Determining clinically relevant neuropsychological change in an epilepsy sample

Marnie Cumner

A thesis
submitted in partial fulfilment of the degree of
Doctor of Psychology (Clinical Neuropsychology)

School of Social Sciences and Psychology
Faculty of Arts, Education and Human Development
Victoria University
Melbourne, Australia

August 2009

ABSTRACT

Accurate interpretation of pre- and post-surgical cognitive assessments is essential for patients undergoing surgery for temporal lobe epilepsy (TLE). However, potentially confounding issues such as practice effects and low test-retest reliability may influence a person's test score when they are given the same neuropsychological test on a second occasion, such as following surgery. Determining how much change in test scores is due to methodological issues and how much is due to genuine post-operative cognitive improvement or decline can better inform both clinical and client decisions regarding surgical intervention for TLE. Such data has been reported in North American TLE populations but the generalisability of this data to other centres is unknown. Using data obtained from unoperated patients with TLE, the current study utilised standardised regression-based change (SRB) and reliable change index (RCI) methodologies to determine change on retest across cognitive domains. Both the local and North American RCIs and SRB change norms were then applied to an Australian post-operative sample to evaluate cognitive outcomes, as well as to investigate the generalisability of change data derived from different demographic backgrounds.

"I, Marnie Cumner, declare that the Doctor of Psychology (Clinical Neuropsychology) thesis entitled *Determining Clinically Relevant Neuropsychological Change in an Epilepsy Sample* is no more than 40,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

Signature: Date:

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr Fiona Bardenhagen, who, despite recruitment challenges, methodological mountains and statistical conundrums, has provided unfailingly positive support and encouragement over the last four years. Fiona's emphasis on best-practice and self-reflection has provided me with a professional framework which I will continue to use throughout my career – thank you.

My sincere thanks also go to the St Vincent's Hospital Neurosciences Department and the Neuropsychology Unit for giving me the opportunity to undertake this research project. Seeing research patients in a busy department at times required flexible and creative thinking, and for this I offer my thanks to both the clinical and administrative staff.

I would also like to offer my sincere thanks to all the participants who were involved in this study. They gave up their free time, travelled great distances and concentrated hard on challenging tasks in the name of epilepsy research. Their selfless and enthusiastic attitudes towards the research gave me motivation when the going got tough and for this I thank each and every one of them.

To my friends, who never complained when my only topics of conversation were thesis-related, an enormous thank you. Despite my increasingly hermit-like habits, you never failed to lend a sympathetic ear or provide opportunities for some most-welcome procrastination. A special thank you to Reegs for understanding when I was too busy to leave the house, to Shazi and Lisa for being my personal cheerleaders, and to Kirst for providing fellow academic-hermit solidarity. A huge thank you also to my "uni" friends, the most supportive and lovely bunch one could ever hope to meet.

Most importantly, my love and thanks go to my family, Michael, Deann and Meg, and to my fiancé, Tom. Without your unconditional love, support, encouragement, patience, optimism, humour, reassurance, understanding and sheer fabulousness, I would not have been able to even entertain my dream of becoming a neuropsychologist. I simply can not express how grateful I am to have you all in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AD	Auditory Delayed
AI	Auditory Immediate
ARD	Auditory Recognition Delayed
ARI	Absolute Risk Increase
ARR	Absolute Risk Reduction
BNT	Boston Naming Test
FSIQ	Full Scale Intelligence Quotient
GM	General Memory
IM	Immediate Memory
IQ	Intelligence Quotient
NNH	Number Needed to Harm
NNT	Number Needed to Treat
PIQ	Performance Intelligence Quotient
POI	Perceptual Organisation Index
PSI	Processing Speed Index
r_{12}	Test-retest interval
RCI	Reliable Change Index
SRB	Standardised Regression-Based
SD	Standard Deviation
SE_e	Standard Error of Estimation
SE_{diff}	Standard Error of the Difference
SE_m	Standard Error of Measurement
SE_{pred}	Standard Error of Prediction
TLE	Temporal Lobe Epilepsy
VCI	Verbal Comprehension Index
VD	Visual Delayed
VI	Visual Immediate
VIQ	Verbal Intelligence Quotient
WAIS-III	Wechsler Adult Intelligence Scale – Third Edition
WMI	Working Memory Index
WMS-III	Wechsler Memory Scale – Third Edition

CHAPTER 1

Determining Clinically Relevant Cognitive Change in an Epilepsy Sample

For people undergoing surgery for temporal lobe epilepsy (TLE), neuropsychological assessment of memory, language and other cognitive functioning before surgery is essential, as is accurate interpretation of change after surgery. However, cognitive test scores can change post-operatively in this population for reasons other than surgical intervention. Confounding issues, such as practice effects and test-retest stability, may influence a person's score when they are given the same neuropsychological test on a second occasion (Chelune, Naugle, Luders, Sedlak, & Awad, 1993; Hermann, Seidenberg, Schoenfeld, Peterson, Leveroni, et al., 1996). It is therefore important to know how much change in test scores is due to methodological issues and how much is due to genuine improvement or decline in a patient's cognitive abilities. While data from the United States of America is available to help determine change in an epilepsy population (Chelune et al, 1993; Hermann et al., 1996; Martin, Sawrie, Gilliam, Mackey, Faught, et al., 2002b; Sawrie, Chelune, Naugle, & Luders, 1996), it has been suggested that researchers should gather and use local data (Martin et al., 2002b; Sawrie et al., 1996). This investigation used test-retest data from a pre-surgical TLE sample in Australia to derive change norms for this population and to determine if local data enhances the accuracy of these methods.

The current chapter reviews different change methodologies and discusses their advantages and disadvantages, while Chapter Two considers cognition in TLE and reviews the literature to identify relevant variables for predicting cognitive change on retest. Chapter Three reports on the generation of Reliable Change Indices (RCI) and Standardised Regression-Based (SRB) change norms based on the test-retest data of a sample of Australian adults with unoperated TLE. The fourth chapter describes the application of local RCI and SRB change methodologies to a sample of Australian adults following surgical intervention for TLE. Finally, in Chapter Five, the post-operative cognitive outcomes are examined in the context of theoretical frameworks regarding the relationship between the temporal lobes and cognition, particularly memory.

Basic Psychometrics– Determining Cognitive Change

Serial cognitive assessments require an appropriate method for determining genuine change, whilst also controlling for methodological confounds, such as test-retest stability, practice effects and regression to the mean (Chelune, 2003). Given the potential clinical implications of interpreting measurement error rather than actual change, much effort has been devoted to deriving appropriate statistical approaches to control for these confounds.

The accurate interpretation of neuropsychological test results in clinical practice relies on sound psychometric properties of the neuropsychological measures themselves, including acceptable levels of reliability and validity, as well as appropriate normative data. However, tests with perfect psychometric properties do not exist and normative data is not always available for all clinical populations. Nonetheless, the clinician has a responsibility to strive to improve clinical practice and patient outcomes by utilising the best possible statistical methodologies.

The following section discusses the potential statistical and methodological pitfalls of working with psychometric data, such as test-retest stability, the effects of practice and regression to the mean. Evaluating what constitutes clinically relevant cognitive change has been greatly advanced by the introduction of RCI and SRB equations, both of which will also be discussed in this section. The term "clinically relevant change" (Martin et al., 2002b, p. 1554) refers to statistically significant change scores at the individual level – using RCI and SRB methods – and is the preferred term over "clinically meaningful change", which implies validation against ecological markers, such as occupational performance.

Jacobson, Follette and Revenstorf (1984) conceptualised clinical significance as the return to normal functioning. The clinical significance of a change in a person's cognitive test score over time should reflect the standards of efficacy determined by patients, clinicians and researchers (Jacobson & Truax, 1991). In contrast, statistical significance refers to real differences between data; differences which are not simply chance findings (Jacobson & Truax, 1991). A *statistically* significant test finding

should be corroborated by non-test information about that person's everyday functioning, in order for that test finding to be also considered *clinically* significant (Matarazzo & Hermann, 1984).

Test-Retest Reliability

Test-retest reliability– also referred to as test stability – refers to the consistency of a particular neuropsychological test score over time (Barker, Pistrang, & Elliot, 2002). The scores obtained on two separate testing occasions are correlated to calculate test-retest reliability, or the stability coefficient (r_{12}). The degree to which the two scores correlate indicates the measurement consistency of a particular test. For example, a high correlation suggests the results are less likely to be caused by error and that differences in a person's test scores are more likely due to actual change in the domain being measured (Groth-Marnat, 2003).

Neuropsychological tests with perfect test stability (e.g., $r_{12} = 1.00$) would simply require the subtraction of the initial test score from the retest score to calculate change. The higher the stability of a neuropsychological measure, the less measurement error associated with administering that test a second time. However, neuropsychological tests are neither perfectly reliable nor perfectly stable, and so consideration must be given to the effects of bias and error. Bias refers to a systematic change in a person's performance on a test, the most common of which is practice effects (Chelune, 2003).

The Effects of Practice

Practice effects refer to an improvement in test scores as a result of taking the test previously and do not reflect improvement in the skills being assessed (Kaufman, 2003). The magnitude of a practice effect may be influenced by factors such as a person's baseline test performance and age, the duration of the test-retest interval and the type of cognitive measure (Dikmen, Machamer, Temkin, & McLean, 1990; Temkin, Heaton, Grant, & Dikmen, 1999). To determine changes in a person's cognition over time, it is important to have an understanding of the expected

magnitude of practice effects on different neuropsychological measures for that specific sample (Mitrushina & Satz, 1991).

Retest gains are evident on most neuropsychological measures and occur in both clinical and non-clinical samples (Chelune et al., 1993; Dikmen, Heaton, Grant, & Temkin, 1999; Hermann et al., 1996; Ivnik, Smith, Lucas, Petersen, Boeve, et al., 1999; Kneebone, Andrew, Baker, & Knight, 1998; Matarazzo, Carmody, & Jacobs, 1980; McCaffrey, Ortega, Orsillo, Nelles, & Haase, 1992; Rapport, Brines, Axelrod & Theisen, 1997; Sawrie et al., 1996). In an early review of the literature, the subtest scores of healthy adults on the Wechsler Adult Intelligence Scale (WAIS) were found to increase up to seven points on retesting, over retest intervals of up to 20 weeks (Matarazzo et al., 1980). Motivational differences, practice effects, test stability and random error were among a variety of factors thought to be potentially contributing to these retest gains. In addition, the length of the retest interval also had a clear role to play in the magnitude of the observed practice effects. However, the effects of practice decrease slowly and gains may remain evident up to one to two years after initial testing, particularly on visually-based or speeded tasks (Chelune, 2003; Dodrill & Troupin, 1975). Consequently, based on their own data and that of others preceding them (e.g., Dodrill & Troupin, 1975), Matarazzo et al. (1980) suggested the Full Scale IQ of the WAIS be routinely corrected by five points on retest, solely to account for the effects of practice. However, adjustments based on group means do not necessarily reflect individual change and valuable information regarding inter-individual variability may be lost (Chelune, 1993; Jacobson & Truax, 1991).

Subsequent examination of the WAIS-R standardisation data found that statistically significant differences between the Verbal IQ and Performance IQ were so common in the normal population that base-rate data should also be generated for different patient samples and routinely employed in clinical practice (Matarazzo & Hermann, 1984). Base-rates are increasingly utilised in clinical and research settings as an appropriate means of conveying test-retest change (Matarazzo & Herman, 1984). As an example, a retest increase of seven points on the Verbal IQ measure may reach statistical significance, but if 30% of the normal population shows such an increase, it becomes clinically redundant. In contrast, if a 7-point difference is evident in only

4% of the normal population, then this would be considered clinically unusual and therefore, clinically meaningful. Using the WAIS-R and WMS-R, Sawrie and colleagues (1996) twice assessed patients with TLE (mean age = 31.5 years) over a period of approximately eight months and found practice effects evident for 15 of the 32 measures administered. Most significantly, the average improvement on the PIQ of the WAIS-R was 5 IQ points, leading the authors to suggest that patients who do not show improvement on retest may, in fact, have deteriorated slightly (Sawrie et al., 1996).

Overall, the literature suggests adults continue to benefit from practice effects until their seventh decade, beyond which their scores on retest may be expected to decline (Ivnik et al., 1999; Mitrushina & Satz, 1991). This pattern of findings is consistent with an increasing loss of grey matter, particularly in the prefrontal cortex, which constitutes part of the normal adult cognitive ageing process (Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003).

Measurement Error and Regression to the Mean

Consideration must be given to the measurement error associated with the tests themselves. The *standard error of measurement* (SE_m) refers to the band of error which surrounds the observed score and is related to the error of the test (Chelune, 2003). Different neuropsychological measures have different stability coefficients and, subsequently, different standard errors of measurement. The measurement error of a test plays an important role in determining the clinical significance of change scores; low test stability may lead to a large test-retest difference simply due to random fluctuations in the measurement, while a small test-retest difference on a test with high stability may be clinically significant (Chelune, 2003).

Regression to the mean is a statistical phenomenon that describes the propensity for a person's test score to move toward the group mean upon retesting (Martin et al., 2002b). It is caused by a combination of the person's baseline test score and the stability of the measure (Barnett, van der Pols, & Dobson, 2005; Chelune et al., 1993; Hermann et al., 1991; Speer, 1992). For example, baseline scores which are

especially low or high are more likely to regress towards the group mean on retest (Speer, 1992). This is because a person's true score is frequently closer to the mean of the population than their observed score (Chelune, 2003). The estimated true score is based on obtained results but regressed towards the mean (Atkinson, 1991). The *standard error of estimation* (SE_e) is associated with estimated true scores and represents an unbiased estimate of population measurement error based on the population mean (Lord & Novick, 1968). A further type of standard error of measurement is the *standard error of prediction* (SE_p; Lord & Novick, 1968), which is used to determine whether variation in test scores at retest is due to measurement error or another intervening variable (Atkinson, 1991).

Methods for Calculating Cognitive Change over Time

There is an abundance of literature examining the most accurate way to determine test-retest score change; however, a definitive answer remains difficult due to the wide variability in samples studied, the measures used, and the change methodologies employed (Bruggemans, Van de Vijver, & Huysmans, 1997; Chelune et al., 1993; Chelune, 2003; Dudek, 1979; Edwards, Yarvis, Mueller, Zingale, & Wagman, 1978; Hagemann & Arrindell, 1993; Hsu, 1989; Iverson, 2001; Lewis, Maruff, Silbert, Evered & Scott, 2007; Lord & Novick, 1968; Jacobson & Truax, 1991; Martin et al., 2002b; McSweeny, Naugle, Chelune, & Luders, 1993; Rasmussen, Larsen, Houx, Skovgaard, Hanning, et al., 2001; Sawrie et al., 1996). The advantages and disadvantages of the predicted true score method (PTS) method, reliable change indices and standardised regression-based change norms will be examined in detail below.

Predicted True Score Methodology

At its simplest, change may be calculated using the predicted true score method, which regresses a person's obtained score towards the mean (Dudek, 1979; Lord and Novick, 1968). PTS methodology is based on classical test theory, where a person's obtained score (X) is the sum of both the true score (T) and measurement error (E) (Cohen & Swerdick, 2005, p. 129):

$$X = T + E$$

The true score can be conceptualised as the average of observed scores over an infinite number of repeat test administrations using the same test (Wilmes, 2003). Confidence intervals derived from the SE_m and the SE_p, discussed above, can then be placed around the person's PTS to allow for regression to the mean (Atkinson, 1991; Dudek, 1979). A limitation of the PTS approach is its failure to consider either the effects of practice or a person's baseline performance. As a result, statistical approaches which also incorporate these sources of measurement error have been developed.

Reliable Change Index Scores

Statistical analyses, such as *t*-tests and repeated measures, can determine whether a difference between two scores is significant, but, as discussed above, such differences may be common in the normal population (Chelune, 2003; Matarazzo & Hermann, 1984). Based on early work aimed at defining clinically significant change in psychotherapy (Jacobson et al., 1984), Jacobson and Truax (1991) developed the *Reliable Change Index* (RCI).

Jacobson and Truax's (1991) RCI approach allowed clinicians to determine the extent to which change after intervention is clinically meaningful; it provides an estimate of the probability that a particular difference score would not be obtained by chance alone (Iverson, 2001). The difference between a person's test and retest score is evaluated against the distribution of change scores from a comparable group of individuals (Chelune, 2003). The method uses summary statistics (mean, standard deviation, stability coefficient) to estimate the distribution of test-retest change scores likely to be obtained, given no intervening variable. Confidence intervals are placed around the mean change score and represent the limits of reliable change, beyond which a change score may be considered clinically meaningful (Chelune, 2003). RCIs can be used to measure change over time on any continuous scale for any clinical issue (Parabiaghi, Barbato, D'Avanzo, Erlicher, & Lora, 2005). The increased importance of serial neuropsychological assessments as part of routine

clinical practice – including epilepsy surgery (Chelune et al., 1993; Engman, Andersson-Roswell, Saumelsson, & Malmgren, 2006; Hermann et al., 1996; Lo Galbo, Sawrie, Roth, Kuzniecky, Knowlton, et al., 2005; Martin et al., 1998; Martin et al., 2002b; Sawrie et al., 1996; Sherman, Slick, Connolly, Steinbok, Martin, et al., 2003), cardiac bypass graft surgery (Bruggemans et al., 1997; Raymond, Hinton-Bayre, Radel, Ray, & Marsh, 2006), diagnosis of neurodegenerative conditions (Troster, Woods, & Morgan, 2007), recovery from traumatic brain injury (Collie, Maruff, Makdissi, McStephen, Darby et al., 2004), as well as in mental health (Evans, Margison, & Barkham, 1998; Medalia & Richardson, 2005; Parabiaghi et al., 2005) and older adult (Frerichs & Tuokko, 2006), and medical populations (Woods, Childers, Ellis, Guaman, Grant et al., 2006) – has also seen increased use of RCI methodology in research settings (e.g., Bauer, Lambert & Nielsen, 2004; Chelune & Goldstein, 1991; Heaton, Temkin, Dikmen, Avitable, Taylor et al., 2001; Hermann et al., 1996; Kay & Kane, 1991; Levine, Miller, Becker, Selnes & Cohen, 2004; Maassen, 2000; Martin et al., 2002b; Speer, 1992; Wise, 2004).

The original RCI method (Jacobson and Truax, 1991) involved dividing the difference between the pre- and post-treatment scores by the standard error of the difference between the two test scores. The SE_{diff} is calculated using the test's standard error of measurement (SE_m; Dudek, 1979), as follows:

$$SE_m = SD * (1\text{-}r)^{1/2}$$

$$SE_{diff} = (2 * (SE_m)^2)^{½}$$

The SE_{diff} represents the distribution of change scores that would be expected when no change has occurred (Jacobson and Truax, 1991). However, Jacobson and Truax's (1991) model was designed for evaluating psychotherapy outcomes and did not consider the effects of practice and systematic bias. As a result, the original RCI model has undergone numerous refinements, largely based on different measures of dispersion (Bruggemans et al., 1996; Chelune et al., 1993; Chelune, 2003; Edwards et al., 1978; Hageman & Arrindell, 1993; Hsu, 1989; Iverson, 2001; Lewis et al., 2007; Rasmussen et al., 2001; Sawrie et al., 1996). Much of the subsequent literature on

RCIs has emerged from the epilepsy surgery literature, where patients routinely undergo pre- and post-surgical neuropsychological assessment (Chelune et al., 1993; Sawrie et al., 1996).

Chelune and colleagues (1993) generated the S_{diff} from a control, or non-surgical, group of epilepsy patients. To account for systematic bias, Chelune et al (1993) centred the expected differences around the group's average practice score, rather than around a mean of zero, as originally suggested by Jacobson and Truax (1991). The S_{diff} between the baseline and the follow-up test scores determined whether an observed change in a person's test-retest scores was significant (Chelune, 2003). The test-retest change score was divided by the SE_{diff} and a constant representing the effects of practice was then added. This correction factor was calculated by simply subtracting the retest mean from the test mean and rounding the answer up to the nearest whole number. After correcting for practice, the SE_{diff} was multiplied by ± 1.64 – generating the 90% confidence interval – to determine whether patients had significantly improved, declined or remained unchanged on retest. Scores falling outside the 90% confidence interval were considered uncommon in an unoperated epilepsy population. Similarly, scores lying within the 90% confidence intervals were interpreted as a relatively common occurrence and not representative of clinically relevant change (Hermann et al., 1996; Martin et al., 2002b; Sawrie et al., 1996).

Chelune and colleagues' (1993) practice-adjusted RCI method assumed practice was constant at test and retest and therefore multiplied the standard error of measurement by two. However, the standard deviation of scores is likely to be different from test to retest due to practice effects. Iverson (2001) therefore suggested the calculation of separate standard errors of measurement for test and retest using the respective standard deviations and stability coefficients. The modified practice-adjusted RCI method (Iverson, 2001) became:

$$S_{diff} = [(SE_{m1})^2 + [(SE_{m2})^2]^{1/2}$$

However, the above RCI method still does not control for regression to the mean (Bruggemans, et al., 1997). In response to this limitation, Chelune (2003) proposed a

further refinement, using the standard error of prediction (Lord & Novick, 1968), also known as the standard error of estimate (Garrett, 1958). As discussed above, the SE_{pred} represents the standard error of a retest score which has been predicted from a baseline score using a regression equation (Dudek, 1979; Chelune, 2003);

$$SE_{pred} = SD_y * (1-r_{12}^2)^{1/2}$$

where SE_{pred} equals the standard error of retest scores predicted from baseline scores, r_{12} represents the stability coefficient and SD_y refers to the standard deviation of the observed retest scores. Using baseline scores, the SE_{pred} estimates how well the predicted retest scores correspond to the actual retest scores (Chelune, 2003).
Using the test-retest data from the WMS-III technical manual (Wechsler, 1997) to illustrate the different results obtained using the different RCI methods, consider a 65-year-old adult who initially scores 100 (mean = 98.0; SD = 16.9) on the WMS-III Immediate Memory Index ($r_{12} = 0.84$) and 117 on retest (mean = 110.2; SD = 19.8). As outlined above, the standard error of measurement for time one is calculated by subtracting the stability coefficient from one (1 – 0.84 = 0.16), then taking the square root of this number ($\sqrt{0.16} = 0.4$) and multiplying it by the standard deviation of the test (0.4 * 16.9 = 6.76). The standard error of measurement for time two is then, 0.4 * 19.8 = 7.92.

Table 1
Comparison of RCI formulas using test-retest data from the WMS-III technical manual

	Formulas	**Example** $r_{12} = 0.84$	**90% Confidence Intervals** (S_{diff} * ±1.64)	**Change** (retest score = 117)
Original RCI (Jacobson & Truax, 1991)	$S_{diff} = (2 * (SE_m)^2)^{1/2}$ $S_{diff} = (2* (6.76)^2)^{1/2}$	$S_{diff} = 9.56$	± 15.7 (84.3-115.7)	Yes
Practice-adjusted RCI (Chelune et al., 1993)	$S_{diff} = [(SE_{m1})^2 + [(SE_{m2})^2]^{1/2}$ $S_{diff} = [(6.76)^2 + [(7.92)^2]^{1/2}$	$S_{diff} = 10.41$	± 17.1 (82.9-117.1)	No
RCI using the SE_{pred} (Iverson, 2001)	$SE_{pred} = SD_y * (1-r_{12}^2)^{1/2}$ $SE_{pred} = 19.8 * (1-0.84^2)^{1/2}$	$SE_{pred} = 10.66$	± 17.5 (82.5-117.5)	No

After the S_{diff} and SE_{pred} are obtained, they are multiplied by the desired band of error (Table 1 used ± 1.64 to represent the 90% confidence intervals), which are then placed around the observed baseline score to delineate the boundaries of reliable change. Each RCI method produces different, but comparable, confidence intervals for reliable change; however, as is demonstrated in Table 1, these small numerical differences may nonetheless result in different clinical conclusions.

The capacity of the RCI model to determine whether a person's retest scores represent meaningful clinical change is an important tool for outcomes research. However, RCI methods have been criticised for considering only a person's baseline performance and for assuming practice is a constant source of bias (Bruggemans et al., 1997; Chelune, 2003). Employing a correction factor assumes every person will benefit to the same extent from practice, whereas evidence suggests the magnitude of practice effects may vary according to factors such as baseline performance, the length of the test-retest interval and demographic variables (McSweeny et al., 1993; Rapport et al., 1997; Sawrie et al, 1996). Finally, although RCIs calculate whether the difference between test and retest is significantly different, they do not indicate how unusual this difference is in the normative sample (Atkinson, 1991).

Standardized Regression-Based Change Scores

Based on the foundations of RCI methodology is a more recent statistical approach to determining change, SRB change norms, which subsume the advantages and address the limitations of their predecessor (McSweeny et al., 1993). Like RCIs, SRB change scores are based on measures of distribution (SD, S_{diff}, SE_{pred}) and account for measurement error and regression to the mean. However, they also consider the differential impact of practice effects and demographic variables (e.g., age, education, length of retest interval) on test performance (Chelune, 2003; McSweeny, Naugle, Chelune, & Luders, 1993). A further advantage of SRB change methodology is the capacity to determine not only the direction of change, but also the magnitude of that change, with either individual or group data (Martin et al., 2002b). Finally, SRB change scores can be converted to a common metric (e.g., *z*-scores, *t*-scores) which allows for the comparison of scores across different neuropsychological measures.

McSweeny and colleagues (1993) illustrated a standardised regression-based approach for determining change using WAIS-R and WMS-R test and retest data from both surgical and non-surgical epilepsy samples. Based on the non-surgical group's scores, McSweeny and colleagues (1993) used a linear regression analysis to generate a formula for predicting retest scores from baseline scores:

$$Y_p = \beta * X_o + C,$$

where Y_p = the predicted retest score; β = the unstandardised beta weight of the regression equation; X_o = the observed baseline score; and C = the constant of the regression formula. The predicted retest scores were then converted to standardised *t*-scores – with a mean of 50 and a standard deviation of 10 – using the prediction equations from the regression analyses.

Using SRB methodology, confidence intervals can be placed around a predicted retest score to define the boundaries of change using standard scores (e.g., *t*- or *z*-scores). The magnitude of change between a person's observed and predicted retest scores can be expressed as standard deviations (e.g., for 95% confidence intervals, *z*-score = $SE_{est} * \pm 1.96$). For example, a *z*-score of greater than zero indicates a person's observed retest score is above the predicted level. In contrast, a negative *z*-score represents a retest score that is below expectation. For determining individual-level change using *z*-scores, a person's predicted retest score (Y_p) is subtracted from their observed retest score (Y_o) and the result is then divided by the SE_{est} (e.g., $z = (Y_o - Y_p)/ SE_{est}$). Alternatively, to express change using *t*-scores (with a mean of 50 and standard deviation of 10), the formula is, $t = 50 + [10 * (Y_o - Y_p)/ SE_{est}]$ (McSweeny et al., 1993).

The SRB approach also allows for the consideration of multiple predictors of retest scores by using a stepwise linear regression model (Chelune, 2003; Temkin et al., 1999). As a result, the regression analyses not only consider the effect of a person's baseline test performance, but also consider the contribution of other factors such as retest interval and demographic variables to the prediction equation. Chapter Two

reviews the literature on cognition in TLE and identifies appropriate variables for inclusion in these prediction equations.

CHAPTER 2

Cognition in Temporal Lobe Epilepsy

Interpreting cognitive changes following surgery for TLE is important both at the individual and the group level. An understanding of the incidence with which adverse post-operative cognitive outcomes occurs assists in pre-surgical decision-making; a delicate weighing of the risks associated with ongoing poorly controlled seizures versus the physical and cognitive risks inherent in neurosurgery. At the individual level, a person considering surgical intervention is making a life-changing decision and requires answers about the risks and advantages particular to their own unique circumstances. Moreover, where adverse cognitive outcomes do occur, the patient deserves the best possible understanding of what has changed, by how much and what can be done to help. This information, in turn, allows clinicians to offer individually-tailored recommendations to help compensate for lost function.

To aid in clinical and personal decision-making regarding surgical intervention for TLE, it is important to have an understanding of the aetiology of epilepsy, as well as those factors which most influence post-surgical cognitive outcomes. The following section explores the physiological mechanisms underpinning seizures, as well as the demographic and seizure variables which are thought to contribute to cognitive status in people with TLE.

The Aetiology and Symptomatology of TLE

Epilepsy refers to episodes of altered behaviour or consciousness due to seizures caused by electrical discharges in the brain (Cull & Goldstein, 1997). An estimated 50 million people worldwide are affected by epilepsy (World Health Organisation, 2001) and approximately 1 in 200 Australians have epilepsy (Howard Florey Institute, 2007). Temporal lobe epilepsy is the most common type of focal epilepsy, associated

with early onset in life, resistance to medication and significant impact on cognitive, social and psychiatric functioning (Hermann, Seidenberg, & Jones, 2008).

A person's seizure profile – type, severity, frequency and age at seizure onset – is determined by the underlying epileptogenic process or lesion (Helmstaedter & Kockelmann, 2006; Jokeit & Schacher, 2004). The pathology determines the seizure-onset region, while a person's seizure profile mediates the type, degree and course of cognitive impairment in epilepsy (Elger, Helmstaedter, & Kurthen, 2004). TLE is not a static condition and may continue to evolve throughout the lifespan as a function of genes, developmental mechanisms and neuronal plasticity (Scharfman, 2007). The aetiology of TLE is varied and ranges from clear precipitating events, such as head trauma and intracerebral infections, to underlying structural abnormalities, such as hippocampal sclerosis (Fish, 1998; Van Paesschen & Revesz, 1998).

The hippocampus is folded into the mesial temporal lobe (see Figure 1) and, together with the adjacent rhinal cortex, forms part of the limbic system, which is essential to declarative memory (Zillmer & Spiers, 2001). Hippocampal sclerosis, also called Ammon's horn sclerosis or mesial temporal sclerosis, is the most frequently observed physiological abnormality in TLE (Fish, 1998; Vinters, Armstrong, Babb, Daumas,Duport, Robitaille, et al., 1993). It is characterised by neuronal cell loss and gliosis and may be caused by congenital mechanisms which prevent the normal maturation, migration or densities of hippocampal neurons (Luders & Comair, 2001). The association between sclerotic atrophy of medial temporal structures and seizures was first described in 1825 by Bouchet and Cazauvieilh (as cited in Van Paesschen & Revesz, 1998, pg. 501). Histopathological studies have suggested that up to 70% of individuals with TLE have hippocampal sclerosis (Vinters et al., 1993).

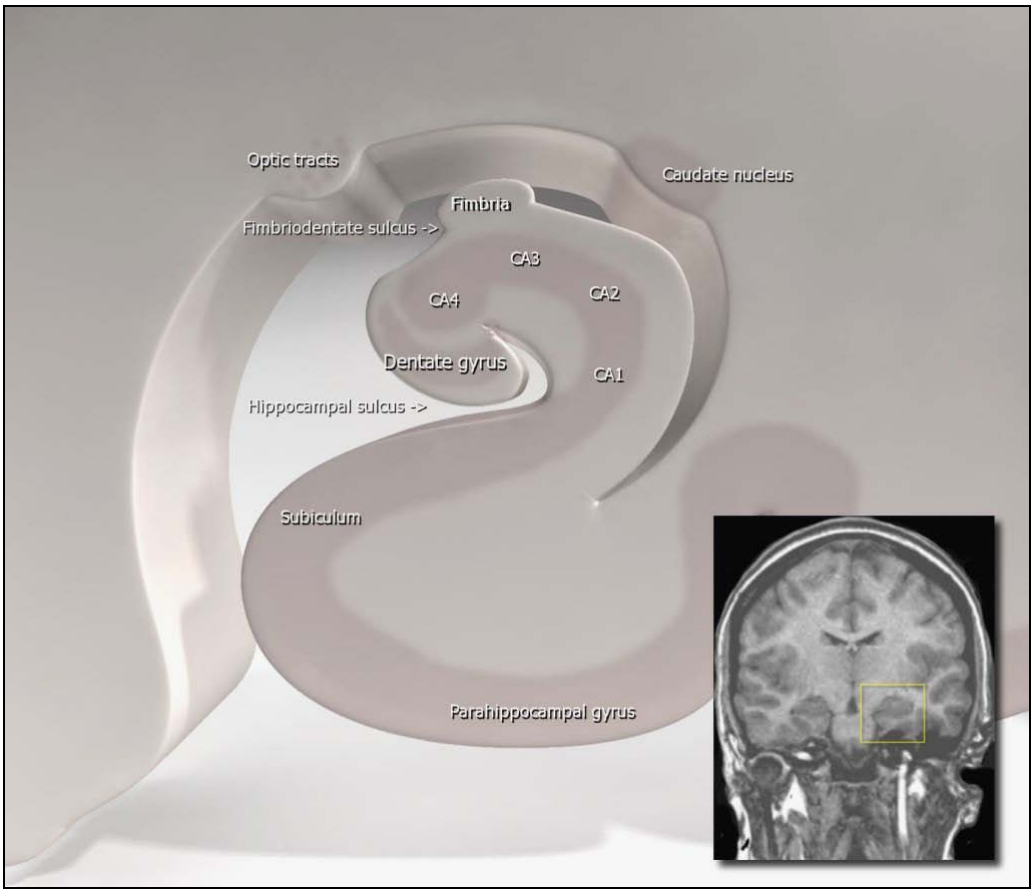

Figure 1. Hippocampal structures
http://commons.wikimedia.org/wiki/File:Hippocampus_(brain).jpg

The development of neuroimaging techniques has allowed an unprecedented *in vivo* window into the location and nature of underlying structural brain abnormalities associated with TLE. Volumetric magnetic resonance imaging (MRI) is a reliable test of the presence of unilateral mesial temporal sclerosis (Berkovic, Andermann, Olivier, Ethier, Melanson, Robitaille, et al., 1991; Cook, Fish, Shorvon, Straughan, & Stevens, 1992) and provides a means by which to characterise the nature of brain abnormalities in people with TLE (Bonilha, Rorden, Castellano, Cendes, & Li, 2005; Engel & Shewmon, 1998). A different imaging technique, positron emission tomography (PET), has identified hypometabolism in the lateral and mesial temporal lobes in adults with TLE, and has associated these metabolic changes with disturbances in cognition and behaviour (Henry, Mazziotta, & Engel, 1993; Rauch, Dougherty, Cosgrove, Cassem, Alphert, et al., 2001).

Quantitative MRI studies with TLE patients have found atrophy of both the hippocampus and adjacent structures (e.g., Bernasconi, Bernasconi, Caramanos,

Antel, Andermann, et al., 2003; Fuerst, Shah, Shah, & Watson, 2003; O'Brien, Bowden, Bardenhagen, & Cook, 2003). Further to this, people with intractable unilateral TLE and hippocampal sclerosis may experience progressive volume loss in the hippocampus on the side of seizure onset (Fuerst et al., 2003). Both hypometabolism and volumetric abnormalities have been associated with cognitive impairment in people with TLE (Baxendale, van Paesschen, Thompson, Connelly, Duncan, et al., 1998; Griffith, Pzyalski, Seidenberg, & Hermann, 2004; Hermann, Seidenberg, Bell, Rutecki, Sheth, et al., 2002; Hermann et al., 2006; Marques, Caboclo, da Silva, Noffs, Carrete, et al., 2007; Oyegbile, Dow, Jones, Bell, Rutecki, et al., 2004).

Factors Affecting Cognition in TLE

It seems undisputed in the literature that the cognitive trajectory for people with TLE differs from the general population (Aikia, Salmenpera, Partanen, Kalvianen, Akanuma, et al., 2001; Aldenkamp, Baker, & Meador, 2004; Andersson-Roswall et al., 2004; Glosser, Cole, French, Saykin, & Sperling, 1997; Helmstaedter, Kurthen, Lux, Reuber, & Elger, 2003; Hermann, Seidenberg, Dow, Jones, Rutecki, et al., 2006; Jokeit & Ebner, 1999; Loring, Barr, Hamberger, Helmstaedter, 2008; Marques et al. 2007; Martin et al., 2002; Oyegbile et al., 2004; Piazzini et al., 2006; Seidenberg, Pulsipher, & Hermann, 2007; Thompson & Duncan, 2005; Vingerhoets, 2006). However, as discussed above, the cognitive profile of a person with TLE is heterogeneous and may be influenced by any combination of a number of factors, including aetiology, gender, age of onset, seizure type and severity, epileptogenic region and anti-epilepsy medications (e.g., Aldenkamp & Arends, 2004; Dodrill, 2004; Elger et al., 2004; Helmstaedter & Kurthen, 2001; Jokeit & Ebner, 1999; Trenerry, Jack, Cascino, Sharbrough, & Ivnik, 1995). Adverse cognitive outcomes in TLE have been associated with both lower baseline performance on cognitive tests (Hermann et al., 2006; Oyegbile et al., 2004), longer duration of epilepsy (Helmstaedter, Kurthen, Lux, Johansen, Quiske, et al., 2000; Marques et al., 2007; Oyegbile et al., 2004; Piazzini et al., 2006) and age at onset (Hermann et al., 2006; Piazzini et al., 2006).

The most commonly reported cognitive deficit in TLE is memory impairment – due to the effects of the primary epileptogenic lesion in the temporal lobe (Chelune, 1995; Hermann et al., 2008) – with declarative memory difficulties reported in approximately 70% of adults with TLE (Hermann, Seidenberg, Schoenfeld, & Davies, 1997). Unilateral damage to the hippocampus is traditionally thought to result in material-specific memory difficulties. This means that the language dominant and non-dominant temporal lobes hold differential responsibility for learning and memory of verbal and nonverbal material, respectively (Milner, 1972). Under this theory, the language-dominant left temporal lobe receives, encodes, stores and retrieves verbal information, whilst the non-dominant right temporal lobe is involved in non-verbal memory processing. This theory of material-specificity has found much support in the epilepsy literature; verbal memory difficulties have frequently been associated with lateralisation of the epileptic focus to the language dominant left hemisphere (e.g., Aikia et al., 2001; Glosser, Deutsch, Cole, Corwin, & Saykin, 1997; Golby, Poldrack, Brewer, Spencer, Desmond, et al., 2001; Helmstaedter, Grunwald, Lehnertz, Gleissner, & Elger, 1997; Hermann et al., 1997). However, there is less consistent evidence for a relationship between visual memory impairment and seizures arising in the non-dominant – usually right – hemisphere (e.g., Alessio, Damasceno, Camargo, Kobayashi, Guerreiro, et al., 2004; Gleissner, Helmstaedter & Elger, 1998; Helmstaedter, Pohl, & Elger, 1995; Hermann et al., 1997; Lee, Yip, & Jones-Gotman, 2002; Majdan, Sziklas & Jones-Gotman, 1996; Milner, 2003; Wieser, Engel, Williamson, Babb, & Gloor, 1993). It has been speculated that a possible explanation for this may be due to difficulties finding a “pure” test of visual memory functioning; that is, people may use verbal strategies to complete non-verbal tasks (Baxendale, 2008; Kneebone, Miller, Bowden, Lah, & Lee, 2008).

In addition to memory difficulties, adults with TLE often have cognitive deficits which are more extensive than might be expected from a focal epileptogenic lesion (Hermann, Seidenberg, Lee, Chan, & Rutecki, 2007). A generalised pattern of cognitive dysfunction has been observed in people with TLE, who demonstrate poorer performance compared to controls across all cognitive domains, including language, speed of processing, executive function, attention and memory (Hermann et al., 1997; Jokeit & Ebner, 2002; Martin, Griffith, Faught, Gilliam, Mackey, et al., 2005;

Oyegbile et al., 2004). Generalised cognitive difficulties are thought to be due to the multifactorial aetiology of TLE and may be underpinned by widespread structural, EEG, and metabolic brain changes, due to the effect of longstanding seizure activity and medication use (Aldenkamp et al., 2004; Elger et al., 2004; Fuerst et al., 2001; Helmstaedter & Kockelmann, 2006; Hermann, Seidenberg, Bell, Rutecki, Sheth, et al., 2002; Hermann et al., 2008; Worrell, Sencakova, Jack, Flemming, Fulgham, et al., 2002). In addition, the mechanisms underpinning cognitive change in TLE may also include the effects of normal, age-related cognitive decline (Helmstaedter et al., 2003; Jokeit & Ebner, 2002), neuro-developmental insult at critical periods (Glosser et al., 1997; Hermann et al., 2006) and cognitive reserve capacity (Helmstaedter et al., 2003). However, whilst epileptic activity may cause generalised cognitive changes, these can be halted, and even reversed, when seizures are controlled (Helmstaedter et al., 2003).

Age

Cognitive change in people with TLE has been increasingly viewed in the context of a lifespan model (Dennis, 2000; Hermann et al., 2002; Hermann et al., 2006; Helmstaedter et al., 2003). People whose seizures first occur in childhood are more likely to have experienced neurological compromise during critical stages of development, leading to widespread cognitive dysfunction (Dennis, 2000; Glosser et al., 1997; Loring et al., 2008; Vingerhoets, 2006). However, earlier onset of seizures may also provide the opportunity for redistribution of function (Loring et al., 2008). Evidence for this cerebral functional plasticity can be found in the 30% of patients with TLE who have some pattern of atypical hemispheric language dominance (Janszky, Jokeit, Heinemann, Schulz, Woermann et al., 2003). Nonetheless, the advent of seizures at an early age may compromise early cognitive development and the effects of seizures and medications may accumulate on this vulnerable foundation (Dennis, 2000; Elger et al., 2004; Hermann et al., 2006).

Investigations with children have contributed to a lifespan model of TLE. Children with recent-onset epilepsy have been found to show a pattern of mild diffuse cognitive impairment and academic underachievement (Aldenkamp, Weber, Wilhelmina,

Overweg-Plandsoen, Reijs et al. 2005; Hermann et al., 2006). This evidence of baseline difficulties suggest it is the underlying aetiology which affects cognitive functioning, not the seizures themselves (Aldenkamp et al., 2005; Hermann et al., 2006). Further evidence for this can be found in volumetric studies reporting an altered structure-function relationship in children with recent-onset epilepsy (Hermann et al., 2006). Finally, a study of newly-diagnosed adults reported impaired verbal memory at baseline but no evidence of progression, suggesting cognitive deficits are not exclusively due to the effects of chronic seizures or AEDs (Aikia et al., 2001).

At the other end of the lifespan, there is evidence for age-related cognitive decline in people with TLE (Helmstaedter et al., 2003), with a suggested rate of one standard deviation of decline over 30 years, across cognitive measures (Helmstaedter et al., 2002). Advancing age is known to be associated with cognitive decrements, particularly in the areas of attention, memory and speed of processing (Bopp & Verhaeghen, 2005; Christensen, 2001; Glisky, 2007). Further to this, additional damage due to epilepsy may accelerate the ageing process (Dennis, 2000; Helmstaedter & Kockelmann, 2006). Helmstaedter and colleagues (2002) reported that surgery may result in an exacerbation of pre-existing memory deficits as well as detrimentally affecting memory performance in older age. Normal ageing is associated with a decline in the fronto-temporal aspects of memory and removal of the temporal neocortex during surgery for TLE may put older patients at particular risk of cognitive impairment due to reduced ability to compensate for losing functional neocortex (Helmstaedter et al., 2002).

In a longitudinal study investigating normal cognitive ageing processes, Mitrushina & Satz (1991) examined cognitive changes over time in a sample of healthy older adults (57-85 years). Although practice effects on PIQ subtests (visually-based tasks) were evident for the youngest of the three age groups (57-65 years), these effects diminished with advancing age and the performance of people in the oldest age group (76-85 years) actually declined on retest. In contrast, they found evidence of uniform decline (defined as no practice effects) across all ages of the older adults on the VIQ subtests (verbally-based tasks) and suggested this pattern reflected the normal ageing

process. In a cross-sectional study of ageing and cognition in TLE, Jokeit & Ebner (2002) reported cognitive decline was evident over time, but occurred slowly and was difficult to distinguish from normal ageing. Nonetheless, memory decline may start from a lower initial level in people with TLE, which means poor performance may emerge earlier in the lifespan than in healthy controls (Helmstaedter & Elger, 1998; Jokeit & Ebner, 2002).

Duration of Seizures

A particularly salient issue in the TLE literature is whether seizures *per se* are associated with cognitive decline over time, in the absence of surgical intervention. This idea is encompassed by the term "progression" (pg. 445, Seidenberg et al., 2007) and investigation of this concept is fraught with methodological difficulties. Duration of seizures is considered an important indicator of chronicity as it is associated with a number of variables that can cause adverse cognitive functioning. These include focal or generalised seizures, inter-ictal discharges, metabolic disturbances, long-term use of AEDs and increased risk of seizure-related head injuries, as well as psychosocial and psychological sequelae (Jokeit & Ebner, 1999; Oyegbile et al., 2004; Vingerhoets, 2006). The contribution of these factors is not easily disentangled, and the relationship between epilepsy duration and cognitive decline remains clouded.

As an example, the frequency with which seizures occur is not necessarily constant, but may vary with time, medications, hormonal levels, stress and so forth. In a review of the literature, Dodrill (2004) found the relationship between frequency of seizures and cognition difficult to determine due to limited studies and those studies which were reviewed often lacked a control group against which to compare the impact of developmental and ageing influences. In addition, self-report and carer reports regarding seizure frequency are vulnerable to subjective distortions and often conflict with medical records. These methodological difficulties are likely to be at least partly responsible for the inconsistent evidence regarding the cognitive impact of chronic seizures over time. With this caveat in mind, a relationship between frequency of seizures and cognitive decline has been nonetheless reported in several studies (Dodrill, 2002; Helmstaedter et al., 2003; Hermann et al., 2006; Holmes, Dodrill,

Wilkus, Ojemann, & Ojemann, 1998; Piazzini et al., 2006; Thompson & Duncan, 2005).

Studies which follow the same, unoperated subject cohort over time are best placed to investigate the impact of seizures on cognition and an increased number of longitudinal research designs have emerged in recent years (e.g., Helmstaedter et al., 2000; Hermann et al., 2006; Holmes et al., 1998; Piazzini et al., 2006; Selwa, Berent, Giordani, Henry, Buchtel et al., 1994; Thompson & Duncan, 2005). Piazzini and colleagues (2006) assessed the cognitive function of adults (mean age = 37.2 years) with TLE who had not undergone surgery and a healthy control group (mean age = 37.1 years), at baseline and following a five-year interval. They reported a progressive decline in attention and psychomotor speed in the TLE group compared to controls, but no evidence of decline in other cognitive domains. Hermann and colleagues (2006) assessed unoperated TLE and control groups over a four-year period and identified adverse cognitive outcomes in a subset of patients, with memory found to be particularly vulnerable. In both studies, cognitive decline was associated with abnormal cerebral volumetrics, history of tonic-clonic seizures, lower baseline intellectual functioning, longer duration of epilepsy and older age at onset (Hermann et al., 2006; Piazzini et al., 2006).

Thompson and Duncan (2005) investigated adults with TLE who had been assessed on two occasions (median age at first assessment = 44 years), over ten years apart, with no intervening surgical treatment. The frequency of complex partial seizures was associated with a decline in memory, naming and executive functions, but not in overall intellectual functioning. A ten-year follow-up study of TLE patients treated with AEDs (mean age = 32 years) reported stable verbal memory, with mild reductions in performance on tasks assessing speed of processing and visual-spatial abilities (Holmes et al., 1998). Finally, Helmstaedter and colleagues (2003) followed 102 non-surgical patients over a period of ten years and found 50% demonstrated significant memory declines, but showed little evidence of change in other cognitive domains.

In contrast, Selwa and colleagues (1994) investigated cognition in 28 non-surgical TLE patients (mean age = 31.3 years) with a mean retest interval of approximately two years. Their results indicated no evidence of cognitive decline in this group, yet TLE patients voiced strong concerns of deteriorating memory function. This may be because memory deterioration occurs over longer periods than were investigated, or self-perceived 'memory difficulties' may refer to inconstant memories due to ictal and post-ictal amnesia. Alternatively, self-reported memory disturbances may be due to fluctuating anti-epileptic medication levels or the influence of affective functioning on cognition (Selwa et al., 1994).

In a study of patients with newly diagnosed (mean age = 35.3 years) and longstanding left-sided TLE (mean age = 42.4 years), Aikia and colleagues (Aikia et al., 2001) found evidence for impaired verbal memory in both groups. For newly diagnosed patients, memory impairment was associated with early-onset left TLE and secondarily generalised seizures. Follow-up of these patients after five years revealed no further deterioration in verbal memory performance for either group, suggesting that memory difficulties in left TLE are evident at diagnosis but do not continue to deteriorate (Aikia et al., 2001).

Education and Baseline Ability

Lower education level has been associated with both the onset of seizures at an earlier age and lower current IQ (Oyegbile et al., 2004), suggesting educational level may be a relevant marker for cognitive reserve capacity (Jokeit & Ebner, 1999). Cognitive reserve refers to the possibility that people with higher baseline cognitive abilities benefit from increased plasticity or neuro-protection which may assuage the effects of cerebral disease or insults (Deary, Whiteman, Starr, Whalley, & Fox, 2004; Oyegbile et al., 2004; Stern, 2002). The epilepsy literature provides support both for (Helmstaedter et al., 2003; Hermann et al., 2006; Jokeit & Ebner, 1999; Oyegbile et al., 2004; Piazzini et al., 2006) and against cognitive reserve theory (Andersson-Roswall et al., 2004; Marques et al., 2007; Thompson and Duncan, 2005). An alternative theory – called 'brain battering' – suggests lower educational attainment is a marker of lower socioeconomic status. The latter is in turn associated with poorer

health care and other adverse factors, ultimately resulting in decreased neuropsychological performance over time (Munoz, Ganapathy, Eliasziw, & Hachinski, 2000; Oyegbile et al., 2004).

Research has suggested that patients with TLE may have lower baseline performance across cognitive domains, compared to the general population (Hermann et al., 1997; Hermann et al., 2008; Helmstaedter, 2005; Marques et al., 2007; Oyegbile et al., 2004). Andersson-Roswall et al. (2004) compared patients with TLE and healthy controls over a three to four year period. They found the TLE group performed significantly worse on baseline cognitive testing than controls and showed no significant changes at follow-up, whilst the control group showed higher scores (i.e., practice effects).

Baseline cognitive ability may also affect the magnitude of practice effects seen on retest. Rapport and colleagues (1997) tested healthy adults (mean age = 26.7 years) every two weeks over eight weeks to examine the relationship between baseline IQ and practice effects. People with 'Average' and 'High Average' FSIQ at baseline gained more from previous exposure to the WAIS-R than those with a 'Low Average' baseline FSIQ, suggesting practice effects are mediated by baseline IQ. Regardless of baseline IQ, all adults improved over the four testing occasions; however, those with higher baseline performances improved the most – a case of "the rich get richer" (Rapport et al., 1997, p. 378).

In the surgical context, a person's baseline performance is considered a good indication of both the 'functional adequacy' of the tissue to be resected and reserve capacity (Helmstaedter, 2004). In addition, this functional adequacy is thought to determine the magnitude of post-operative cognitive decline and seizure control (Chelune, 1995; Helmstaedter & Kockelmann, 2006; Kneebone, Chelune, Dinner, Awad, & Naugle, 1995). Patients with higher baseline performance for memory – those with the most 'functional' memory abilities – are significantly more vulnerable to memory decline following surgery than those with lower baseline performance (Baxendale, Thompson, Harkness, & Duncan, 2006; Chelune et al., 1991; Helmstaedter et al., 2003; Hermann, Seidenberg, Haltiner, & Wyler, 1995; Ivnik,

Sharbrough, & Laws, 1988; White, Matchinsky, Beniak, Arndt, Duncan, et al., 2002). Nonetheless, it is important to note that people who perform better at baseline will also perform better post-operatively (Helmstaedter et al., 2003; Hermann et al., 2006; Jokeit & Ebner, 1999; Oyegbile et al., 2004; Seidenberg et al., 2007). These findings support the need to consider differences in baseline performance when determining the magnitude of practice effects.

Anti-Epileptic Medications

Neuropsychological test data in patients with TLE may not only reflect cognitive impairment due to epileptic lesions and associated damage (Baxendale et al., 1998), but also the reversible effects of anti-epileptic medications (Helmstaedter, 2004). Striking the right balance between seizure control and the cognitive side-effects of medications is particularly difficult using the so-called 'old' generation of anti-epileptic drugs (AEDs), including phenobarbitone, phenytoin, carbamazepine and valproate (Ortinski & Meador, 2004). Even those medications considered to have minimal cognitive side-effects show mild, psychomotor slowing when compared with no treatment at all (Vermeulen & Aldenkamp, 1995).

In recent years, the introduction of a 'new' range of AEDs has markedly decreased the incidence of cognitive side-effects. In their review of the literature, Aldenkamp and colleagues (Aldenkamp, De Krom & Reijs, 2003) concluded that although the majority of newer AEDs demonstrate adequate cognitive profiles, this conclusion is tempered by the dearth of controlled studies. A notable exception, however, was the effect of topiramate on cognition, with a number of studies indicating associated attentional difficulties, as well as adverse effects on verbal function and language (Aldenkamp et al., 2003; Meador, Loring, Hulihan, Kamin & Karim, 2003; Thompson, Baxendale, Duncan & Sander, 2000).

Treatment with more than one AED, regardless of type, is associated with cognitive deficits, as are elevated drug serum levels (Kwan & Brodie, 2001; Ortinski & Meador, 2004; Vermeulen & Aldenkamp, 1995). The most common cognitive side-effects of the newer anti-epileptic medications are sedation, somnolence, distractibility,

insomnia and dizziness (Ortinski & Meador, 2004). Physical and emotional side-effects may also be experienced, including nausea, poor balance and affective symptoms. Both cognitive side-effects and poorly controlled seizures affect a person's quality of life, and so alternative interventions, such as surgery, may be considered.

Surgical Intervention for TLE

For approximately 30% of people with TLE, seizures remain non-responsive to anti-epileptic medication (Kwan & Brodie, 2000). Surgery for medically refractory TLE has become an increasingly accepted method of treatment, largely due to enhanced methods of anatomical and functional seizure localisation (Engel, Wiebe, French, Sperling, Williamson, et al., 2003; Fabinyi, 2002). Significant advances in both diagnostic testing and surgical procedures have resulted in more patients undergoing surgical treatment, with safer and more effective outcomes (NIH Consensus Panel, 1990; Wiebe, Blume, Girvin & Eliasziw, 2001). Highly sophisticated neuroimaging, electrophysiological monitoring and micro-neurosurgical techniques have allowed for increasingly accurate diagnosis of structural lesions and for specific tailoring of the surgical resection to the epileptogenic lesion (Clusmann, Schramm, Kral, Helmstaedter, & Ostertun, 2002).

In 1957, Scoville and Milner published a key paper on memory loss associated with bilateral medial temporal lobe resection (republished in 2000; see Figure 2). The seizures of their patient, H.M., became well-controlled by surgical intervention; however, he experienced dense anterograde amnesia that persisted until his death in December 2008, along with a temporally-graded retrograde memory for events up to eleven years prior to the surgery. This dramatic outcome precipitated the understanding that the hippocampus and surrounding structures are fundamentally involved in the formation of new memories (Scoville & Milner, 1957).

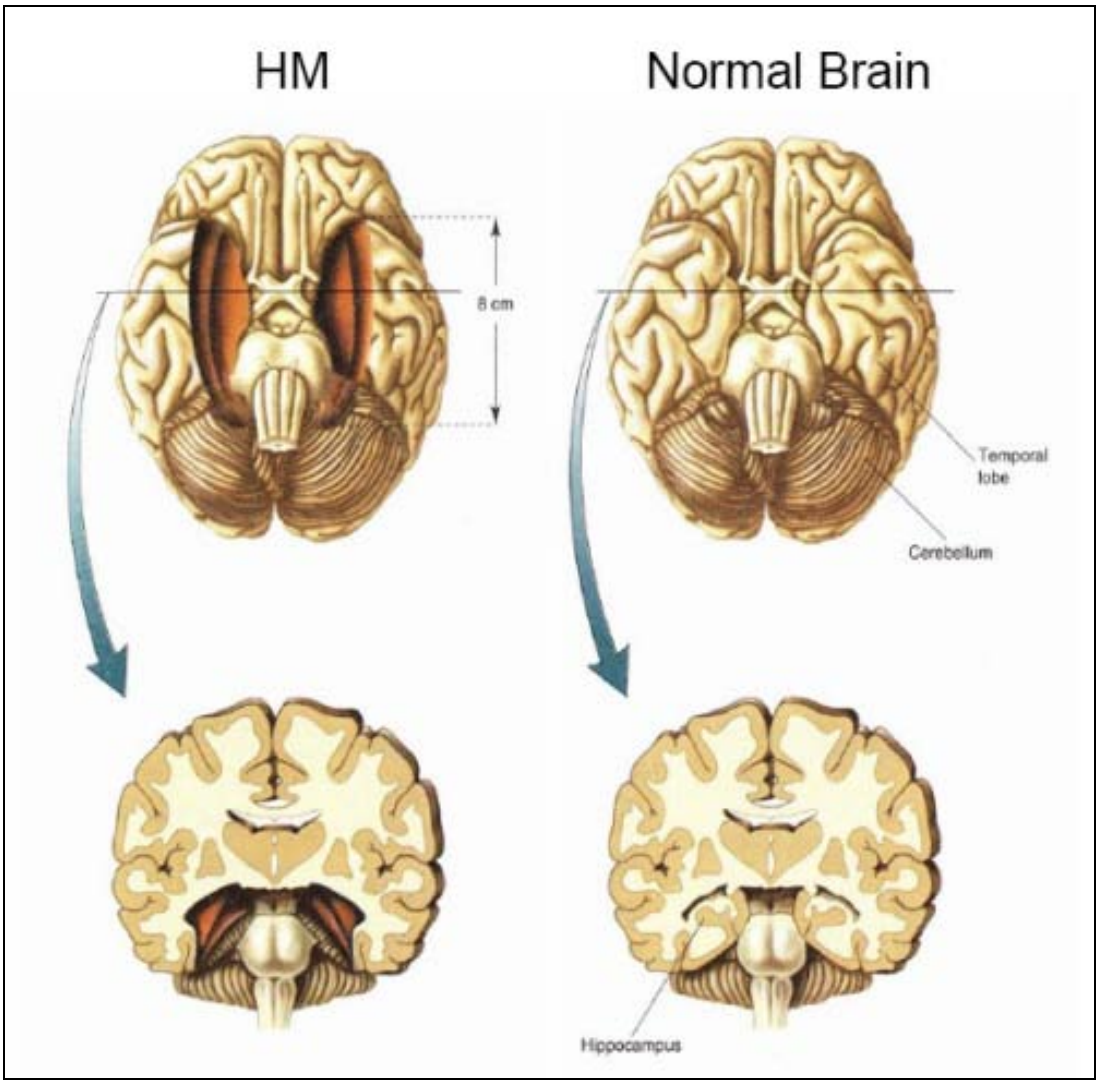

Figure 2. Comparison of normal brain with H.M.'s brain following his bilateral temporal lobectomy. http://scienceblogs.com/neurophilosophy/2007/07/remembering_henry_m.php

The most common surgical procedure for treating mesial TLE is anterior temporal lobectomy (ATL) with hippocampectomy (Fabinyi, 2002). This technique has a high rate of success for eliminating or reducing complex partial seizures in patients with hippocampal sclerosis (Fabinyi, 2002). The literature suggests approximately 70% of TLE patients will become seizure-free following this type of resection (Armon, Radtke, Friedman & Dawson, 1996; Engel Van Ness, Rasmussen & Ojemann, 1993). In contrast, individuals with non-mesial TLE demonstrate no evidence of hippocampal sclerosis and are therefore more likely to undergo an anterior two-thirds temporal lobectomy, which is less targeted and removes more temporal neocortex (Helmstaedter, Reuber & Elger, 2002). A standard two-thirds en bloc temporal lobe resection has been shown to adversely affect both memory acquisition and retention,

while a selective amygdalohippocampectomy does not affect acquisition, but impairs retention (Elger et al., 2004). Nonetheless, evidence suggests either approach may cause neurological or neuropsychological sequelae, particularly involving verbal memory, depending on the side and extent of the resection (Bell & Davies, 1998; Clusmann et al., 2002; Elger et al., 2004; Helmstaedter & Elger, 1996; Helmstaedter et al., 2003; Hermann et al., 1992; Pauli, Pickel & Schulemann, 1997).

Surgical intervention for TLE may not only provide freedom from seizures, but may also bring social and psychological benefits (Wiebe et al., 2001). In addition to this, cost-benefit analysis has put surgical intervention ahead of long-term management with medications (Silvenius, 1999). However, not all patients with TLE can be successfully treated by surgery and stringent pre-operative criteria must be met before a patient is considered an appropriate surgical candidate. Whilst there is no consensus on a pre-surgical evaluation protocol across epilepsy centres, a range of investigations are undertaken and typically include a detailed medical and social history, neurological examination, psychiatric assessment, scalp electroencephalogram (EEG), video-EEG telemetry, neuroimaging and neuropsychological assessment.

Post-Operative Cognitive Outcomes

Despite excellent outcomes for seizure control, surgical intervention for TLE may also be associated with a significant risk of post-operative cognitive decline. Substantial evidence can be found in the post-surgical epilepsy literature to support the theory of material-specific memory deficits, discussed above. Consistent with the assumed functional neuroanatomy of the left temporal lobe, people undergoing left ATL are at increased risk of impaired verbal memory (Alpherts, Vermuelen, van Rijen, Lopes da Silva, & van Veelen, 2006; Bauer, Breier, Crosson, Gilmore, Fennel et al., 1995; Chelune, Naugle, Luders & Awad, 1991; Chelune et al., 1993; Helmstaedter, et al., 2002; Hermann et al., 1992; Ivnik et al., 1987; Lee et al., 2002; McSweeny et al., 1993; Seidenberg et al., 1998; Selwa et al., 1994). Similar to the nonsurgical literature, findings of comparable decrements in nonverbal memory following right-sided resection are less consistent and the relationship does not appear to be as strong as that between left-sided surgery and verbal memory changes (Bauer

et al., 1995; Gleissner et al., 1998; Ivnik et al., 1987; Jones-Gotman, Brulot, McMackin, Cendes, Andermann, et al., 1993; Kneebone et al., 1995; Piggott & Milner, 1993; Rausch, Kraemer, Pietras, Le, Vickrey et al., 2003; Selwa et al., 1994). Given the potential for adverse post-operative cognitive outcomes and the subsequent impact on a person's daily functioning and quality of life, careful pre- and post-operative assessment of cognitive functioning is paramount.

The Role of Neuropsychology in Epilepsy Surgery Settings

Pre-operative neuropsychological assessment of potential surgical candidates is necessary for obtaining a cognitive baseline, contributing to the lateralisation of the epileptogenic focus, and predicting post-surgical neuropsychological outcome (Jones-Gotman et al., 1993; Chelune, 1994). This information is also provided to patients and clinicians to help decide whether to progress to surgery. The post-operative neuropsychological assessment aims to characterise any changes in cognitive functioning as a result of surgery, both positive and negative, and provides recommendations for management of any identified difficulties (Helmstaedter, 2004; Lineweaver, Morris, Naugle, Najm, Diehl, et al., 2006; Sawrie et al., 1998). .

A longstanding role of the pre-operative neuropsychological assessment has also been to assist in the localisation and lateralisation of the epileptogenic lesion (Jones-Gotman, Smith, & Zatorre, 1993). Consistent with material-specific theory, patterns of performance on verbal and non-verbal tasks may allow for lateralisation of cognitive impairment and therefore epileptogenic lesion, although neuropsychological test results alone do not provide sufficient localisation information. Reliable localisation conclusions may only be drawn where there is converging evidence from various clinical sources, including neuroimaging, EEG monitoring, patient history, neurological examination and neuropsychological testing (Jones-Gotman et al., 1993).

The cognitive impact of surgery differs markedly across patients due to the heterogeneous aetiology and effects of TLE. Attention has therefore turned to the identification of pre-operative risk factors which may predict cognitive outcome following surgery. As mentioned earlier, there is evidence to suggest the higher a

person's pre-operative memory abilities, the greater the risks of post-surgical verbal memory decline, particularly in patients with left-sided TLE (Baxendale et al., 2006; Chelune et al., 1991; Dodrill, Jones-Gotman, Loring, & Sass, 1993; Helmstaedter et al., 2003; Hermann et al., 1996; Jokeit, Ebner, Holthausen, Markowitsch, Moch et al., 1997; Rausch et al., 2003; Seidenberg et al., 1998). However, further risk factors have also been identified and include resection of the language-dominant hemisphere, older age at surgery and the extent of the resection (Alpherts et al., 2006; Baxendale, Thompson & Duncan, 2008; Davies, Bell, Bush & Wyler, 1998; Helmstaedter, 2004; Helmstaedter & Elger, 1996; Helmstaedter & Elger, 1998). These wide-ranging factors suggest post-operative memory function relies on broader considerations than simply the integrity of the hippocampus (Baxendale, 2008).

Evaluating Cognitive Change Following Surgery for TLE

Whilst the fundamental aim of surgical intervention for TLE is the elimination of seizure activity, post-operative sequelae may be far-reaching and involve changes in a person's cognitive, emotional and psychosocial functioning, as well as quality of life. Positive post-surgical outcomes for seizure control are high but not absolute, and, as discussed above, negative neuropsychological outcomes are not infrequent. Evaluating a person's cognitive functioning following surgery provides a vehicle for determining what post-operative support is required, as well as suggesting appropriate cognitive and adjustment strategies based on a person's cognitive strengths and weaknesses. The relevance of these clinical decisions and recommendations is reliant on the accurate interpretation of pre- to post-surgical changes in cognitive functioning.

Normative and base-rate data for estimated retest change scores is available for some neuropsychological instruments, including the Wechsler Adult Intelligence Scale – Third Edition (WAIS-III) and Wechsler Memory Scale – Third Edition (WMS-III) (Iverson, 2001). However, the test-retest scores obtained from these standardisation samples is potentially problematic for use with epilepsy populations (Chelune et al., 1993; Martin et al., 2002b). Retest intervals of two to twelve weeks (mean = 35 days) were used by the test publishers to derive stability coefficients for the WAIS-III and

WMS-III indices (Wechsler, 1997). These intervals are brief compared to the average retest interval of six to eighteen months in epilepsy surgery settings, and may result in inflated estimates of practice and stability (Chelune et al., 1993). Further to this, comparing people with epilepsy against the standardisation sample assumes that patients with neurological disorders benefit from practice to the same degree as the neurologically intact population. However, an absence of practice effects has been reported in patients with TLE (Dodrill & Troupin, 1975; Hermann et al., 2006).

Determining Cognitive Outcomes using RCIs

Chelune and colleagues (1993) first applied a reliable change approach in the epilepsy surgery setting with a sample of 40 unoperated TLE patients. Using the methodology of Jacobson and Truax (1991), stability coefficients were calculated and the standard error of measurement was generated for each of the WAIS-R and WMS-R index scores. Cut-off scores for reliable change were delineated using 90% confidence intervals, with scores falling beyond these cut-offs indicating genuine change had occurred. Using the cut-off scores derived from the unoperated sample, Chelune et al. categorised post-operative patients as improved, declined or unchanged on the WAIS-R and WMS-R indices. Consistent with previous studies (e.g., Ivnik et al., 1987; Rausch & Crandall, 1982) modest improvements in IQ were found in patients from both the surgical and non-surgical groups, providing further evidence of practice effects (Chelune et al., 1993).

Helmstaedter and colleagues' (2003) also used RCIs to investigate memory and seizure outcomes for up to ten years following surgical intervention. Cognitive decrements, particularly in verbal memory, were more frequent and severe following left-sided surgery, more extensive resections and in patients who continued to have seizures following surgery. However, the results suggested that ongoing memory decline is associated with uncontrolled TLE and can be halted or even reversed with successful seizure control. Helmstaedter and associates (2003) characterised this dichotomous outcome trend as resulting in "double winners" or "double losers" (p. 431). Specifically, successful seizure control facilitated the recovery of extratemporal functions in the first year following surgery, with further recovery of temporal lobe

function observed over longer follow-up (two to ten years). In contrast, "double losers" suffered the cumulative affect of continuing seizures, unsuccessful surgery and the impact of ongoing seizure activity on cognition (Helmstaedter et al., 2003; Rausch et al., 2003).

Determining Cognitive Outcomes using SRB Norms

Whilst Chelune and colleagues (1993) presented retest change data in TLE using a categorical approach, others have presented their change data using a continuous, normative approach. McSweeny et al. (1993) twice tested a sample of 50 non-surgical epilepsy patients (mean age = 31.6 years) using the WMS-R and WAIS-R. Regression equations were used to predict retest scores based on a person's baseline performance and any other significant predictor variables, such as age or education. These equations were then used to determine whether the observed retest scores of left- and right-temporal lobectomy patients were significantly different from the predicted retest scores. The differences between the observed and predicted retest scores were then converted to *t*-scores, providing an illustration of the magnitude of change across measures, on a common metric. The most significant decline was in verbal memory, with both the right and left temporal lobectomy groups affected, albeit the latter more dramatically.

Seidenberg and colleagues (1998) investigated neuropsychological change following ATL in 88 patients, with and without mesial TLE. They also derived regression-based estimates of retest change from a nonsurgical control group. Surgery was not associated with statistically significant cognitive decline in either the left or right mesial-TLE groups, although a pattern of mild decline was evident, particularly for memory. For the non-mesial TLE group there was clear evidence of a significant decline in cognition, particularly in verbal memory, following both left- and right-sided surgery. This finding is contrary to predictions based on material-specificity hypotheses.

Determining Cognitive Outcomes using RCI and SRB Methodologies

Hermann and colleagues (1996) used the retest data from an unoperated epilepsy sample to calculate both RCIs and SRB change norms for use in the epilepsy surgery setting. Similarly, Sawrie and colleagues (1996) employed both methods to evaluate the cognitive effects of epilepsy surgery across a number of cognitive domains. A sample of 51 unoperated epilepsy patients underwent neuropsychological assessment on two occasions, approximately eight months apart. Using the methodology of Jacobson and Truax (1991), 90% cut-off scores were calculated to mark the boundaries of reliable change on retest. The SRB approach was also used to generate multiple regression analyses for predicting retest scores from baseline performance. Sawrie and colleagues (1996) recommended clinicians in epilepsy surgery settings utilise both RCIs and SRB change norms to evaluate the significance and magnitude of change across cognitive measures.

Following the introduction of the third edition of the Wechsler scales, Martin and colleagues (2002) established RCIs and SRB change scores for the subtests and indices of both the WAIS-III and the WMS-III. Forty-two unoperated epilepsy patients (mean age = 34.8 years) were assessed on two occasions, with a mean retest interval of approximately seven months. The resulting RCI and SRB change score indices controlled for psychometric confounds and provided a standardised method for evaluating change following epilepsy surgery, at both the individual and group levels (Martin et al., 2002b).

The Number Needed to Harm and the Number Needed to Treat

A clear explanation of the potential cognitive impact following epilepsy surgery is imperative for patients considering this treatment option. A patient should be provided with information such as the number of people with similar presentations who have experienced adverse cognitive effects of surgery, and how likely this is to happen to them. A useful method of determining such information is to use base-rates to generate the number needed to treat (NNT) and the number needed to harm (NNH) (Cook, 1995; Laupacis, Sackett, & Roberts, 1988; Sackett, Straus, Richardson,

Rosenberg, & Haynes, 2000*).* The NNT tells us how many people must be treated using a particular intervention for a beneficial outcome (Citrome, 2007). Similarly, the NNH indicates the number of people who must be treated with a particular intervention before harm is caused to one patient (Citrome, 2007; Straus, Richardson, Glasziou, & Haynes, 2005).

The NNT and NNH are calculated using the absolute risk reduction (ARR). This term refers to the difference in probabilities of an event in the control and treatment groups (Cook, 1995). If the event rate is lower in the treatment group, this suggests a potential benefit from the treatment; however, if the event rate is greater in the treatment group (i.e., a negative ARR), this indicates the treatment may be harmful (Cook, 1995). The ARR is calculated by subtracting the experimental event rate (e.g., the percentage of the post-operative TLE sample experiencing decline on retest) from the control event rate (e.g., the percentage of the pre-operative TLE sample experiencing decline on retest). If this calculation generates a positive result, it is referred to as the ARR as it indicates patients will not experience more adverse events associated with the experimental treatment (i.e., surgery) than they would have with the control treatment (i.e., medication). Conversely, if subtracting the experimental rate from the control rate gives a negative result, it is called the absolute risk increase (ARI).

$$\text{ARR/ARI} = \text{control rate} - \text{experimental rate}$$

or,

$$\text{ARR} = \text{control rate} > \text{experimental rate}$$
$$\text{ARI} = \text{control rate} < \text{experimental rate}$$

The number needed to treat is simply the reciprocal of the ARR and can range from 1 (or -1) to infinity (or -infinity) (Lessaffre & Pledger, 1999). It is calculated as follows (Sackett et al., 2000),

$$\text{NNT} = \frac{100}{\text{ARR}}$$

The number needed to harm is calculated in the same way at NNT but is used to ascribe a disadvantage to a possible treatment. For example, if calculating the NNT generates a negative number, this indicates the particular intervention is disadvantageous and the figure subsequently becomes a NNH (Citrome, 2007).

Rationale for the Current Study

The validity of neuropsychological formulation and opinion relies in turn on the reliability and validity of the neuropsychological measures utilised in assessment. Neuropsychological tests must be interpreted using the best available statistical approaches and normative data in order to provide the highest standards of clinical care to the patient (APS, 2007; Wong, 2006).

To accurately evaluate the effects of epilepsy surgery on cognitive functioning, it is important to establish normative and base-rate information using a relevant comparison group that has been tested twice over an extended period and therefore reflects routine pre- and post-surgical neuropsychological assessment procedures (Chelune et al., 1993; Chelune, Sands, Barrett, Naugle, Ledbetter, et al., 1996; Hermann et al., 1996; Sawrie et al., 1996). This is most effectively achieved using a comparable non-surgical epilepsy sample to determine appropriate reliable change indices and regression equations (Dodrill, Hermann, Rausch, Chelune, & Oxbury, 1993; Hermann et al., 1996; Martin, Sawrie, Roth, Morawetz, & Kuzniecky, 1998; Martin et al., 2002b; McSweeny et al., 1993; Smith, Elliot & Lach, 2002; Vingerhoets, 2006). Studies have also emphasised the need to use demographic information, such as age, education and baseline cognitive functioning, as well as seizure onset, duration and control, to ensure the change norms derived are most representative of the local epilepsy population (Helmstaedter et al., 2003; Martin, et al., 2002; McSweeny et al., 1993; Sawrie et al., 1996).

Research has cautioned against the application of change norms derived from non-local epilepsy samples, instead recommending the calculation of norms based on local demographic information and seizure characteristics (Martin et al., 2002b; Sawrie et al., 1996). The current study obtained demographic and cognitive retest data from

unoperated patients with TLE and used SRB and RCI approaches to determine cognitive change on retest in a local Australian sample. Establishing local change norms provides clinicians with a standardized method, based on comparable demographics, for determining cognitive change that might be considered unusual (Martin et al., 2002b). In the context of epilepsy surgery, this enables the clinician to determine the cognitive impact of surgery at an individual level. Further to this, an improved understanding of post-surgical changes in cognitive functioning will more accurately inform patients and families of the potential cognitive changes associated with surgery for TLE.

Aims and Hypotheses

The aims of the current study cover two main areas and these are reflected in the two following experimental chapters. Firstly, Chapter Three aimed to develop local cognitive change norms based on a pre-operative TLE sample, using RCIs and SRB change norms. The aim of Chapter Four was to compare the classification of cognitive change according to local versus North American change data (Martin et al., 2002b), by applying them to a local post-operative TLE sample.

Regarding the generation of local change norms, it was hypothesised that improvements would be evident in the nonsurgical epilepsy population due to known methodological confounds. That is, both the effects of practice and the less-than-perfect stability of the neuropsychological measures were expected to artificially enhance retest scores. The study aimed to generate RCIs based on both the standard error of the difference and the standard error of prediction to evaluate any differences between these two approaches. Based on the literature, it was expected that the SE_{pred} method would hold a mild advantage over the SE_{diff} approach in accurately classifying change on retest (Chelune, 2003).

Using SRB change methodology, it was expected that a person's baseline test score would significantly predict their performance on retest (Dikmen et al., 1999; Heaton et al., 2001; Hermann et al., 1996; Martin et al., 2002b; McSweeny et al., 1993; Sawrie et al., 1996; Sherman et al., 2003). Based on a review of the literature, it was

also expected that a person's age (Helmstaedter et al., 2002; Martin et al., 2002b; Mitrushina & Satz, 1991) and number of anti-epileptic medications (Meador, 1996; Sherman et al., 2003) would be significant predictors variables. In contrast, it was hypothesised that neither a patient's education level (McSweeny et al., 1993; Sawrie et al., 1996), nor the length of their seizure duration (Martin et al., 2002b; McSweeny et al., 1993; Sawrie et al., 1996), would predict their retest performance. Finally, given the extended retest intervals, it was not anticipated that the length of the test-retest interval would significantly predict retest performance (Chelune et al., 1999; Hermann et al., 1996; McSweeny et al., 1993; Sawrie et al., 1996).

The application of both local and North American change norms (Martin et al., 2002b) to a local, post-operative sample was expected to reveal different cognitive change base-rates. Specifically, the local norms were expected to more closely classify the retest data according to the specific confidence intervals applied. Further to this, it was hypothesised that dividing the post-operative sample into right- and left-sided surgery would reveal cognitive differences as a function of laterality following surgical intervention. In particular, it was hypothesised that left-sided surgery would be associated with verbal memory decline, but a less clear relationship was anticipated between right-sided surgery and related visual memory decline. In addition to this, a decline in confrontational naming was expected for those patients undergoing left ATLs (Bell et al., 2000; Hermann et al., 1992; Seidenberg et al., 1998).

Overall, the current study aimed to enhance the interpretation of cognitive outcomes for both surgical and nonsurgical epilepsy patients, firstly by developing local norms, and secondly, by clarifying whether or not local norms are in fact needed. It was hoped that by improving clinical interpretation of post-operative cognitive changes following TLE surgery, patients considering this procedure would be more comprehensively informed of the potential cognitive risks and rewards.

CHAPTER 3

Generating Reliable Change Indices and Standardised Regression-Based Norms

Chapter One reviewed research recommending the generation of RCIs and SRB change norms based on neuropsychological test data collected from a local, comparable population (Martin et al., 2002b; Sawrie et al., 1996). In the case of surgical intervention for TLE, the most appropriate comparison group are people who have a diagnosis of TLE but have not yet progressed to surgical intervention (Dodrill et al., 1993; Hermann et al., 1996; Martin et al., 1998; Martin et al., 2002b; McSweeny et al., 1993; Smith et al., 2002; Vingerhoets, 2006). This chapter describes the calculation of RCIs and SRBs using a local unoperated TLE sample. The proportion of patients classified as declined, unchanged or improved, according to the Australian and US change norms (Martin et al., 2002b), were compared to investigate whether non-local change data is equivalent to local data.

Method

Participants

Ethics approval was obtained from the St Vincent's Hospital and Victoria University Human Research and Ethics Committees (Appendix A). The participants were 17 female and 24 male patients with unoperated TLE, recruited from the St Vincent's Hospital Comprehensive Epilepsy Program. All were aged 16 years or over and had been diagnosed with TLE by experienced consultant neurologists and epileptologists based on clinical symptomatology, EEG findings and neuroimaging results. The determination of local change norms requires the use of a comparison group with similar test-retest intervals to the surgical population (Hermann et al., 1996) and the use of a clinical, unoperated population was deemed most appropriate for this investigation. In previous research, sample sizes of 40 (Hermann et al., 1996) and 42 (Martin et al., 2002b) were found to be adequate in order to determine RCIs and SRBs.

Formal consent was obtained using the relevant Participant Information and Consent

Form (See Appendix B). Patients under the age of 18 years and patients who did not have the capacity to give voluntary and informed consent were excluded from this study, with the exception of one participant aged 16 years at the time of the first assessment, for whom archival data was obtained. Participants were included or excluded according to the following criteria: (a) documented seizures of temporal lobe origin or involving temporal areas, as identified by clinical symptomatology, EEG and neuroimaging findings, (b) no history of neurodegenerative disease, and (c) no history of epilepsy surgery. The latter criterion was allowed one exception; a woman who had undergone unsuccessful surgery for seizures as a young child and continued to experience medically intractable TLE, leading to consideration of surgery for her as an adult. She was included because several patients with similar histories have progressed to surgery in the St Vincent's Hospital epilepsy program. Finally, people with intellectual disabilities (FSIQ <70) are often excluded from the epilepsy literature (see Sherman et al., 2003 for an exception); however, this study included three people with Full Scale IQs under 70, as their inclusion was considered to more accurately represent the full spectrum of cases seen in this clinical population.

There were three sources of data collection for this study. In the first, letters of invitation were sent to all patients who had previously undergone a neuropsychological assessment as part of the investigative work-up for surgery (n = 136), of whom 34 subsequently agreed to participate. Of those who did not participate, 26 had since gone on to have surgery, six were lost to follow-up, four were considered too unwell to participate by their doctors, four lived inter-state, four declined with no reason stated, two were deceased and 56 simply did not respond. For those 34 who agreed to participate, brief screening questions were administered to cover the exclusion criteria and discuss any participant queries, before an appointment time for cognitive re-assessment was arranged.

The second source of data collection involved access to archival data for six patients who had undergone two pre-operative neuropsychological assessments. Two pre-operative assessments had been deemed necessary for these six patients due to lengthy and complex pre-operative work-ups, often because of conflicting lateralising results, or because the patient had re-considered surgical intervention after some time had

passed. Finally, one patient in the epilepsy program who had not yet undergone a pre-surgical neuropsychological assessment was recruited for two neuropsychological assessments, twelve months apart.

All of the 41 participants were either on the waiting list for surgery, had decided against surgical intervention, or had been excluded from surgery due to clinical reasons, such as high risk of lesion excision causing cognitive decline, unclear focus localisation, or seizures not yet deemed definitively intractable. Nineteen participants had been identified as experiencing seizures originating from the left temporal lobe and 12 from the right temporal lobe. Six individuals had non-localising temporal-extemporal involvement and four had bilateral temporal involvement, according to EEG evidence. After data collection and at the time of writing, four of the recruited participants had subsequently undergone ATL, one was on the surgery waiting list and six were still undergoing pre-operative investigations.

The following demographic information and medical details were obtained from the participants - age, gender, number of years of formal education, handedness, age at onset of seizures (years), frequency of seizures (per week), time since last known seizure (weeks), current medications, and current surgical status. In addition to this, patients were questioned regarding any perceived memory or cognitive changes since their previous assessment.

Materials

The following neuropsychological tests were included in the cognitive test battery and together form part of the routine pre- and post-surgical neuropsychological assessment battery utilised by the St Vincent's Hospital Neuropsychology Unit.

Wechsler Adult Intelligence Scale – Third Edition (Wechsler, 1997a).

The WAIS-III is an individually administered battery of tasks used to assess general intellectual function in individuals aged 16 to 89 years. Research has revealed sound psychometric properties and utility for clinical practice, resulting in the WAIS-III becoming one of the most frequently used tests in clinical practice (Camara, Nathan &

Puente, 2000). Of 13 core subtests, six (Vocabulary (Vocab), Similarities (Sim), Arithmetic (Arith), Digit Span (DSp), Information (Info), Comprehension (Comp) and Letter Number Sequencing (LNS)) comprise the *Verbal Intelligence Quotient* (VIQ) and five (Picture Completion (PC), Block Design (BD), Matrix Reasoning (MR), Picture Arrangement (PA) and Digit Symbol (Coding)) make up the *Performance Intelligence Quotient* (PIQ). The overall, or *Full Scale IQ* (FSIQ), results from a composite of 11 Verbal IQ and Performance IQ subtests. Based on a four-factor model of the test, 11 of the 14 subtests are used to generate the *Verbal Comprehension Index* (VCI; Vocab, Sim, Info) and the *Perceptual Organisation* (POI; PC, BD, MR), *Working Memory* (WMI; DSp, Arith, LNS) and *Processing Speed* (PSI; Coding, Symbol Search) indices (Strauss, Sherman, & Spreen, 2006).

Wechsler Memory Scale – Third Edition (Wechsler, 1997b).

The subtests of the WMS-III used for the current study were, Logical Memory I, II (LMI and LMII) and (LMR) Recognition, Faces I and II Recognition (FI and FII), Verbal Paired Associates I, II (VPAI and VPAII) and Recognition (VPAR), Family Pictures I and II Recall (FPI and FPII), Spatial Span (SpSp), Letter-Number Sequencing (LNS) and Digit Span (DSp).

The WMS-III is comprised of 11 core subtests designed to assess declarative and working memory abilities in adults. Eight primary indices are generated from the core subtests: Immediate Memory, comprised of Auditory Immediate (LMI and VPAI) and Visual Immediate (FI and FPI); General Memory, composed of Auditory Delayed Recall (LMII and VPAII), Auditory Recognition Delayed (LMR and VPAR) and Visual Delayed (FII and FPII); and, Working Memory (LNS and SpSp), including both auditory and visual working memory subtests. In the current study, the Auditory Recognition Delayed index was excluded due to unacceptably poor stability (i.e., less than 0.8; Wechsler 1997). The six supplemental subtests are not routinely included in the St Vincent's Hospital pre-operative neuropsychology battery due to time constraints in clinical evaluations, as well as the lower reliability and stability of individual subtests, as opposed to that of index scores.

The WAIS-III and the WMS-III were co-developed using the same standardisation sample (Psychological Corporation, 1997). The Letter-Number Sequencing subtest is included in both the WAIS-III and WMS-III and contributes to the Working Memory Index in both batteries; however, in the standardisation sample the subtest was administered only once due to the effects of practice and the same procedure was followed for the current study (Tulsky & Ledbetter, 2000) The Digit Span subtest was also administered only once, according to this rationale.

Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983).

The BNT is a 60-item measure of visual naming ability using black and white drawings of common objects. Research suggests it is sensitive to left temporal lobe dysfunction in individuals with TLE (Randolph, Lansing, Ivnik, Cullum, & Hermann, 1999; Schefft, Testa, Dulay, Privitera, & Yeh, 2003).

Procedures

Testing for this study was completed at the Neuropsychology Unit at St Vincent's Hospital. The results of the 34 patients who had already undergone neuropsychological assessment were obtained from their neuropsychology files. In order to obtain test-retest data, the WAIS-III, WMS-III and BNT were readministered (mean retest interval = 50.44 months, SD = 32.6). One participant who had not undergone an initial neuropsychological assessment was assessed on two occasions, twelve months apart. The cognitive test results of the six patients who had completed two pre-operative assessments were accessed from the Neuropsychology Unit's archival database.

Following the neuropsychological assessment, participants were provided with a summary of their results, including any changes in their cognitive test results over time, as well as a list of strategies to deal with any areas of weakness. Participants were offered the opportunity to receive a summary of the research project results when completed. Participants were not offered any form of remuneration for their involvement in the study.

Statistical Analyses

The data analysis involved the development of both RCIs and SRB change norms. Using these techniques, the test-retest data from the unoperated TLE group was used to generate 90% and 95% confidence intervals for determining cognitive change.

Reliable change indices.

As discussed in Chapter One, a number of different formulas for calculating RCIs are available. Two RCI methods – based on either the standard error of the difference or the standard error of prediction – were used in the current study and the resulting confidence intervals generated by each were then compared.

The first step involved using the test-retest data of the unoperated epilepsy sample to generate RCIs using the methodology described by Jacobson & Truax (1991). Test-retest coefficients (r_{12}) were calculated for each subtest score and index score of the WAIS-III and the WMS-III, as well as for the total score on the BNT. Following this – according to Iverson's (2001) modification – the test-retest coefficients were used to compute the standard error of measurement at each testing time. The SE_m in turn generated the standard error of the difference, using the formulae detailed in earlier chapters (Chelune et al., 1993; Iverson, 2001; Jacobson & Truax, 1991). For a detailed example of these calculations, refer to page eight of Chapter One.

Consistent with previous studies, a correction factor – based on the mean practice effects at each testing occasion – was then calculated to account for the effects of practice (Chelune et al., 1993; Hermann et al., 1996; Sawrie et al., 2002). However, this method of correction does not account for regression to the mean (Bruggemans et al., 1997). To overcome this problem, the standard error of prediction was also calculated (Chelune, 2003). The SE_{pred} represents the standard error of a retest score which has been predicted from a baseline score using a regression equation. This revised method allows a closer approximation to the theoretical distribution of practice-adjusted test-retest difference scores (Chelune, 2003).

Finally, both the SE_{diff} and the SE_{pred} were multiplied by ±1.64 standard deviations, and the correction factor added, to establish 90% change score confidence intervals for each methodology. To investigate the most accurate classification of the data, 95% confidence intervals were also calculated (i.e., SE_{diff} * ±1.96). Scores falling within these confidence intervals represented the distribution of scores for each measure on retest that would be expected if no real change had occurred, with 90% or 95% confidence. Therefore, scores falling outside these confidence intervals represent a change that would be expected to occur in an unoperated epilepsy population less than 10% or 5% of the time, respectively. Confidence intervals generated from both the above methods were placed around the predicted true score – not the obtained score – due to the effects of regression to the mean on retest (Dudek, 1979; Lord & Novick, 1968). As an example, for a retest score of 100, the PTS would remain 100 as there is no effect of regression to the mean; however, for a retest score of 90, the PTS would be approximately 95, due to the influence of regression to the mean. Similarly, a retest score of 110 would regress slightly towards the mean on retest and the PTS would approximately equal 105.

Standardised regression-based change norms.

SRB equations were generated for each index of the WAIS-III and WMS-III, using the methodology initially developed by McSweeny and colleagues (1993). This approach controls for the effects of both regression to the mean and practice, as well as using linear regression to predict retest scores from baseline performance on a particular neuropsychological measure. The SRB approach also allows for consideration of multiple predictors of retest scores, and so other potential factors which may affect retest performance were examined. Based on the epilepsy literature reviewed in Chapter Two, the relevant predictor variables – baseline test score, test-retest interval (months), age (years), education (years) and duration of seizures (years) – were entered into the regression equation to produce the predicted retest score.

Using these regression equations, retest scores were predicted based on the observed baseline scores, as well as the standard error of the estimate and the constant from the regression equations using McSweeny and colleagues' (1993) formula, discussed in

detail on page 11 of Chapter One. The difference between a person's predicted retest score and their observed retest score was then transformed into a standardised *z*-score. The change-score norms can also be converted to any number of standard scores, including *t*-scores (McSweeny et al., 1993).

Following the convention established in previous studies, change was demarcated using 90% confidence intervals, calculated by multiplying the *z*-scores by ± 1.64 standard deviation units (Hermann et al., 1996; Martin et al., 2002b; Sawrie et al., 1996). In addition, 95% confidence intervals were also calculated – *z*-score multiplied by ±1.96 – for comparative purposes.

Examining right versus left differences in the pre-operative sample.

To investigate any differences in retest performance according to laterality of seizure focus, the pre-operative group was divided into two groups – right- and left-sided TLE – and repeated measures analyses of variance (ANOVAs) were computed.

Examining cognition over time.

Given the extended test-retest intervals in this study (mean = 50.44 months, SD = 32.6), the opportunity was taken to examine changes in cognition over time in the unoperated TLE sample. Repeated measures analyses of variance were performed to examine potential differences between the groups, as well as the effect of age and retest interval on retest scores.

Results

Demographic Characteristics of the Participants

Forty-one adults (mean age 41.2 years; SD 14.2 years) with TLE who had not yet undergone surgical intervention were included in the study. Males were slightly over-represented in this sample, with 24 males and 17 females included overall. All participants had been educated to at least Year Eight, with 20 completing secondary

school and 15 of these continuing on to tertiary studies (mean education 12.29; SD 2.53). A patient's age at seizure onset was determined both from medical records and self-report, with half the sample experiencing the onset of seizures in childhood and adolescence (< 18 years of age) and half with seizure-onset in adulthood (mean age = 18.45; SD 14.08). Based on clinical symptomatology, EEG findings and neuroimaging results, 19 had been identified as experiencing seizures arising from the left temporal lobe, 12 from the right temporal lobe, four were reported to have a bilateral seizure focus and the seizure localisation of six remained unclear. Approximately two-thirds of the patient sample were being treated with a combination of two or more anti-epileptic medications. The length of the test-retest interval varied widely; the shortest interval was only seven months and the longest was ten years. Finally, over half the sample reported subjective memory difficulties since their first assessment (56.1%), whilst 13 (31.7%) reported no perceived memory concerns. The data for the remaining five participants (12.2%) had been obtained from the archival database and no subjective memory reports were available. Table 2 presents the participant demographics in full.

Table 2
Demographic and seizure information for the pre-operative sample

Variable (n = 41)	**Range**	**Mean**	**SD**
Age	16-75	41.2	14.2
Education (years)	8-17	12.29	2.53
Full Scale IQ (N=37)	54-123	92.05	16.06
Age at seizure onset (years)	0.5-64	18.45	14.08
Duration of seizures (years)	2-53	21.76	13.53
Retest interval (months)	7-126	50.44	32.6
Gender (male/female)	24/17		
Handedness (left/right)	5/36		
Left-sided TLE	19		
Right-sided TLE	12		
Bilateral TLE	4		
TLE focus unclear	6		
No/Monotherapy/polytherapy	1/11/29		
Subjective memory changes reported (Y/N/no data)	23/13/5		

Table 3 presents the mean and standard deviations at both test and re-test on the WAIS-III indices for the entire pre-operative sample, as well as the right- and left-sided surgery groups. Table 4 presents the same data for the WMS-III indices and BNT.

Table 3
Comparison of left- and right pre-operative groups on the WAIS-III indices

Index		**N**	**Time 1**		**Time 2**		**F**	**df**	***p***
			Mean	**SD**	**Mean**	**SD**			
FSIQ	All	26	92.4	13.6	93.4	15.9	0.313	24	0.581
	Left	17	91.4	13.6	92.1	15.3			
	Right	9	94.4	14.0	95.8	17.7			
VIQ	All	26	91.2	13.4	90.9	15.2	0.430	24	0.518
	Left	17	90.2	13.4	89.3	14.4			
	Right	9	93.1	14.2	94	17.1			
PIQ	All	26	95.2	13.9	97.8	15.8	0.245	24	0.625
	Left	17	94.1	13.7	96.9	15.6			
	Right	9	97.4	14.9	99.6	17.0			
VCI	All	25	91.7	13.4	92.4	14.9	0.004	23	0.949
	Left	17	92.1	13.3	92.2	14.5			
	Right	8	90.8	14.5	92.8	16.8			
POI	All	25	100.3	13.3	103	14.4	1.304	23	0.265
	Left	17	97.9	11.2	101.2	13.5			
	Right	8	105.3	16.8	106.6	16.3			
WMI	All	25	92.2	15.9	91.1	17.7	2.072	23	0.163
	Left	17	90.3	15.3	86.8	16.5			
	Right	8	96.0	17.6	100.3	17.6			
PSI	All	26	89.4	13.6	92.0	12.6	0.871	24	0.360
	Left	17	87.9	12.9	90.1	12.0			
	Right	9	92.1	15.1	95.7	13.7			

Repeated measures analyses of variance were calculated for the entire pre-operative sample, as well as the right- and left-sided surgery groups, to evaluate any differences in cognitive change between the groups from test to retest. As evident in Table 3, there was no significant difference in amount of change between the three groups on any of the WAIS-III indices.

Table 4
Comparison of left- and right pre-operative groups on the WMS-III indices and BNT

Index		N	Time 1		Time 2		F	df	*p*
			Mean	SD	Mean	SD			
AI	All	29	87.9	20.2	87.7	18.4	0.559	27	0.461
	Left	18	85.3	20.5	86.2	20.6			
	Right	11	92.1	20.0	90.2	14.6			
VI	All	29	84.7	18.6	85.9	15.6	0.552	27	0.464
	Left	18	86.9	21.0	87.2	16.8			
	Right	11	81.2	13.8	83.8	14.0			
IM	All	30	83.9	20.7	84.5	19.0	0.057	28	0.814
	Left	18	83.2	22.6	83.8	21.8			
	Right	12	85.0	18.3	85.5	14.6			
AD	All	30	86.6	21.1	87.8	18.7	0.467	28	0.500
	Left	18	83.8	21.1	86.6	21.1			
	Right	12	90.7	21.2	89.5	15.1			
VD	All	30	85.6	18.2	85.5	17.1	0.993	28	0.328
	Left	18	89.6	19.7	86.4	18.0			
	Right	12	79.8	14.4	84.0	16.2			
GM	All	30	83.9	20.6	85.1	18.6	0.047	28	0.830
	Left	18	83.4	21.9	84.4	20.3			
	Right	12	84.7	19.4	86.3	16.3			
WM	All	30	92.2	11.9	94.0	15.9	2.061	28	0.162
	Left	18	91.4	12.0	89.4	15.3			
	Right	12	93.4	12.1	100.8	14.9			
BNT	All	27	47.9	9.6	45.6	10.2	0.600	25	0.446
	Left	18	47.1	10.0	44.5	10.3			
	Right	9	49.7	9.0	47.8	10.3			

Repeated measures analyses of variance were also calculated for the three groups on the WMS-III indices and BNT. As reported in Table 4, there were no significant differences in the amount between the groups on any of the WMS-III indices. Based on the lack of significant differences between the left- and right-sided TLE groups, it was not deemed necessary to split the pre-operative sample before calculating the RCIs and SRB change norms.

Test Stability and Practice Effects

The test-retest reliability, or stability, coefficients calculated for the subtests and indices of the WAIS-III and WMS-III are presented in Tables 5 and 6, respectively. For comparative purposes, the stability coefficients from the Wechsler standardisation samples are also listed (Wechsler, 1997). The stability coefficients of the WAIS-III tended to be higher than for the WMS-III, with an overall trend towards more stable composite indices.

Due to the administration of an abbreviated test battery to nine of the participants on their first assessment occasion, the generation of statistical change data for these participants was necessarily limited. The number of cases available for analysis at both test and retest is detailed in Tables 5 and 6. In addition to this – as evidenced by the lower subject numbers for the Picture Arrangement and Comprehension subtests – approximately half the sample required the FSIQ of the WAIS-III to be prorated.

Table 5
Stability coefficients for the WAIS-III indices and subtests for the current and standardisation samples

WAIS-III	**N**	**Current Sample (mean retest interval = 50.44 months)**	**Wechsler Sample (mean retest interval = ~35 days)**
Verbal IQ	35	0.92	0.97
Performance IQ	35	0.91	0.94
Full Scale IQ	35	0.94	0.98
VCI	34	0.95	0.96
POI	34	0.85	0.93
WMI	33	0.77	0.94
PSI	35	0.87	0.88
Picture completion	34	0.49	0.83
Vocabulary	33	0.88	0.93
Coding	35	0.87	0.84
Similarities	36	0.84	0.86
Block design	34	0.75	0.86
Arithmetic	33	0.78	0.88
Matrix reasoning	35	0.67	0.90
Digit span	33	0.68	0.90
Information	33	0.97	0.91
Picture arrangement	23	0.79	0.74
Comprehension	27	0.91	0.84
Symbol Search	34	0.82	0.77
LNS	32	0.52	0.82

Overall, the WAIS-III stability coefficients (Wechsler, 1997a) in Table 5 tended to be higher than the coefficients generated from the current sample. However, with the exception of the Working Memory Index which was markedly lower in the current sample, all the WAIS-III indices showed comparable stability coefficients. As expected, a number of individual subtests showed low levels of stability (e.g., picture completion, letter-number sequencing, digit span, matrix reasoning), lending support to the idea that interpreting test profiles at the subtest-level is prone to error (Glutting, McDermott, Watkins, Kush, & Konold, 1997; Iverson, 2001; Livingston, Jennings, Reynolds, & Gray, 2003; Matarazzo & Hermann, 1984; Wechsler, 1997). Nonetheless, low subtest stability does not necessarily equate to low composite index stability (Livingston et al., 2003) and this pattern was evident in the current results. For example, the stability coefficients of the PC (0.49), MR (0.67) and BD (0.75) subtests were relatively poor; however; the overall stability of the POI composite measure (0.85) was adequate.

Table 6
Stability coefficients for the WMS-III indices and subtests and the BNT for the current and standardisation samples

WMS-III	N	Stability	Wechsler Stability
Auditory Immediate	38	0.87	0.93
Visual Immediate	38	0.75	0.82
Immediate Memory	39	0.88	0.91
Auditory Delayed	39	0.84	0.87
Visual Delayed	39	0.78	0.83
General Memory	39	0.86	0.91
Working Memory	38	0.62	0.86
Logical Memory I	39	0.75	0.88
Logical Memory II	39	0.78	0.79
Faces I	39	0.64	0.74
Faces II	38	0.62	0.74
VPA I	38	0.82	0.93
VPA II recall	38	0.74	0.83
Family Pictures I	38	0.60	0.81
Family Pictures II	38	0.60	0.84
Spatial Span	38	0.64	0.79
Boston Naming Test	34	0.8	0.8

Consistent with previous findings, the AI, AD, IM and GM indices of the WMS-III had stability coefficients greater than 0.8 (Iverson, 2001). Overall, the stability coefficients of the WMS-III indices of the current sample were only mildly lower than

the standardisation sample; however, the Working Memory Index was again significantly poorer. Similarly, the WMS-III subtests also showed lower stability coefficients than those from the standardisation sample. Once again, although a number of the visual memory subtests of the WMS-III had low stability (e.g., Faces I and II, Family Pictures I and II, Spatial Span), the stability of the composite memory indices was largely adequate.

The test-retest means and standard deviations for the WAIS-III indices and subtest scaled scores, as well as the WMS-III indices and subtest scaled scores are presented in Tables 7 and 8, respectively. The difference between the means of the two testing occasions was calculated using paired samples *t*-tests to determine whether any significant practice effects were evident. Bonferroni corrections were not done as they are highly conservative and may mask actual differences (Moore & Baker, 2002).

Table 7
Test and retest scores for the WAIS-III

Variable	**N**	**Test**		**Retest**		***t***	***p***
		Mean	**SD**	**Mean**	**SD**		
WAIS-III:							
Verbal IQ	35	91.89	15.97	91.02	16.82	0.784	0.438
Performance IQ	35	93.97	16.16	96.74	17.82	-2.271	**0.030**
Full Scale IQ	35	92.17	16.34	92.83	17.83	-0.651	0.520
VCI	34	93.15	17.90	92.71	17.81	0.464	0.646
POI	34	97.79	15.63	101	17.44	-1.993	**0.055**
WMI	33	92.73	14.87	91.39	16.46	0.713	0.481
PSI	35	88.43	15.09	91.57	13.26	-2.458	**0.019**
Picture completion	35	8.4	3.2	9.8	3.2	-2.385	**0.023**
Vocabulary	34	8.7	3.2	8.4	3.8	0.804	0.427
Coding	36	7.4	2.9	7.5	2.8	-0.976	0.336
Similarities	36	8.6	3.5	8.2	3.3	1.09	0.283
Block design	34	9.6	2.8	10.3	3.3	-2.052	**0.048**
Arithmetic	34	8.6	3.2	8.5	3.2	0.425	0.674
Matrix reasoning	34	10.7	3.0	10.2	3.7	0.897	0.376
Digit span	34	8.6	2.9	8.6	3.0	0.135	0.894
Information	33	9.1	3.5	9.2	3.6	-1.299	0.203
Picture arrangement	23	9.3	3.6	9.0	3.5	0.828	0.417
Comprehension	27	8.1	3.3	8.2	3.8	-1.188	0.245
Symbol Search	34	8.4	3.5	9.7	4.1	-2.272	**0.030**
LNS	31	8.6	2.7	8.4	3.4	0.305	0.763

Despite a mean test-retest interval of 50 months, statistically significant (p <0.05) practice effects were found for the Performance IQ and PSI. In addition to this, there was a trend towards significant practice effects on the POI (p = 0.055). Consistent with the literature, the Picture Completion, Block Design and Symbol Search subtests also showed statistically significant practice effects on retest. However, the majority of the measures did not show evidence of significant effects of practice.

Table 8
Test and retest scores for the WMS-III and BNT

Variable	N	Test		Retest		*t*	*p*
		Mean	SD	Mean	SD		
WMS-III:							
Auditory Immediate	38	88.71	19.94	89.47	19.42	-0.461	0.647
Visual Immediate	38	85.00	18.95	86.95	17.17	-0.938	0.355
Immediate Memory	39	84.51	21.02	86.23	20.40	-1.042	0.304
Auditory Delayed	39	87.67	21.00	88.62	19.59	-0.512	0.612
Visual Delayed	39	86.21	17.79	86.31	16.90	-0.055	0.956
General Memory	39	85.05	20.92	86.13	19.31	-0.627	0.534
Working Memory	38	92.24	11.92	93.89	15.62	-0.822	0.416
Logical Memory I	40	7.7	3.7	7.6	3.8	0.791	0.434
Logical Memory II	40	7.1	3.5	7.5	3.9	-0.699	0.489
Faces I	38	8.6	3.2	9.3	3.3	-2.113	**0.041**
Faces II	37	9.3	3.2	9.5	2.9	-0.879	0.385
VPA I	37	8.1	3.7	8.5	3.6	-1.753	0.088
VPA II recall	37	7.9	3.7	8.2	3.7	-1.489	0.145
Family Pictures I	37	6.7	3.3	6.2	3.1	0.679	0.502
Family Pictures II	37	6.7	3.6	6.0	3.4	0.947	0.350
Spatial Span	37	8.6	2.9	8.7	3.0	-0.788	0.436
Boston Naming Test	34	47.71	10.72	45.59	11.28	1.787	0.083

As listed in Table 8, none of the WMS-III composite indices showed statistically significant practice effects; however, practice effects were evident on the Faces I subtest. This interesting result is comparable to the findings of Martin et al. (2002), who reported a practice effect increase of one subtest point on the Faces I subtest. A potential explanation for evidence of practice effects on this particular task may be related to its recognition memory structure; patients are provided with prompts to recognise faces seen previously, but are not required to spontaneously recall the information.

Calculation of Reliable Change Indices

RCIs were calculated using both the standard error of the difference (SE_{diff}) and the standard error of prediction (SE_{pred}). This was done to determine whether the differing formulas resulted in significantly different RCI change cut-off scores.

Using the standard error of the difference.

The standard error of the difference (SE_{diff}) was generated based on the standard error of measurement using the formula,

$$SE_m = SD\ (\sqrt{1-r}).$$

Using the stability coefficients and standard deviations for each index or subtest score, separate standard errors of measurement were calculated for both the test and retest occasions to account for the effects of practice (Iverson, 2001). For a more detailed description regarding the calculation of this statistic, refer to page eight of Chapter One. The data used to generate Reliable Change Indices using the SE_{diff} is summarised in Tables 9 and 10.

Table 9
RCI data for the WAIS-III based on the SE_{diff}

Variable	Mean change	Practice correction	SE_{diff}	90% cut-off s ±	Adjusted 90% cut-offs	95% cut-off s ±	Adjusted 95% cut-off s
WAIS-III:							
VIQ	-0.87	0	6.56	10.69	± 11	12.86	±13
PIQ	2.77	3	7.22	11.78	(-9) (+15)	14.15	(-12)(+18)
FSIQ	0.66	1	5.93	9.68	(-9) (+11)	11.62	(-11)(+13)
VCI	-0.44	0	5.64	9.71	± 10	11.05	±12
POI	3.21	4	9.07	15.89	(-12) (+ 20)	17.78	(-14)(+22)
WMI	-1.34	0	10.64	17.3	± 18	20.85	±21
PSI	3.14	4	7.24	11.84	(-8)(+16)	14.19	(-11)(+19)
Picture comp.	0.76	1	3.23	5.30	(-4) (+8)	6.33	(-5)(+9)
Vocabulary	-0.07	0	1.72	2.82	±3	3.37	±4
Coding	0.55	1	1.45	2.38	(-2)(+4)	2.85	(-2)(+4)
Similarities	-0.05	0	1.92	3.16	±3	3.77	±4
Block design	2.53	3	2.16	3.55	(-3)(+5)	4.24	(-4)(+6)
Arithmetic	-0.18	0	2.12	3.48	±4	4.16	±5
MR	-0.97	0	2.74	4.49	±5	5.36	±6
Digit span	-0.33	0	2.36	3.87	±4	4.63	±5
Information	0.49	1	0.87	1.43	(-1)(+3)	1.70	(-1)(+3)
Picture arr.	-0.56	0	2.30	3.77	±4	4.51	±5
Comprehension	0.93	1	1.51	2.48	(-2)(+4)	2.96	(-2)(+4)
Symbol Search	1.48	2	2.29	3.75	(-2)(+6)	4.48	(-3)(+7)
LNS	-0.28	0	3.01	4.93	±5	5.90	±6
BNT	-2.12	0	6.96	±11.63	±12	±13.64	±14

Note: 90% cut-off scores = SE_{diff} * ± 1.64
95% cut-off scores = SE_{diff} * ± 1.96

Consistent with previous investigations, the RCI cut-off values were adjusted using a correction factor based on the mean practice effects for each measure (Chelune et al., 1993; Hermann et al., 1996; Martin et al., 2002b; Sawrie et al., 1996). This involved subtracting the mean retest score from the mean baseline score and rounding the resulting answer up to a whole number. Where the resulting number was a negative, the number was rounded up to zero and no correction factor was added. This practice correction was then added to the 90% and 95% confidence intervals to give 'adjusted' confidence intervals. To calculate the 90% and 95% confidence intervals, both the SE_{diff} and SE_{pred} were multiplied by the conventional confidence band of ±1.64 or ±1.96 standard deviations, respectively. This established a 90% or 95% confidence interval around the mean practice effects and delineated the boundaries of test-retest change. Those scores which fell outside the 90% confidence intervals (i.e., < or > 5%, or 2.5%, on either end of the change distribution) were considered an uncommon

occurrence within an unoperated TLE sample. The adjusted confidence intervals – or cut-off scores – are listed in Tables 9 to 12, for both the SE_{diff} and the SE_{pred}, respectively.

Table 10
RCI data for the WMS-III based on the SE_{diff}

Variable	Mean change	Practice correction	SE_{diff}	90% cut-offs ±	Adjusted 90% cut-offs	95% cut-off s ±	Adjusted 95% cut-offs
WMS-III:							
AI	0.76	1	10.03	16.45	(-16)(+18)	19.66	(-19)(+21)
VI	1.95	2	12.79	20.98	(-19) (+23)	25.07	(-24)(+28)
IM	1.72	2	10.14	16.63	(-15)(+19)	19.87	(-18)(+22)
AD	0.95	1	11.49	18.84	(-18)(+20)	22.52	(-22)(+24)
VD	0.1	1	11.51	18.88	(-18)(+20)	22.56	(-22)(+24)
GM	1.08	2	10.65	17.47	(-16)(+20)	20.87	(-19)(+23)
WM	1.65	2	12.11	19.86	(-18)(+22)	23.74	(-22)(+26)
LM I	-0.67	0	2.65	4.35	±5	5.20	±6
LM II	-0.8	0	2.46	4.03	±5	4.82	±5
Faces I	0.72	1	2.76	4.52	(-4)(+6)	5.41	(-5)(+7)
Faces II	0.45	1	2.66	4.37	(-4)(+6)	5.22	(-5)(+7)
VPA I	0.79	1	2.19	3.59	(-3)(+5)	4.29	(-5)(+7)
VPA II recall	-0.03	0	2.67	4.38	(-4)(+6)	5.23	±6
FP I	-1.87	0	2.86	4.70	±5	5.61	±6
FP II	-2.47	0	3.13	5.14	±6	6.14	±7
Spatial Span	0.18	1	2.50	4.11	(-4)(+6)	4.91	(-4)(+6)

Note: 90% cut-off scores = SE_{diff} * ± 1.64
95% cut-off scores = SE_{diff} * ± 1.96

Importantly, RCIs based on Iverson's (2001) formula do not take into account regression to the mean – a phenomenon which is exacerbated by poor test stability (Basso et al., 1999; Bruggemans et al., 1997). This is a key consideration given the less-than-perfect stability of many of the measures and so an alternative formula – the standard error of prediction – was also calculated (Chelune, 2003).

Using the standard error of prediction.

As discussed in Chapter One, the SE_{pred} represents the standard error of a retest score predicted from a baseline score in a regression equation (Chelune, 2003). This method not only takes into account measurement error and practice effects but also considers the tendency of scores to regress towards the mean on re-test (Basso et al.,

1999; Bruggemans et al., 1997). The RCIs calculated using the SE_{pred} are presented in Tables 11 and 12.

Table 11
RCI data for the WAIS-III based on the SE_{pred}

Variable	Practice correction	SE_{pred}	90% cut-offs	Adjusted 90% cut-offs	95% cut-offs	Adjusted 95% cut-offs
WAIS-III:			±		±	
Verbal IQ	0	6.59	10.81	± 11	12.92	±13
Performance IQ	3	7.39	12.12	(-10) (+16)	14.48	(-12)(+18)
Full Scale IQ	1	6.08	9.98	(-9) (+11)	11.92	(-11)(+13)
VCI	0	5.56	9.12	± 10	10.90	±11
POI	4	9.19	15.07	(-12) (+20)	18.01	(-15)(+23)
WMI	0	10.50	17.22	± 18	20.58	±21
PSI	4	6.54	10.72	(-7)(+15)	12.81	(-9)(+17)
Picture completion	1	2.79	4.57	(-4)(+8)	5.47	(-4)(+8)
Vocabulary	0	1.80	2.96	±3	3.54	±4
Coding	1	1.38	2.26	(-2)(+4)	2.71	(-2)(+4)
Similarities	0	1.79	2.94	±3	3.51	±4
Block design	3	2.18	3.58	(-3)(+5)	4.28	(-4)(+6)
Arithmetic	0	2.00	3.28	±4	3.92	±4
Matrix reasoning	0	2.75	4.50	±5	5.38	±6
Digit span	0	2.20	3.61	±4	4.31	±5
Information	1	0.88	1.44	(-1)(+3)	1.72	(-1)(+3)
Picture arrangement	0	2.15	3.52	±4	4.21	±5
Comprehension	1	1.58	2.58	(-2)(+4)	3.09	(-3)(+5)
Symbol Search	2	2.35	3.85	(-2)(+6)	4.60	(-3)(+7)
LNS	0	2.90	4.76	±5	5.69	±6

Table 12
RCI data for the WMS-III and BNT based on the SE_{pred}

Variable	Practice correction	SE_{pred}	90% cut-offs	Adjusted 90% cut-offs	95% cut-offs	Adjusted 95% cut-offs
WMS-III:			±		±	
Auditory Immediate	1	9.58	15.71	(-15)(+17)	18.78	(-18)(+20)
Visual Immediate	2	11.36	18.63	(-17)(+21)	22.27	(-21)(+25)
Immediate Memory	2	9.69	15.89	(-14)(+18)	18.99	(-17)(+21)
Auditory Delayed	1	10.63	17.43	(-16)(+20)	20.83	(-20)(+22)
Visual Delayed	1	10.58	17.34	(-16)(+20)	20.73	(-20)(+22)
General Memory	2	9.85	16.15	(-15)(+19)	19.31	(-18)(+22)
Working Memory	2	12.26	20.10	(-19)(+23)	24.02	(-23)(+27)
Logical Memory I	0	2.51	4.12	±5	4.93	±5
Logical Memory II	0	2.44	4.00	±4	4.78	±5
Faces I	1	2.54	4.16	(-4)(+6)	4.97	(-4)(+6)
Faces II	1	2.28	3.73	(-3)(+5)	4.46	(-4)(+6)
VPA I	1	2.06	3.38	(-3)(+5)	4.04	(-5)(+7)
VPA II recall	0	2.49	4.08	(-4)(+6)	4.88	±5
Family Pictures I	0	2.48	4.07	±5	4.86	±5
Family Pictures II	0	2.72	4.46	±5	5.33	±6
Spatial Span	1	2.31	3.78	(-3)(+5)	4.52	(-4)(+6)
Boston Naming Test	0	6.77	11.10	±12	13.27	±14

The individual patient data was then classified according to both the 90% and 95% cut-offs generated by these different methods in order to ascertain which methodology provided the closest approximation to the expected theoretical distribution (e.g., 5%, 90%, 5%). The percentage of people in the pre-operative group experiencing a change in retest score outside the 90% and 95% confidence intervals generated by the two different RCI methods – based on the SE_{diff} and the SE_{pred} – are presented in Tables 13 and 14.

Table 13
Percentages of pre-operative group experiencing change on retest on WAIS-III indices

Index	N	SE_{diff} - adjusted method				SE_{pred} - adjusted method			
		Prediction interval	**% declined**	**% unchanged**	**% improved**	**Prediction interval**	**% declined**	**% unchanged**	**% Improved**
Hypothesised distributions		***90%*** ***95%***	***5*** ***2.5***	***90*** ***95***	***5*** ***2.5***	***90%*** ***95%***	***5*** ***2.5***	***90*** ***95***	***5*** ***2.5***
VIQ	35								
90%		**±11**	5.7	91.4	2.9	**±11**	5.7	91.4	2.9
95%		**±13**	5.7	94.3	0	**±13**	2.8	94.3	2.8
PIQ	35								
90%		**-9, +15**	0	94.3	5.7	**-10, +16**	2.9	91.4	5.7
95%		**-12, +18**	0	97.1	2.9	**-12, +18**	0	97.1	2.9
FSIQ	35								
90%		**-9, +11**	8.6	82.9	8.6	**-9, +11**	8.6	82.9	8.6
95%		**-11, +13**	5.7	88.6	5.7	**-11, +13**	5.7	88.6	5.7
VCI	34								
90%		**±10**	8.8	91.2	0	**±10**	8.8	91.2	0
95%		**±12**	8.8	91.2	0	**±11**	8.8	91.2	0
POI	34								
90%		**-11, +19**	5.9	94.1	0	**-12, +20**	5.9	94.1	0
95%		**-15, +23**	5.9	94.1	0	**-15, +23**	5.9	94.1	0
WMI	33								
90%		**±18**	6.1	90.9	3	**±18**	6.1	90.9	3
95%		**±21**	6.1	90.9	3	**±21**	6.1	90.9	3
PSI	35								
90%		**-8, +16**	8.6	85.7	5.7	**-7, +15**	11.4	82.9	5.7
95%		**-12, +18**	5.7	88.6	5.7	**-9, +17**	8.6	85.7	5.7
BNT	34								
90%		**±12**	8.8	91.2	0	**±12**	8.8	91.2	0
95%		**±14**	5.9	94.1	0	**±14**	5.9	94.1	0

Using the 90% confidence intervals, the two RCI methods resulted in highly similar classification rates, with both largely conforming to hypothesised classification rates (i.e., 5%, 90%, 5%). Similarly, the 95% confidence intervals for both the SE_{diff} and SE_{pred} RCI approaches classified similar percentages of people as declined, unchanged and improved on retest. These results suggest that RCIs derived using both methods are comparable in predicting retest scores on the WAIS-III indices and BNT.

Table 14
Percentages of pre-operative group experiencing change on retest on WMS-III indices

Index	N	RCI - SE_{diff} - adjusted method				RCI - SE_{pred} - adjusted method			
		Prediction interval	**% declined**	**% unchanged**	**% improved**	**Prediction interval**	**% declined**	**% unchanged**	**% improved**
Hypothesised		***90%***	***5***	***90***	***5***	***90%***	***5***	***90***	***5***
distributions		***95%***	***2.5***	***95***	***2.5***	***95%***	***2.5***	***95***	***2.5***
AI	38								
90%		**-16, +18**	2.6	89.5	7.9	**-15, +17**	2.6	89.5	7.9
95%		**±22**	2.6	92.1	5.3	**-18, +20**	2.6	92.1	5.3
VI	38								
90%		**-19, +23**	7.9	89.5	2.6	**-17, +21**	7.9	92.1	0
95%		**-24, +28**	2.6	97.4	0	**-21, +25**	2.6	97.4	0
IM	39								
90%		**-15, +19**	5.1	84.6	10.3	**-14, +18**	5.1	87.2	7.7
95%		**-21, +23**	5.1	92.3	2.6	**-17, +21**	5.1	89.8	5.1
AD	39								
90%		**-18, +20**	5.1	87.2	7.7	**-16, +20**	0	92.3	7.7
95%		**±23**	0	92.3	7.7	**-20, +22**	0	92.3	7.7
VD	39								
90%		**-18, +20**	10.3	82	7.7	**-16, +20**	5.1	87.2	7.7
95%		**±23**	2.6	92.3	5.1	**-20, +22**	5.1	89.8	5.1
GM	39								
90%		**-16, +20**	5.1	87.2	7.7	**-15, +19**	5.1	87.2	7.7
95%		**±20**	0	94.9	7.7	**-18, +22**	0	92.3	7.7
WM	38								
90%		**-18, +22**	7.9	89.5	2.6	**-19, +23**	7.9	86.8	5.3
95%		**-23, +25**	7.9	92.1	0	**-23, +27**	5.3	94.7	0

As with the WAIS-III indices, the two RCI methods provided similar classifications of change on retest for the WMS-III indices. The only notable differences in classification rates were for the 90% confidence intervals on the Auditory Delayed and Visual Delayed memory indices. The SE_{pred} method did not identify any significant decline on retest in the pre-operative group on the AD index, whereas the SE_{diff} approach identified 5.1% of the group as demonstrating significant decline. On the VD memory index, the SE_{diff} again identified significant change (10.3%) beyond that classified by the SE_{pred} approach. However, overall, the classification rates

achieved by the SE_{pred} and SE_{diff} methods were comparable, using 95% confidence intervals.

Calculation of SRB Change Norms

Regression equations to predict retest scores were generated for each WAIS-III and WMS-III index and subtest using a stepwise regression method. Relevant predictor variables were identified from the literature – age at first testing occasion, level of education, duration of seizures, test-retest interval and baseline performance - and included in the regression equations. A regression equation was calculated for the BNT; however the data for this test was so positively skewed as to limit the clinical applicability of SRB change norms and it was therefore not reported. The multiple *R*, standard error of the estimate, constant and unstandardised beta weights (regression coefficients) of the regression equations are summarised in Table 15.

Table 15
Regression equations for each index of the WAIS-III and the WMS-III

Index	R	SE est.	Constant	*F*	Sig. of model	β baseline	β age	β edu	β duration	β retest interval	β AEDs
WAIS-III											
VIQ	0.923	6.55	1.66	191.25	<0.0005	**0.97***					
PIQ	0.914	7.33	1.97	168.19	<0.0005	**1***					
FSIQ	0.943	6.05	-1.97	262.65	<0.0005	**1.03***					
VCI	0.959	5.18	10.46	179.55	<0.0005	**0.92***				**-0.08#**	
POI	0.845	9.48	8.85	79.64	<0.0005	**0.94***					**-4.52#**
WMI	0.802	10.15	6.74	27.08	<0.0005	**0.8***	**0.28#**				
PSI	0.91	5.8	22.15	49.94	<0.0005	**0.73***	**0.26#**		**-0.23***		
WMS-III											
AI	0.905	8.63	11.22	44.41	<0.0005	**0.79#**	**0.38***		**-0.27#**		
VI	0.753	11.45	28.94	44.46	<0.0005	**0.68***					
IM	0.877	9.95	14.3	105.53	<0.0005	**0.85***					
AD	0.867	10.01	14.77	53.51	<0.0005	**0.71***	**0.31***				
VD	0.778	10.77	22.63	56.57	<0.0005	**0.74***					
GM	0.861	9.97	18.55	122.94	<0.0005	**0.8***					
WM	0.694	11.56	9.19	16.24	<0.0005	**0.78***	**0.33#**				

* $p < 0.05$; # $p < 0.01$

The sign (i.e., positive or negative) of the regression coefficient indicates the direction of the relationship between the variables. For example, if the regression coefficient is positive, then the relationship of that variable with the dependent variable will also be positive and the opposite relationship holds true for a negative regression coefficient.

Baseline test performance was a significant predictor of retest performance across all the WAIS-III indices – Verbal IQ ($F_{(1,33)} = 191.25$, $p < 0.0005$), Performance IQ ($F_{(1,33)} = 168.2$, $p < 0.0005$), Full Scale IQ ($F_{(1,33)} = 262.65$, $p < 0.0005$), Verbal Comprehension Index ($F_{(2,31)} = 179.55$, $p < 0.0005$), Perceptual Organisation Index ($F_{(1,32)} = 79.64$, $p < 0.0005$), Processing Speed Index ($F_{(3,31)} = 49.94$, $p < 0.0005$), and Working Memory Index ($F_{(2,30)} = 27.08$, $p < 0.0005$). Retest performance was also significantly predicted by baseline performance for all the WMS-III indices – Auditory Immediate ($F_{(3,34)} = 44.4$, $p < 0.0005$), Visual Immediate ($F_{(1,36)} = 44.46$, $p < 0.0005$), Immediate Memory ($F_{(1,37)} = 105.53$, $p < 0.0005$), Auditory Delayed ($F_{(2,36)} = 53.51$, $p < 0.0005$), Visual Delayed ($F_{(1,37)} = 56.57$, $p < 0.0005$), General Memory ($F_{(1,37)} = 122.94$, $p < 0.0005$) and Working Memory ($F_{(1,36)} = 22.64$, $p < 0.0005$).

Age was a significant predictor variable for the Working Memory ($F_{(2,31)} = 51.92$, $p = 0.045$) and Processing Speed Indices ($F_{(2,31)} = 51.92$, $p = 0.002$) of the WAIS-III, as well as for the Auditory Immediate ($F_{(2,31)} = 51.92$, $p = 0.002$), Auditory Delayed ($F_{(2,31)} = 51.92$, $p = 0.013$) and Working Memory Indices ($F_{(2,31)} = 51.92$, $p = 0.016$) of the WMS-III. Seizure duration significantly enhanced the prediction equation for the Processing Speed Index ($F_{(2,31)} = 51.92$, $p = 0.009$) and the Auditory Immediate Index ($F_{(2,31)} = 51.92$, $p = 0.029$) of the WMS-III, while the length of the test-retest interval was significant for only the Verbal Comprehension Index ($F_{(2,31)} = 51.92$, $p = 0.023$).

The regression equations were run a second time with an additional predictor variable – number of anti-epileptic medications – to investigate the impact of polypharmacy on cognitive functioning. The AED predictor variable was a significant negative predictor of retest performance for only the Perceptual Organisation Index of the WAIS-III ($F_{(2,31)} = 51.92$, $p < 0.0005$), indicating that the lower the number of AEDs a person is taking, the better their performance on this measure.

Using the unstandardised beta weights of the significant predictor variables, the predicted retest scores of the unoperated epilepsy group were calculated using

McSweeny and colleagues' (1993) equation ($Y_p = \beta * X_o + C$). As an example, the regression equation used to predict a person's retest score on the Auditory Delayed Index would be: Y_p = (baseline score * 0.71) + (age * 0.31) + constant (14.77). This provided a predicted retest score for that person on that measure, based on the identified significant predictor variables from the unoperated TLE control group (See page 11 for a more detailed discussion).

To determine how well the regression equations predicted the retest scores of the pre-operative TLE group, they were applied to individual data to determine whether there had been decline, improvement or no change on retesting. The percentage of people in the pre-operative group experiencing a change in retest score outside the 90% and 95% confidence intervals are listed in Table 16.

Table 16
Classifications of retest change – application of SRBs to entire pre-operative sample

		90% CI – pre-operative sample				95% CI – pre-operative sample			
Index	**N**	**Prediction Interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**
Hypothesised		***90%***	***5***	***90***	***5***	***95%***	***2.5***	***95***	***2.5***
VIQ	35	±10.74	5.7	88.6	5.7	±12.84	5.7	91.4	2.9
PIQ	35	±12.02	0	94.3	5.7	±14.37	0	94.3	5.7
FSIQ	35	±9.92	5.7	88.6	5.7	±11.86	2.9	91.4	5.7
VCI	34	±8.5	8.8	82.4	8.8	±10.15	8.8	91.2	0
POI	34	±15.55	5.9	94.1	0	±18.58	2.9	97.1	0
WMI	33	±16.65	6.1	90.9	3	±19.89	6.1	90.9	3
PSI	35	±9.51	5.7	94.3	0	±11.37	0	97.1	2.9
AI	38	±14.15	2.6	89.5	7.9	±16.91	0	94.7	5.3
VI	38	±18.78	2.6	92.1	5.3	±22.44	2.6	97.4	0
IM	39	±16.32	5.1	92.3	2.6	±19.5	5.1	94.9	0
AD	39	±16.42	5.1	87.2	7.7	±19.62	0	94.9	5.1
VD	39	±17.66	2.6	89.7	7.7	±21.11	0	94.9	5.1
GM	39	±16.35	7.7	89.7	2.6	±19.54	2.55	94.9	2.55
WM	38	±18.96	5.3	94.7	0	±22.66	5.3	94.7	0

The regression equations were adequate in classifying the retest scores of the pre-operative sample, according to the chosen confidence intervals. However, for both the RCIs and SRBs, it is important to note that the expected distribution (e.g., 2.5% decline, 95% unchanged and 2.5% improvement) may not have been clearly

delineated due to limited sample sizes in the current investigation, and this would be expected to ameliorate as sample size increased.

Right versus Left Differences in the Classification of Change on Retest

Finally, to investigate any differences in retest performance according to laterality of seizure focus, the pre-operative group was divided into right- and left-sided TLE (the four participants with bilateral TLE focus and six with unclear focus were excluded) and between-subjects multivariate ANOVAs were run using both RCI-generated and SRB-based change classifications (declined, no change, improved). Using the RCI methods (both SE_{diff} and SE_{pred}), the change scores of the right- and left-TLE pre-operative groups were not significantly different on the majority WAIS-III and WMS-III indices, with the exception of the Processing Speed Index (SE_{diff}, $F_{(3,25)} = 2.996$, $p = 0.05$ and SE_{pred}, $F_{(3,25)} = 2.996$, $p = 0.05$). Similarly, the SRB approach revealed no significant differences between the right- and left pre-operative TLE groups on any of the WAIS-III or WMS-III indices.

Cognition over Time

The test-retest interval in the current study was markedly longer than many other similar studies and so the pre-operative retest data was also analysed to examine the longstanding question of whether chronic seizures over time impact negatively upon cognition. Repeated measures analyses of variance were generated, with age and length of the retest interval as covariates. The results of these analyses are listed in Table 17.

Table 17
Effect of retest interval and age on retest scores, for the WAIS-III and WMS-III indices

Index	N	Test		Retest		Age	Retest interval	F	Partial eta^2
		Mean	SD	Mean	SD	*p*	*p*		
WAIS-III:									
VIQ	37	91.86	15.72	90.50	16.88	-	-		
PIQ	37	93.78	15.82	96.11	17.97	-	-		
FSIQ	37	92.05	16.06	92.19	17.98	-	-		
VCI	37	93.54	17.59	91.91	18.16	-	-		
POI	37	96.92	15.39	100.23	17.78	-	-		
PSI	36	92.19	14.52	90.49	16.53	-	-		
WMI	37	88.65	15.04	91.22	13.24	-	-		
WMS-III:									
AI	39	89.13	19.85	88.18	19.77	-	0.039	4.613	0.116
VI	39	85.08	18.71	85.90	17.39	-	-		
IM	39	84.51	21.02	84.80	20.90	-	-		
AD	39	87.67	21.01	87.05	20.36	-	-		
VD	39	86.21	17.79	85.34	17.03	-	-		
GM	39	88.97	19.30	89.88	18.56	-	-		
WM	39	85.05	20.92	84.61	20.03	0.023	-	5.642	0.139
BNT	34	47.14	11.15	45.59	11.28	-	0.038	4.705	0.132

Note: - indicates non-significant result

There was a small but significant effect of the length of the retest interval on retest performance for both the BNT ($F_{(1,31)} = 4.705$, p = 0.038; Wilk's Lambda = 0.868; partial eta square = 0.132), with the results suggesting an association between length of the retest interval and poorer confrontational naming performance on retest. Unexpectedly, a longer retest interval was significantly associated with improved performance on retest for the AI memory index ($F_{(1,35)} = 4.613$, p = 0.039; Wilk's Lambda = 0.884; partial eta square = 0.116). However, although statistically significant, the proportion of the variance accounted for by retest interval was moderate – 13.2% for the BNT and 11.6% for the AI index. There was also a small but significant effect of age on the WM index ($F_{(1,35)} = 5.642$, p = 0.023; Wilk's Lambda = 0.861; partial eta square = 0.139), with age associated with improved retest performance, although again this variable only accounted for 13.9% of the overall variance.

Discussion

Consistent with recommendations from the epilepsy literature, the present research aimed to generate RCI and SRB change norms based on an Australian TLE sample, to assist with determining cognitive change following surgery for this population (Chelune et al., 1993; Hermann et al., 1996; Martin et al., 1998; Martin et al., 2002b; McSweeny, 1993; Sawrie et al., 1996).

Test Stability and the Effects of Practice

It was hypothesised that the stability coefficients generated from the local, unoperated TLE sample would be comparable to those reported in previous TLE research (Martin et al., 2002b; Sawrie et al., 1996). This hypothesis was upheld with the WAIS-III indices tending to have the highest test stability, followed by the WMS-III indices and, finally, the individual subtests (Martin et al., 2002b).

It was hypothesised that practice effects would have a significant impact on the retest scores of the non-surgical group in the current study (Chelune et al., 1993; Martin et al., 2002b; Sawrie et al., 1996). This hypothesis was partially supported with modest but statistically significant improvements evident on the Performance IQ (PIQ), Perceptual Organisation Index (POI) and Processing Speed Index (PSI) of the WAIS-III. Evidence of significant practice effects on retest has been reported in previous research – particularly for visually-based tasks on both the WAIS-III and the WMS-III (Dikmen et al., 1999; Hermann et al., 1996; Kaufman, 2003; Martin et al., 2002b; Rapport et al., 1997; Sawrie et al., 1996). Specifically, the scoring structure of several PIQ subtests scales rewards bonus points for swift responses, thereby potentially inflating the overall PIQ (Kaufman, 2003; Rapport et al., 1997). Consistent with this, practice effects were seen in the current study on the Picture Completion, Block Design and Symbol Search subtests.

Despite significantly longer retest intervals (mean = 50.44 months; SD = 32.6) than the one month retest interval reported in the WAIS-III manuals, the mean increase in re-test score on the PIQ in the current study was 2.8 points. Interestingly, although longer retest intervals were employed, the magnitude of practice effects in the current study were comparable to those reported by Martin and colleagues (2002; mean retest interval = 7 months).

Reliable Change Indices

The current study aimed to establish base-rates for cognitive change on retest – in an unoperated TLE sample – using two different reliable change methodologies, based on the standard error of the difference (Chelune et al., 1993) and the standard error of prediction (Iverson et al., 2001). Based on the literature, it was hypothesised that the SE_{pred} would more closely reflect the theoretical distribution of change scores (e.g., 5% declined 90% unchanged, 5% improved) on the WMS-III, but that there would no difference in the classifications of change between SE_{pred} and SE_{diff} on the WAIS-III (Atkinson, 1991; Chelune, 2003). The first hypothesis was partially supported, with the SE_{pred} resulting in a distribution of change scores which was slightly closer to the expected theoretical distribution for three of the WMS-III indices (i.e., IM, AD & VD), although the SE_{diff} more accurately classified change on two of the indices (VI and WM). The second part of the hypothesis was upheld, with the SE_{pred} and SE_{diff} RCIs producing either identical, or highly similar, classification rates for the majority of the WAIS-III indices.

Overall, the results indicated no marked differences in the classification rates of the SE_{diff} and SE_{pred} formulas, which suggests regression to the mean was not a marked influence in the current TLE sample. Nonetheless, the SE_{pred} approach generated change classifications which were marginally closer to the chosen theoretical distribution for change (i.e., 5%, 90%, 5%) across the majority of the WMS-III indices, whilst the SE_{diff} method only held this advantage for two WAIS-III indices. As a result, recommendations to use the SE_{pred} for determining change in cognitive retest scores are supported by the results of the current study (Atkinson, 1991; Chelune, 2003).

Standardised Regression-Based Change Norms

SRB change norms were also generated using data from the local, unoperated TLE sample. The factors chosen to be included in the regression analysis were informed by previous research and included age, education, duration of seizures, length of the test-retest interval and baseline performance (Helmstaedter et al., 2003; Lange, Chelune, Taylor, Woodward & Heaton, 2006; McSweeny et al., 1993; Sawrie et al., 1996; Sherman et al., 2003; Martin et al., 2002b). In turn, these regression equations were applied to the neuropsychological data of patients to predict retest scores which were then compared to the person's observed score to determine the direction and magnitude of change (Sawrie et al. 1996).

It was hypothesised that baseline test performance would be a significant predictor of a person's score on retest and this was supported by the current results. Consistent with previous studies, baseline performance was by far the strongest predictor of retest performance across all indices of the WAIS-III and WMS-III (Dikmen et al., 1999; Heaton et al., 2001; Hermann et al., 1996; McSweeny et al., 1993; Martin et al., 2002b; Sawrie et al., 1996; Sherman et al., 2003; Temkin et al., 1999).

It was anticipated that age would also represent a significant predictor variable in the regression equation and this hypothesis was partially supported. Including a person's age at the first assessment significantly improved the predictive value of the regression equations for the WMI and PSI of the WAIS-III, as well as the AI, AD and WM indices of the WMS-III. These results suggest that the older a person is at the time of their first assessment – and presumably consideration for surgery – the more likely they are to experience a significant increase on these indices at retest. This finding is comparable to some previous investigations with unoperated TLE samples (Helmstaedter et al, 2003; Martin et al., 2002b); however, age was not identified as a significant predictor variable in other TLE studies (Hermann et al., 1996; McSweeny et al., 1993; Thompson & Duncan, 2005; Sawrie et al., 1996).

The number of years a person had experienced seizures was expected to significantly contribute to predicting retest scores and this hypothesis was again partially supported. A person's seizure duration was a significant negative regression coefficient for both the Processing Speed Index and the Auditory Immediate Index. This finding suggests a shorter duration of lifetime seizures is significantly associated with a stronger performance on speed and auditory memory measures. These results are consistent the findings of some previous investigations (Jokeit & Ebner, 1999; Kaaden & Helmstaedter, 2009; Marques et al., 2007; Pitkanen & Sutula, 2002), but not others (Martin et al., 2002b; McSweeny et al., 1993; Sawrie et al., 1996; Selwa et al., 1994; Strauss et al., 1995). It is important to observe that studies are often hampered by methodological restrictions – limited time ranges, unrecognised cohort biases, confounded variables, inconsistent reports of seizure onset between patients and others – which may serve to cloak the effects of seizure duration on cognition (Jokeit & Ebner, 1999).

The contribution of a person's level of education to their performance on retest remains equivocal in the literature, with studies reporting both significant (Hermann et al., 1996; Martin et al., 2002b; Temkin et al., 1999) and non-significant results (McSweeny et al., 1996; Pai & Tsai, 2005; Sawrie et al., 1996), depending on the intelligence or memory index investigated. In the current study, it was hypothesised that education would contribute significantly to the regression equation; however this hypothesis was not supported and education did not significantly predict retest performance on any of the WAIS-III or WMS-III indices. This result conflicts with several previous studies, which have reported associations between a person's level of education and retest performance for both the VIQ and PIQ of the WAIS (Temkin et al., 1999), the FSIQ of the WAIS-R (Hermann et al., 1996) and the Auditory Immediate and Delayed indices of the WMS-III (Martin et al., 2002b). The current TLE sample was reasonably well-educated (mean education = 12.3 years) and might therefore be expected to show significantly higher performances on retest due to the effects of practice. However, it is possible the extended retest interval negated this potential advantage.

The length of time between assessments was not expected to be a significant predictor of retest scores for the current sample, due to the extended period retest intervals. However, this hypothesis was only partially supported. Unexpectedly, the length of the retest interval was a significant predictor variable for the VCI, indicating a shorter retest interval was associated with higher retest scores on this measure. This result finds some potential support in the literature, with Temkin and colleagues (1999) reporting the retest interval as a mild but significant predictor of retest performance on the VIQ and PIQ of the original WAIS. However, the length of the retest interval was not a significant predictor of retest performance in several other investigations (Chelune et al., 1999; Hermann et al., 1996; McSweeny et al., 1993; Sawrie et al., 1996).

It was hypothesised that the number of anti-epileptic drugs a patient was taking would have an adverse effect on cognition. This hypothesis was again partially supported with a lower number of prescribed AEDs associated with a significantly higher score on the Perceptual Organisation index only. This indicates the more AEDs a person is prescribed for TLE, the more likely they are to adversely affect POI performance. The finding partially replicates the work of Sherman and colleagues (2003) with a paediatric epilepsy sample. They reported that for each AED taken at baseline, retest IQ performance was expected to drop by approximately 4.5 points. However, a number of previous studies have not reported cognitive changes due to polytherapy (Hermann et al., 2006; Piazzini et al., 2006).

In the current investigation, 27 of the 41 participants were prescribed at least one ‘older’ AED (i.e., phenobarbitone, phenytoin, carbamazepine or valproate) as part of their treatment regime and four participants were taking topiramate at the time of retest. These patients may have performed more slowly due to the influence of medications on processing speed, therefore not gaining bonus points on subtests of the POI and subsequently achieving lower scores. Exclusion of those participants on either topiramate (n=4) or phenobarbital (n=1) was considered due to their potential cognitive side-effects (Aldenkamp, 2001); however, pre-surgical patients with TLE are prescribed these AEDs on occasion and their inclusion was considered to be representative of the wider pre-surgical TLE population.

Are Seizures Related to Adverse Cognitive Changes?

The question of whether seizures contribute to progressive cognitive decline finds varying responses in the literature and it is clear that the answer is more complex than a simple 'yes' or 'no'. The mean retest interval for the current study was approximately four years (50 months) - a three to four year period has been suggested as sufficient in determining significant cognitive change over time (Seidenberg et al., 2007) - and therefore provided the opportunity to contribute to this body of literature. It was anticipated that there would be some evidence of cognitive decline in the current sample on retest, particularly in the areas of memory, naming, attention and speed of processing (Helmsteadter et al., 2003; Hermann et al., 2006; Holmes et al., 1998; Marques et al., 2007; Oyegbile et al., 2004; Seidenberg et al., 2007; Seidenberg, O'Leary, Giordani, Berent, & Boll, 1981). This hypothesis was only partially supported. Consistent with previous work, a significant interaction between the length of the retest interval and lower retest performance was evident for confrontational naming (Thompson & Duncan, 2005). However, there were no other significant associations between retest interval and lower retest performance for any of the WAIS-III or WMS-III indices. Several previous studies have also reported no ongoing cognitive deterioration in people with unoperated TLE over extended timeframes (Rausch et al., 2003; Selwa et al., 1994). An unexpected finding in the current study was that the effects of retest interval and age were associated with mild, but significant, improvements at retest on the AI and WM indices. This counter-intuitive finding is perhaps due to the utilisation of age-corrected indices and future investigations may instead consider utilising raw scores, as using age-corrected indices risks overlooking seizure-related cognitive changes over time.

Memory has frequently been reported as the cognitive domain most vulnerable to progressive adverse changes (Andersson-Roswall et al., 2004; Helmstaedter et al., 2003; Vingerhoets, 2006), although this finding was not replicated in the current study. In a direct comparison of cognition in medically-treated and surgically-treated

TLE samples, Helmstaedter and colleagues (2003) found evidence of progressive memory impairment in both groups, for those patients whose seizures remained uncontrolled. The latter caveat may go some way toward explaining the absence of significant cognitive deterioration evident in the current pre-operative sample; the majority of these participants experienced at least some reduction in the frequency of their seizures with AED treatment, although almost half the sample (15 adults) continued to experience seizures at least weekly.

The varying results reported in the literature regarding the impact of seizures on cognition may, in part, be attributable to the differing methodologies employed (e.g., length of the retest interval, neuropsychological measures, samples investigated). Moreover, the interpretation of these is complicated by the heterogeneity of the epilepsy population, exemplified by differing aetiologies, varying seizure duration, frequency and severity, numerous antiepileptic medications, as well as social and mood factors (Piazzini et al., 2006). Nonetheless, the possibility of ongoing for cognitive decline where seizures are uncontrolled has led several studies to recommend earlier surgical intervention for TLE, wherever possible (Baxendale et al., 2008; Helmstaedter et al., 2002; Seidenberg et al., 1998; Thompson & Duncan, 2005; Wiebe et al., 2001)

Interestingly, while 56% of the patient sample reported concerns of a decline in memory abilities since their first assessment, the results outlined above indicate the percentage of patients who actually experienced such a decline was markedly lower than this, at no time exceeding 10.3% of the sample. Subjective reports of memory impairment may reflect the impact of depression on cognition (Banos, LaGory, Sawrie, Faught, Knowlton, et al., 2004; Brand, Jolles, & Gispen-de Wied, 1992; Marino, Meador, Loring, Okun, Fernandez, et al., 2009; Paradiso, Hermann, Blumer, Davies, & Robinson, 2001), rather than genuine cognitive change, and future research should consider including mood state as a potential predictor variable.

Future Directions

The extended retest intervals employed in the current study represented both a strength and a limitation. Determining change norms based on such long retest intervals (i.e., approximately four years) risks underestimating practice effects in post-operative patients who are typically re-assessed over much shorter timeframes. Conversely, the longer retest intervals in the current study allowed the investigation of the course of cognition in unoperated TLE patients; an analysis which would not have been possible over shorter periods. Evidence suggests that verbal memory decline may continue to decline for up to two years following left ATL (Alpherts, et al., 2006) and standard prolonged neuropsychological review is therefore recommended (Baxendale, 2008; Binnie & Polkey, 2000). The long retest intervals in the current study may therefore provide a good reflection of cognitive change over similar timeframes (i.e., two or more years) in a pre-operative control sample. However, ideally future studies should consider employing the use of a control group over extended timeframes (e.g., greater than 10 years) to investigate cognitive deterioration over time in unoperated TLE patients (Piazzini et al., 2006; Vingerhoets, 2006).

Finally, analysing the impact of right versus left TLE on retest performance was limited by the small sample size for right TLE patients. Although no differences on retest performance were found between the right and left TLE groups, future studies would benefit from using larger sample sizes to replicate this finding. Further to this, future research may consider generating RCI and SRB change norms for right and left TLE samples in order to more closely evaluate any pre-operative differences between the groups.

CHAPTER 4

Determining Cognitive Change in a Local Post-Operative Sample

In the part of the current study, reported in Chapter Three, an unoperated TLE sample was assessed twice pre-operatively using the WAIS-III and WMS-III. The retest data was used to generate RCIs and SRB norms against which to evaluate post-operative cognitive change.

The current chapter reports on the application of these RCIs and SRBs to a sample of post-operative TLE patients, also from the St Vincent's Hospital epilepsy program, in order to assess their utility for clinical practice. To determine whether change norms are generalisable across TLE populations, previously-published North American RCI and SRB data was also applied to the local post-operative sample (Martin et al., 2002b). The application of local and non-local RCI and SRB change data to the same post-operative sample was intended to determine whether they resulted in different classification rates of post-operative change.

It was hypothesised that the local change norms would better classify the retest data than the US-derived norms. Further, it was expected that the left and right ATL patients would differ on post-operative test scores as a function of laterality following surgical intervention. In particular, it was hypothesised that material-specific patterns of cognitive change – left-sided surgery associated with verbal memory decline and right-sided surgery associated with visual memory decline – would be evident in the post-operative Australian sample. Finally, a decline in confrontational naming was expected for those patients undergoing left ATLs.

Method

Participants

The neuropsychological data of 79 female and 72 male patients who had undergone surgical intervention for TLE over the past 11 years was retrieved from the St Vincent's Hospital Neuropsychology Unit's archival database. Of these, 75 had been assessed pre- and post-operatively using the WAIS-R and WMS-R, while 76 had been assessed pre- and post-operatively using the WAIS-III and WMS-III. Given the current RCI and SRB change norms were derived from WAIS-III and WMS-III test scores only, those patients who had been assessed post-operatively using the WAIS-R and WMS-R were excluded from the current study. Access to archival data, without seeking consent, was approved for this study by the St. Vincent's Hospital and Victoria University Human Research and Ethics Committees.

Materials

The 76 participants had been administered the full neuropsychological assessment battery utilised by St Vincent's Hospital Neuropsychology Unit – described in detail in the previous chapter – as part of the routine neuropsychological work-up and follow-up for patients involved in the comprehensive epilepsy program.

Procedures

All neuropsychological test scores are routinely entered into an archival database at the Neuropsychology Unit at St Vincent's Hospital, along with demographic and diagnosis details. A search of the database revealed 76 adults with a diagnosis of TLE who had undergone both pre- and post-operative neuropsychological assessment using the WAIS-III and WMS-III.

The archival retest results of the post-operative patient group were classified using the local RCI- and SRB-based parameters of change derived in Chapter Three and those in the previously published North American data (Martin et al., 2002b). Classifications of change in the post-operative sample, for the Australian and North American analyses, were then compared to investigate whether the new, local change norms produced significantly different classifications of change.

Statistical Methods

As in the previous chapter, there were two components to the statistical analysis of the post-operative data. Firstly, the RCIs derived from the pre-surgical TLE sample using the SE_{diff} (Iverson, 2001) were applied to the retest data of the post-operative sample, according to the methodology described in detail in Chapter Three. In addition, the RCIs published by Martin et al. (2001), using the SE_{diff} method, were also applied to the post-surgical sample. For this reason, the SE_{diff} approach was chosen rather than the SE_{pred} method (Chelune, 2003) to allow for direct comparison of the resulting classifications of change.

The SE_{diff} from both Martin et al.'s (2001) study and the current investigation were multiplied by ±1.64 to establish 90% confidence intervals. 95% confidence intervals were also calculated (i.e., $SE_{diff} * \pm 1.96$) for comparative purposes. The sample of 76 post-operative patients was then classified – according to the reliable change cut-offs – as either significantly declined, unchanged, or significantly improved. To ascertain if the RCIs derived from the Australian and North American samples produced significantly different change classifications, Cohen's kappa and Kendall's tau analyses were conducted.

The kappa value provides a measure of the agreement between two data sources. Whilst there are no absolute cut-offs for interpreting kappa coefficients, some approximate guidelines have been provided. Fleiss (1981) suggested kappa values over 0.75 indicate strong agreement above the level of chance and Gardener (1995) suggested that a kappa coefficient of at least 0.7 must be obtained before calculating

further analyses. Landis and Koch (1977) have also provided a scale for interpreting kappa coefficients, listed in Table 18.

Table 18
Kappa interpretation scale (Landis & Koch, 1977)

Kappa value	Agreement between two sources
< 0.00	Poor
0.00 – 0.20	Slight
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Substantial
0.81 – 1.00	Almost perfect

Deciding which level of agreement is clinically significant is essentially dependent upon the clinical implications of making a Type I error. For example, say the current study accepted a moderate agreement between the classification rates of the Australian and North American change norms as being indicative of the generalisability of the North American data to non-local TLE samples. However, the accurate determination of cognitive change following TLE surgery relies on the minimisation of measurement error. Indeed, minimising the impact of measurement error on interpretation of cognitive change is the main aim of the current study. As a result, it was decided to use a highly conservative kappa coefficient level of greater than 0.80 to delineate the extent of agreement between the Australian and North American RCIs and SRB change norms.

The second phase of analysis applied the SRB equations generated using the demographic and seizure characteristics of the pre-operative sample – baseline test score, test-retest interval (months), age (years), education (years) and duration of seizures (years) – to the post-surgical TLE sample. This formula calculated a predicted retest score for each person. The SRB equations generated by Martin et al. (2001) were also applied to the post-surgical sample. The 90% and 95% confidence intervals were calculated, as before, to delineate the boundaries of clinically relevant cognitive change from pre- to post-surgical assessment. The retest scores were then transformed into standardised *z*-scores by subtracting the predicted retest score from

the observed retest score and dividing the result by the residual standard deviation to examine the magnitude of any observed changes (McSweeny et al., 1993; see Chapter One for formulae). Once again, the resulting classifications of change were evaluated using Kendall's tau to determine whether there was any significant difference between the change cut-offs derived from the two different TLE samples (i.e., Australian and North American).

A further analysis examined the effect of seizure laterality on post-operative cognitive outcomes. The post-operative sample was divided into those who had undergone right and left ATLs. The presence or absence of change on retest was determined for these groups using the RCI and SRB approaches, described above.

Number needed to harm and number needed to treat.

The number of people in the pre-operative sample showing decline on retest was used as the control event rate in order to calculate the number needed to harm (NNH) and number needed to treat (NNT) in the post-operative sample. A NNT was calculated for both the right- and left-sided surgery groups on each index of the WAIS-III and WMS-III, using the formulae outlined in Chapter Two (see page 33). As reported in Chapter Two, the left and right pre-operative groups did not differ in change over time, and so the same control event rate was assumed across the left and right post-operative groups.

Results

Demographic Characteristics and Group Comparisons

Seventy-six patients (mean age 33.93 years; SD 10.5 years) who had undergone surgical intervention for TLE were included in the analysis. There was an approximately equal gender division, with 36 males and 40 females included in the post-operative sample. With the exception of two, all had been educated to at least year eight. Sixty-four had completed secondary school, but only 10 studied at the tertiary level (mean education 11.06; SD 1.7). According to medical records, 39

patients had undergone left ATL and 37 had received right ATLs. Due to the retrospective nature of this data collection, data regarding age at seizure onset and the duration of seizures had to be gained from inspection of medical records. Within a single file, there were often more than one estimate of age at seizure onset by clinicians and the patients themselves, making it difficult to determine both age of seizure onset and duration of seizures. As a result, this predictor variable is potentially more unreliable than the other variables. Table 19 presents the participant demographics in full.

Table 19
Demographic and seizure information for the post-operative sample

Variable	Range	Mean	SD
(N=76)			
Age at first assessment	16-57	33.93	10.5
Education (years)	6-16	11.06	1.7
Full Scale IQ (N=73)	56-121	92.63	13.24
Retest interval (months)	6-62	15.56	10.07
Age at seizure onset (years)	n/a		
Duration of seizures (years)	n/a		
Gender (male/female)	36/40		
Handedness (right/left/ambidextrous)	67/7/2		
Left-sided surgery	39		
Right-sided surgery	37		
Pathology			
TLE with hippocampal sclerosis	33		
TLE no hippocampal sclerosis	26		
Other TL pathology (e.g., tumour)	13		
Extra-temporal seizure focus	4		

As detailed in Table 20, the demographic and clinical characteristics of the pre- and post-operative TLE groups were compared. Using the Shapiro-Wilk test of normality, the variables age and Full Scale IQ were found to be normally distributed for both groups; however, neither the test-retest intervals nor education levels of either group were normally distributed. Independent *t*-tests were used to evaluate group differences in the normally distributed variables and Mann-Whitney tests were used for the non-normally distributed variables. As the post-operative data was obtained from archival records, missing data was at times unavoidable. This resulted in some data sets with slightly fewer than 76 participants and these are listed in Table C1, Appendix C.

Patients who have surgery through the St Vincent's Comprehensive Epilepsy Program are routinely reviewed in the six to twelve months following surgical intervention. As expected, the length of the test-retest interval was significantly longer for the pre-operative group gathered for this study (U = 446, N_1 = 41, N_2 = 73, p = <0.0005, two-tailed). Education was also significantly different between groups (U = 1180.5, N_1 = 41, N_2 = 73, p = 0.028, two-tailed), with the pre-operative sample having been educated to a significantly higher level, although the difference in means was only one year. According to Levene's test of equality of variance, age violated the assumption of equal variance and so a Mann-Whitney test was run for this variable. According to the analysis, age did not differ significantly between the pre- and post-surgical groups (U = 1413, N_1 = 41, N_2 = 73, p = 0.407, two-tailed). Similarly, an independent *t*-test found no significant difference between the mean Full Scale IQs of the two groups (t = -0.2, df = 108, p = 0.841). The percentages for handedness and gender were also comparable for each group.

Table 20
Pre-operative and post-operative demographic and seizure information for the Australian sample

Group	**Pre-op (n=41)**		**Post-op (n=73)**		
	Mean	*SD*	*Mean*	*SD*	*p*
Age (years)	37.1	14.2	33.9	10.5	0.219
Retest interval (months)	50.4	32.6	15.6	10.1	<0.0005*
Education (years)	12.3	2.5	11.1	1.7	0.028*
FSIQ	92.1	16.1	92.6	13.2	0.841
	Percentage		*Percentage*		
Sex - male	58.5		47.4		
-female	41.5		52.6		
Handedness					
- right	87.8		88.2		
- left	9.8		9.2		
- ambidext.	2.4		2.6		

***Note:* *** indicates *p* values obtained using Mann-Whitney U test, all other results refer to independent *t*-tests

Due to the differences between the pre- and post-operative samples on the length of the retest interval and level of education, an analysis of covariance was completed to determine whether these differences had a significant effect on retest scores. There was a small but significant effect of length of the retest interval on the pre- and post-operative groups for both the WMI ($F_{(2,93)}$ = 3.265, p = 0.043) and the GM index ($F_{(2,103)}$ = 3.763, p = 0.026), with an overall reduction in the mean retest scores of the

WMI index but, unexpectedly, an increase in the mean retest scores on the GM index. There was no significant effect of either retest interval or education level for any of the other WAIS-III or WMS-III indices. These results suggest that caution must be used when applying the RCIs and SRB change norms generated from the pre-operative sample to the post-operative sample for the WMI and GM index, but that these change data are appropriate for all other indices.

Determining Reliable Change in the Post-Operative Sample

Table 21 lists the demographic information for the post-operative sample following division into groups according to seizure pathology – left anterior temporal lobectomy and right anterior temporal lobectomy.

Table 21
Demographic variables for post-operative sample, grouped according to seizure laterality

	Left ATL (n = 37)		**Right ATL (n = 36)**	
Variable	**Mean**	**SD**	**Mean**	**SD**
Age at first assessment	32.2	8.4	36.1	12.2
Education (years)	11.1	1.6	11.1	1.9
Full Scale IQ (N=73)	91.5	14	94.1	13
Retest interval (months)	14.8	10.3	16.3	11
Gender (male/female)	12/25		21/5	
Handedness (right/left/ambi)	32/3/2		33/3/0	

Although the groups were not matched, the quasi-randomised experimental design, in which all consecutive cases were used, should have approximated randomisation. Independent *t*-tests were used to examine the differences between the demographic and clinical characteristics of the groups. The left and right ATL groups did not differ significantly in age (t = -1.691, df = 65, p = 0.096), education (t = 0.309, df = 71, p = 0.758), length of the test-retest interval (t =-0.347, df = 68, p = 0.73) or baseline FSIQ (t = -0.893, df = 68, p = 0.375).

Reliable change indices.

The Reliable Change Indices (RCI) – using the SE_{diff} method – derived from the pre-surgical control group in Chapter Three were applied to the post-surgical sample. The 90% and 95% confidence intervals, derived from both the Australian and North American (Martin et al., 2001) pre-operative samples, were applied to the Australian post-surgical sample. This allowed investigation into whether RCIs are generalisable across different demographic backgrounds.

The classifications of change on retest according to 95% confidence intervals (see Tables C2 and C3, Appendix C) were at times too conservative and resulted in an under-estimation of change on retest. This tendency was evident for both the local and North American norms. However, this inclination improved with the application of less conservative (90%) confidence intervals and therefore the analysis of cognitive change below is based on the base-rates generated by 90% confidence intervals. Finally, the RCIs were also applied to the entire post-surgical sample; however, this obscured laterality effects by averaging out significant differences across the left and right post-operative groups (see Tables C4 and C5, Appendix C).

The prediction intervals listed in Tables 22 and 23 refer to the values around the predicted retest score that would be expected to include 90% of the non-surgical control group. Therefore, the remaining 5% at either end of the distribution represent those scores which are considered unusual in the normative sample and are therefore unlikely to have occurred by chance alone.

Finally, to examine the level of agreement between the Australian and North American RCIs regarding the classifications of post-operative change, Cohen's kappa and Kendall's tau analyses were employed. Cohen's kappa was used preferentially; however, this analysis requires identical categories from the two sources – at least one patient in each of the declined, unchanged and improved groups. Where identical categories between the Australian and North American data sources did not exist, Kendall's tau analyses were utilised.

Table 22
Classification of cognitive change on WAIS-III indices using RCIs, according to surgery side

		Australian RCIs – SE_{diff} method				North American RCIs – SE_{diff} method				
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Cohen's kappa**
Hypothesised Distribution			*5*	*90*	*5*		*5*	*90*	*5*	
VIQ		**±11**				**-8, +8**				
Left	34		2.9	91.2	5.9		5.9	82.4	11.8	0.633
Right	33		3	81.8	15.2		3	72.7	24.2	0.753
PIQ		**-9, + 15**				**-7, +13**				
Left	34		5.9	91.2	2.9		8.8	79.4	11.8	0.560
Right	33		3	90.9	6.1		6.1	84.8	9.1	0.726
FSIQ		**-9, + 11**				**-6, +8**				
Left	34		2.9	85.3	11.8		8.8	73.5	17.6	0.663
Right	33		0	84.8	15.2		9.1	66.7	24.2	*0.648*^
VCI		**±10**				**-7, +9**				
Left	34		5.9	88.2	5.9		17.6	76.5	5.9	0.622
Right	33		0	81.8	18.2		3	72.7	24.2	*0.787*^
POI		**-11, +19**				**-8, +14**				
Left	33		6.1	90.9	3		12.1	75.8	12.1	0.495
Right	32		3.1	93.8	3.1		12.5	81.3	6.3	0.462
WMI		**±18**				**-17, +14**				
Left	33		3	93.9	3		3	81.8	15.2	0.463
Right	33		3	87.9	9.1		3	78.8	18.2	0.687
PSI		**-8, +16**				**n/a**				
Left	31		22.6	64.5	12.9	-	-	-	-	-
Right	30		6.7	86.7	6.7	-	-	-	-	-
BNT		**±12**				**n/a**				
Left	35		11.4	88.6	0	-	-	-	-	-
Right	35		0	100	0	-	-	-	-	-

Note: See Table 9 for qualitative descriptors of agreement, as indicated by Cohen's kappa
^ *Kendall's tau-b analysis used*

Consistent with the literature, left ATL resulted in a decline in naming abilities on retest for a number of people (11.4%); a trend which was not evident for the right

ATL group. In contrast, some significant retest improvements in speed of processing performance (12.9%) were evident following left ATL. However, processing speed also had the highest incidence of decline for the left-sided group, with one in five people (22.6%) experiencing a decline in performance on this measure. In stark contrast, only 6.7% of the right group demonstrated a significant reduction in processing speed performance. Moreover, there was evidence of improvement following right ATL in verbal abilities (15.2%) and FSIQ (15.2%).

As discussed above, it was decided to use highly conservative kappa coefficients of greater than 0.81 to reflect the extent of the agreement between the Australian and North American change classifications. Using this conservative criterion, the classifications of change between the Australian and North American RCIs, for both the left and right surgery groups, were significantly different, suggesting the North American RCIs were poorly generalisable to an Australian post-operative TLE sample. However, this finding is solely reliant on the chosen agreement criterion – in this case highly conservative – and accepting a lower agreement criterion (e.g., 0.61-0.8, or "substantial"; Landis & Koch, 1977) would result in comparable change classifications across the majority of the indices.

Table 23
Classification of cognitive change on WMS-III indices using RCIs, according to surgery side

		Australian RCIs – SE_{diff} method				**North American RCIs – SE_{diff} method**				
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Cohen's kappa**
Hypothesised Distribution			*5*	*90*	*5*		*5*	*90*	*5*	
AI		**-16, +18**				**-14, +16**				
Left	36		16.7	77.8	5.6		19.4	75	5.6	0.927
Right	35		8.6	80	11.4		11.4	74.3	14.3	0.849
VI		**-19, +23**				**-15, +21**				
Left	36		0	94.4	5.6		2.9	80.6	16.7	*0.503*^
Right	35		5.7	94.3	0		8.6	91.4	0	0.785
IM		**-15, +19**				**-12, +16**				
Left	36		8.3	88.9	2.8		19.4	63.9	16.7	0.392
Right	35		2.9	97.1	0		11.4	88.6	0	0.371
AD		**-18, +20**				**-19, +15**				
Left	36		5.6	80.6	13.9		5.6	72.2	22.2	0.783
Right	35		0	85.7	14.3		0	80	20	0.800
VD		**-18, +20**				**-13, +23**				
Left	36		2.8	80.6	16.7		5.6	86.1	8.3	0.618
Right	35		5.7	91.4	2.9		8.6	91.4	0	*0.670*^
GM		**-16, +20**				**-16, +18**				
Left	36		8.3	83.3	8.3		8.3	75	16.7	0.765
Right	33		0	97	3		0	93.9	6.1	0.653
WM		**-18, +22**				**-22, +18**				
Left	32		6.3	84.4	9.4		6.3	78.1	15.6	0.806
Right	35		2.9	97.1	0		2.9	94.3	2.9	1.000

Note: ^ *Kendall's tau-b analysis used*

Consistent with expectation, a greater percentage (16.7%) of the left ATL group showed significant decline on the Auditory Immediate memory index than the right ATL group. However, a number of people (8.6%) in the right ATL group also experienced significant verbal memory decline. Further, the anticipated material-specific pattern of visual memory decline following right ATL was seen in only a

small percentage of people (5.7%). In summary, the left-sided surgery group demonstrated greater cognitive morbidity, particularly in the areas of auditory memory, processing speed and confrontational naming.

The kappa coefficients generated for the left ATL group on the WMS-III indicated significant agreement regarding change classifications on the AI and WM indices. A kappa coefficient was not able to be calculated for the VI index – due to non-identical categories, discussed above. Instead, Kendall's tau – a nonparametric measure of correlation – was used and indicated a significant correlation between the two sets of RCIs on the VI index (tau = 0.503, N = 36, p = 0.024). This means that the Australian and North American RCIs did not produce significantly different change classifications on this measure. For the right ATL group, the classification rates of change on the AI, AD and WM indices all demonstrated strong agreement.

Overall, these results indicate that – using a very conservative agreement criterion – RCIs are unlikely to be accurately generalisable across different demographic samples for the majority of WAIS-III and WMS-III indices, with the exception of the AI, VI and WM memory indices. As discussed above, deciding on an acceptable criterion for agreement, in this case, was dependent on the clinical implications of generalising non-local change norms which may not accurately capture genuine cognitive change. Evidence of the potential risk involved in setting too low a criterion for agreement is seen in Tables 23 and 24; for example, despite a Cohen's kappa of 0.783 ("substantial agreement", see Table 18, page 74), the change classifications dictated by the Australian and North American RCIs on the AD index of the WMS-III show poor apparent agreement.

Standardised regression-based change norms.

The SRB approach to measuring post-operative change was also employed. The predicted retest scores for the post-operative TLE group were generated using the formula described by McSweeny et al. (1993). The SE_{est} was multiplied by the relevant standard deviations (i.e., ±1.64 or ±1.96 respectively) to obtain 90% and 95% (see Tables C6 and C7, Appendix C for the latter) confidence intervals, which were placed around each person's predicted retest scores to delineate the boundaries of

reliable change. The SRB change norms were also applied to the entire post-operative sample and these are listed in Tables C8 and C9, Appendix C.

As outlined above, information regarding each person's lifetime duration of seizures was not available for the post-operative sample. However, in the pre-operative sample, duration of seizures was a significant predictor variable for both the Processing Speed Index of the WAIS-III and the Auditory Immediate Index of the WMS-III. As a result, a further regression equation was generated from the pre-operative data, which omitted seizure duration as a predictor variable. The resulting equations are listed in Table 24 and were utilised in the application of SRBs to the post-operative sample in subsequent analyses.

Table 24
Regression equations for the Auditory Immediate Memory and Processing Speed indices, with and without seizure duration

Index	R	R^2	SEest	Constant	Sig.	β baseline	β age	β edu	β duration	β retest interval	β AEDs
WAIS-III											
PSI + seizure duration	0.910	0.829	5.75	22.15	<0.0005	**0.73#**	**0.25#**		**-0.23#**		
PSI – seizure duration	0.887	0.786	6.32	19.57	<0.0005	**0.73#**	**0.19***				
WMS-III											
AI + seizure duration	0.905	0.818	8.63	11.22	<0.0005	**0.79#**	**0.38***		**-0.27#**		
AI – seizure duration	0.873	0.762	9.89	10.02	<0.0005	**0.78#**	**0.3***				

* $p < 0.05$; # $p < 0.01$

Whilst the removal of seizure duration as a predictor variable for the PSI resulted in slightly lower *R* value, the revised equation continued to account for 79% of the variance, compared to 83% for the original model. Similarly, removing seizure duration from the regression equations for the AI memory index resulted in a slightly less strong correlation between the observed and predicted values, but the model nonetheless continued to account for 77% of the variance, compared to 82% in the initial equation. Given the different SE_{est} associated with the revised regression equations, the prediction intervals also differed slightly for the PSI (e.g., 6.32 * ± 1.64 = ± 10.36) and AI indices (e.g., 9.89 * ± 1.64 = ±16.22).

A comparison of the classifications of change generated by the Australian and North American SRBs was once again calculated using Cohen's kappa or Kendall's tau

analyses and these are listed in Tables 25 and 26. As discussed above, the prediction intervals were placed around a person's predicted retest score to define the boundaries of change.

Table 25
Classification of cognitive change on WAIS-III indices using SRBs, according to surgery side

		Australian SRBs				North American SRBs				
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Cohen's kappa**
Hypothesised Distribution			*5*	*90*	*5*		*5*	*90*	*5*	
VIQ		**±10.75**				**±7.95**				
Left	34		2.9	91.2	5.9		5.9	82.4	11.8	0.633
Right	33		3	81.8	15.2		3	75.8	21.2	0.825
PIQ		**±12.02**				**±9.64**				
Left	34		5.9	91.2	2.9		5.9	79.4	14.7	0.561
Right	33		3	90.9	6.1		3	90.9	6.1	0.696^
FSIQ		**±9.92**				**±7.66**				
Left	34		2.9	85.3	11.8		20.6	73.5	5.9	0.346
Right	33		0	84.8	15.2		12.1	69.7	18.2	*0.716*^
VCI		**±8.5**				**±7.66**				
Left	34		8.8	88.2	2.9		11.8	79.4	8.8	0.692
Right	33		0	81.8	18.2		0	72.7	27.3	*0.652*^
POI		**±15.55**				**±10.34**				
Left	33		3	93.9	3		6.1	78.8	15.2	0.402
Right	32		3.1	93.8	3.1		6.3	81.3	12.5	0.462
WMI		**±16.65**				**±14.3**				
Left	33		3	81.8	15.2		3	81.8	15.2	*-0.153*^
Right	33		3	87.9	9.1		3	81.8	15.2	1.000
PSI		**±10.36**								
Left	31		12.9	61.3	25.8		-	-	-	-
Right	30		3.3	86.7	10.0		-	-	-	-

Note: ^ *Kendall's tau-b analysis used*

Although the base-rates of cognitive change differed between the RCI and SRB change methodologies; however, the overall trends were mostly similar. Once again, those people who underwent left ATL showed greater cognitive morbidity post-operatively but improvement was also evident in this group following surgery. Following left ATL, a significant number of people experienced post-operative decline on the PIQ (5.9%), VCI (8.8%) and PSI (12.9%). In contrast, there was also

evidence of post-operative improvement in a significant number of people on the VIQ (5.9%) and FSIQ (11.8%) following left ATL.

Interestingly, following right ATL, a significant proportion of people demonstrated improvements on the FSIQ (15.2%), VIQ (15.2%) and almost one in five people in this group improved significantly on the VCI post-operatively (18.2%). In the right ATL group, no patient in the right-sided group experienced significant decline on retest on the FSIQ, VIQ, PIQ, VCI, POI, WMI or PSI, beyond that which would be expected from a normal distribution of change scores.

The kappa coefficients generated for the left ATL group on the WAIS-III indicated moderate levels of agreement between the Australian and North American change classifications for the majority of the indices, with the exception of the WMI (tau = -0.153, N = 33, $p = 0.531$), which showed identical classifications of change. For the right ATL group, the Australian and North American change classifications showed high levels of agreement on the VIQ and WMI. It is important to note that for the right ATL group, the kappa coefficient was 1.000, indicating perfect agreement between the two sets of change classifications. However, these change classifications are clearly not identical and the strong kappa coefficient illustrates the potential for this statistic to over-state the strength of agreement. Kendall's tau analyses also indicated strong agreement between the two sets of SRBs on the FSIQ (tau = 0.716, N = 33, $p < 0.0005$), PIQ (tau = 0.821, N = 33, $p = 0.002$), and VCI (tau = 0.652, N = 33, $p = 0.001$).

Table 26
Classification of cognitive change on WMS-III indices using SRBs, according to surgery side

		Australian SRBs				North American SRBs				
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Cohen's kappa**
Hypothesised Distribution			*5*	*90*	*5*		*5*	*90*	*5*	
AI		**±16.22**				**±14.6**				
Left	36		16.7	77.8	5.6		19.4	75	5.6	1.000
Right	35		8.6	80	11.4		8.6	74.3	17.1	0.469
VI		**±18.78**				**±18**				
Left	36		0	86.1	13.9		0	86.1	13.9	1.000
Right	35		5.7	94.3	0		8.6	91.4	0	1.000
IM		**±16.32**				**±14.65**				
Left	36		8.3	88.9	2.8		16.7	75	8.3	0.562
Right	35		2.9	97.1	0		5.7	94.3	0	0.653
AD		**±16.42**				**±14.12**				
Left	36		16.7	69.4	13.9		16.7	69.4	13.9	1.000
Right	35		0	80	20		2.9	65.7	31.4	0.449
VD		**±17.66**				**±19.98**				
Left	36		0	86.1	13.9		5.6	94.4	0	*0.097*^
Right	35		5.7	91.4	2.9		9.6	91.4	0	*0.670*^
GM		**±16.35**				**±17.27**				
Left	36		8.3	86.1	5.6		8.3	88.9	2.8	0.877
Right	33		0	97	3		3	93.9	3	0.788
WM		**±18.96**				**±17.2**				
Left	32		6.3	78.1	15.6		6.3	71.9	21.9	0.844
Right	35		2.9	97.1	0		0	85.7	14.3	0.533

Note: ^ *Kendall's tau-b analysis used*

Based on the literature, it was hypothesised that patients who had left ATLs would demonstrate a higher incidence of decline on post-operative verbal memory measures. This expectation was supported, with 16.7% of patients in the left ATL group showing significant verbal memory decline on retest, compared with 8.6% in the right ATL group. Regarding visual memory outcomes, a significant proportion (13.9%) of the left ATL group demonstrated improvement on retest.

Patients who underwent right ATL were expected to demonstrate adverse visual memory outcomes. However, there was no strong material-specific effect, with only 5.7% of people in this group experiencing a significant decline on the Visual Immediate and Visual Delayed indices. A finding which further clouded material-specific patterns of change was that whilst 11.4% of the right ATL group showed significantly improved immediate auditory memory on retest, 8.6% of this group also experienced significant decline.

For the left-sided surgery group, the Australian and North American SRBs produced identical change classifications on the AD and VI indices, as well as highly similar classifications on the GM and WM indices. In contrast, the remaining indices showed limited levels of agreement and the classifications of change on the VD differed significantly (tau = 0.097, N = 36, p = 1.000) indices. In the right ATL group, the classifications of change were identical on the VI index, with all other WMS-III indices showing only moderate agreement between classification rates. Kendall's tau was again utilised for the VD index (tau = 0.67, N = 35, p = 0.005), but revealed similar change classifications resulting from the Australian and North American SRB change norms.

For comparison purposes, a summary of the classification rates derived from both the local and US samples, using both the RCI and SRB change approaches is listed in Tables 27 and 28.

Table 27
Classification of cognitive change on WAIS-III indices using both RCIs and SRB approaches, according to surgery side

	Australian RCIs		North American RCIs		Australian SRBs		North American SRBs	
Index	**% declined**	**% improved**	**% declined**	**% improved**	**% declined**	**% improved**	**% declined**	**% improved**
Hypothesised distribution	*5%*	*5%*	*5%*	*5%*	*5%*	*5%*	*5%*	*5%*
VIQ								
Left	2.9	5.9	5.9	11.8	2.9	5.9	5.9	11.8
Right	3	15.2	3	24.2	3	15.2	3	21.2
PIQ								
Left	5.9	2.9	8.8	11.8	5.9	2.9	5.9	14.7
Right	3	6.1	6.1	9.1	3	6.1	3	6.1
FSIQ								
Left	2.9	11.8	8.8	17.6	2.9	11.8	20.6	5.9
Right	0	15.2	9.1	24.2	0	15.2	12.1	18.2
VCI								
Left	5.9	5.9	17.6	5.9	8.8	2.9	11.8	8.8
Right	0	18.2	3	24.2	0	18.2	0	27.3
POI								
Left	6.1	3	12.1	12.1	3	3	6.1	15.2
Right	3.1	3.1	12.5	6.3	3.1	3.1	6.3	12.5
WMI								
Left	3	3	3	15.2	3	15.2	3	15.2
Right	3	9.1	3	18.2	3	9.1	3	15.2
PSI								
Left	22.6	12.9	-	-	12.9	25.8	-	-
Right	6.7	6.7	-	-	3.3	10.0	-	-

Table 28
Classification of cognitive change on WMS-III indices using both RCIs and SRB approaches, according to surgery side

	Australian RCIs		North American RCIs		Australian SRBs		North American SRBs	
Index	**% declined**	**% improved**	**% declined**	**% improved**	**% declined**	**% improved**	**% declined**	**% improved**
Hypothesised distribution	*5%*	*5%*	*5%*	*5%*	*5%*	*5%*	*5%*	*5%*
AI								
Left	16.7	5.6	19.4	5.6	16.7	5.6	19.4	5.6
Right	8.6	11.4	11.4	14.3	8.6	11.4	8.6	17.1
VI								
Left	0	5.6	2.9	16.7	0	13.9	0	13.9
Right	5.7	0	8.6	0	5.7	0	8.6	0
IM								
Left	8.3	2.8	19.4	16.7	8.3	2.8	16.7	8.3
Right	2.9	0	11.4	0	2.9	0	5.7	0
AD								
Left	5.6	13.9	5.6	22.2	16.7	13.9	16.7	13.9
Right	0	14.3	0	20	0	20	2.9	31.4
VD								
Left	2.8	16.7	5.6	8.3	0	13.9	5.6	0
Right	5.7	2.9	8.6	0	5.7	2.9	9.6	0
GM								
Left	8.3	8.3	8.3	16.7	8.3	5.6	8.3	2.8
Right	0	3	0	6.1	0	3	3	3
WM								
Left	6.3	9.4	6.3	15.6	6.3	15.6	6.3	21.9
Right	2.9	0	2.9	2.9	2.9	0	0	14.3

Overall, the performance of the left- and right-sided surgery groups revealed marked individual variability. Nonetheless, some discernible trends in cognitive outcomes did emerge. For those people undergoing left ATL, verbal memory and speed of processing appeared most vulnerable to post-operative decline. In the case of right ATL, there was a small but definite chance of post-operative decline in both visual

and verbal memory. However, favourable cognitive outcomes were more likely in the right ATL group, particularly for language abilities and verbal memory.

As evidenced in Tables 27 and 28, the classification rates of cognitive change were comparable between both RCI and SRB approaches, as well as between Australian and North American data. In particular, the Australian RCIs and SRB norms produced comparable or identical classification rates across the majority of the WAIS-III and WMS-III indices, with the exception of the Auditory Delayed index, on which a larger number of people were classified as declined on retest by the SRB approach.

Number needed to harm and number needed to treat.

The absolute risk reduction/increase (ARR/ARI) was calculated by subtracting the experimental event rate (i.e., the percentage of the post-operative sample experiencing decline on retest) from the control event rate (i.e., percentage of the pre-operative sample experiencing decline on retest). The ARR/ARI was then divided into 100 to calculate the NNT or NNH. As an example, an ARR/ARI of zero indicates no difference between the two groups and that the level of harm attributable to the two treatments therefore does not differ.

The NNT is rounded up to the next highest whole number, as it is not possible to "treat a fraction of a person" (Citrome, 2007, p.883). In contrast, conservative calculation would round a NNH down to the nearest person (Schwartz, 2007). A smaller NNT indicates a greater difference in the outcomes associated with medical (i.e., AEDs) and surgical intervention for TLE. In contrast, a large NNT of 100 or more would suggest little difference between medical and surgical intervention (Citrome, 2007).

The NNHs and NNTs for the right- and left-sided surgery groups – based on the 90% RCIs (SE_{diff}) and SRB change norms described above – are listed in Table 29 for the WAIS-III indices and Table 30 for the WMS-III indices. The NNT and NNH based on the 95% confidence intervals were also calculated and these are listed in Tables C10 and C11, Appendix C.

Table 29
Absolute Risk Reduction and Number Needed to Treat / Number Needed to Harm values for left and right ATL groups on the WAIS-III indices, using 90% change norms

		SRBs				RCIs (SE_{diff})			
Index	**N**	**Control event rate**	**Exp. event rate**	**ARR/ ARI**	**NNT (+) or NNH (-)**	**Control event rate**	**Exp. event rate**	**ARR/ ARI**	**NNT (+) or NNH (-)**
				Control – Exp.	**100/ARR**			**Control – Exp.**	**100/ARR**
VIQ									
LATL	34	5.7	2.9	2.8	36	5.7	2.9	2.8	36
RATL	33	5.7	3	2.7	37	5.7	3	2.7	37
PIQ									
LATL	34	0	5.9	-5.9	-16	0	5.9	-5.9	-16
RATL	33	0	0	0	No difference	0	3	-3	-33
FSIQ									
LATL	34	5.7	2.9	2.8	36	8.6	2.9	5.7	18
RATL	33	5.7	0	5.7	18	8.6	0	8.6	12
VCI									
LATL	34	8.8	8.8	0	No difference	8.8	5.9	2.9	35
RATL	33	8.8	3	5.8	18	8.8	0	8.8	12
POI									
LATL	34	5.9	2.9	3	34	5.9	6.1	-0.2	-500
RTL	33	5.9	3	2.9	35	5.9	3.1	2.8	36
WMI									
LATL	34	6.1	2.9	3.2	32	6.1	3	3.1	33
RATL	33	6.1	3	3.1	33	6.1	3	3.1	33
PSI									
LATL	31	5.7	12.9	-7.2	-13	8.6	22.6	-14	-7
RATL	30	5.7	3.3	2.4	42	8.6	6.7	1.9	53
BNT									
LATL	34	-	-	-	-	8.8	11.4	-2.6	-38
RATL	35	-	-	-	-	8.8	0	8.8	12

Using the SRB data on the VIQ for the left ATL group as an example, the ARR indicated 2.8% of operated patients will not experience the decline on this measure that they would have experienced if only treated with AEDs. If a positive number

results when the experimental rate is subtracted from the control rate, it is referred to as the Absolute Risk Reduction. In this same example, the subsequent NNT figure indicates 36 people would have to undergo surgical intervention to prevent one adverse event that would have happened in the AED-treated (control) group.

In contrast, if subtracting the experimental rate from the control rate results in a negative number, this figure is referred to as the Absolute Risk Increase. Using the SRB data on the PIQ for the left ATL group as an example, 5.9% of these patients will experience a significant deterioration in their PIQ performance that would not have occurred if they did not have surgery. This means that the NNT becomes a NNH – as it signifies a disadvantage to the treatment. In this example, for every 16 patients who undergo left ATL, one person – beyond those that would have occurred with medical treatment only – will experience a post-operative decline in their PIQ performance.

As evident in Table 29, the cognitive domain most likely to suffer adverse effects following left ATL was processing speed, with one in seven patients demonstrating a decline in this domain following surgery. In contrast, the relatively high NNT values for the right ATL group indicated minimal harm to speed of processing associated with surgery in this group. Indeed, the results suggested that surgery is associated with positive outcomes, compared to AED treatment, for every 12 people in the right ATL group on the FSIQ, VCI and BNT.

Table 30
Absolute Risk Reduction and Number Needed to Treat / Number Needed to Harm values for left and right ATL groups on the WMS-III indices, using 90% change norms

		SRBs				RCIs (SE_{diff})			
Index	**N**	**Control event rate**	**Exp. event rate**	**ARR/ ARI**	**NNT (+) or NNH (-)**	**Control event rate**	**Exp. event rate**	**ARR/ ARI**	**NNT (+) or NNH (-)**
AI									
LATL	36	2.6	19.4	-16.8	-6	2.6	16.7	-14.1	-7
RATL	35	2.6	8.6	-6	-16	2.6	8.6	-6.0	-16
VI									
LATL	36	2.6	0	2.6	39	7.9	0	2.6	39
RATL	35	2.6	8.6	-6	-16	7.9	5.7	2.2	45
IM									
LATL	36	5.1	8.3	-3.2	-31	5.1	8.3	-3.2	-31
RATL	35	5.1	2.9	2.2	46	5.1	2.9	2.2	46
AD									
LATL	36	5.1	16.7	-11.6	-8	5.1	5.6	-0.5	-200
RATL	35	5.1	5.7	-0.6	-166	5.1	0	5.1	20
VD									
LATL	36	2.6	0	2.6	39	10.3	2.8	7.5	14
RATL	35	2.6	5.7	-3.1	-32	10.3	5.7	-3.1	-32
GM									
LATL	36	7.7	8.3	-0.6	-166	5.1	8.3	-0.6	-166
RATL	33	7.7	3	4.7	22	5.1	0	5.1	20
WM									
LATL	32	5.3	6.3	-1	-100	7.9	6.3	-1	-100
RATL	35	5.3	0	5.3	19	7.9	2.9	5.0	20

Note: negative NNTs represent NNHs

As expected, people who underwent left ATL experienced more post-operative cognitive morbidity in verbal memory skills. Specifically, Table 30 shows that for the left ATL group, 16.8% of patients experienced a significant deterioration in their post-operative Auditory Immediate memory index score, which would not have occurred had they been treated with AEDs instead. Reported in NNT terms, this means that for every six patients who underwent left ATL, one experienced a significant

deterioration in their Auditory Immediate index score, beyond that which would have occurred had they been treated only with medication. The association between left ATL and verbal memory decline was further supported by a relatively small NNH on the Auditory Delayed index. However, right ATL was also significantly associated with post-operative verbal memory decline. For example, for every 16 people who underwent right ATL, one experienced a significant decline in their post-operative performance on the Auditory Immediate memory index. These values can be used to predict outcomes in similar samples.

Discussion

The post-operative classification rates resulting from both the RCI and SRB change equations in the current study were expected to differ significantly from those of Martin and colleagues' (2002) and this hypothesis was partially supported. The change classifications resulting from the Australian and North American RCIs were not highly similar for the majority of the WAIS-III and several WMS-III indices. However, the local and non-local RCIs resulted in similar change classifications on the AI and AD for the left ATL group and the VI, VD, GM and WM for the right ATL group, suggesting some generalisability of the North American RCIs to Australian samples for memory indices.

In contrast, the Australian and North American SRB change norms generated highly similar classifications of change for the left ATL group on the AI, AD, VI, GM, WM indices of the WMS-III and the WMI of the WAIS-III. Similar change classifications in the right ATL group were also evident on the VIQ and WMI of the WAIS-III and the VI and GM indices of the WMS-III. The generalisability of the SRB change norms across more indices than the RCIs was somewhat unexpected as the former included local demographic variables and seizure characteristics, which would intuitively suggest that these resulting regression equations were more specifically tailored for the target population. In contrast, perhaps demographic variables, such as age and education, are more universally predictive of test performance, leaving RCIs – which do not account for these variables – more highly dependent on the use of local data.

The above results suggest that whilst it is ideal to derive local change norms wherever possible, it may nonetheless be appropriate to apply RCIs and SRB change norms generated from North American TLE samples to Australian TLE samples, on certain measures. Potential support for these results is found in the work of ` and colleagues (Bowden, Lissner, McCarthy, Weiss, & Holdnack, 2007). The authors compared an Australian community sample against the WAIS-III US standardisation data and found equivalence of measurement of core cognitive abilities, suggesting no significant differences in the cognitive abilities of Australians and Americans. Further prospective support comes from the work of Woods et al. (2006), who found that RCI data obtained from healthy controls generalised to a non-matched sample of patients with HIV (who were older, higher educated, more cognitively impaired and had a longer retest interval).

Based on the literature, it was expected that patients who had undergone left ATL would experience more frequent adverse verbal memory outcomes than their right-sided counterparts (Baxendale et al., 2008; Chelune et al., 1991; Chelune, 1995; Helmstaedter & Elger, 1996; Helmstaedter et al., 2004; Hermann et al., 1992; Hermann et al., 1995; Hermann et al., 1997; Hoppe, Elger, & Helmstaedter, 2007; Ivnik et al., 1988; Lee et al., 2002; Martin et al., 1998; Mueller, Kaaden, Scorzin, Urbach, Fimmers, et al., 2009; Ojemann & Dodrill, 1985; York, Rettig, Grossman, Hamilton, Armstrong, et al., 2003) and this hypothesis was supported. Based on both RCI and SRB data, patients in the left ATL group more frequently demonstrated auditory immediate memory decline (16.7-19.4%) than those who had undergone right ATL (8.6%). Verbal delayed memory was also more frequently adversely affected in the left ATL group (5.6-16.7%) than the right ATL group (0%). However, previous outcome studies have reported higher rates of post-operative verbal memory decline, in both left ATL patients (40-60%) and right ATL patients (10-30%) (Chelune et al., 1993; Hoppe et al., 2007; Martin et al., 1998; Martin et al., 2002b) than was evident in the current study. These differences may be accounted for by slight variations in the formulas applied (e.g., Chelune et al., 1993), different predictor variables included in the regression equations (e.g., Martin et al., 2002b), or potential heterogeneity of TLE aetiology in the different samples under investigation.

In contrast, the right ATL group was expected to show a significant, but less robust, association with visual memory difficulties (Helmstaedter et al., 2004; Hoppe et al., 2007; Lee et al., 2002). However, unexpectedly, the risk to visual memory following right ATL was no greater than the risk to verbal memory. A small percentage of the right ATL group showed significant decline on both the Visual Immediate and Visual Delayed indices (5.7%); however, the incidence of decline was less than that seen for this group on the Auditory Immediate Index (8.6%). An equivocal association between right ATL and visual memory is consistent with several previous studies (Gleissner et al., 1998; Hermann et al., 1995; Ivnik et al., 1988; McDermid Vaz, 2004; Mueller et al., 2009; Seidenberg et al., 1998).

The literature has also reported evidence for decline in verbal memory following right ATL (Baxendale et al., 2008; Bell & Davies, 1998; Hermann et al., 1992; Hoppe et al., 2007; Martin et al., 1998). Baxendale and colleagues (2005) reported one in ten right ATL patients experienced verbal learning deficits and the results of the current study are consistent with this. Conversely, there was no evidence of significant decline in visual memory following left ATL, a result consistent with some studies (Ivnik et al., 1987; Saykin, Robinson, Stafiniak, Kester, Gur, et al., 1992), but not others (Helmstaedter et al., 2003; Hoppe et al., 2007). The results of the current study, together with those of previous studies, suggests that right ATL poses a similar level of risk to both immediate verbal and visual memory.

It was also hypothesised that the left ATL group would experience a greater incidence of post-operative decline in confrontational naming abilities and this was supported by the results. Of the left ATL group, 11.4% demonstrated significant post-operative decline on the BNT. In contrast, no person in the right ATL group declined. This pattern of results is supported by the literature, which suggests post-operative naming deficits occur in a small percentage of left ATL patients (Bell et al., 2000; Hermann et al., 1994; Seidenberg et al., 1998). However, Bell colleagues (2000) reported significantly higher rates of decline (40%) in their left ATL sample.

Overall, as reported in the literature, there was considerable variability in individual outcomes for the post-operative sample (Engman et al., 2006; Hermann et al., 1992; Hermann et al., 1995; Martin, Kretzmer, Palmer, Sawrie, Knowlton, et al., 2002a; Sherman et al., 2003). However, the majority remained cognitively stable pre- to post-operatively, as reported previously (Lineweaver et al., 2006; Martin et al., 2002b). In the current study, the frequency of memory decline ranged from only 5.7% on the Visual Immediate and Delayed memory indices to 16.7% on the Auditory Immediate memory index. These percentages were comparable to those seen in the non-surgical control group. In addition to the expected decline in verbal memory following left ATL, a number of patients also demonstrated post-operative improvement on this measure (5.6%), although this incidence was not markedly greater than would be expected for a normal distribution, using 90% confidence intervals. Nonetheless, this result is consistent with findings that a minority of patients show significant post-operative cognitive improvement in memory skills following left ATL (Baxendale et al., 2006; Baxendale et al., 2008; Helmstaedter, Brosch, Kurthen, & Elger, 2004; Martin et al., 1998; Paglioli, Palmini, Portuguez, Paglioli, Azambuja, et al., 2006; Wachi, Tomikawa, Fukuda, Kameyama, Kasahara, et al., 2001).

Also consistent with previous reports, the post-operative cognitive outcomes in right ATL were similarly heterogeneous, with nominal rates of both improvement and decline on visual memory measures (Baxendale, 2008). These results suggest that visual memory is less vulnerable to adverse effects following ATL than verbal memory, regardless of surgery side. Evidence for verbal memory decline in patients who have undergone right ATL and infrequent improvements in visual memory for this same group challenges a strict interpretation of material-specific theory and is discussed in further detail in Chapter Five.

Strengths and Limitations

The data for the post-operative sample was obtained from archival records of consecutive post-operative assessments, making the current study a quasi-randomised experimental design, which resulted in good matching of left and right post-operative

patient characteristics. Nonetheless, there were significant differences in education level and the length of the retest interval between the pre- and post-operative samples. In particular, the pre-operative group had slightly more education than the post-operative group and their retest data was obtained following a much longer retest interval. Subsequent analyses of covariance revealed a statistically significant effect of retest interval on both the WMI and the GM index, but there were no significant effects of retest interval or education level on the other WAIS-III and WMS-III indices. The longer retest intervals in the pre-operative sample than the post-operative sample on the WMI and GM index – and subsequently lower likely magnitude of practice effects due to dissipation with time – may therefore have led to an underestimation of practice effects for these indices. This in turn may have influenced the classification of change on retest, with a potential bias towards overestimating change. Nonetheless, the magnitude of practice effects calculated from the pre-operative sample on the WMI and GM index were comparable to Martin et al.'s (2002) study, which used retest intervals more akin to those in the current post-operative group.

As discussed above, seizure control is likely to be an important factor in predicting cognitive change over time. However the frequency of seizures was difficult to reliably ascertain from participants and medical records, and indeed varied significantly for individuals across the course of their illness due to factors such as medication changes and psychosocial factors. Future studies should consider including seizure frequency where possible, as it represents an important marker for seizure control. Further to this, the retrospective data collection for the post-operative sample precluded the collection of reliable information regarding a person's seizure duration. As a result, the regression equations for the Auditory Immediate memory index and Processing Speed index excluded seizure duration as a predictor variable, despite results from the pre-operative sample indicating it significantly contributed to the regression equation. Nonetheless, the revised regression equations (minus seizure duration) continued to account for a large proportion of the variance.

A further limitation of the current study was the heterogeneity of the pre-operative sample. Although they were all surgical candidates, only four had gone on to have

surgery at the time of writing. Reasons for not progressing to surgery included significant physical or cognitive risks associated with surgery, as well as personal preference. As noted by Sherman and colleagues (2003), a heterogeneous sample may result in large standard errors of prediction for the regression model and subsequently lead to underestimation of post-operative change. In addition to this, a number of the people in the pre-operative group had been identified as experiencing affective symptoms – particularly anxiety – which might potentially influence their suitability for surgery. However, affective symptoms in TLE are well-recognised (Hermann, Seidenberg, & Bell, 2000; Hixson & Kirsch, 2009; Kanner, 2006; Paradiso et al., 2001; Swinkels, Kuyk, van Dyck, & Spinhoven, 2005) and the presence of these in the pre-operative group was likely mirrored in the post-operative sample. Nonetheless, future investigations should consider the contribution of affective symptoms in predicting cognitive scores on retest.

In summary, the current results support previous recommendations to derive local change norms wherever possible (Martin et al., 2002b; Sawrie et al., 1996), but that US RCIs and SRB change norms are nonetheless adequate for determining change in Australian TLE patients on certain cognitive domains. Further to this, for Australian clinical settings unable to collect local data, it is recommended the current change norms be applied. In deciding which change approach to utilise, both the RCI and SRB formulas demonstrated similar trends in cognitive change following ATL, although when compared directly, the local RCI data (both SE_{diff} and SE_{pred}) held a slight advantage over local SRB change norms. The application of RCI methods is reasonably uncomplicated and the importance of determining meaningful cognitive change following surgery for TLE far outweighs the few additional calculations required.

CHAPTER 5

General Discussion

The current study used a local, unoperated TLE sample to derive RCIs and SRB change norms for determining cognitive change in Australian epilepsy patients. Previous research has cautioned against the application of RCI and SRB change scores derived from non-local demographic data (Martin et al., 2002b; Sawrie et al., 1996) and the current results support this recommendation. Nonetheless, the Australian and North American RCIs and SRBs produced identical, or highly similar, classifications for several of the WAIS-III and WMS-III indices. This indicated that the North American change norms are adequately generalisable to an Australian post-operative sample on certain cognitive domains (see page 96 for further discussion). However, a comprehensive pre- and post-operative neuropsychological assessment should investigate more than simply visual and verbal memory. Rather, such assessments should evaluate a person's general intellectual capacity, learning and retention, language abilities and executive functioning, as well as their affective functioning *(*Baker & Goldstein, 2004; Thompson, 2003). As a result, it is recommended that local change norms are most appropriate for determining cognitive change following surgical intervention for TLE in Australian patients.

The current study compared the adequacy of both RCIs and SRB change norms in determining cognitive change in a post-operative TLE sample and found little difference in the resulting classification rates, with the exceptions of PSI and the AD memory index (see Tables 28 and 29). With the latter caveats in mind, this is a promising finding, as RCIs require less complex statistical calculations than SRBs. RCIs can be generated simply by using the means and standard deviations of a cognitive measure administered on two occasions, as well as the test's stability coefficient (Chelune, 2003). However, given the differences in the classification rates for these two indices, it is recommended SRBs also be generated for these indices for comparative purposes and to ensure the most conservative estimate of post-operative decline is made available to clinicians and patients deciding on surgical intervention.

Findings and Outcomes from the Pre-operative Sample

RCIs and SRB change norms have clear clinical utility for any setting where determining cognitive change over time is required. For epilepsy patients, these methodologies are able to accurately investigate any perceived or actual changes over time, even without progression to surgery. Consistent with previous research, the results of the current study found evidence of a decline in confrontational naming skills was associated with longer retest intervals (Thompson & Duncan, 2005). However, no further evidence of cognitive decline associated with continuing seizures was uncovered. This finding is at odds with several previous studies which have reported a relationship between ongoing seizures and declines in memory, attention and speed of processing domains (Helmsteadter et al., 2003; Hermann et al., 2006; Holmes et al., 1998; Marques et al., 2007; Oyegbile et al., 2004; Seidenberg et al., 2007; Seidenberg, O'Leary, Berent, & Boll, 1981). However, several studies have also suggested seizures are not associated with ongoing cognitive decline in TLE patients (Rausch et al., 2003; Selwa et al., 1994). Please refer to page 69 of Chapter Three for a more detailed discussion.

The SRB change methodology used a stepwise regression approach, where the baseline scores of a control group were regressed against their retest scores to create a formula for predicting retest scores (McSweeny et al., 1993). This method allowed for consideration of multiple predictors of retest scores and so other factors which may affect retest performance were also included in the regression model. Linear regression using only baseline scores provides reasonable predictive accuracy when looking at homogenous populations with average baseline performance levels (Levine et al., 2004; Temkin et al., 1999; Salinsky, Storzbach, Dodrill, & Binder, 2001). However, epilepsy patients are not a homogeneous group – salient predictor variables include seizure characteristics and baseline performance – and so the multiple regression method was employed in the current study.

Consistent with the literature, a person's baseline performance was overwhelmingly the strongest predictor of their score on retest (Dikmen et al., 1999; Heaton et al., 2001; Hermann et al., 1996; McSweeny et al., 1993; Martin et al., 2002b; Sawrie et

al., 1996; Sherman et al., 2003; Temkin et al., 1999). Other significant variables included age, education, seizure duration and number of AEDs, although these variables accounted for less than 9% of the statistical variance. This result is similar to that of Martin et al. (2002), who reported these variables accounted for less than 7% of the variance in their study. Interestingly, education did not meet the statistical significance for inclusion for any of the regression equations. Research suggests a higher level of education is associated with higher baseline performance and "cognitive reserve" (Jokeit & Ebner, 1999; Oyegbile et al., 2004) and it may be that the predictive contribution of education level was absorbed by the strong predictive value of baseline performance in the current study.

90% versus 95% confidence intervals: Drawing a line in the sand

Although it is conventional to use 90% confidence intervals to determine significant change, it is not obligatory and the current study calculated 90% and 95% confidence intervals for both the RCI and SRB change approaches. Previous studies have also reported change data for 70% and 80% confidence intervals (Hermann et al., 1996). The stability of a test and the criterion chosen for change may result in cut-offs which are either too liberal or too conservative, resulting in a high number of false positives or false negatives, respectively (Sawrie et al., 1996). For example, the practice-adjusted RCI model proposed by Chelune and colleague (1993) and revised by Iverson (2001) has been criticised for increasing the risk of a Type I statistical error due to multiple comparisons (Keith, Puente, Malcolmson, Tartt, Coleman et al., 2002). However, the clinician can partly control this risk by adjusting the confidence intervals accordingly (Woods et al., 2006). The question of which confidence intervals to apply is inherently tied to Type I and II statistical error rates and the clinical implications of false positive or false negative errors (Sato, 1996; Woods et al., 2006).

In the current study, the 90% confidence intervals were found to provide classifications of change which largely adhered to the expected theoretical distribution of retest scores in the pre-operative group (i.e., 5% declined, 90% unchanged, 5% improved). In contrast, the 95% confidence intervals were at times too conservative

and did not identify change in either direction on retest. In these instances, the risk of making a Type II error increased and the use of 90% confidence intervals was therefore preferred. However, decreasing the risk of a Type II error in this way may have resulted in higher rates of misclassification of change on retest. Subsequently, in deciding which confidence intervals to employ, it is important to consider the clinical implications of identifying change where there is none, or over-looking genuine cognitive change.

As an example, say a person's baseline IQ score on the Auditory Immediate Index is 90 and they decline 22 points on this measure following surgery. According to the 95% RCI confidence intervals, this decline would not be deemed clinically significant. However, the person in question would likely notice differences in their daily functioning and not validating their experience may be counterproductive. In addition to this, the person would be misclassified as unchanged, leading to inaccurate estimates of cognitive morbidity for clinicians and patients deciding whether to proceed to surgery. The reciprocal argument might also be made, wherein a person declines 22 points on retest but notices no functional change in their day-to-day lives. In this instance, informing them they have experienced a significant deterioration in their verbal memory skills is unlikely to convey any benefits and may instead result in heightened anxiety. These examples reflect the importance of considering a statistically significant change in a person's test results against the broader context of all postoperative outcomes, as well as their day-to-day functioning, in order to ensure the result is also clinically significant, with adequate ecological validity (Baxendale & Thompson, 2005).

Ecological validity refers to the degree to which neuropsychological test performance reflects a person's real world functioning (Chaytor, Temkin, Machamer, & Dikmen, 2007). The relationship between neuropsychological tests and measures of everyday functioning is moderate and influenced by varying factors, such as emotional issues, premorbid functioning, motor abilities and health-related issues (Chaytor & Schmitter-Edgecombe, 2003; Sbordone & Guilmette, 1999). At times in the current study, the difference between the 90% and 95% confidence intervals was only two to three IQ points, yet this small margin influenced who was, and was not, classified as

significantly changed on retest. The clinical and functional meaning of statistically significant cognitive change may depend on a person's ability to compensate for deficits, their environmental demands and their psychosocial resources (Baxendale et al., 2006). It will be important for future research to include measures of quality of life and possibly work performance measures to help determine the ecological validity of post-operative classifications of cognitive change.

Cognitive Change Following Surgery for TLE

The current study applied RCIs and SRB change norms to investigate post-operative cognitive outcomes following TLE surgery in a local sample. The left ATL group was expected to experience more frequent verbal memory decline than the right ATL group (Alpherts et al., 2006; Bauer et al., 1995; Chelune et al., 1991; Chelune et al., 1993; Helmstaedter, et al., 2002; Hermann et al., 1992; Ivnik et al., 1987; Lee et al., 2002; McSweeny et al., 1993; Seidenberg et al., 1998; Selwa et al., 1994), whilst it was anticipated that significant visual memory decline would occur in only a minority of the right ATL group. The results of the current study supported these hypotheses, with the left ATL group found to be more vulnerable to verbal memory deficits. Further, a clear pattern of visual memory deficits following right ATL was not evident, but rather this group also showed evidence of verbal memory decline. In addition to this, the rate of visual memory decline associated with right ATL (5.7%) was not markedly greater than expected levels (e.g., 5% decline).

This pattern of results is not unprecedented (e.g., Bauer et al., 1995; Dobbins, Kroll, Tulving, Knight, & Gazzaniga, 1998; Kneebone et al., 1995; Jones-Gotman et al., 1993; Gleissner et al., 1998; Helmstaedter et al., 2003; Hoppe et al., 2007; Ivnik et al., 1987; Piggott & Milner, 1993; Rausch et al., 2003; Selwa et al., 1994). In addition, several studies have reported significant memory improvements in a minority of patients following both left and right ATL (Baxendale et al., 2006; Baxendale et al., 2008; Helmstaedter et al., 2004; Martin et al., 1998; Paglioli et al., 2006; Wachi et al., 2001). The results raise questions about the continued clinical utility of the material-specific hypothesis, particularly for right-sided TLE.

While both the literature and the results of the current study support a relationship between dominant temporal lobe resection and verbal memory decline, a dearth of convincing evidence to support the converse argument – visual memory decline following non-dominant ATL – necessitates consideration of alternative theories to the traditional material-specific model. Memory assessment tasks themselves have come under scrutiny and it has been suggested the means by which memory is assessed is crucial for identifying deficits (Baxendale, 2008). Dulay and colleagues (Dulay, Levin, York, Mizrahi, Verma, et al., 2009) reported different visual memory outcomes following right ATL, depending on the type of visual memory measure used and suggested memory for spatial locations as particularly vulnerable to decline following surgery.

Although the WMS-III verbal memory tasks appear sensitive to left temporal lobe dysfunction, visual memory tasks – including the visual reproduction subtests not used in this study – appear to have limited lateralising value and therefore contribute little to diagnosis (Baker, Austin & Downes, 2003; Lee et al., 2002; Wilde, Strauss, Chelune, Hermann, Hunter, et al., 2003). For example, a person may complete a nonverbal task – such as remembering geometric shapes – using a verbal labelling strategy (Baxendale, 2008; Helmstaedter et al., 2004). Consequently, it has been suggested that pre- and post-operative neuropsychological assessment should include both learning and retention tests, as well as using nonverbal tasks which are not easily verbalised, wherever possible (Lee et al., 2002). The latter represents an important area for future research, particularly in the context of developing assessment paradigms for use with functional magnetic resonance imaging (fMRI).

Using the Number Needed to Harm and Number Needed to Treat

An important role of the pre-operative neuropsychological assessment is to inform surgical candidates of the potential cognitive risks and advantages associated with surgical treatment for TLE. In order to clearly convey the differences between medical (i.e., AEDs) and surgical treatment, the NNH and NNT values were also calculated for the current study. NNH and NNT values help to determine the clinical significance – or effect size – of a difference in outcome (Citrome, 2007). As

discussed above, the results of the current study indicated verbal memory was most vulnerable to post-operative decline in patients undergoing left ATL and this was reflected by small NNH values, for both the RCI and SRB methods (-7 and -6 respectively). These values indicate that for every six or seven people who undergo a left ATL, one will experience a significant decline in their verbal memory.

Where NNH values are very disparate, depending on the statistical change method employed, it is suggested the most conservative value be utilised for clinical decision-making purposes. For example, for the AD index, the RCI approach generated a NNH value of -200, whilst the SRB change norms calculated a NNH value of -8. With such high cognitive stakes, it is recommended that the lower NNH value – which equates to a greater risk of harm – be the value upon which clinical decisions and patient information are based.

Strengths and Limitations of the Current Study

A strength of the current study was its generation of cognitive change norms for TLE patients, utilising evidence-based statistical methodologies and locally collected data. A further strength was the calculation of base-rates of pre- and post-operative cognitive change, which provides data on the incidence of adverse cognitive outcomes following surgery for TLE. This, in turn, enables important information regarding the potential cognitive sequelae of TLE surgery to be communicated to patients, caregivers and clinicians (Helmstaedter & Kurthen, 2001).

The results of the current study are tempered by a limitation in the comparison of the local and US change norms. The criterion chosen to indicate adequate agreement between the two sets of change classifications – based on Cohen's kappa coefficient – is essentially arbitrary and therefore vulnerable to misinterpretation (Brennan & Silman, 1992). For example, choosing a low criterion for agreement would mean that more widely differing rates of change classification between the Australian and North American data are accepted as similar enough to generalise across demographic backgrounds. However, to reduce the risk of a Type II error, a conservative criterion was chosen for the current study, resulting in similar classifications of change on only

a handful of indices. Nonetheless, statistical methods for comparing different sets of data derived from the same sample are limited and future studies may wish to pursue alternative approaches to this statistical challenge.

The pursuit of psychometric best-practice for analysing change over time continues with Maassen and colleagues only very recently suggesting a modified formula for calculating reliable change (Maassen, Bossema, & Brand, 2008). They criticised previous methods (e.g., Chelune et al., 1993; McSweeny et al., 1996) for calculating a standard error which does not account for fluctuations due to sampling, therefore resulting in estimation intervals which are too small. Further to this, the methodology of McSweeny et al. was described as too lenient, incorrectly designating patients as improved or declined. The current study utilised the RCI-adjusted formula proposed by Iverson (2001), and based on the formula of Chelune et al. (1993), which calculated the standard error of measurement at both test and retest. The current investigation also used the SRB formula suggested by McSweeny et al. (1996) but used stepwise, rather than linear, regression analyses to allow for the inclusion of multiple predictor variables. As a result, should Maassen et al.'s criticisms prove founded, the current study may also suffer from the outlined limitations. Future research may consider further comparisons of the varying RCI and SRB formulas cited in the literature.

A final limitation of the current investigation was that it did not include mood at the time of neuropsychological assessment as a potential predictor variable. The effects of mood on cognition and AEDs on mood are well-recognised (Helmstaedter et al., 2004; Hixson & Kirsch, 2009; Loring et al., 2007; Schmitz, 2006) and characterising the relationship between cognitive performance, affective functioning and seizure aetiology represents an important area for future research. This recommendation is particularly salient given 56% of the pre-operative sample reported concerns of a decline in their memory abilities since the initial assessment.

Future Directions

Advances in neuroimaging and electrophysiological techniques have largely overtaken the lateralisation and localisation role historically ascribed to pre-surgical neuropsychological assessment. This is not unreasonable given the poor lateralising value of neuropsychological measures, discussed above, and suggests a new framework for the role of neuropsychology in epilepsy surgery settings is needed. To insist on maintaining merely a localisation role for neuropsychology in this setting risks obsolescence in the face of more accurate and specialised localising approaches, such as volumetric MRI and fMRI. In addition to this, the relatively stable performance of patients on the majority of IQ measures pre- to post-operatively potentially raises the question of whether it is necessary to include IQ measures in a pre-operative assessment battery.

However, the contribution of neuropsychology to the epilepsy surgery setting is far broader than simply identification, diagnosis and prediction of cognitive deficits. It provides a baseline measure of cognitive, behavioural and affective functioning against which cognitive change – due either to surgery, medication or other, psychosocial factors – can be evaluated. Given the increasing evidence against a simple verbal-nonverbal information processing dichotomy, it becomes increasingly important to assess widespread cognitive functioning and move beyond a wholly temporal lobe focus (Barr, 2007). As a result, a comprehensive assessment across cognitive domains, as well as affective and psychosocial functioning, remains clearly warranted. Finally, the advances in functional neuroimaging techniques offer an exciting opportunity for neuropsychologists to be involved in the development of appropriate and targeted assessment paradigms.

Neuropsychological assessment in the epilepsy surgery setting also plays an important role in providing pre-operative counselling – including potential risks, advantages, outcomes and adjustment issues – to patients, families and clinicians. With this in mind, the change methodologies and NNH/NNT values discussed above are an important tool for communicating the potential risks and benefits to patients considering surgical treatment for TLE. Moreover, the neuropsychologist is solely

responsible for measuring post-operative cognitive change and this is most accurately accomplished using the change methodologies employed in the current study. Post-operative cognitive outcomes reflect everyday realities and, in the event of decline, the neuropsychologist can offer tailored recommendations and compensatory strategies informed by the patient's strengths, weaknesses, psychosocial supports, and personality style.

The cognitive change approaches employed in the current study have clear clinical utility for determining cognitive change over time in epilepsy patients, regardless of whether they are surgical candidates. In addition to this, the potential application of RCIs and SRBs extend beyond the epilepsy setting and can provide a valuable clinical tool in any setting where the assessment of cognitive change over time is required.

References

Aikia, M., Salmenpera, T., Partanen, K., & Kalvianen, R. (2001). Verbal memory in newly diagnosed patients and patients with chronic left temporal lobe epilepsy. *Epilepsy and Behavior, 2*, 20-27.

Aldenkamp, A. P., & Arends, J. (2004). Effects of epileptiform EEG discharges on cognitive function: Is the concept of "transient cognitive impairment" still valid? *Epilepsy and Behavior, 5*(Sup 1), S25-S34.

Aldenkamp, A. P., Baker, G. A., & Meador, K. J. (2004). The neuropsychology of epilepsy: What are the factors involved? *Epilepsy and Behavior, 5*(Sup 1), S1-S2.

Aldenkamp, A. P., De Krom, M., & Reijs, R. (2003). Newer antiepileptic drugs and cognitive issues. *Epilepsia, 44*(supp 4), 21-29.

Aldenkamp, A. P., Weber, B., Overweg-Plandsoen, W. C. G., Reijs, R., & van Mil, S. (2003). Educational underachievement in children with epilepsy: A model to predict the effects of epilepsy on educational achievement. *Journal of Child Neurology, 20*, 175-180.

Alessio, A., Damasceno, B. P., Camargo, C. H. P., Kobayashi, P., Guerreiro, C. A. M., & Cendes, F. (2004). Differences in memory performance and other clinical characteristics in patients with mesial temporal lobe epilepsy with and without hippocampal atrophy. *Epilepsy and Behavior, 5*, 22-27.

Alpherts, W. C. J., Vermeulen, J., van Rijen, P. C., Lopes da Silva, F. H., & van Veelen, C. W. M. (2006). Verbal memory decline after temporal epilepsy surgery? *Neurology, 67*, 626-631.

Andersson-Roswall, L., Engman, E., Samuelsson, H., Sjoberg-Larson, C., & Malmgren, K. (2004). Verbal memory decline and adverse effects on cognition in adult patients with pharmacoresistant partial epilepsy: A longitudinal controlled study of 36 patients. *Epilepsy and Behavior, 5*, 677-686.

Armon, C., Radtke, R. A., Friedman, A. H. & Dawson, D. V. (1996). Predictors of outcome of epilepsy surgery: Multivariate analysis with validation. *Epilepsia, 37,* 814-821.

Atkinson, L. (1991). Three standard errors of measurement and the WMS-R. *Psychological Assessment, 3,* 136-138.

Australian Psychological Society. (September, 2007). *Code of ethics.* Retrieved May 12, 2009, from http://www.psychology.org.au/about/ethics/

Banos, J. H., LaGory, J., Sawrie, S., Faught, E., Knowlton, R., Prasad, A., Kuzniecky, R., & Martin, R. C. (2004). Self-report of cognitive abilities in temporal lobe epilepsy: Cognitive, psychosocial and emotional factors. *Epilepsy and Behavior, 5,* 575-579.

Baker, G. A., Austin, N. A., & Downes, J. J. (2003). Validation of the Wechsler Memory Scale-III in a population of people with intractable temporal lobe epilepsy. *Epilepsy Research, 53,* 201-206.

Baker, G. A., & Goldstein, L. H. (2004). The dos and don'ts of neuropsychological assessment in epilepsy. *Epilepsy and Behavior, 5,* s77-s80.

Barker, C., Pistrang, N., & Elliot, R. (2002). *Research Methods in Clinical Psychology: An Introduction for Students and Practitioners.* (2nd Edition ed.). London: John Wiley and Sons, Ltd.

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology, 34,* 215-220.

Barr, W. B. (2007). Epilepsy and neuropsychology: Past, present, and future. *Neuropsychology Review, 17,* 381-383.

Bauer, R. M., Breier, J., Crosson, B., Gilmore, R., Fennell, E. B. & Roper, S. (1995). Neuropsychological functioning before and after unilateral temporal lobectomy

for intractable epilepsy. *Journal of the International Neuropsychological Society, 1*, 362.

Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment*, *82*, 60-70.

Baxendale, S. A., van Paesschen, W., Thompson, P. J., Connelly, A., Duncan, J. S., Harkness, & Shorvon, S. D. (1998). The relationship between quantitative MRI and neuropsychological functioning in temporal lobe epilepsy. *Epilepsia, 39*, 158-166.

Baxendale, S. & Thompson, P. J. (2005). Defining meaningful postoperative change in epilepsy surgery patients: Measuring the immeasurable? *Epilepsy and Behavior, 6*, 207-211.

Baxendale, S., Thompson, P. J. & Duncan, J. S. (2008). Improvements in memory function following anterior temporal lobe resection for epilepsy. *Neurology, 71*, 1319-1325.

Baxendale, S., Thompson, P., Harkness, W., & Duncan, J. (2006). Predicting memory decline following epilepsy surgery: A multivariate approach. *Epilepsia, 47*, 1887-1894.

Bell, B. D. & Davies, K. G. (1998). Anterior temporal lobectomy, hippocampal sclerosis, and memory: Recent neuropsychological findings. *Neuropsychology Review, 8*, 25-41.

Bell, B. D., Davies, K. G., Hermann, B. P., & Walters, G. (2000). Confrontation naming after anterior temporal lobectomy is related to age of acquisition of the object names. *Neuropsychologia, 38*, 83-92.

Benton, A. L., Hammer, K., & Sivan, A. B. (1983). *Multilingual Aphasia Examination.* Iowa City: AJA Associates.

Berkovic, S. F., Andermann, F., Olivier, A., Ethier, R., Melanson, D., Robitaille, Y., Kuzniecky, R., Peters, T., & Feindel, W. (1991). Hippocampal sclerosis in temporal lobe epilepsy demosntrated by magnetic resonance imaging. *Annals of Neurology, 29,* 175-182.

Bernasconi, N., Bernasconi, A., Caramanos, Z., Antel, S. B., Andermann, F., & Arnold, D. L. (2003). Mesial temporal damage in temporal lobe epilepsy: A volumetric MRI study of the hippocampus, amygdala and parahippocampal region. *Brain, 126,* 462-469.

Binnie, C. D., & Polkey, C. E. (2000). Commission on neurosurgery of the international league against epilepsy (ILAE) 1993-1997: Recommended standards. *Epilepsia, 41,* 1346-1349.

Bonilha, L., Rorden, C. Castellano, G., Cendes, F., & Li, L. M. (2005). Voxel-based morphometry of the thalamus in patients with refractory medial temporal lobe epilepsy. *Neuroimage, 25,* 1016-1021.

Bopp, K. L. & Verhaegen, P. (2005). Ageing and verbal memory span: A meta-analysis. *Journal of Gerontology, 5,* 223-233.

Bowden, S. C., Lissner, D., McCarthy, K. A. L., Weiss, L. G., & Holdnack, J. A. (2007). Metric and structural equivalence of core cognitive abilities measured with the Wechsler Adult Intelligence Scale-III in the United States and Australia. *Journal of Clinical and Experimental Neuropsychology, 29,* 768-780.

Brand, A. N., Jolles, J., & Gispen-de Weid, C. (1992). Recall and recognition memory deficits in depression. *Journal of Affective Disorders, 25,* 77-86.

Brennan, P., & Silman, A. (1992). Statistical methods for assessing observer variability in clinical measures. *British Medical Journal, 304,* 1491-1494.

Bruggemans, E. F., Van de Vijver, F. J. R., & Huysmans, H. A. (1997). Assessment of cognitive deterioration in individual patients following cardiac surgery: Correcting for measurement error and practice effects. *Journal of Clinical and Experimental Neuropsychology, 19,* 543-559.

Camara, W. J., Nathan, J. S., & Puente, A. E. (2000). Psychological test usage: Implications in professional psychology. *Professional Psychology: Research and Practice, 31,* 141-154.

Chaytor, N. & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review, 13*, 181-197.

Chaytor, N., Temkin, N., Machamer, J., & Dikmen, S. (2007). The ecological validity of neuropsychological assessment and the role of depressive symptoms in moderate to severe traumatic brain injury. *Journal of the International Neuropsychological Society, 13*, 377-385.

Chelune, G. J. (1994). The role of neuropsychological assessment in the presurgical evaluation of the epilepsy candidate. In A. R. Wyler & B. P. Hermann (Ed.), *The Surgical Management of Epilepsy.* (pp. 78-89). Boston: Butterworth-Heinemann.

Chelune, G. J. (1995). Hippocampal adequacy versus functional reserve: Predicting memory functions following temporal lobectomy. *Archives of Clinical Neuropsychology, 10*, 413-432.

Chelune, G. J. (2003). Assessing reliable neuropsychological change. In R. D. Franklin (Ed.), *Prediction in Forensic and Neuropsychology: Sound Statistical Practices.* (pp. 123-147). New Jersey: Lawrence Erlbaum Associates.

Chelune, G. J. & Goldstein, G. (1991). Interpreting test-retest changes in neuropsychological practice. *The Clinical Neuropsychologist, 5*, 262-264.

Chelune, G. J., Naugle, R. I., Luders, H. & Awad, I. A. (1991). Prediction of cognitive change as a function of preoperative ability level among temporal lobectomy patients at six months follow-up. *Neurology, 41*, 399-404.

Chelune, G. J., Naugle, R. I., Luders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology, 7*, 41-52.

Chelune, G. J., Sands, K., Barrett, J., Naugle, R. I., Ledbetter, M., & Tulsky, D. (1999). Test-retest characteristics and measures of meaningful change for the Wechsler Memory Scale-III. *Journal of the International Neuropsychological Society,* (5), 109.

Christensen, H. (2001). What cognitive changes can be expected with normal ageing? *Australian and New Zealand Journal of Psychiatry, 35*(6), 768-775.

Citrome, L. (2007). Show me the evidence: Using number needed to treat. *Southern Medical Journal, 100*, 881-884.

Clusmann, H., Schramm, K., Kral, T., Helmstaedter, T., Ostertun, B., Fimmers, R., Haun, D., & Elger, C. E. (2002). Prognostic factors and outcome after different types of resection for temporal lobe epilepsy. *Journal of Neurosurgery, 97,* 1131-1141.

Cohen, R. J. & Swerdlik, M. E. (2005). *Psychological Assessment and Testing.* (6th ed.). New York: McGraw-Hill.

Collie, A., Maruff, P., Makdissi, M., McStephen, M., Darby, D. G., & McCrory, P. (2004). Statistical procedures for determining the extent of cognitive change following concussion. *British Journal of Sports Medicine, 38,* 273-278.

Cook, R. J. (1995). The number needed to treat: A clinically useful measure of treatment effect. *British Medical Journal, 310,* 452-454.

Cook, M. J., Fish, D. R., Shorvon, S. D., Straughan, K., & Stevens, J. M. (1992). Hippocampal volumetric and morphometric studies in frontal and temporal lobe epilepsy. *Brain, 115*, 1001-1015.

Cull, C. G., & Goldstein, L. H. (1997). *The Clinical Psychologist's Handbook of Epilepsy: Assessment and Management.* (1st edition ed.). London: Routledge.

Davies, K. G., Bell, B. D., Bush, A. J., & Wyler, A. R. (1998). Prediction of verbal memory loss in individuals after anterior temporal lobectomy. *Epilepsia, 39,* 820-828.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the scottish mental surveys of 1932 and 1947. *Journal of Personality and Social Psychology, 86*, 130-147.

Dennis, M. (2000). Developmental plasticity in children: The role of biological risk, development, time, and reserve. *Journal of Communication Disorders, 33*, 321-331.

Dikmen, S. S., Heaton, R. K., Grant, I., & Temkin, N. R. (1999). Test-retest reliability and practice effects of expanded Halstead-Reitan neuropsychological test battery. *Journal of the International Neuropsychological Society, 5*, 346-356.

Dikmen, S. S., Machamer, J., Temkin, N., & McLean, A. (1990). Neuropsychological recovery in patients with moderate to severe head injury: 2 year follow-up. *Journal of Clinical and Experimental Neuropsychology, 12*, 507-519.

Dobbins, I. G., Kroll, N. E. A., Tulving, E., Knight, R. J., & Gazzaniga, M. S. (1998). Unilateral medial temporal lobe impairment: Type deficit, function deficit, or both? *Neuropsychologia, 36*, 115-127.

Dodrill, C. B. (2002). Progressive cognitive decline in adolescents and adults with epilepsy. *Progress in Brain Research, 135*, 399-407.

Dodrill, C. B. (2004). Neuropsychological effects of seizures. *Epilepsy and Behavior, 5*(Sup 1), S21-S24.

Dodrill, C. B., Hermann, B. P., Rausch, R., Chelune, G. J., & Oxbury, S. (1993). Neuropsychological testing for assessing prognosis following surgery for epilepsy. In J. J. Engel (Ed.), *Surgical Treatment of the Epilepsies.* (2nd ed., pp. 272-284).

Dodrill, C. B., Jones-Gotman, M., Loring, D. W., & Sass, K. J. (1993). Contributions of neuropsychology. In J. J. Engel (Ed.), *Surgical Ttreatment of the Epilepsies.* (2nd ed., pp. 263-271). New York: Raven Press, Ltd.

Dodrill, C. B., & Troupin, A. S. (1975). Effects of repeated administration of a comprehensive neuropsychological battery among chronic epileptics. *Journal of Nervous and Mental Disease, 161*, 185-190.

Dudek, F. J. (1979). The continuing misrepresentation of the standard error of measurement. *Psychological Bulletin, 86*, 335-337.

Dulay, M. F., Levin, H. S., York, M. K., Mizrahi, E. M., Verma, A., Goldsmith, I., Grossman, R. G., & Yoshor, D. (2009). Predictors of individual visual memory decline after unilateral anterior temporal lobe resection. *Neurology, 72*, 1837-1842.

Edwards, D. W., Yarvis, R. M., Mueller, D. P., Zingale, H. C., & Wagman, W. J. (1978). Test-taking and the stability of adjustment scales: Can we assess patient deterioration? *Education Quarterly, 2*, 275-292.

Elger, C. E., Helmstaedter, C., Kurthen, M. (2004). Chronic epilepsy and cognition. *Lancet Neurology, 3*, 663-672.

Engel, J. J., Van Ness, P. C., Rasmussen, T. B., & Ojemann, L. M. (1993). Outcome with respect to seizures. In J. J. Engel (Ed.), *Surgical Treatment of the Epilepsies.* (2nd ed. ed., pp. 609-621). New York: Raven Press.

Engel, J., & Shewmon, D. A. (1998). Who should be considered a surgical candidate? In J. Engel (Ed.), *Surgical Treatment of the Epilepsies.* (2nd ed., pp. 23-34). New York: Raven Press.

Engel, J., Wiebe, S., French, J., Sperling, M., Williamson, P, et al. (2003). Practice parameter: Temporal lobe and localized neocortical resections for epilepsy. *Neurology, 60*, 538-547.

Engman, E., Andersson-Roswall, L., Samuelsson, H., & Malmgren, K. (2006). Serial cognitive change patterns across time after temporal lobe resection for epilepsy. *Epilepsy and Behavior, 8*, 765-772.

Evans, C., Margison, F., & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence Based Mental Health, 1*, 70-72.

Fabinyi, G. C. A. (2002). Surgery for epilepsy. *The Medical Journal of Australia, 176*, 410-411.

Fish, D. R. (1998). Epilepsy: Epidemiology, presentation and classification. In F. Scaravilli (Ed.), *Neuropathology of Epilepsy* (pp. 11-43). Singapore: World Scientific Publishing Company.

Frerichs, R. J., & Tuokko, H. A. (2006). Reliable change scores and their relation to perceived change in memory: Implications for the diagnosis of mild cognitive impairment. *Archives of Clinical Neuropsychology, 21*, 109-116.

Fuerst, D., Shah, J., Shah, A., & Watson, C. (2003). Hippocampal sclerosis is a progressive disorder: Longitudinal volumetric MRI study. *Annals of Neurology, 53*, 413-416.

Gardener, W. (1995). On the reliability of sequential data: Measurement, meaning, and correction. In John M. Gottman (Ed.), *The Analysis of Change.* (pp. 339-359). Mahwah, N. J.: Erlbaum.

Gleissner, U., Helmstaedter, C., & Elger, C. E. (1998). Right hippocampal contribution to visual memory: A presurgical and postsurgical study in patients with temporal lobe epilepsy. *Journal of Neurology, Neurosurgery, and Psychiatry, 65*, 665-669.

Glisky, E. L. (2007). Changes in cognitive function in human aging. In D. R. Riddle (Ed.), *Brain aging: Models, Methods, and Mechanisms* (pp. 3-20). Boca Raton, FL, US: CRC Press.

Glosser, G., Cole, L. C., French, J. A., Saykin, A. J., & Sperling, M. R. (1997). Predictors of intellectual performance in adults with intractable temporal lobe epilepsy. *Journal of the International Neuropsychological Society, 3*, 252-259.

Glosser, G., Deutsch, G. K., Cole, L. C., Corwin, J., & Saykin, A. J. (1998). Differential lateralisation of memory discrimination and response bias in temporal lobe epilepsy patients. *Journal of the International Neuropsychological Society, 4*, 502-511.

Glutting, J., McDermott, P., Watkins, M., Kush, J., & Konold, T. (1997). The base rate problem and its consequences for interpreting children's ability profiles. *School Psychology Review, 26*, 176-188.

Golby, A. J., Poldrack, R. A., Brewer, J. B., Spencer, D., Desmond, J. E., & Gabrieli, J. D. E. (2001). Material-specific lateralization in the medial temporal lobe and prefrontal cortex during memory encoding. *Brain, 124*, 1851-1854.

Griffith, H. R., Pyzalski, R. W., Seidenberg, M., & Hermann, B. P. (2004). Memory relationships between MRI volumes and resting PET metabolism of medial temporal lobe structures. *Epilepsy and Behavior, 5*, 669-676.

Groth-Marnat, G. (2003). *Handbook of Psychological Assessment.* (4th ed.). New Jersey: John Wiley & Sons, Inc.

Hagemann, W. J. J. M. & Arindell, W. A. (1993). A further refinement of the reliable change (RC) index by *i*mproving the pre-post *d*ifference score: Introducing RD_{ID}. *Behavior Research and Therapy, 31*, 693-700.

Hauser, W. A. & Annegers, J. F. (1983). Epidemiology of epilepsy. In Laidlaw, J., Richens, A., & Chadwich, D. (Ed.), *A Textbook of Epilepsy.* Edinburgh: Churchill Livingstone.

Heaton, R. K., Temkin, N., Dikmen, S., Avitable, N., Taylor, M. J., Marcotte, T. D., & Grant, I. (2001). Detecting change: A comparison of three neuropsychological methods, using normal and clinical samples. *Archives of Clinical Neuropsychology, 16*, 75-91.

Helmstaedter, C. (2004). Neuropsychological aspects of epilepsy surgery. *Epilepsy and Behavior, 5*, 545-555.

Helmstaedter, C. (2005). Effects of chronic temporal lobe epilepsy on memory functions. In Arzimanoglou, A., Aldenkamp, A., Cross, H., Lassonde, M., Moshe, S., & Schmitz, B. (Ed.), *Cognitive Dysfunction in Children with Temporal Lobe Epilepsy.* (pp. 13-30). Montrouge: John Libbey Eurotext.

Helmstaedter, C. & Elger, C. E. (1996). Cognitive consequences of two-thirds anterior temporal lobectomy on verbal memory in 144 patients: A three-month follow-up study. *Epilepsia, 37,* 171-180.

Helmstaedter, C. & Elger, C. E. (1998). Functional plasticity after left anterior temporal lobectomy: Substitution and compensation of verbal memory impairment. *Epilepsia, 39,* 399-408.

Helmstaedter, C. & Kurthen, M. (2001). Memory and epilepsy: Characteristics, course, and influence of drugs and surgery. *Current Opinion in Neurology, 14*, 211-216.

Helmstaedter, C., & Kockelmann, E. (2006). Cognitive outcomes in patients with chronic temporal lobe epilepsy. *Epilepsia, 47*(Supp. 2), 96-98.

Helmstaedter, C., Brosch, T., Kurthen, M., & Elger, C. E. (2004). The impact of sex and language dominance on material-specific memory before and after left temporal lobe surgery. *Brain, 127,* 1518-1525.

Helmstaedter, C., Grunwald, T., Lehnertz, K., Gleissner, U., Elger, C. E. (1997). Differential involvement of left temporolateral and temporomesial structures in verbal declarative learning and memory: Evidence from temporal lobe epilepsy. *Brain Cognition, 35*, 110-131.

Helmstaedter, C., Kurthen, M., Lux, S., Johansen, K., Quiske, A., Schramm, J. & Elger, C. E. (2000). Long-term clinical, neuropsychological, and psychosocial follow-up in surgically and nonsurgically treated patients with drug-resistant temporal lobe epilepsy. *Nervenarzt, 71*, 629-642.

Helmstaedter, C., Kurthen, M., Lux, S., Reuber, M., & Elger, C. E. (2003). Chronic epilepsy and cognition: A longitudinal study in temporal lobe epilepsy. *Annals of Neurology, 54,* 425-432.

Helmstaedter, C., Pohl, C., & Elger, C. E. (1995). Relations between verbal and nonverbal memory performance: Evidence of confounding effects particularly in patients with right temporal lobe epilepsy. *Cortex, 31,* 345-355.

Helmstaedter, C., Reuber, M., & Elger, C. C. E. (2002). Interaction of cognitive aging and memory deficits related to epilepsy surgery. *Annals of Neurology, 52,* 89-94.

Henry, T. R., Maziotta, J. C., & Engel, J., Jr. (1993). Interictal metabolic anatomy of mesial temporal lobe epilepsy. *Archives of Neurology, 50*, 582-589.

Hermann, B. P., Seidenberg, & Bell, B. (2000). Psychiatric comorbidity in chronic epilepsy: Identification, consequences, and treatment of major depression. *Epilepsia, 41*(supp 2), s31-s41.

Hermann, B. P., Seidenberg, M., & Bell, B. (2002). The neurodevelopmental impact of childhood onset temporal lobe epilepsy on brain structure and function and the risk of progressive cognitive effects. *Progress in Brain Research, 135*, 429-438.

Hermann, B. P., Seidenberg, M., Dow, C., Jones, J., Rutecki, P., Bhattacharya, M. S., & Bell, B. (2006). Cognitive prognosis in chronic temporal lobe epilepsy. *Annals of Neurology, 60*, 80-87.

Hermann, B. P., Seidenberg, M., Lee, E., Chan, F., & Rutecki, P. (2007). Cognitive phenotypes in temporal lobe epilepsy. *Journal of the International Neuropsychological Society, 13*, 12-20.

Hermann, B. P., Seidenberg, M., Schoenfeld, J., & Davies, K. (1997). Neuropsychological characteristics of the syndrome of mesial temporal lobe epilepsy. *Archives of Neurology, 54*, 369-376.

Hermann, B. P., Seidenberg, M., Schoenfeld, J., Peterson, J., Leveroni, C., & Wyler, A. R. (1996). Empirical techniques for determining the reliability, magnitude,

and pattern of neuropsychological change after epilepsy surgery. *Epilepsia, 37*, 942-950.

Hermann, B. P., Wyler, A. R., Bush, A. J., & Tabatabai, F. R. (1992). Differential effects of left and right anterior temporal lobectomy on verbal learning and memory performance. *Epilepsia, 33*, 289-297.

Hermann, B., Seidenberg, M., & Jones, J. (2008). The neurobehavioural comorbidities of epilepsy: Can a natural history be developed? *The Lancet Neurology, 7*, 151-160.

Hermann, B., Seidenberg, M., Bell, B., Rutecki, P., Sheth, R., Ruggles, K., Wendt, G., O'Leary, D., & Magnotta, V. (2002). The neurodevelopmental impact of childhood-onset temporal lobe epilepsy on brain structure and function. *Epilepsia, 43*, 1062-1071.

Hixson, J. D., & Kirsch, H. E. (2009). The effects of epilepsy and its treatments on affect and emotion. *Neurocase (Epub ahead of print)*, 1-11.

Holmes, M. D., Dodrill, C. B., Wilkus, R. J., Ojemann, L. A. & Ojemann, G. M. (1998). Is partial epilepsy progressive? Ten-year follow-up of EEG and neuropsychological changes in adults with partial seizures. *Epilepsia, 39*, 1189-1193.

Hoppe, C., Elger, C. E., & Helmstaedter, C. (2007). Long-term memory impairment in patients with focal epilepsy. *Epilepsia, 48*, 26-29.

Howard Florey Institute. (2007). *Epilepsy.* Retrieved December 12, 2008, from http://www.florey.edu.au/the-brain/brain-disorders/epilepsy/

Hsu, L. M. (1989). Reliable changes in psychotherapy: Taking into account regression toward the mean. *Behavioral Assessment, 11*, 459-467.

Iverson, G. L. (2001). Interpreting change on the WAIS-III/WMS-III in clinical samples. *Archives of Clinical Neuropsychology, 16*, 183-191.

Ivnik, R. J., Sharbrough, F. W., & Laws, E. R. (1987). Effects of anterior temporal lobectomy on cognitive function. *Journal of Clinical Psychology, 43,* 128-137.

Ivnik, R. J., Sharbrough, F. W., & Laws, E. R. (1988). Anterior temporal lobectomy for control of complex partial seizures: Information for counseling patients. *Mayo Clinic Proceedings, 63,* 783-791.

Ivnik, R. J., Smith, G. E., Lucas, J. A., Petersen, R. C., Boeve, R. F., Kokmen, E., & Tanalos, E. G. (1999). Testing normal older people three or four times at 1- to 2-year intervals: Defining normal variance. *Neuropsychology, 13,* 121-127.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy. *Journal of Consulting and Clinical Psychology, 59,* 12-19.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15,* 336-352.

Janszky, J., Jokeit, H., Heinemann, D., Schulz, R., Woermann, F. G., & Ebner, A. (2003). Epileptic activity influences the speech organisation in medial temporal lobe epilepsy. *Brain, 126,* 2043-2051.

Jokeit, H., & Ebner, A. (1999). Long term effects of refractory temporal lobe epilepsy on cognitive abilities: A cross-sectional study. *Journal of Neurology, Neurosurgery, and Psychiatry, 67,* 44-50.

Jokeit, H., & Schacher, M. (2004). Neuropsychological aspects of type of epilepsy and etiological factors in adults. *Epilepsy and Behavior, 5*(Sup 1), S14-S20.

Jokeit, H., Ebner, A., Holthausen, H., Markowitsch, H. J., Moch, A., Pannek, H., Schulz, R. & Tuxhorn, I. (1997). Individual prediction of change in delayed recall of prose passages after left-sided anterior temporal lobectomy. *Neurology,* (49), 481-487.

Jones-Gotman, M., Brulot, M., McMackin, D., Cendes, F., Andermann, F., Evans, A. & Peters, T. (1993). Word and design list learning deficits related to side of hippocampal atrophy as assessed by volumetric MRI measurements. *Epilepsia, 34*, 71.

Jones-Gotman, M., Smith, M. L., & Zatorre, R. J. (1993). Neuropsychological testing for localizing and lateralizing the epileptogenic region. In J. J. Engel (Ed.), *Surgical Treatment of the Epilepsies.* (2nd ed., pp. 245-261). New York: Raven Press, Ltd.

Kaaden, S. & Helmstaedter, C. (2009). Age at onset of epilepsy as a determinant of intellectual impairment in temporal lobe epilepsy. *Epilepsy and Behavior, 15*, 213-217.

Kaplan, E. F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test.* (2nd ed.). Philadelphia: Lea & Febiger.

Kaufman, A. S. (2003,). Practice effects. Message posted to http://www.speechandlanguage.com/cafe/13.asp

Kay, G. and Kane, R. L. (1991). Repeated measures in neuropsychology: Use of serial testing to measure change in cognitive functioning. *Journal of Clinical and Experimental Neuropsychology, 13*, 49-50.

Keith, J. R., Puente, A. E., Malcolmson, K. L., Tartt, S., Colemen, A. E., & Marks, H. F. (2002). Assessing postoperative cognitive change after cardiopulmonary bypass surgery. *Neuropsychology, 16*, 411-421.

Kneebone, A. C., Andrew, M. J., Baker, R. A., & Knight, J. L. (1998). Neuropsychologic changes after coronary artery bypass grafting: Use of reliable change indices. *Annals of Thoracic Surgery, 65*, 1320-1325.

Kneebone, A. C., Chelune, G. J., Dinner, D., Awad, I. A. & Naugle, R. I. (1995). Use of the intracarotid amobarbital procedure to predict material specific memory change following anterior temporal lobectomy. *Epilepsia, 36*, 857-865.

Kneebone, A. C., Miller, L. A., Bowden, S. C., Lah, S., & Lee, G. P. (2008). The right temporal lobe and visuospatial memory: Experiences and insights from epilepsy surgery. *14th Annual Conference of the APS College of Clinical Neuropsychologists*. Adelaide, Australia.

Kwan, P. & Brodie, M. J. (2000). Early identification of refractory epilepsy. *New England Journal of Medicine, 342*, 314-319.

Kwan, P., & Brodie, M. J. (2001). Neuropsychological effects of epilepsy and antiepileptic drugs. *Lancet, 357*, 216-222.

Landis, J., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

Lange, R. T., Chelune, G. J., Taylor, M. J., Woodward, T. S., & Heaton, R. K. (2006). Development of demographic norms for four new WAIS-III/WMS-III indexes. *Psychological Assessment, 18*, 174-181.

Laupacis, A., Sackett, D. L., & Roberts, R. S. (1988). An assessment of clinically useful measures of the consequences of treatment. *New England Journal of Medicine, 318*, 1728-1733.

Lee, T. M. C., Yip, J. T. H., & Jones-Gotman, M. (2002). Memory deficits after resection from left or right anterior temporal lobe in humans: A meta-analytic review. *Epilepsia, 43*, 283-291.

Lesaffre, E., & Pledger, G. (1999). A note on the number needed to treat. *Controlled Clinical Trials, 20*, 439-447.

Levine, A. J., Miller, E. N., Becker, J. T., Selnes, O. A., & Cohen, B. A. (2004). Normative data for determining significance of test-retest differences on eight common neuropsychological instruments. *The Clinical Neuropsychologist, 18*, 373-384.

Lewis, M. S., Maruff, P., Silbert, B. S., Evered, L. A., & Scott, D. A. (2007). The influence of different error estimates in the detection of postoperative cognitive

dysfunction using reliable change indices with correction for practice effects. *Archives of Clinical Neuropsychology, 22*, 249-257.

Lineweaver, T. T., Morris, H. H., Naugle, R. I., Najm, I. M., Diehl, B., & Bingaman, W. (2006). Evaluating the contributions of state-of-the-art assessment techniques to predicting memory outcome after unilateral anterior temporal lobectomy. *Epilepsia, 47*, 1895-1903.

Livingston, R. B., Jennings, E., Reynolds, C. R., & Gray, R. M. (2003). Multivariate analysis of the profile stability of intelligence tests: High for IQs, low to very low for subtest analyses. *Archives of Clinical Neuropsychology, 18*, 487-507.

LoGalbo, A., Sawrie, S., Roth, D. L., Kuzniecky, R., Knowlton, R., Faught, E., & Martin, R. (2005). Verbal memory outcome in patients with normal preoperative verbal memory and left mesial temporal sclerosis. *Epilepsy and Behavior, 6*, 337-341.

Lord, F. M., & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores.* Reading, MA: Addison-Wesley.

Loring, D. W., Barr, W., Hamberger, M., & Helmstaedter, C. (2008). Neuropsychology evaluation - adults. In Engel, J., Pedley, T. A., Aicardi, J., Dichter, M. A., Moshe, S., Perucca, E., & Trimble, M. (Ed.), *Epilepsy: A Comprehensive Textbook* (2nd ed., pp. 1057-1066). Philadelphia: Lippincott Williams & Wilkins.

Luders, H., & Comair, Y. G. (2001). *Epilepsy Surgery* (2nd ed.). Philadelphia: Lippincott, Williams & Wilkins.

Maassen, G. H. (2000). Principles of defining reliable change indices. *Journal of Clinical and Experimental Neuropsychology, 22*, 622-632.

Maassen, G. H., Bossema, E., & Brand, N. (2008). Reliable change and practice effects: Outcomes of various indices compared. *Journal of Clinical and Experimental Neuropsychology, 30*, 1-14.

Majdan, A., Sziklas, V. & Jones-Gotman, M. J. (1996). Performance of healthy subjects and patients with resection from the anterior temporal lobe on matched tests of verbal and visuoperceptual learning. *Journal of Clinical and Experimental Neuropsychology, 18*, 416-430.

Marino, S. E., Meador, K. J., Loring, D. W., Okun, M. S., Fernandez, H. H., Fessler, A. J., Kustra, R. P., Miller, J. M., Ray, P. G., Roy, A., Schoenberg, M. R., Vahle, V. J., & Werz, M. A. (2009). Subjective perception of cognition is related to mood and not performance. *Epilepsy & Behavior, 14*, 459-464.

Marques, C. M., Caboclo, L. O. S. F., da Silva, T. I., Noffs, M. H., Carrete, Jr., H., Lin, K., Lin, J., Sakomoto, A. C., & Yacubian, E. M. T. (2007). Cognitive decline in temporal lobe epilepsy due to unilateral hippocampal sclerosis. *Epilepsy and Behavior, 10*, 477-485.

Martin R. C., Kretzmer T., Palmer C., Sawrie S., Knowlton R., Faught E., Morawetz R., & Kuzniecky R. (2002a). Risk to verbal memory following anterior temporal lobectomy in patients with severe left-sided hippocampal sclerosis. *Archives of Neurology, 59*, 1895-1901.

Martin, R. C., Griffith, H. R., Faught, E., Gilliam, F., Mackey, M., Vogtle, L. (2005). Cognitive functioning in community dwelling older adults with chronic partial epilepsy. *Epilepsia, 46*, 298-303.

Martin, R. C., Sawrie, S. M., Gilliam, F., Mackey, M., Faught, E., Knowlton, R., & Kuzniecky, R. (2002b). Determining reliable cognitive change after epilepsy surgery: Development of reliable change indices and standardised regression-based change norms for the WMS-III and WAIS-III. *Epilepsia, 43*, 1551-1558.

Martin, R. C., Sawrie, S. M., Roth, D. L., Morawetz, R. B., & Kuzniecky, R. (1998). Individual memory change after anterior temporal lobectomy: A base rate analysis using regression-based outcome methodology. *Epilepsia, 39*, 1075-1082.

Matarazzo, J. D., & Hermann, D. O. (1984). Base rate data for the WAIS-R: Test-retest stability and VIQ-PIQ differences. *Clinical Neuropsychology*, (6), 351-366.

Matarazzo, J. D., Carmody, T. P. & Jacobs, L. D. (1980). Test-retest reliability and stability of the WAIS: A literature review with implications for clinical practice. *Journal of Clinical Neuropsychology, 2*, 89-105.

McCaffrey, R. J., Ortega, A., Orsillo, S. M., Nelles, W. B., & Haase, R. F. (1992). Practice effects in repeated neuropsychological assessments. *The Clinical Neuropsychologist, 6*, 32-42.

McDermid Vaz, S. A. (2004). Nonverbal memory functioning following right anterior temporal lobectomy: A meta-analytic review. *Seizure, 13*, 446-452.

McSweeny, A. J., Naugle, R. I., Chelune, G. J., & Luders, H. (1993). "*T* scores for change": An illustration of a regression approach to depicting change in clinical neuropsychology. *The Clinical Neuropsychologist, 7*, 300-312.

Meador, K. J. (1996). Cognitive effects of epilepsy and of anti-epileptic medications. In E. Wyllie (Ed.), *The Treatment of Epilepsy: Principles and Practice.* (pp. 1121-1130). Baltimore: Williams and Wilkins.

Meador, K. J., Loring, D. W., Hulihan, J. F., Kamin, M. & Karim, R. (2003). Differential cognitive and behavioral side effects of topiramate and valproate. *Neurology, 60*, 1483-1488.

Medalia, A. & Richardson, R. (2005). What predicts a good response to cognitive remediation interventions? *Schizophrenia Bulletin, 41*, 942-953.

Milner, B. (1958). Psychological defects produced by temporal lobe excision. *Research Publications - Association for Research in Nervous and Mental Disease., 36*, 244-257.

Milner, B. (2003). Visual recognition and recall after right temporal-lobe excision in man. *Epilepsy and Behavior, 4*, 799-812.

Mitrushina, M. & Satz, P. (1991). Effect of repeated administration of a neuropsychological battery in the elderly. *Journal of Clinical Psychology, 47*, 790-801.

Mueller, C., Kaaden, S., Scorzin, J., Urbach, H., Fimmers, R., Helmstaedter, C., Zenter, J., Lehmann, T., & Schramm, J. (2009). Shrinkage of the hippocampal remnant after surgery for temporal lobe epilepsy: Impact on seizure outcome and neuropsychology. *Epilepsy and Behavior, 14*, 379-386.

Munoz, D. G., Ganapathy, G. R., Eliasziw, M., & Hachinski, V. (2000). Educational attainment and socioeconomic status of patients with autopsy-confirmed Alzheimer's disease. *Archives of Neurology, 57*, 85-89.

NIH Consensus Panel. (1990). Consensus conference on surgery for epilepsy. *The Journal of the American Medical Association, 264*, 729-733.

O'Brien, C. E., Bowden, S. C., Bardenhagen, F. J., & Cook, M. J. (2003). Neuropsychological correlates of hippocampal and rhinal cortex volumes in patients with mesial temporal sclerosis. *Hippocampus, 13*, 892-904.

Ojemann, G., & Dodrill, C. (1985). Verbal memory deficits after left temporal lobectomy for epilepsy. *Neurosurgery, 62*, 101-107.

Ortinski, P. & Meador, K. J. (2004). Cognitive side effects of antiepileptic drugs. *Epilepsy and Behavior, 5*, s60-65.

Oyegbile, T. O., Dow, C., Jones, J., Bell, B., Rutecki, P., Sheth, R., Seidenberg, M., & Hermann, B. P. (2004). The nature and course of neuropsychological morbidity in chronic temporal lobe epilepsy. *Neurology, 62*, 1736-1742.

Paglioli, E., Palmini, A., Portuguez, M., Paglioli, E., Azambuja, N., da Costa, J. C., da Silva Filho, H. F., Martinez, J. V., & Hoeffel. J. R. (2006). Seizure and memory outcome following temporal lobe surgery: Selective compared with nonselective approaches for hippocampal sclerosis. *Journal of Neurosurgery, 104*, 70-78.

Pai, M. C., & Tsai, J. J. (2005). Is cognitive reserve applicable to epilepsy? the effect of educational level on the cognitive decline after onset of epilepsy. *Epilepsia, 46*(Suppl. 1), 7-10.

Parabiaghi, A., Barbato, A., D'Avanzo, B., Erlicher, A., & Lora, A. (2005). Assessing reliable and clinically significant change on health of the nation outcome scales: Method for displaying longitudinal data. *Australian and New Zealand Journal of Psychiatry, 39,* 719-725.

Paradiso, S. Hermann, B. P., Blumer, D., Davies, K., & Robinson, R. G. (2001). Impact of depressed mood on neuropsychological status in temporal lobe epilepsy. *Journal of Neurology, Neurosurgery, and Psychiatry, 70*, 180-185.

Pauli, E., Pickel, S., & Schulemann, H. (1997). Neuropsychologic findings depending on the type of the resection in temporal lobe epilepsy. *Advances in Neurology, 81*, 373-377.

Piazzini, A., Turner, K., Chifari, R., Morabito, A., Canger, R., & Canevini, M. P. (2006). Attention and psychomotor speed decline in patients with temporal lobe epilepsy: A longitudinal study. *Epilepsy Research, 72,* 89-96.

Piggott, S. & Milner, B. (1993). Memory for different aspects of complex visual scenes after unilateral-temporal or frontal-lobe resection. *Neuropsychologia, 13*, 1-15.

Pitkanen, A. & Sutula, T. P. (2002). Is epilepsy a progressive disorder? Prospects for new therapeutic approaches in temporal lobe epilepsy. *The Lancet Neurology, 1*, 173-181.

Psychological Corporation. (1997). *The WAIS-III and WMS-III Technical Manual.* San Antonio, Texas: The Psychological Corporation.

Randolph, C., Lansing, A. E., Ivnik, R. J., Cullum, C. M., & Hermann, B. P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology, 14*, 489-496.

Rapport, L. J., Brines, D. B., Axelrod, B. N., Theisen, M. E. (1997). Full Scale IQ as mediator of practice effects: The rich get richer. *The Clinical Neuropsychologist, 11*, 375-380.

Rasmussen, L. S., Larsen, K., Houx, P., Skovgaard, L. T., Hanning, C. D., Moller, J. T., & the ISPOCD group. (2001). The assessment of postoperative cognitive function. *Acta Anaesthesiologica Scandinavica, 45*, 275-289.

Rauch, S. L., Dougherty, D. D., Cosgrove, G. R., Cassem, E. H., Alpert, N. M., Price, B. H., Nierenberg, A. A., Mayberg, H. S., Baer, L., Jenike, M. A., & Fischman, A. J. (2001). Cerebral metabolic correlates as potential predictors of response to anterior cingulotomy for obsessive compulsive disorder. *Biological Psychiatry, 50*, 659-667.

Rausch, R. & Crandall, P. H. (1982). Psychological status related to surgical control of temporal lobe seizures. *Epilepsia, 23*, 191-202.

Rausch, R., Kraemer, S., Pietras, C. J., Le, M., Vickrey, B. G., Passaro, E. A. (2003). Early and late cognitive changes following temporal lobe epilepsy for surgery. *Neurology, 60*, 951-959.

Raymond, P. D. Hinton-Bayne, A. D., Radel, M., Ray, M. J., Marsh, N. A. (2006). Assessment of statistical change criteria used to define significant change in neuropsychological test performance following cardiac surgery. *European Journal of Cardio-Thoracic Surgery, 29*, 82-88.

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of adults: A shrinking brain. *The Journal of Neuroscience, 23*, 3295-3301.

Richardson, M. P., Strange, B. A., Duncan, J. S., & Dolan, R. J. (2003). Preserved verbal memory function in left medial temporal pathology involves reorganisation of function to right medial temporal lobe. *Neuroimage, 20*, s112-s119.

Ring, H. A., Moriarty, J., & Trimble, M. R. (1998). A prospective study of the early postsurgical psychiatric associations of epilepsy surgery. *Journal of Neurology, Neurosurgery and Psychiatry, 64*, 601-604.

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-Based Medicine: How to Practice and Teach EBM.* (2nd ed.). Edinburgh: Churchill Livingstone.

Salinksy, M., Storzbach, D., Dodrill, C., & Binder, L. (2001). Test-retest bias, reliability, and regression equations for neuropsychological measures repeated over a 12-16 week period. *Journal of the International Neuropsychological Society, 7*, 597-605.

Sato, T. (1996). Type I and type II error in multiple comparisons. *The Journal of Psychology, 130*, 293-302.

Sawrie, S. M., Chelune, G. J., Naugle, R. I., & Luders, H. O. (1996). Empirical methods for assessing meaningful neuropsychological change following epilepsy surgery. *Journal of the International Neuropsychological Society, 2*, 556-564.

Sawrie, S. M., Martin, R. C., Gilliam, F. G., Roth, D. L., Faught, E., & Kuzniecky, R. (1998). Contribution of neuropsychological data to the prediction of temporal lobe epilepsy surgery outcome. *Epilepsia, 39*, 319-325.

Saykin, A. J., Robinson, L. J., Stafiniak, P., Kester, D. B., Gur, R. C., O'Connor, M. J., & Sperling, M. R. (1992). Neuropsychological changes after anterior temporal lobectomy: Acute effects on memory, language, and, music. In T. L. Bennett (Ed.), *The Neuropsychology of Epilepsy.* (pp. 263-290). New York: Plenum.

Sbordone, R. J. & Guilmette, T. J. (1999). Ecological validity: Prediction of everyday and vocational functioning from neuropsychological test data. In J. Sweet (Ed.), *Forensic Neuropsychology: Principles and Practice.* (pp.227-254). Lisse: Swets & Zeitlinger.

Scharfman, H. E. (2007). The neurobiology of epilepsy. *Current Neurology and Neuroscience Reports, 7*, 348-354.

Schefft, B. K., Testa, S. M., Dulay, M. F., Privitera, M. D., & Yeh, H. (2003). Preoperative assessment of confrontational naming ability and interictal

paraphasia production in unilateral temporal lobe epilepsy. *Epilepsy and Behavior, 4*, 161-168.

Schmitz, B. (2006). Effects of antiepileptic drugs on mood and behavior. *Epilepsia, 47*(Supp. 2), 28-33.

Schwartz, A. (2007). *EBM and Decision Tools.* Retrieved May 25, 2009, from http://araw.mede.uic.edu/~alansz/tools.html

Scoville, W. B. & Milner, B. (2000). Loss of recent memory after bilateral hippocampal lesions, 1957. *The Journal of Neuropsychiatry and Clinical Neurosciences, 12*, 103-113.

Seidenberg, M., Herman, B., Wyler, A. R., Davies, K., Dohan, F. C., & Leveroni, C. (1998). Neuropsychological outcome following anterior temporal lobectomy in patients with and without the syndrome of mesial temporal lobe epilepsy. *Neuropsychology, 12*, 303-316.

Seidenberg, M., Pulsipher, D. T., & Hermann, B. (2007). Cognitive progression in epilepsy. *Neuropsychology Review, 17*, 444-454.

Seidenberg, M., O'Leary, D. S., Giordani, B., Berent, S., & Boll, T. J. (1981). Test-retest IQ changes of epilepsy patients: Assessing the influence of practice effects. *Journal of Clinical Neuropsychology, 3*, 237-255.

Selwa, L. M., Berent, S., Giordani, B., Henry, T. R., Buchtel, H. A., & Ross, D. A. (1994). Serial cognitive testing in temporal lobe epilepsy: Longitudinal changes with medical and surgical therapies. *Epilepsia, 35*, 743-749.

Sherman, E. M. S., Slick, D. J., Connolly, M. B., Steinbok, P., Martin, R., Strauss, E., Chelune, G. J., & Farrell, K. (2003). Re-examining the effects of epilepsy surgery on IQ in children: Use of regression-based change scores. *Journal of the International Neuropsychological Society, 9*, 879-886.

Silvenius, H. (1999). Cost and cost-effectiveness of epilepsy surgery. *Epilepsia, 40*, 32-39.

Smith, M. L., Elliot, I. M., & Lach, L. (2002). Cognitive skills in children with intractable epilepsy: Comparison of surgical and non-surgical candidates. *Epilepsia, 43*, 631-637,

Speer, D. C. (1992). Clinically significant change: Jacobson and truax (1991) revisited. *Journal of Consulting and Clinical Psychology, 60*, 402-408.

Stern, Y. (2002). What is cognitive reserve? theory and research application of the reserve concept. *Journal of the International Neuropsychological Society, 8*, 448-460.

Straus, S. E., Richardson, W. S., Glasziou, P., & Haynes, R. B. (2005). *Evidence-Based Medicine: How to practice and teach EBM* (3rd ed.). Edinburgh: Elsevier.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary.* (3rd ed.). Oxford: University Press.

Swinkels, W. A. M., Kuyk, J., van Dyck, R., & Spinhoven, P. (2005). Psychiatric comorbidity in epilepsy. *Epilepsy and Behavior, 7*, 37-50.

Temkin, N. R., Heaton, R. K., Grant, I., & Dikment, S. S. (1999). Detecting significant change in neuropsychological test performance: A comparison model. *Journal of the International Neuropsychological Society, 5*, 357-369.

Thompson, P. J. (2003). Neuropsychological assessment and treatment of epilepsy. In P. W. Halligan, Kischka, U., & Marshall, J. C. (Eds.), *Handbook of Clinical Neuropsychology.* (pp. 585-605). Oxford: University Press.

Thompson, P. J., & Duncan, J. S. (2005). Cognitive decline in severe intractable epilepsy. *Epilepsia, 46*, 1780-1787.

Thompson, P. J., Baxendale, S. A., Duncan, J. S., & Sander, J. W. A. S. (2000). Effects of topiramate on cognitive function. *Journal of Neurology, Neurosurgery and Psychiatry, 69*, 636-641.

Trenerry, M. R., Jack, C. R., Cascino, G. D., Sharbrough, F. W., & Ivnik, R. J. (1995). Gender differences in post-temporal lobectomy verbal memory and relationships between MRI hippocampal volumes and preoperative verbal memory. *Epilepsy Research, 20*, 69-76.

Troster, A. I., Woods, S. P., & Morgan, E. E. (2007). Assessing cognitive change in parkinson's disease: Development of practice effect-corrected reliable change indices. *Archives of Clinical Neuropsychology, 22*, 711-718.

Tulsky, D.S. & Ledbetter, M. F. (2000). Updating to the WAIS-III and WMS-III: Considerations for research and clinical practice. *Psychological Assessment, 12*, 253-262.

Van Paesschen, W. & Revesz, R. (1998). Hippocampal sclerosis. In F. Scaravilli (Ed.), *Neuropathology of Epilepsy* (pp. 501-573). Singapore: World Scientific Publishing Company.

Vermeulen, J. & Aldenkamp, A. P. (1995). Cognitive side-effects of chronic antiepileptic drug treatment: A review of 25 years of research. *Epilepsy Research, 22*, 65-95.

Vingerhoets, G. (2006). Cognitive effects of seizures. *Seizure, 15*, 221-226.

Vinters, H. V., Armstrong, D. L., Babb, T. L., Daumas-Duport, C., Robitaille, Y., Bruton, C. J., & Farrell, M. A. (1993). The neuropathology of human symptomatic epilepsy. In J. Engel (Ed.), *Surgical Treatment of the Epilepsies* (2nd ed., pp. 593-608). New York: Raven Press.

Wachi, M., Tomikawa, M., Fukuda, M., Kameyama, S., Kasahara, K., Sasagawa, M., Shirane, S., Kanazawa, O., Yoshino, M., Aoki, S., & Sohma, Y. (2001). Neuropsychological changes after surgical treatment for temporal lobe epilepsy. *Epilepsia, 42*, 4-8.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale - III.* San Antonio, Texas: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Memory Scale - III.* San Antonio, Texas: The Psychological Corporation.

White, J. R., Matchinsky, D., Beniak, T. E., Arndt, R. C., Walczak, T., Leppik, I. E., Rarick, J., Roman, D. D., & Gumnit, R. J. (2002). Predictors of postoperative memory function after left anterior temporal lobectomy. *Epilepsy & Behavior, 3*, 383-389.

Wiebe, S., Blume, W. T., Girvin, J. P., & Eliasziw, M. (2001). For the effectiveness and efficiency of surgery for temporal lobe epilepsy study group. A randomized, controlled trial of surgery for temporal-lobe epilepsy. *New England Journal of Medicine, 345*, 311-318.

Wilde, N. J., Strauss, E., Chelune, G. J., Hermann, B. P., Hunter, M., Loring, D. W., Martin, R. C., & Sherman, E. M. S. (2003). Confirmatory factor analysis of the WMS-III in patients with temporal lobe epilepsy. *Psychological Assessment, 15*, 56-63.

Wieser, H. G., Engel, J., Williamson, P. D., Babb, T. L., & Gloor, P. (1993). Surgically remediable temporal lobe syndromes. In J. Engel (Ed.), *Surgical Treatment of the Epilepsies* (2nd ed., pp. 49-63). New York: Raven Press Ltd.

Wilmes, K. (2003). The methodological and statistical foundations of neuropsychological assessment. In Halligan, P. W., Kischka, U. & Marshall, J. C. (Ed.), *Handbook of Clinical Neuropsychology.* (pp. 27-47). Oxford: Oxford University Press.

Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment, 82*, 50-59.

Wong, T. M. (2006). Ethical controversies in neuropsychological test selection, adminstration and interpretation. *Applied Neuropsychology, 13*, 68-76.

Woods, S. P., Childers, M., Ellis, R. J., Guaman, S., Grant, I., & Heaton, R. K. (2006). A battery approach for measuring neuropsychological change. *Archives of Clinical Neuropsychology, 21*, 83-89.

World Health Organisation. (2001). *Epilepsy: Etiology, Epidemiology and Prognosis* (Fact sheet 165 ed.). Geneva: World Health Organisation.

Worrell, G. A.D, Sencakova, D., Jack, C. R., Flemming, K. D., Fulgham, J. R., & So, E. L. (2002). Rapidly progressive hippocampal atrophy: Evidence for a seizure-induced mechanism. *Neurology, 58*, 1553-1556.

York, M. K., Rettig, G. M., Grossman, R. J., Hamilton, W. J., Armstrong, D. D., Levin, H. S., Mizrahi, E. M. (2003). Seizure control and cognitive outcome after temporal lobectomy: A comparison of classic Ammon's Horn sclerosis, atypical mesial temporal sclerosis, tumoral pathologies. *Epilepsia, 44*, 387-398.

Zillmer, E. A., & Spiers, M. V. (2001). *Principles of Neuropsychology*. Belmont, CA: Wadsworth.

APPENDIX A

Ethics Approval Forms

FILE COPY

Research and Grants Unit
Ph: (03) 9288 3930 Fax: (03) 9288 3205

Friday, 13 October 2006

Dr F Bardenhagen
Clinical Neurosciences
SVH

PO Box 2900
Fitzroy Victoria 3065 Australia
Telephone 03 9288 2211
www.svhm.org.au

Dear Dr Bardenhagen

Protocol No: HREC-A 092/06

'Determining clinically meaninful neuropsychological change in an epilepsy population'.

Dr F Bardenhagen A/Prof S Bowden Prof M Cook Ms M Cumner

The Professional Secretariat of Human Research Ethics Committee-A (HREC-A) has agreed that your latest correspondence dated Tuesday 26 September 2006, has satisfied the conditions imposed and granted full approval for this project to be undertaken at St. Vincent's Health.

HREC-A is constituted and operates in accordance with the NHMRC National Statement on Ethical Conduct in Research Involving Humans 1999 (including supplementary note 7 dated November 1992).

HREC-A has a policy of granting approval for four years. Ethical approval is valid for four years from the date of this letter. Approval may be renewed at the end of this period by resubmission to HREC-A.

Approval is subject to:

1. immediate notification to HREC-A and sponsor of any serious adverse effects on participants;
2. immediate notification of any unforeseen events that may affect the continuing ethical acceptability of the project;
3. notification and reasons for ceasing the project prior to its expected date of completion;
4. the completion of an annual report on progress of the project;
5. HREC-A approval of any proposed modification to the project; and
6. the submission of a final report and papers published on completion of project.

St Vincent's

Continuing the Mission of the Sisters of Charity

SV44812

This approval is for the following participant information and consent form(s):

- Participant Information and Consent Form version 2 dated 24 August 2006

The following documents are enclosed:

- One signed copy of page 1 of the application form
- 1 copy(s) of the Approval to Examine Medical Records form #
- One copy of the Health Information Services (HIS) Research Guidelines

Please note that this approval includes the cover letter to potential participants version 2 dated 24 August 2006.

Approval does NOT include the participant brochure (version 1 dated 26 September 2006) submitted with your letter dated 26 September 2006. The cover letter and participant information and consent form (PICF) are considered to be sufficient. If there is information in the brochure that is essential to informed consent and is not already in the PICF, it should be incorporated into a revised PICF and submitted as an amendment.

Please note that the proposed follow-up letter needs to be submitted to HREC-A for approval, as an amendment, before it can be implemented.

Yours sincerely

Jill Hambling

Secretary, Human Research Ethics Committee-A

#cc: Health Information Services

VICTORIA UNIVERSITY | A NEW SCHOOL OF THOUGHT

MEMO

TO	Dr Fiona Bardenhagen Department of Psychology St. Albans Campus	DATE	15/1/2007
FROM	Professor Michael Polonsky Chair Victoria University Human Research Ethics Committee		
SUBJECT	Ethics Approval - HRETH. 06/122		

Dear Dr. Fiona Bardenhagen,

Thank you for submitting this application for ethical approval of the project:

HRETH06/122 Determining Clinically Meaningful Neuropsychological Change in an Epilepsy Population

The proposed amendments have been accepted by the Chair, Victoria University Human Research Ethics Committee and approval for application HRETH06/243 has been granted from 05/01/2007 to 29/11/2009.

Please note that the Human Research Ethics Committee must be informed of the following: any changes to the approved research protocol, project timelines, any serious or unexpected adverse effects on participants, and unforeseen events that may effect continued ethical acceptability of the project. In these unlikely events, researchers must immediately cease all data collection until the Committee has approved the changes.

If you have any queries, please do not hesitate to contact me on 9919 4625.

On behalf of the Committee, I wish you all the best for the conduct of the project.

Professor Michael Polonsky
Chair
Victoria University Human Research Ethics Committee

APPENDIX B

Invitation to Participate and the Participant Information Sheet and Consent Form

Name
Address
Address

Date

Dear Ms/Mr

Re: Invitation to participate in research project

Some time ago, you were sent an invitation to participate in a research project looking at changes in neuropsychological test scores in people with epilepsy, being undertaken at St. Vincent's Hospital, Melbourne. To date, we have not heard from you about whether or not you wish to partipate.

You were identified as a potentially suitable participant in this research project because you have already had neuropsychological testing at St Vincent's, or because you may be interested in having a neuropsychological assessment.

At the moment, we need some more people to participate in this research project. We plan to telephone you between the **date** and **date** to see if you wish to participate and to answer any questions you may have about the project. If you do not wish to be telephoned, please either leave a message at the Neuropsychology Unit at St. Vincent's Hospital on **9288 3559** or send an email to **marnie.cumner@live.vu.edu.au**

Please find enclosed a detailed participant information and consent form. If you would like to discuss the project before the above dates, please contact Marnie Cumner on **0423 176 895** or at St. Vincent's Hospital on **9288 3559.**

Please be assured that your decision to participate or not will in no way affect your ongoing clinical care at St. Vincent's Hospital.

Thank you kindly for taking the time to consider this project.

Yours Sincerely,

Jasmine O'Rafferty
Secretary, Neuropsychology Unit
Enc: Participant information and consent form

ST. VINCENT'S HEALTH

PARTICIPANT INFORMATION AND CONSENT FORM

Version 2 Dated 24 August 2006

PROTOCOL NO. (SVH): #092/06

NAME OF PARTICIPANT:

U.R. NO:

FULL PROJECT TITLE:

Determining clinically meaningful neuropsychological change in an epilepsy population.

NAME/S OF INVESTIGATOR/S: Dr. Fiona Bardenhagen; Associate Professor Stephen Bowden; Professor Mark Cook

Student Researcher: Marnie Cumner

This Participant Information and Consent Form is **6** pages long. Please make sure you have all the pages.

1. Your Consent

You are invited to take part in this research project and your participation is voluntary.

This Participant Information contains detailed information about the research project. Its purpose is to explain to you as openly and clearly as possible all the procedures involved in this project before you decide whether or not to take part in it.

Please read this Participant Information carefully. Feel free to ask questions about any information in the document. You may also wish to discuss the project with a relative or friend or your local health worker, and please feel free to do this.

Once you understand what the project is about and if you agree to take part in it, you will be asked to sign the Consent Form. By signing the Consent Form, you indicate that you understand the information and that you give your consent to participate in the research project.

You will be given a copy of the Participant Information and Consent Form to keep as a record.

2. Purpose and Background

The purpose of this project is to obtain information about how performance on a number of tests may change over time in people who have temporal lobe epilepsy, who have not undergone surgery for treatment of their epilepsy.

A total of 50 people will participate in this project.

Epilepsy patients are often given a series of neuropsychological tests before and after surgery in order to see if the surgery has resulted in any change in their memory or thinking skills.

Previous research has found that there are some difficulties associated with comparing test results obtained at two different times. Researchers have noticed that issues such as practice effects may influence a person's score when they are given the same tests a second time. Practice effects refer to an improvement in scores due to retesting on the same cognitive test, and do not reflect improvement in the skills being assessed. It is therefore important for clinicians to know how much change in test scores is due to issues such as practice effects, and how much change is due to genuine improvement or decline in the abilities measured by the test.

You are invited to participate in this research because:
(a) you have already been assessed with the neuropsychological battery which is part of pre-surgery procedure, but you have not yet undergone surgery for epilepsy. By undergoing a second neuropsychological assessment, the researchers can look at the changes in scores over the two assessments.
(b) you have been identified by your neurologist as a person with temporal lobe epilepsy who may be interested in having a neuropsychological assessment. In this case, you will be tested twice, with the second testing session approximately 6 months after the first.

Looking at changes in test scores will then let us know how stable these tests are, so that during post-surgery neuropsychological assessment, the clinician is better able to tell if the patient's abilities have genuinely changed or if practice effects have produced improvement.

The results of this research may be used to help Marnie Cumner to obtain a Masters in Clinical Neuropsychology.

3. Procedures

Information will be obtained from the neuropsychological testing that you have already completed and from a second assessment that will be administered at St. Vincent's Hospital. The neuropsychological testing will take approximately 2-3 hours, and will be the same tests of thinking and memory that are used for assessment before epilepsy surgery. The results of the two assessment sessions will then be compared to obtain a better understanding of how individuals perform on these tests over time.

Some personal and health information, such as age, education and seizure history, will also be collected from you.

By participating in this study, you are giving consent for us to access your previous neuropsychological test results from St. Vincent's Hospital Melbourne, and to undertake further neuropsychological testing.

4. Possible Benefits

We cannot guarantee or promise that you will receive any benefits from this project. However, all participants will be given oral and written feedback on the results of their assessment, and a list of strategies to help with memory,

attention, and word-finding will be provided to all participants. In addition to this, if the research is successful, a better understanding of the effects contributing to changes in neuropsychological test results will allow for improved interpretation of patients' cognitive outcomes after surgery for temporal lobe epilepsy.

5. Possible Risks

The procedures used in this study are not harmful to you; the major risk is of some fatigue from completing the 2-3 hours of testing. You can, at any time, ask to finish the testing if you wish.

6. Alternatives to Participation

The alternative to participation is choosing not to participate.

7. Privacy, Confidentiality and Disclosure of Information

Any information obtained in connection with this project and that can identify you will remain confidential and secure in the Victorian Epilepsy Centre and the Neuropsychology Unit in the department of Clinical Neurosciences. Only the researchers associated with this project will have access to this information. Electronic data will be kept secure through the use of password protection. Personal data will only be disclosed with your permission, except as required by law. If you give us your permission by signing the Consent Form, we plan to publish the results in a collated, coded format in an international medical journal.

In any publication, information will be provided in such a way that you cannot be identified. The consent form that you sign will be kept separately and securely in the Neuropsychology Unit for a period of ten years.

8. New Information Arising During the Project

During the research project, new information about the risks and benefits of the project may become known to the researchers. If this occurs, you will be told about this new information. In all cases, you will be offered all available care to suit your needs and medical condition.

9. Results of Project

If you would like to receive information about the results of this project, please advise the student researcher. Upon completion of the project, participants who have registered their interest will be provided with a brief written summary of the results.

10. Further Information or Any Problems

If you require further information or if you have any problems concerning this project, you can contact the principal researcher, Dr. Fiona Bardenhagen, on 9288 3559.

11. Complaints

If you have any complaints about any aspect of the study or the way in which it is being conducted you may contact the Patient Representative at St. Vincent's Health on Telephone: 9288 2211. You will need to tell the Patient Representative the name of the person who is noted above as principal investigator. As this study has also been approved by Victoria University, the Patient Representative will discuss all complaints with the Secretary of the Victoria University Ethics Committee. If you

prefer, you may contact the University directly, by contacting the Secretary, University Human Research Ethics Committee, Victoria University of Technology, PO Box 14428 MCMC, Melbourne, 8001 (telephone: 9677 4710). A complaint to either the hospital or the university will be discussed with the other party.

12. Research Participant Rights

If you have any questions about your rights as a research participant, then you may contact Jill Hambling, Executive Officer Research at St. Vincent's Health on Telephone: 9288 3930.

13. Participation is Voluntary

Participation in any research project is voluntary. If you do not wish to take part you are not obliged to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage.

Your decision whether to take part or not to take part, or to take part and then withdraw, will not affect your routine treatment, your relationship with those treating you or your relationship with St. Vincent's Hospital.

Before you make your decision, a member of the research team will be available to answer any questions you have about the research project. You can ask for any information you want. Sign the Consent Form only after you have had a chance to ask your questions and have received satisfactory answers.

If you decide to withdraw from this project, please notify a member of the research team before you withdraw.

14. Ethical Guidelines

This project will be carried out according to the *National Statement on Ethical Conduct in Research Involving Humans* (June 1999) produced by the National Health and Medical Research Council of Australia. This statement has been developed to protect the interests of people who agree to participate in human research studies.

The ethical aspects of this research project have been approved by the Human Research Ethics Committee of St. Vincent's Hospital, Melbourne and Victoria University, St. Albans.

15. Reimbursement for your costs

You will not be paid for your participation in this project. However, you will be given detailed oral and written feedback on the results of your assessment, and a list of strategies which may help you with word-finding, memory, and concentration.

CONSENT FORM
Version: 2 Dated 24 August 2006
Site: St. Vincent's Hospital, Melbourne.
Protocol No. (SVH): #092/06

Full Project Title:

Determining clinically meaningful neuropsychological change in an epilepsy population.

I have read, and I understand the Participant Information version *2* dated 24 August 2006.

I freely agree to participate in this project according to the conditions in the Participant Information.

I will be given a copy of the Participant Information and Consent Form to keep.

The researcher has agreed not to reveal my identity and personal details if information about this project is published or presented in any public form.

Participant's Name (printed) ……

Signature Date

Name of Witness to Participant's Signature (printed) ……

Signature Date

Researcher's Name (printed) ……

Signature Date

Note: All parties signing the Consent Form must date their own signature.

REVOCATION OF CONSENT FORM

Revocation of Consent Form

Full Project Title:

Determining clinically meaningful neuropsychological change in an epilepsy population.

I hereby wish to WITHDRAW my consent to participate in the research proposal described above and understand that such withdrawal WILL NOT jeopardise any treatment or my relationship with St. Vincent's Hospital, Melbourne.

Participant's Name (printed) …….

Signature Date

APPENDIX C

Statistics

Table C1
Missing Values for the Pre-operative Sample, Identified using SPSS MVA.

Variable	N	Number of missing cases	% of missing cases
Education	41	0	0.0
Age	41	0	0.0
Seizure duration	41	0	0.0
Retest interval length	41	0	0.0
TLE side	41	0	0.0
Gender	41	0	0.0
Handedness	41	0	0.0
First assessment:			
VIQ	37	4	9.8
PIQ	37	4	9.8
FSIQ	37	4	9.8
VCI	37	4	9.8
POI	37	4	9.8
WMI	36	5	12.2
PSI	37	4	9.8
AI	39	2	4.9
VI	39	2	4.9
IM	39	2	4.9
AD	39	2	4.9
VD	39	2	4.9
GM	39	2	4.9
WM	38	3	7.3
BNT	36	5	12.2
Second assessment			
VIQ	36	5	12.2
PIQ	36	5	12.2
FSIQ	36	5	12.2
VCI	35	6	14.6
POI	35	6	14.6
WMI	35	6	14.6
PSI	36	5	12.2
AI	40	1	2.4
VI	40	1	2.4
IM	41	0	0.0
AD	41	0	0.0
VD	41	0	0.0
GM	41	0	0.0
WM	41	0	0.0
BNT	34	7	17.1

Table C2
Missing Values for the Post-operative Sample, Identified using SPSS MVA.

Variable	N	Number of missing cases	% of missing cases
Education	76	0	0.0
Age	76	0	0.0
Seizure duration	76	0	0.0
Retest interval length	73	0	0.0
TLE side	76	0	0.0
Gender	76	0	0.0
Handedness	76	0	0.0
First assessment:			
VIQ	73	3	3.9
PIQ	73	3	3.9
FSIQ	73	3	3.9
VCI	73	3	3.9
POI	73	3	3.9
WMI	73	3	3.9
PSI	69	7	9.2
AI	76	0	1.0
VI	76	0	0.0
IM	76	0	0.0
AD	76	0	0.0
VD	76	0	0.0
GM	75	1	1.3
WM	76	0	0.0
BNT	73	3	3.9
Second assessment			
VIQ	72	4	5.3
PIQ	72	4	5.3
FSIQ	72	4	5.3
VCI	72	4	5.3
POI	70	6	7.9
WMI	71	5	6.6
PSI	69	7	9.2
AI	74	2	2.6
VI	74	2	2.6
IM	74	2	2.6
AD	74	2	2.6
VD	74	2	2.6
GM	73	3	3.9
WM	70	6	7.9
BNT	73	3	3.9

Table C3
Classification of cognitive change on retest (WAIS-III indices) according to laterality, using 95% RCIs

Index	N	Australian RCIs – SEdiff method				North American RCIs – SEdiff method			
		Prediction interval	% declined	% no change	% improved	Prediction interval	% declined	% no change	% improved
Hypothesised			*2.5%*	*95%*	*2.5%*		*2.5%*	*95%*	*2.5%*
VIQ		**±13**				**±10**			
Left	34		2.9	91.2	5.9		5.9	88.2	5.9
Right	33		3	81.8	15.2		3	81.8	15.2
PIQ		**-12, +18**				**-8, +16**			
Left	34		5.9	94.1	0		8.8	88.2	2.9
Right	33		0	97	3		3	90.9	6.1
FSIQ		**-11, +13**				**-7, +11**			
Left	34		2.9	91.2	5.9		2.9	85.3	11.8
Right	33		0	97	3		6.1	78.8	15.2
VCI		**±12**				**-9, +11**			
Left	34		2.9	91.2	5.9		8.8	85.3	5.9
Right	33		0	93.9	6.1		0	84.8	15.2
POI		**-14, +22**				**-10, +16**			
Left	33		3	93.9	3		6.1	9.9	3
Right	32		3.1	96.9	0		9.4	87.5	3.1
WMI		**±21**				**±19**			
Left	33		3	93.9	3		3	93.9	3
Right	33		3	90.9	6.1		3	90.9	6.1
PSI		**-11, +19**				**n/a**			
Left	31		19.4	74.2	6.5		-	-	-
Right	30		0	96.7	3.3		-	-	-
BNT		**±14**				**n/a**			
Left	35		5.7	94.3	0		-	-	-
Right	35		0	100	0		-	-	-

Table C4
Classification of cognitive change on retest (WMS-III indices) according to laterality, using 95% RCIs

		Australian RCIs – SEdiff method				North American RCIs – SEdiff method			
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**
Hypothesised			*2.5%*	*95%*	*2.5%*		*2.5%*	*95%*	*2.5%*
AI		**-19, +21**				**-17, +19**			
Left	36		13.9	80.6	5.6		16.7	77.8	5.6
Right	35		8.6	88.6	2.9		8.6	85.7	5.7
VI		**-24, +28**				**-18, +26**			
Left	36		0	97.2	2.8		0	94.4	5.6
Right	35		5.7	94.3	0		5.7	94.3	0
IM		**-18, +22**				**-14, +20**			
Left	36		8.3	88.9	2.8		11.1	86.1	2.8
Right	35		2.9	97.1	0		2.9	97.1	0
AD		**-22, +24**				**±20**			
Left	36		2.8	88.9	8.3		2.8	83.3	13.9
Right	35		0	91.4	8.6		0	85.7	14.3
VD		**-22, +24**				**-17, +27**			
Left	36		0	94.4	5.6		2.8	97.2	0
Right	35		0	100	0		0	100	0
ARD		**-23, +29**				**±25**			
Left	36		8.3	88.9	2.8		8.3	88.9	2.8
Right	33		0	93.9	6.1		0	93.9	6.1
GM		**-19, +23**				**-19, +21**			
Left	36		8.3	88.9	2.8		8.3	83.3	8.3
Right	33		0	100	0		0	97	3
WM		**-22, +26**				**±24**			
Left	32		6.3	84.4	9.4		6.3	84.4	9.4
Right	35		2.9	97.1	0		0	100	0

Table C5
Classification of cognitive change on retest (WAIS-III indices) in an Australian post-surgical sample using Australian and North American 90% RCIs

		Australian RCIs – SEdiff method				North American RCIs – SEdiff method			
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**
Hypothesised			*5%*	*90%*	*5%*		*5%*	*90%*	*5%*
VIQ		**±11**				**-8, +8**			
All	67		3	86.6	3		4.5	77.6	17.9
PIQ		**-9, + 15**				**-7, +13**			
All	67		4.5	91	4.5		7.5	82.1	10.4
FSIQ		**-9, + 11**				**-6, +8**			
All	67		1.5	85.1	13.4		9	70.1	20.9
VCI		**±10**				**-7, +9**			
All	67		3	85.1	11.9		10.4	74.6	14.9
POI		**-11, +19**				**-8, +14**			
All	65		4.6	92.3	3.1		10.8	80	9.2
WMI		**±18**				**-17, +14**			
All	66		3	90.9	6.1		3	80.3	16.7
PSI		**-8, +16**				**n/a**			
All	61		16.4	73.8	9.8	-	-	-	-
BNT		**±12**				**n/a**			
All	70		5.7	94.3	0	-	-	-	-

Table C6
Classification of cognitive change on retest (WMS-III indices) in an Australian post-surgical sample using Australian and North American 90% RCIs

		Australian RCIs – SEdiff method				North American RCIs – SEdiff method			
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**
Hypothesised			*5%*	*90%*	*5%*		*5%*	*90%*	*5%*
AI		**-16, +18**				**-14, +16**			
All	71		12.7	78.9	8.5		15.5	74.6	9.9
VI		**-19, +23**				**-15, +21**			
All	71		2.8	94.4	2.8		5.6	85.9	8.5
IM		**-15, +19**				**-12, +16**			
All	71		5.6	93	1.4		15.5	76.1	8.5
AD		**-18, +20**				**-19, +15**			
All	71		2.8	83.1	14.1		2.8	76.1	21.1
VD		**-18, +20**				**-13, +23**			
All	71		4.2	85.9	9.9		7	88.7	4.2
ARD		**-19, +25**				**-22, +20**			
All	69		10.1	85.5	4.3		10.1	79.7	10.1
GM		**-16, +20**				**-16, +18**			
All	69		4.3	89.9	5.8		4.3	84.1	11.6
WM		**-18, +22**				**-22, +18**			
All	67		4.5	91	0		4.5	86.6	9

Table C7
Classification of cognitive change on retest (WAIS-III) according to laterality, using 95% SRBs

		Australian SRBs				North American SRBs			
Index	**N**	**Predictio n interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**
Hypothesised			*2.5%*	*95%*	*2.5%*		*2.5%*	*95%*	*2.5%*
VIQ		**±12.84**				**±9.51**			
Left	34		2.9	91.2	5.9		5.9	88.2	5.9
Right	33		3	90.9	6.1		3	81.8	15.2
PIQ		**±14.37**				**±11.52**			
Left	34		5.9	94.1	0		5.9	85.3	8.8
Right	33		0	93.9	6.1		0	93.9	6.1
FSIQ		**±11.86**				**±9.15**			
Left	34		2.9	91.2	5.9		14.7	79.4	5.9
Right	33		0	97	3		9.1	87.9	3
VCI		**±10.15**				**±9.15**			
Left	34		5.9	91.2	2.9		8.8	85.3	5.9
Right	33		0	90.9	9.1		0	81.8	18.2
POI		**±18.58**				**±12.37**			
Left	33		3	93.9	3		3	90.9	6.1
Right	32		0	100	0		3.1	90.6	6.3
WMI		**±19.89**				**±17.11**			
Left	33		3	93.9	3		3	78.8	18.2
Right	33		3	87.9	9.1		3	87.9	9.1

Table C8
Classification of cognitive change on retest (WMS-III) according to laterality, using 95% SRBs

		Australian SRBs				North American SRBs			
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**
Hypothesised			*2.5%*	*95%*	*2.5%*		*2.5%*	*95%*	*2.5%*
AI		**n/a**				**±17.44**			
Left	36		11.1	86.1	2.8		8.3	86.1	5.6
Right	35		8.3	86.1	5.6		8.6	74.3	17.1
VI		**±22.44**				**±21.52**			
Left	36		0	94.4	5.6		0	94.4	5.6
Right	35		2.9	97.1	0		5.7	94.3	0
IM		**±19.5**				**±17.5**			
Left	36		8.3	88.9	2.8		8.3	88.9	2.8
Right	35		2.9	97.1	0		0	100	0
AD		**±19.62**				**±16.88**			
Left	36		8.3	77.8	13.9		11.1	75	13.9
Right	35		5.7	85.7	8.6		2.9	68.6	28.6
VD		**±21.11**				**±23.87**			
Left	36		0	88.9	11.1		2.8	97.2	0
Right	35		0	100	0		8.6	91.4	0
ARD		**±23.44**				**±21.56**			
Left	36		5.6	91.7	2.8		19.4	77.8	2.8
Right	33		0	93.9	6.1		6.1	87.9	6.1
GM		**±19.54**				**±20.64**			
Left	36		5.6	91.7	2.8		5.6	91.7	2.8
Right	33		0	100	0		0	100	0
WM		**±22.66**				**±21.56**			
Left	32		3.1	81.3	15.6		3.1	84.4	12.5
Right	35		0	100	0		0	94.3	5.7

Table C9
Classification of cognitive change on retest (WAIS-III indices) for entire sample, using 90% SRBs

		Australian SRBs				North American SRBs			
Index	**N**	**Predictio n interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**
Hypothesised			*5%*	*90%*	*5%*		*5%*	*90%*	*5%*
VIQ		**±10.75**				**±7.95**			
All	67		3	88.6	10.4		3	80.6	16.4
PIQ		**±12.02**				**±9.64**			
All	67		4.5	91	4.5		4.5	85.1	10.4
FSIQ		**±9.92**				**±7.66**			
All	67		1.5	85.1	13.4		14.9	73.1	11.9
VCI		**±8.5**				**±7.66**			
All	67		3	85.1	11.9		6	76.1	17.9
POI		**±15.55**				**±10.34**			
All	65		4.6	92.3	3.1		6.2	80	13.8
WMI		**±16.65**				**±14.3**			
All	66		3	81.8	15.2		3	80.3	16.7

Table C10
Classification of cognitive change on retest (WMS-III indices) for entire sample, using 90% SRBs

		Australian SRBs				North American SRBs			
Index	**N**	**Prediction interval**	**% declined**	**% no change**	**% improved**	**Prediction interval**	**% declined**	**% no change**	**% improved**
Hypothesised			*5%*	*90%*	*5%*		*5%*	*90%*	*5%*
AI		**±19.38**				**±14.6**			
All	71		9.7	86.1	4.2		14.1	76.1	9.9
VI		**±18.78**				**±18**			
All	71		2.8	94.4	2.8		4.2	90.1	5.6
IM		**±16.32**				**±14.65**			
All	71		5.6	93	1.4		9.9	87.3	2.8
AD		**±16.42**				**±14.12**			
All	71		2.8	83.1	14.1		9.9	69	21.1
VD		**±17.66**				**±19.98**			
All	71		4.2	85.9	9.9		7	93	0
ARD		**±19.61**				**±18.04**			
All	69		10.1	85.5	4.3		15.9	79.7	4.3
GM		**±16.35**				**±17.27**			
All	69		4.3	89.9	5.8		5.8	91.3	2.9
WM		**±18.96**				**±17.2**			
All	67		4.5	91	4.5		3	79.1	17.9

Table C11
Absolute Risk Reduction and Number Needed to Harm/ Number Needed to Treat for right and left ATL groups on the WAIS-III indices, using 95% SRB change norms

Index	Group	N	Control event rate	Experimental event rate	Absolute risk reduction/increase	Number needed to treat (+) or harm (-)
VIQ	LATL	34	5.7	2.9	2.8	36
	RATL	33	5.7	3	2.7	37
PIQ	LATL	34	0	5.9	-5.9	-17
	RATL	33	0	0	0	No difference
FSIQ	LATL	34	2.9	2.9	0	No difference
	RATL	33	2.9	0	2.9	35
VCI	LATL	34	8.8	5.9	2.9	35
	RATL	33	8.8	0	8.8	12
POI	LATL	34	2.9	2.9	0	No difference
	RATL	33	2.9	0	2.9	35
WMI	LATL	34	6.1	2.9	3.2	32
	RATL	33	6.1	3	3.1	33

Table C12
Absolute Risk Reduction and Number Needed to Treat/ Number Needed to Harm for right and left ATL groups on the WMS-III indices, using 95% SRB change norms

Index	Group	N	Control event rate	Experimental event rate	Absolute risk reduction/increase	Number needed to treat (+) or harm (-)
AI	LATL	36	2.6	11.1	-8.5	-11
	RATL	35	2.6	8.3	-5.7	-17
VI	LATL	36	0	2.6	2.6	39
	RATL	35	2.9	2.6	-0.3	-334
IM	LATL	36	8.3	5.1	-3.2	-32
	RATL	35	2.9	5.1	2.2	46
AD	LATL	36	8.3	0	-8.3	-12
	RATL	35	0	0	0	0
VD	LATL	36	0	0	0	-*
	RATL	35	0	0	0	-*
GM	LATL	36	5.6	2.5	-3.1	-33
	RATL	33	0	2.5	2.5	40
WM	LATL	32	3.1	5.3	2.2	46
	RATL	35	0	5.3	5.3	19

Note: ^ NNH for AI index based on RCI calculations