAMCOG-R,

five participants had to be excluded based on individual circumstances between the first testing interval (*n* = 49*)* and the last interval (*n* = 43). These included withdrawal from the study, significantly lower cognitive scores suggesting their data was no longer valid, diagnostic change, or severe health decline. These exclusions were only for the calculation of rate of cognitive change over time and for all other analyses the baseline data from interval one was used (*n* = 49).

Results

Descriptive Statistics and Hypotheses

From table 1. we can see that the clinical and control group had largely similar distributions for sex with no significant differences. The average difference in age between groups was only just significant difference with an average difference of 3.91 years. Difference in total years of education were significantly different between groups with an average difference of 1.18 years suggesting that most participants in the control group had achieved a matric whilst most participants in the clinical group had achieved either a grade ten or eleven pass. No significant differences in GDS or income were found between groups. Significant differences were found for both measures of cognitive functioning at both the baseline first interval of MMSE, final interval of MMSE, first interval of CAMCOG-R,and final interval of CAMCOG-R.

From table 2. We can see that for raw mean scores that precede IIV extraction, AD scores were significantly lower than that of controls across the board. The first accuracy formula (A), the second formula (B) , the total adjusted error rate, mean reaction time, and list learning recognition scores were all significant in their differences between the control and AD group (table 2.).

Table 1
Descriptive statistics for between group differences in demographic, affective, and cognitive scores.

		Group				
Variable	Total	Control (n = 26)	AD (n = 23)	*t/u/X²*	*P*	*Df*
Age	71.76 (6.89)	69.92 (7.27)	73.83 (5.19)	-2.05	.046*	47
Sex (M:F)	14:35	8:20	8:15	.819	.37	1
Education (years)	10.86 (2.87)	11.58 (2.64)	10.04 (2.95)	1.92	.03*	47
GDS	5.9(3.29)	5.58 (3.35)	6.26 (3.25)	-0.72	.47	47
Income	7877.05 (5392.93)	9191.81 (5511.04)	6390.8 (4958.82)	217	.09	
MMSE						
First Interval (n=49)	25.49 (4.32)	28 (1.62)	22.65 (4.68)	90	<.001**	-
Final Interval (n = 44)	24.5 (5.19)	28.48 (1.34)	20.14 (4.22)	9	<.001**	-
CAMCOG-R						
First Interval (n=49)	80.24 (17.48)	91.15 (6.53)	67.91 (17.86)	44.5	<.001**	-
Final Interval (n=44)	93.61 (4.77)	93.61 (4.77)	62.71 (12.14)	10.912	<.001**	25.58

*The second, third, and fourth columns represent means for each variable with standard deviations in brackets. AD is Alzheimer's Disease, GDS is Geriatric Depression Scale, MMSE is the Mini-Mental State Examination, and CAMCOG-R is the Cambridge Cognitive Examination for Mental Disorders of the Elderly-Revised. Education was the highest level attained, and Income was monthly household income. Mann-Whitney's tests were run for non-parametric data (Income, both intervals of the MMSE, the first interval of the CAMCOG-R), t-tests were run for parametric data (Age, Education, GDS, and the final interval of the CAMCOG-R), and a Chi-Squared test of contingency was run for the categorical Sex variable. *p<.05, **p<.001.*

Table 2.
Descriptive statistics for between group differences in predictor variables for raw mean scores and intra-individual variability (iSD's)

		Group				
Variable	Total	Control (n = 26)	AD (n = 23)	*t/u/X²*	*p*	*df*
Means						
PAL						
Accuracy (A)	5.34 (2.63)	6.97 (2.03)	3.50 (1.93)	73	<.001**	-
Accuracy (B)	2.22 (.84)	2.74 (.60)	1.64 (.69)	5.945	<.001**	47
Errors	49.30 (25.42)	33.03 (17.70)	67.70 (19.68)	73.5	<.001**	-
Reaction Time	3048.86 (1686.31)	2153.24 (985.32)	4061.29 (1755.41)	53	<.001**	-
List Learning Recognition	16.87 (2.93)	18.86 (1.07)	14.62 (2.74)	41.5	<.001**	-
iSD's						
PAL						
Accuracy (A)	9.15 (7.09)	11.51 (7.60)	6.50 (5.50)	192	.032*	-
Accuracy (B)	10.67 (5.13)	10.82 (5.76)	10.52 (4.44)	.201	.841	47
Errors	25.42 (7.23)	11.39 (8.90)	7.12 (3.70)	192	.032*	-
RT	6.33 (9.58)	3.17 (4.33)	9.90 (12.41)	147	.002*	-
List Learning Recognition	9.60 (6.73)	7.50 (5.68)	11.99 (7.13)	160	.005*	-

*The second, third, and fourth columns represent means for each variable with standard deviations in brackets. PAL refers to the Paired Associates Learning Task, Accuracy A and B refer to the two formulas derived for this study respectively. Mann-Whitney tests were run for non-parametric data (Accuracy A, Total Adjusted Error Rate, Reaction Time, and List learning for both iSD's and raw mean scores), and t-tests were used for parametric data (Accuracy B). *p<.05, **p<.001.*

Thus, we can reject our first null hypothesis (**H_1**) that there are no significant differences between the clinical AD and control groups for accuracy-based measures. For

IIV, all were significant except for the second accuracy formula (B). IIV for the first accuracy formula (A), for the total adjusted error rate, for mean reaction time, and for list learning recognition were all statistically significant. Interestingly however, IIV for our first accuracy formula (A) as well as for total adjusted error rate were higher in the control groups than in the clinical AD group. Thus, we can reject our second null hypothesis (**H_2**) that there are not significant differences between the clinical AD and control group for accuracy-based measures of IIV with the exception of the second accuracy-based formula (B).

The third hypothesis (**H_3**) was that accuracy based measures of IIV can predict clinical group membership. Forward logistic stepwise regression models were run for the IIV predictors that showed statistically significant differences between clinical AD and control groups (accuracy formula A, total adjusted error rate, reaction time, and list recognition; see table 2). In the first step only the IIV predictor was included in the model, in step two the two covariates that showed statistically significant differences between clinical AD and control group (age and education) were includes to see if the IIV accuracy measure retained unique statistical significance by predicting group membership, and finally step the raw mean score for the predictor was added to see if IIV retained significance. This final step was to see if accuracy-based measures of IIV could predict clinical group membership over and above the mean of that same score.

The first logistic regression model (table 4.) using the first accuracy formula (A) IIV showed that initially our accuracy-based IIV measure could predict group membership quite well with statistical significance shown. This significance was retained when significant covariates were added , but by the final step only the mean score for the measure had significance with the IIV measure retaining none. Variance explained was highest in step 3 suggesting that raw mean score was the best predictor of group membership and not IIV.

Table 4.
IIV accuracy formula (A) logistic forward stepwise bivariate regression predicting group membership

	Predictor	R^2	β	*SE*	*p*	CI (95%) LB	CI (95%) UP
Step 1		.17					
	iSD		-.12	.05	.018*	.81	.98
Step 2		.31					
	iSD		-.1	.05	.047*	.82	.999
	Age		.11	.06	.05	.999	1.24
	Education		-.18	.12	.15	.66	1.09
Step 3		.54					
	iSD		.02	.08	.78	.88	1.19
	Age		.01	.06	.94	.89	1.14
	Education		.04	.16	.8	.76	1.44
	Mean		-.78	.29	.008*	.26	.82

*iSD is the measure of IIV. Nagelkerke's R-squared was used as a measure of variance explained. β is the unstandardized beta coefficient. All degrees of freedom for all models were 1. *p<.05*

The second logistic regression model (table 5.) using total adjusted error-rate IIV also showed initial promise in predicting group membership in step 1, but by step 2 this statistical significance was lost with the introduction of the covariates. By step 3 only the raw mean score showed statistical significance with the IIV measure showing none. Again variance explained was highest in step 3 suggesting that raw mean score was the best predictor of group membership and not IIV.

Table 5
IIV total error-rate logistic forward stepwise bivariate regression predicting group membership

	Predictor	R^2	β	*SE*	*p*	CI (95%) LB	CI (95%) UP
Step 1		.155					
	iSD		-.15	.08	.04*	.74	1
Step 2		.307					
	iSD		-.13	.08	.06	.996	1.23
	Age		.1	.05	.06	.996	1.23
	Education		-.2	.12	.1	.65	1.04
Step 3		.539					
	iSD		-.11	.09	.23	.75	1.07
	Age		.01	.06	.94	.89	1.14
	Education		.1	.18	.58	.78	1.57
	Mean		.08	.02	.001**	1.03	1.13

.

Table 6.
IIV mean reaction time logistic forward stepwise bivariate regression predicting group membership

	Predictor					CI (95%)	
		R^2	β	*SE*	*p*	LB	UP
Step 1		.23					
	iSD		.17	.08	.03*	1.01	1.4
Step 2		.29					
	iSD		.13	.08	.14	.96	1.34
	Age		.07	.06	1.54	.96	1.19
	Education		-.15	.13	.22	.67	1.1
Step 3		.49					
	iSD		.01	.05	.78	.92	1.11
	Age		.02	.06	.71	.91	1.15
	Education		-.07	.15	.66	.7	1.25
	Mean		.001	<.001	.01*	1	1

The third logistic regression model (table 6.) using mean reaction time IIV also had initial predictive power in predicting group membership in step 1 (β = .17, p =.03). However, by step 2 this statistical significance was lost with the introduction of the covariates (β= .13, p = .14) and by step 3 again only the raw mean score showed statistical significance (β = .001, p = .01) with the IIV measure showing none (β = .01, p = .78). Again variance explained was highest in step 3 (R^2= .49) suggesting that raw mean score was the best predictor of group membership and not IIV.

Table 7.
IIV list recognition logistic forward stepwise bivariate regression predicting group membership

	Predictor					CI (95%)	
		R^2	β	*SE*	*p*	LB	UP
Step 1		.16					
	iSD		.12	.06	.03*	1.01	1.13
Step 2		.29					
	iSD		.1	.06	.09	.94	1.24
	Age		.09	.06	.1	.98	1.22
	Education		-.2	.12	.1	.64	1.04
Step 3		.72					
	iSD		.02	.08	.84	.88	1.18
	Age		-.07	.09	.42	.78	1.11
	Education		-.06	.17	.73	.68	1.31
	Mean		-1.28	.44	.004*	.12	.66

The final logistic regression model (table 7.) using list recognition IIV again also had initial promise in predicting group membership in step 1 and again by step 2 this statistical

significance was lost with the introduction of the covariates. By step 3 similarly only the raw mean score showed statistical significance with the IIV measure showing no statistical significance. Variance explained was once again highest in step 3 suggesting that raw mean score was the best predictor of group membership and not IIV.

With all the logistic forward stepwise bivariate regression models above and the lack of statistical significance of our IIV measures in their ability to predict group membership above covariates or the associated raw mean score, we failed to reject our third null hypothesis (**H_3**) that accuracy-based measures of IIV are able to predict clinical group membership.

Our final hypothesis (**H_4**) was that cognitive decline in AD patients can be predicted over time using accuracy-based measures of IIV using a multiple linear regression model. The first step in such an analysis is to run a correlation matrix to identify significant correlations, and table 3 shows the results from this analysis.

Table 3.
Correlations for raw mean and IIV predictors with measures of cognitive change over time

	Rate of Change			
	CAMCOG-R		MMSE	
Mean	*r*	*p*	*r*	*p*
Accuracy (A)	-0.09	0.29	-0.36	0.008*
Accuracy (B)	-0.16	0.21	-0.35	0.009*
Error	0.11	0.24	0.4	0.003*
Reaction Time	0.24	0.06	0.34	0.013*
List Recognition	-0.26	0.04*	-0.46	<0.001**
ISD				
Accuracy (A)	0.09	0.28	-0.14	0.19
Accuracy (B)	-0.002	0.5	-0.1	0.26
Error	0.06	0.36	0.02	0.5
Reaction Time	0.21	0.08	0.2	0.1
List Recognition	0.003	0.49	0.08	0.31

*The correlation measure used here (r) was Pearson's correlation coefficient. The rate of change for the CAMCOG-R and MMSE are quantifications of the rate of change in general cognitive functioning between baseline scores and the final testing stage for participants where most were measured in interval 3 but some in interval 2 and our calculation accounts for this. *p<.05, **p<.001.*

The raw mean correlations with CAMCOG-R rate of change for our first accuracy formula (A), our second accuracy formula, total adjusted error rate, and mean reaction time were not significant with only list recognition showing a significant correlation. However, the

raw mean correlations with MMSE rate of change for our first accuracy formula (A), our second accuracy formula (B), total adjusted error rate, mean reaction time and list recognition were all statistically significant suggesting that accuracy based measures are correlated with rate of cognitive change. However, no significant correlations were found between the IIV of our predictors and the rate of change in CAMCOG-R and MMSE (refer to table 1) with the correlation between IIV reaction time and CAMCOG-R rate of change approaching the closest to statistical significance. As a result, the multiple regression analysis was halted and as a result we failed to reject our final null hypothesis (**H_4**) that accuracy based measures of IIV can predict cognitive change over time.

Discussion

The aim of this study was to investigate the overall utility of accuracy based measures of IIV in the identification and capture of AD as well as to contribute to a better understanding of predictive methods in AD as a supplement to existing clinical diagnostic tools in its treatment, specifically in the South African context. The central research question asked by this study was whether accuracy-based measures of IIV are useful in identifying AD, capturing AD, and predicting cognitive decline in AD. We thus has the following four hypotheses: : (**H_1**) there is a significant difference between the clinical and control groups for accuracy-based measures; (**H_2**) There is a significant difference between the clinical and control groups for accuracy-based measures of IIV; (**H_3**) Cognitive decline in AD can be predicted by accuracy-based measures of IIV; and (**H_4**) Clinical group membership can be predicted with accuracy-based measures of IIV.

It was found in our analyses that there was a significant difference in raw accuracy-based mean scores between the AD group and control group in all of our predictor variables. Our two accuracy-based formula (A & B) for measuring performance on the PAL and the list learning recognition scores were all significantly higher in the control group than the AD

group suggesting better performance on this task. The control group also had significantly lower error rates as well as reaction times than the AD group, further suggesting better performance on this task. Thus we were able to reject our first null hypothesis (**H_1**) there is a significant difference between the clinical and control groups for accuracy-based measures. These findings are in line with what one would expect as these measures capture visual memory and learning and are affiliated with ideal functioning of the medial and temporal lobe (Sahakian & Owen, 1992) and these are faculties that are distinctly affected in the presence of AD (De Jager et al., 2015; Rasmussen & Langerman, 2019). This suggests that performance on accuracy-based measures of learning and memory have clinical utility in capturing AD at early stages of the disease process.

Additionally it was found that for all accuracy-based measures of IIV there was a significant difference between the clinical AD and control group, but not in an identical direction to the raw mean scores of the associated predictor variables. Our first accuracy-based formula (A) IIV as well as total adjusted error IIV (both derived from the PAL) were both statistically higher in the control group than the AD group which was a surprise finding since other accuracy-based measures of IIV such as list-learning have been found to be higher in AD groups than control groups (Christ et al., 2018). This type of result could be explained by the fact that tasks higher in strategic processing could create more IIV and as a result variability becomes a function of learning and thinking rather than of neuropathology and dysfunction (Christ e al., 2018; Allaire & Masiske, 2005). Reaction time IIV as well as list learning recognition were statistically lower in the control group than the AD group which is more in line with what we would expect. This is because such differences in IIV are indicative of less optimal performance in the given measure which one would expect from the AD group over the controls as has been found in previous studies (Holtzer et al., 2008; Kälin et al., 2014; Kliegel & Sliwinski, 2004; Morgan et al., 2012; Murphy et al., 2007;

Tractenberg & Pietrzak, 2011). We thus were able to reject our second null hypothesis (**H_2**) that there were significant differences in accuracy-based measured of IIV between control and AD groups. Overall these findings suggests that measures of IIV for reaction time in the PAL and list learning recognition have some utility in identifying cognitive impairment to a certain degree, but that accuracy-based measures of IIV from the PAL cannot be used to identify or capture AD at its earliest stages.

This determination is congruent with our next finding as neither of these accuracy-based measures of IIV (formula A or error rate) were able to predict clinical membership above the ability of the raw mean score for both these measures. The same was found for all other IIV predictor variables suggesting that accuracy-based measures of IIV have no ability to predict the presence of AD. Thus we failed to reject our third null hypothesis (**H_3**) that accuracy-based measures of IIV are unable to predict clinical group membership. Whilst disappointing, this finding is in line with previous research that found that that accuracy-based measures of IIV lack predictive power (Christ et al., 2018; Kälin et al., 2014; Murphy et al., 2007; Tractenberg & Pietrzak, 2011). This can most likely be explained by the fact that accuracy-based measured of IIV have inherently less temporal resolution than RT measures since far fewer trials are required to produce iSD's.

Finally, it was found that accuracy-based measures of IIV were unable to predict cognitive decline over time. The initial indicator of any multiple regression's predictive power is correlation between predictor variables and outcome variables, and in our case no statistically significant correlations between accuracy-based measures of IIV and rate of change in cognitive performance over time were found thus indicating that no predictive relationship exists. As a result we failed to reject our final null hypothesis (**H_4**) that accuracy-based measures of IIV cannot predict cognitive change over time. To our knowledge this was the first study that sought to find such a predictive relationship.

Limitations

There were a number of limitations present in this study. As previously noted, this study was a secondary data analysis and part of a larger ongoing study (Christ et al., 2018). Because this study was a secondary data analysis, we had no control over the sample size beyond reducing the number of participants used in the analysis. The parent study's original sample of N =52 was reduced to N = 49 for this study. Due to not having a larger sample size, this study may not have had sufficient statistical power to achieve its aims, and this may have impacted our ability to find statistically significant results for two of our hypotheses. Thus, a smaller sample size could have led to incorrectly accepting the null hypothesis. Hence, any future studies that may wish to replicate this study should recruit a larger sample size.

We found a statistical difference between the clinical and control groups for the age demographic variable. Age is likely to have been a confounder for the clinical group which resulted in this statistical difference between the control and clinical groups in age. It is also furthermore impossible to know to what extent this statistical difference in age for groups is actually as a result of normal aging related changes or if it is due to AD for the clinical group. To overcome this limitation in future studies, the sample for the clinical and control groups could be better matched in terms of age. As there was also a significant difference between the groups in terms of education, this limitation could be overcome in future studies by adequately matching the sample groups in terms of education.

Summary and Conclusion

Awareness of AD in South Africa is exceptionally low which can lead to the appropriate preventative measures not being sought out. There is a need for sensitive and accurate cognitive measures that can predict cognitive decline in AD patients in order to

supplement existing clinical diagnostic tools in AD's treatment. Hence, the overall aim of this study was to contribute to the literature on the overall utility of IIV in accuracy-based measures in the identification and capturing of AD. Statistical analyses detected significant between-group differences for both accuracy-based measures, and accuracy-based measures of IIV with IIV being higher in two accuracy-measured derived from the PAL, an unexpected result to say the least.

Overall, accuracy-based measures have significant clinical utility in the identification and capture of AD, but accuracy-based measures of IIV in the PAL seem to be less powerful in this regard when compared both to list learning recognition and the overall reaction time measures of IIV. Future research into other accuracy-based measures of IIV should be conducted to assess if they are able to predict clinical group membership as well as change in cognitive performance over time.

Acknowledgements

The authors of this study would like to thank Professor Bjorn Christ as the supervisor for this project. Without him we would not have been able to do this project at all. We would also like to thank Professor Lauren Wild and Catherine Wild at the University of Cape Town for their help in guiding he psychology honours class of 2022 (the present authors included) through their research projects. Additionally we would like to thank Mr. Milton Gering for his instrumental help with developing the two statistical formulas employed in this study for capturing accuracy rates in the paired associates learning task.

References

Anderson, E. D., Wahoske, M., Huber, M., Norton, D., Li, Z., Koscik, R. L., Umucu, E., Johnson, S. C., Jones, J., Asthana, S., & Gleason, C. E. (2016). Cognitive variability—A marker for incident MCI and AD: An analysis for the Alzheimer's Disease Neuroimaging Initiative. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 4(1), 47-55. https://doi.org/10.1016/j.dadm.2016.05.003.

Bangen, K. J., Weigand, A. J., Thomas, K. R., Delano-Wood, L., Clark, L. R., Eppig, J., Werhane, M. L., Edmonds, E. C., & Bondi, M. W. (2019). Cognitive dispersion is a sensitive marker for early neurodegenerative changes and functional decline in nondemented older adults. *Neuropsychology*, *33*(5), 599–608. https://doi.org/10.1037/neu0000532.

Bielak, A. A., Hultsch, D. F., Strauss, E., MacDonald, S. W., & Hunter, M. A. (2010). Intraindividual variability in reaction time predicts cognitive outcomes 5 years later. *Neuropsychology*, *24*(6), 731. https://doi.org/10.1037/a0019802.

Cambridge Cognition. (2004). CANTAB eclipse test administration guide. Cambridge: UK Cambridge Cognition. https://www.cambridgecognition.com/cantab/.

Christ, B. U., Combrinck, M. I., & Thomas, K. G. (2018). Both reaction time and accuracy measures of intraindividual variability predict cognitive performance in Alzheimer's disease. *Frontiers in Human Neuroscience*, *12*, 124. https://doi.org/10.3389/fnhum.2018.00124.

Costa, A. S., Dogan, I., Schulz, J. B., & Reetz, K. (2019). Going beyond the mean: Intraindividual variability of cognitive performance in prodromal and early neurodegenerative disorders. *The Clinical Neuropsychologist*, *33*(2), 369-389. https://doi.org/10.1080/13854046.2018.1533587.

Duff, K., Beglinger, L. J., Schoenberg, M. R., Patton, D. E., Mold, J., Scott, J. G., & Adams, R. L. (2005). Test-retest stability and practice effects of the RBANS in a community dwelling elderly sample. Journal of Clinical and Experimental Neuropsychology, 27(5), 565-575. https://doi.org/10.1080/13803390490918363.

De Jager, C. A., Joska, J. A., Hoffman, M., Borochowitz, K. E., & Combrinck, M. I. (2015). Dementia in rural South Africa: A pressing need for epidemiological studies. SAMJ: *South African Medical Journal*, *105*(3), 189-190. https://doi.org/10.7196/SAMJ.8904.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12(3), 189-198. https://doi.org/10.1016/0022-3956(75)90026-6.

Fowler, K. S., Saling, M. M., Conway, E. L., Semple, J. M., & Louis, W. J. (1995). Computerized delayed matching to sample and paired associate performance in the early detection of dementia. Applied Neuropsychology, 2(2), 72-78. https://doi.org/10.1207/s15324826an0202_4.

Galea, M., & Woodward, M. (2005). Mini-mental state examination (MMSE). Australian Journal of Physiotherapy, 51(3), 198. https://doi.org/10.1016/s0004-9514(05)70034-9.

Gorus, E., De Raedt, R., Lambert, M., Lemper, J. C., & Mets, T. (2008). Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology*, *21*(3), 204-218. https://doi.org/10.1177/0891988708320973.

Golay, P., Fagot, D., & Lecerf, T. (2013). Against coefficient of variation for estimation of intraindividual variability with accuracy measures. Tutorials in Quantitative Methods for Psychology, 9(1), 6-14. http://archive-ouverte.unige.ch/unige:26621.

Halliday, D. W., Stawski, R. S., Cerino, E. S., DeCarlo, C. A., Grewal, K., & MacDonald, S. W. (2018). Intraindividual variability across neuropsychological tests: Dispersion and disengaged lifestyle increase risk for Alzheimer’s disease. *Journal of Intelligence*, *6*(1), 12. https://10.3390/jintelligence6010012.

Haynes, B. I., Bauermeister, S., & Bunce, D. (2017). A systematic review of longitudinal associations between reaction time intraindividual variability and age-related cognitive decline or impairment, dementia, and mortality. *Journal of the International Neuropsychological Society*, *23*(5), 431-445. https://doi.org/10.1017/S1355617717000236.

Holtzer, R., Verghese, J., Wang, C., Hall, C. B., & Lipton, R. B. (2008). Within-person across-neuropsychological test variability and incident dementia. *Jama*, *300*(7), 823-830. https://doi.org/10.1001/jama.300.7.823.

Jackson, J. D., Balota, D. A., Duchek, J. M., & Head, D. (2012). White matter integrity and reaction time intraindividual variability in healthy aging and early-stage Alzheimer

disease. *Neuropsychologia*, *50*(3), 357-366. https://doi.org/10.1016/j.neuropsychologia.2011.11.024.

James, K. A., Grace, L. K., Thomas, K. G., & Combrinck, M. I. (2015). Associations between CAMCOG-R subscale performance and formal education attainment in South African older adults. International Psychogeriatrics, 27(2), 251-260. https://doi.org/10.1017/S1041610214002233.

Kälin, A. M., Pflüger, M., Gietl, A. F., Riese, F., Jäncke, L., Nitsch, R. M., & Hock, C. (2014). Intraindividual variability across cognitive tasks as a potential marker for prodromal Alzheimer's disease. *Frontiers in Aging Neuroscience*, *6*(1), 147. https://doi.org/10.3389/fnagi.2014.00147.

Karlsen, R. H., Karr, J. E., Saksvik, S. B., Lundervold, A. J., Hjemdal, O., Olsen, A., ... & Skandsen, T. (2020). Examining 3-month test-retest reliability and reliable change using the Cambridge Neuropsychological Test Automated Battery. Applied Neuropsychology: Adult, 1-9. https://doi.org/10.1080/23279095.2020.1722126.

Kieffer, K. M., & Reese, R. J. (2002). A reliability generalization study of the Geriatric Depression Scale. Educational and Psychological Measurement, 62(6), 969-994. https://doi.org/10.1177%2F0013164402238085.

Kliegel, M., & Sliwinski, M. (2004). MMSE cross-domain variability predicts cognitive decline in centenarians. *Gerontology*, *50*(1), 39-43. https://doi.org/10.1159/000074388.

Lima, D. A., & Lourenço, R. A. (2010). Cross-cultural adaptation of section A of the Cambridge Examination for Mental Disorders of the Elderly-Revised Version (CAMDEX-R) for dementia diagnosis. Cadernos de Saúde Pública, 26(7), 1345-1354. https://www.scielo.br/j/csp/a/H48szTGP6Nw3zd6yKvbvSkg/?format=pdf&lang=pt.

Lin, S. S. H., & McDonough, I. M. (2021). Intra-individual cognitive variability in neuropsychological assessment: a sign of neural network dysfunction. *Aging, Neuropsychology, and Cognition*, *1*(1), 1-25. https://doi.org/10.1080/13825585.2021.2021134.

Lopez, O. L., Schwam, E., Cummings, J., Gauthier, S., Jones, R., Wilkinson, D., Waldemar, G., Zhang, R., & Schindler, R. (2010). Predicting cognitive decline in Alzheimer's

Field Code Changed

disease: an integrated analysis. Alzheimer's & Dementia, 6(6), 431-439. https://doi.org/10.1016/j.jalz.2010.04.003.

Lowe, C., & Rabbitt, P. (1998). Test\ re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: theoretical and practical issues. Neuropsychologia, 36(9), 915-923. https://doi.org/10.1016/S0028-3932(98)00036-0.

MacDonald, S. W., & Stawski, R. S. (2020). Longitudinal changes in response time mean and inconsistency exhibit predictive dissociations for risk of cognitive impairment. *Neuropsychology*, *34*(3), 264. http://dx.doi.org/10.1037/neu0000608.

MacDonald, S. W., Nyberg, L., & Bäckman, L. (2006). Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. *Trends in Neurosciences*, *29*(8), 474-480. https://doi.org/10.1016/j.tins.2006.06.011.

Murphy, K. J., West, R., Armilio, M. L., Craik, F. I., & Stuss, D. T. (2007). Word-list-learning performance in younger and older adults: Intra-individual performance variability and false memory. *Aging, Neuropsychology, and Cognition*, *14*(1), 70-94. https://doi.org/10.1080/138255890969726.

Morgan, E. E., Woods, S. P., & Grant, I. (2012). Intra-individual neurocognitive variability confers risk of dependence in activities of daily living among HIV-seropositive individuals without HIV-associated neurocognitive disorders. *Archives of Clinical Neuropsychology*, *27(*3), 293-303. https://doi.org/10.1093/arclin/acs003.

Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), Best methods for the analysis of change: Recent advances, unanswered questions, future directions (pp. 92–105). American Psychological Association. https://doi.org/10.1037/10099-006.

Olayinka, O. O., & Mbuyi, N. N. (2014). Epidemiology of dementia among the elderly in sub-Saharan Africa. *International Journal of Alzheimer's disease, 2014*(1). https://doi.org/10.1155/2014/195750.

Prince, M., Comas-Herrera, A., Knapp, M., Guerchet, M., & Karagiannidou, M. (2016). World Alzheimer report 2016: improving healthcare for people living with dementia: coverage, quality and costs. http://eprints.lse.ac.uk/67858/.

Field Code Changed

Randolph, C., Tierney, M. C., Mohr, E., & Chase, T. N. (1998). The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. *Journal of Clinical and Experimental Neuropsychology, 20*(3), 310-319. https://doi.org/10.1076/jcen.20.3.310.823.

Rasmussen, J., & Langerman, H. (2019). Alzheimer's disease–why we need early diagnosis. *Degenerative Neurological and Neuromuscular Disease, 9*(1), 123. https://doi.org/10.2147/DNND.S228939.

Roalf, D. R., Quarmley, M., Mechanic-Hamilton, D., Wolk, D. A., Arnold, S. E., & Moberg, P. J., & Alzheimer's Disease Neuroimaging Initiative. (2016). Within-individual variability: An index for subtle change in neurocognition in mild cognitive impairment. *Journal of Alzheimer's Disease, 54*(1), 325-335. https://doi.org/10.3233/JAD-160259.

Sahakian, B. J., & Owen, A. M. (1992). Computerized assessment in neuropsychiatry using CANTAB: discussion paper. Journal of the Royal Society of Medicine, 85(7), 399-402. https://ncbi.nlm.nih.gov/pmc/articles/PMC1293547.

Salthouse, T. A., Nesselroade, J. R., & Berish, D. E. (2006). Short-term variability in cognitive performance and the calibration of longitudinal change. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61*(3), 144-151. https://doi.org/10.1093/geronb/61.3.P144.

Schutte, C. M., Tsikane, M., & Nchoe, K. (2021). Developing and testing a South African Brief Cognitive Score in literate and illiterate people of mixed language groups. South African Journal of Psychiatry, 27(1), 1-6. https://doi.org/10.4102/sajpsychiatry.v27i0.1327.

Stawski, R. S., MacDonald, S. W., Brewster, P. W., Munoz, E., Cerino, E. S., & Halliday, D. W. (2019). A comprehensive comparison of quantifications of intraindividual variability in response times: A measurement burst approach. *The Journals of Gerontology: Series B, 74*(3), 397-408. https://doi.org/10.1093/geronb/gbx115.

Tales, A., Leonards, U., Bompas, A., Snowden, R. J., Philips, M., Porter, G., Haworth, J., Wilcock, G., & Bayer, A. (2012). Intra-individual reaction time variability in amnestic mild cognitive impairment: A precursor to dementia? Journal of Alzheimer's Disease, 32(2), 457–466. https://doi.org.10.3233/JAD-2012-120505.

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: a comprehensive review. Journal of the American Geriatrics Society, 40(9), 922-935. https://doi.org/10.1111/j.1532-5415.1992.tb01992.x.

Tractenberg, R. E., & Pietrzak, R. H. (2011). Intra-individual variability in Alzheimer's disease and cognitive aging: definitions, context, and effect sizes. *PLoS One*, *6*(4), e16973. https://doi.org/10.1371/journal.pone.0016973.

Troyer, A. K., Vandermorris, S., & Murphy, K. J. (2016). Intraindividual variability in performance on associative memory tasks is elevated in amnestic mild cognitive impairment. *Neuropsychologia*, *90*(1), 110-116. https://doi.org/10.1016/j.neuropsychologia.2016.06.011.

Vaughan, L., Leng, I., Dagenbach, D., Resnick, S. M., Rapp, S. R., Jennings, J. M., Brunner, R. L., Simpson, S. L., Beavers, D. P., Coker, L. H., Gaussoin, S. A., Sink, K. M., & Espeland, M. A. (2013). Intraindividual variability in domain-specific cognition and risk of mild cognitive impairment and dementia. Current Gerontology and Geriatrics Research, 2013(1). https://doi.org/10.1155/2013/495793.

Yesavage, J. A., Brooks 3rd, J. O., Taylor, J., & Tinklenberg, J. (1993). Development of aphasia, apraxia, and agnosia and decline in Alzheimer's disease. The American Journal of Psychiatry, 150(5), 742-747. https://doi.org/10.1176/ajp.150.5.742

Appendix A:

Clinical Diagnostic Criteria (NINCDS)

NINCDS / ADRDA (Alzheimer's Disease)

McKhann G *et al. Neurology* 1984; 34: 939-944

0. Negative

1. Possible

   i) Presence of a dementia syndrome, in absence of other neurological, psychiatric or systemic disorders capable of causingdementia, but with atypical features, such as variations in the onset, presentation or clinical course of the illness.

   ii) Presence of a second systemic disease or brain disorder sufficient to produce dementia, but not considered to be *the* cause of the dementia.

   iii) Single, gradually progressive, severe cognitive deficit (eg. worseningamnestic syndrome), in the absence of another identifiable cause.

2. Probable

   i) Dementia, established by history & clinical examination, and documented with, or confirmed by, cognitive or neuropsychologicaltests e.g. MMSE (<23), CAMCOG (<80).

   ii) Deficits in 2 or more areas of cognition.

   iii) Progressive worsening of memory and other cognitive functions.

   iv) No disturbance of consciousness.

   v) Age of onset > 40; usually > 65.

   vi) Absence of systemic disorders or other brain diseases [or psychiatric disorders] that could in themselves account for the progressive

deficits in memory & cognition.

3. Definite

i) Probable AD on clinical criteria.

ii) Histopathological evidence (biopsy, autopsy).

Supportive evidence for the diagnosis of Probable AD:

i) Progressive deterioration of specific cognitive functions such as language (aphasia), motor skills (apraxia), and perception (agnosia);

ii) impaired activities of daily living and altered patterns of behaviour;

iii) family history of similar disorders, particularly if confirmedneuropathologically; and

iv) laboratory results of:

a) normal lumbar puncture as evaluated by standard techniques;

b) normal pattern or non-specific change in EEG, such as increasedslow wave activity, and

c) evidence of cerebral atrophy on CT with progression documentedby serial observation.

Clinical features consistent with the diagnosis of Probable AD:

i) Plateaus in the course of progression of the illness;

ii) Associated symptoms of depression, insomnia, incontinence, delusions, illusions, hallucinations, catastrophic verbal, emotional orphysical outbursts, sexual disorders and weight loss;

iii) Other neurological abnormalities in some patients, especially those with more advanced disease, including motor signs such as increasedmotor tone, myoclonus or a gait disorder;

iv) Seizures in advanced disease;

v) CT normal for age.

Features that make the diagnosis of Probable AD uncertain or unlikely include:

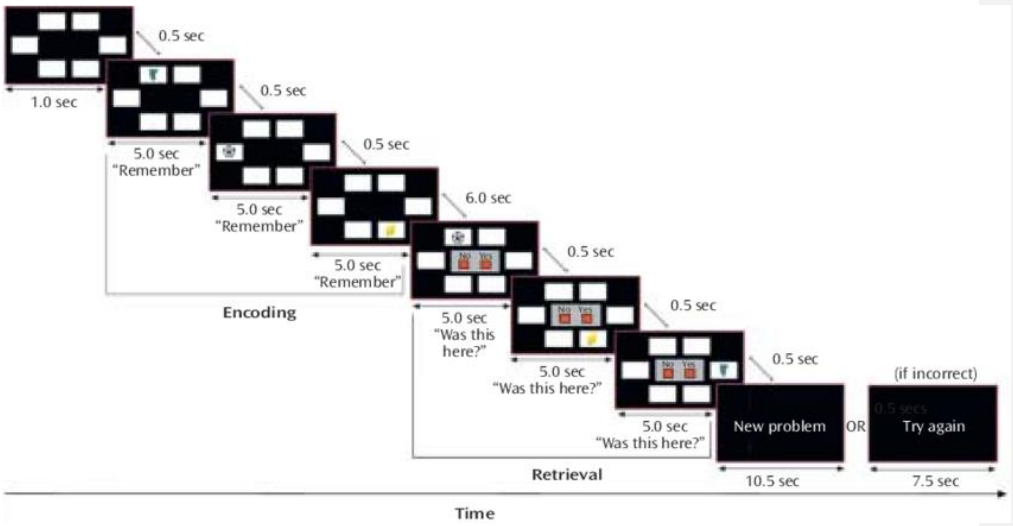
i) Sudden apoplectic onset;

ii) Focal neurological signs such as hemiparesis, sensory loss, visualfield deficits, and incoordination early in the course of the illness;

iv) Seizures or gait disturbances at the onset or very early in the courseof the illness.

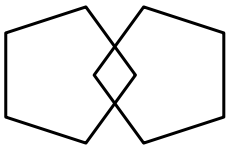
Appendix B:

Paired Associates Learning Task (PAL)



The test begins with a blank screen and white boxes on the periphery of the screens border, each of which are then randomly opened and then subsequently closed. A pattern is contained in one or more of each of the boxes and is displayed when a box containing said pattern is opened. Following all the boxes' 'contents' being revealed, the patterns that were shown to the participant are then exhibited in the middle of the screen sequentially. Participants are directed to match the patterns to the specific box from which they originated on the periphery of the screen. If an error is made, the patterns are redisplayed to remind the participant of their original placements. The test itself is comprised of eight different stages that get progressively more difficult with the first seven stages having six boxes and the eighth having eight boxes. In the first two stages only one pattern is shown, with two being shown in the third and fourth stages, three in the fifth and sixth stages, six in the seventh, and finally eight in the final stage. Each stage consists of 10 trial

Appendix C:
Mini-Mental State Examination (MMSE)

One point for each answer **DATE:**			
ORIENTATION Year Season Month Date Time	……/ 5	……/ 5	...../ 5
Country Town District Hospital Ward/Floor	……/ 5	……/ 5	...../ 5
REGISTRATION Examiner names three objects (e.g. apple, table, penny) and asks the patient to repeat (1 point for each correct. THEN the patient learns the3 names repeating until correct).	……/ 3	……/ 3	...../ 3
ATTENTION AND CALCULATION Subtract 7 from 100, then repeat from result. Continue five times: 100, 93, 86, 79, 72, 65 (Alternative: spell "WORLD" backwards: DLROW).	……/ 5	……/ 5	...../ 5
RECALL Ask for the names of the three objects learned earlier.	……/ 3	……/ 3	...../ 3
LANGUAGE Name two objects (e.g. pen, watch).	……/ 2	……/ 2	...../ 2
Repeat "No ifs, ands, or buts".	……/ 1	……/ 1	...../ 1
Give a three-stage command. Score 1 for each stage. (e.g. "Place index finger of right hand on your nose and then on your left ear").	……/ 3	……/ 3	...../ 3
Ask the patient to read and obey a written command on a piece of paper. The written instruction is: "Close your eyes".	……/ 1	……/ 1	...../ 1
Ask the patient to write a sentence. Score 1 if it is sensible and has a subject and a verb.	……/ 1	……/ 1	...../ 1
COPYING: Ask the patient to copy a pair of intersecting pentagons 	……/ 1	……/ 1	...../ 1
TOTAL:	……/ 30	……/ 30	...../ 30

MMSE scoring

24-30: no cognitive impairment 18-23: mild cognitive impairment 0-17: severe impairment

Appendix D:

Geriatric Depression Scale

NAME: **DATE:**

1	Are you basically satisfied with your life?	**No** Yes
2	Have you dropped many of your activities or interests?	**Yes** No
3	Do you feel that your life is empty?	**Yes** No
4	Do you often feel bored?	**Yes** No
5	Are you in good spirits most of the time?	**No** Yes
6	Are you afraid that something bad is going to happen to you?	**Yes** No
7	Do you feel happy most of the time?	**No** Yes
8	Do you often feel helpless?	**Yes** No
9	Do you prefer to stay at home, rather than going out and doing new things?	**Yes** No
10	Do you feel you have more problems with your memory than most?	**Yes** No
11	Do you think it is wonderful to be alive?	**No** Yes
12	Do you feel pretty worthless the way you are now	**Yes** No
13	Do you feel full of energy?	**No** Yes
14	Do you feel that your situation is hopeless?	**Yes** No
15	Do you think that most people are better off than you are?	**Yes** No
> 5 problems (answers in BOLD) indicates probable depression TOTAL:		

THE GERIATRIC DEPRESSION SCALE (GDS)

1. The GDS short form (15 questions) has been derived from the 30 question version. It has been designed for the assessment of depressive symptomatology in elderly people and excludes any questions relating to the physical symptoms of depression common in old age.

2. The GDS is a screening device and should not be used as a diagnostic tool. It can be used to monitor the client's emotional state in relation to treatment or change in physical health. The questionnaire can guide further clinical interviews and when used this way has been found very acceptable to clients.

3. **The questions are read out** and the patient is asked how they have felt over the past week using a Yes/No response format. No further explanation or interpretation should be given to the questions.

4. Each answer indicating depression (bold 'yes' or 'no) counts one point. Scores greater than 5 are indicative of probable depression.

Appendix E:

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E52-24 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6492 • Facsimile [021] 406 6411
Email: Sumayah.ariefdien@uct.ac.za
Website: www.health.uct.ac.za/research/humanethics/forms

03 June 2014

HREC/REF: 167/2014

Dr K Thomas
Psychology
Room no:2.17
PD Hahn Building
Upper Campus -UCT

Dear Dr Thomas

Project Title: INTRAINDIVIDUAL VARIABILITY IN THE PROGRESSION OF ALZHEIMER'S DISEASE: A LONGITUDINAL TRAJECTORY OF COGNITIVE DECLINE-(Doctorate-Bjorn Christ)

Thank you for your letter dated 02 June 2014, addressing the issues raised by the Human Research Ethics Committee.

It is a pleasure to inform you that the HREC has **formally approved** the above mentioned study.

Approval is granted for one year until the 30 June 2015.

Please submit a progress form, using the standardised Annual Report Form, if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.

We acknowledge that the following student:- Bjorn Christ is also involved in this project.

Please note that the on-going ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC REF in all your correspondence.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, HSF HUMAN ETHICS

Hrec/ref:167/2014

Appendix E:

PARTICIPANT INFORMATION LEAFLET AND CONSENT FORM

TITLE OF THE RESEARCH PROJECT: Intraindividual Variability in the Progression of Alzheimer'sdisease: A longitudinal trajectory of cognitive decline

PROTOCOL NUMBER: 167/2014

PRINCIPAL INVESTIGATOR: Mr. Bjorn U. Christ

New text: SECONDARY INVESTIGATORS: Ms Kara Engelbrecht and Ms MelindaSimon.

ADDRESS: Department of Psychology, PD Hahn Psychology Building, UniversityAvenue, University of Cape Town, Rondebosch, Cape Town, 7701

CONTACT NUMBER: +27 72 0710 346

I am inviting you to participate in a research project. Please take some time to read the information presented here. It explains the details of the project. If there are any aspects of the project you do not understand, please do not hesitate to ask the study staff or doctor. It isimportant that you are fully satisfied that you clearly understand what this research entails and how you could be involved. Your participation in the study is entirely voluntary and youare free to decline to participate. If you say no, this will not affect you negatively in any waywhatsoever. It will not affect any future medical treatment you may need. You are also free to withdraw from the study at any point, even if you did initially agree to take part. You do not have to give a reason for withdrawing.

This study has been approved by the Research Ethics Committee of the Faculty of HealthSciences of the University of Cape Town. It will be conducted according to the ethical guidelines and principles of the International Declaration of Helsinki, the South African Guidelines for Good Clinical Practice and the Medical Research Council (MRC) Ethical Guidelines for Research.

This trial is being run by the Applied Cognitive Sciences and Experimental NeuropsychologyTesting (ACSENT) laboratory and

the Divisions of Neurology and Geriatric Medicine in the Department of Medicine at the University of Cape Town. I aim to recruit a total of 90 participants over a period of 18 months.

What is this research study all about?

Some people develop memory problems as they get older. Many elderly people have mild memory difficulties. However, in a few, the problem may be more severe. I am interested infinding out more about how the difficulties with memory and other higher brain functions change over time. In order to do so, I should like to investigate the course of these changes using a small number of methods. These include questions you would need to answer about yourself and tests of memory and other higher brain functions.

I am interested in testing people both with memory difficulties and those without, so that wecan compare the two groups. In this way I might be able to better understand the progressionof change in brain function associated with memory impairment. My research findings may aid in the early detection and treatment of these conditions and help improve the design of drug intervention trials associated with these conditions in the future.

Procedures

If you agree to take part in the study you will be required to partake in a short telephonic interview about your medical history. This is done to ensure you meet all the conditions required to enter the study. You will then be invited to visit our clinic on three separate daysover a two week period. At these visits to our clinic I shall:

(1) interview your relative/friend (someone who knows you well) to find out whetherhe/she thinks you have any memory difficulties.

(2) ask you to complete a short questionnaire about your mental and emotionalfunctioning.

(3) perform tests of your memory and other higher mental functions. These will be conducted in a quiet, relaxed atmosphere. I expect that these tests will be about two hours duration. However, there will be opportunities to rest in-between tests.

The questionnaires and the tests will be administered during the first visit, however for the subsequent two visits you will only be required to complete the tests. After the three baseline visits I would like to re-assess your memory and other higher functions again after six monthsand twelve months, respectively, provided you continue to consent to participation in the study.

If I find that you or your relative/friend has a significant memory problem that is interfering with your daily living activities, we shall refer you to a Memory Clinic. Your permission willalways be sought first.

What will your responsibilities be?

You will be required to attend the study visit at the appropriate time and to participate as fullyas you can with the tests and questionnaires. You should answer the questions as fully and honestly as you can. If there are any questions that you cannot, or do not wish to answer, you should tell us so.

Will you benefit from taking part in this study?

You will receive little direct benefit from the study. However, you will undergo a range ofcognitive tests. As previously indicated, we shall, with your permission, refer you to the appropriate medical services if any treatable abnormalities are found.

Are there any risks in your taking part in this research?

You may feel uncomfortable about answering some of the questions about yourself or yourfriend/relative. Some people don't like talking, or knowing about, problems related to memory or thinking. You should feel free to mention your feelings or concerns to any member of the study team.

If you do not agree to take part, what alternatives do you have?

You are free not to participate in the study or to refuse parts of the study.

Who will have access to your medical records?

The information collected about you, which includes your Groote Schuur Hospital medical history (e.g. the records of the Geriatric Unit and the Memory Clinic), will be treated as confidential and protected. If it is used in a publication or thesis, your identity will remain anonymous. Only the direct study team will have full access to the information. If we need torefer you to a clinic for treatment, we will provide them with the relevant information needed to treat your condition.

Will you be paid to take part in this study and are there any costs involved?

You will not be paid to take part in the study but your transport costs will be covered for thestudy visit. You will be reimbursed for the sum of R50-00 at each visit to the research site. There will be no costs involved for you, if you do take part.

Is there anything else that you should know or do?

- You should inform your family practitioner or usual doctor that you are taking part ina research study.
- You can contact me on +27 72 0710 346 if you have any further queries or encounterany problems.
- You can contact the Research Ethics Committee of the Health Sciences Faculty of theUniversity of Cape Town 021-4066338 if you have any concerns or complaints that have not been adequately addressed by your study doctor.
- You will receive a copy of this information and consent form for your own records.

Declaration by participant and/or friend/relative/guardian

By signing below, I..... , hereby agree to take part in
the research study entitled: "Intraindividual Variability in the Progression of Alzheimer'sdisease: A longitudinal trajectory of cognitive decline"

I declare that:

- I have read or had read to me this information and consent form and it is writtenin a language with which I am fluent and comfortable.
- I have had a chance to ask questions and all my questions have been adequatelyanswered.
- I understand that taking part in this study is **voluntary** and I have not beenpressurised to take part.
- I may choose to leave the study at any time and will not be penalised orprejudiced in any way.
- I may be asked to leave the study before it has finished, if the study doctor or researcher feels it is in my best interests, or if I do not follow the study plan, as agreed to.

Signed at (place) .....…….. on (date) 2015

....

Signature of participant Signature of witness

....

Signature of relative/friend/guardian Signature of witness Declaration by investigator

I (name)……..declare that:

- **I explained the information in this document to ……..**
- **I encouraged him/her to ask questions and took adequate time to answer them.**

- **I am satisfied that he/she adequately understands all aspects of the research, as discussed above**

Signed at (place) .....….....…….. on (date) 2015

....

Signature of investigator Signature of Witness