Clinician Judgement of Intellectual Function.

Michelle M. Lindsay

Being the report of an investigation
Submitted for the degree of
Doctor of Psychology (Clinical Neuropsychology)
January 2005.

Department of Psychology
Victoria University
Victoria, Australia

Clinician Judgement of Intellectual Function.

STUDENT DECLARATION

I, *Michelle Lindsay*, declare that the thesis entitled *Clinician Judgement of Intellectual Function* contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma and except where otherwise indicated, and that this thesis is my own work. This research project was formally approved by the Ethics Committee of the Department of Psychology, Victoria University. The ethical principles specified have been followed in this report of the research project.

Signed:

Date:

ABSTRACT

Estimation of premorbid IQ is an integral part of many neuropsychological assessments, but there has been little research on the efficacy of clinical judgement versus actuarial methods in this area, and only one study has examined the methods used by clinicians to estimate IQ. Experienced and novice neuropsychologists' estimation of intellectual function were compared to objective IQ data and a number of actuarial methods. A survey of methods for estimating premorbid IQ used by experienced Australian, North American and British neuropsychologists and novice Australian neuropsychologists was also conducted. Overall, consistent with past research, experienced clinicians were no more accurate than novices in estimating IQ, and clinicians confidence in their estimates was not related to judgement accuracy. In addition, there was no difference between the accuracy of a regression estimate (the Crawford Index, chosen for its applicability to an Australian sample) and the clinician raters' estimates of IQ, but North American methods (Barona and OPIE) were more accurate in predicting IQ. Clinician inter-rater reliability was found to be good, but clinicians were found to overestimate the relationship between education and IQ. Findings from the survey of methods of estimating premorbid IQ were broadly consistent with past research, with little

use of actuarial methods specifically designed to estimate premorbid IQ reported by clinicians.

ACKNOWLEDGEMENTS

To my supervisors, Dr Fiona Bardenhagen and Ms Izabela Walters, I would like to express my sincere thanks for all the advice, support and guidance they have given me.

I would also like to thank all the clinicians who dedicated their time to assist me in collecting this data.

To my fellow Doctoral students, I am forever grateful for their help in completing this thesis by not only completing the judgement tasks and the questionnaire themselves, but by also helping to recruit other subjects.

I am eternally grateful to my friends and family for their words of encouragement and support throughout the process of completing this thesis. To my parents, Margaret and Bruce Lindsay, I am forever indebted for their encouragement and unconditional support, both financially and emotionally.

To my friends Shannon, Louise and Kari, without whose support, encouragement and great sense of humors I would have struggled to get to the end. Thanks also to Cecelia for her encouragement.

And finally, to Anthony, without whose emotional and financial support and encouragement, this thesis could never have been completed. I can't find the words to thank you enough for all you have done for me over the past few years.

TABLE OF CONTENTS

ABSTRACT	ii
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	v
LIST OF TABLES	xii
Introduction	1
Limitations To Human Judgement	2
Reliability.	3
Representative and Availability Heuristics.	3
Illusory Correlation.	4
Labelling.	5
Anchoring.	5
Confirmatory Bias.	6
Over-Reliance on Salient Data.	6
Information Overload and Increased Confidence.	7
Lack of Feedback.	8
Under Utilization of Base Rates.	8
Limited Insight.	8
Hindsight Bias.	9
The Role of Clinical Experience in Clinical Judgement.	10

Studies Not Supporting The Role Of Experience In Clinical Judgement. 11

Studies Supporting The Role Of Experience In Clinical Judgement. 13

The Role of Experience in Clinical Judgement in Neuropsychology. 14

Research In The Area of Clinician Confidence. 17

Research Into Inter-Rater Reliability. 19

Judgement Research - Clinical Versus Actuarial Methods 21

Studies Which Support The Use Of Actuarial Formulas. 22

Studies Which Support The Use Of Clinical Judgement. 24

Studies Which Support The Use of Actuarial Formulas in Neuropsychology. 27

Studies Which Support The Use of Clinical Judgement in Neuropsychology. 28

Estimation of Premorbid IQ. 29

Methods of IQ Estimation 30

Best Performance Method. 30

Clinician Judgement. 31

Hold – Don't Hold Method. 31

Reading Tests. 32

Actuarial and Demographic Based Methods. 35

Studies Investigating The Accuracy of IQ Estimation. 38

Current Practices Within the Area of Judgement. 40

Past Methodologies in This Area. 42

Drawbacks of Current Research in This Area. 43

Aims and Objectives 45

Hypotheses 46

Study One 47

Method 47

Participants 47

WAIS-III Data Group. 47

Clinician Rater Groups. 48

Demographic Information Clinician Rater Group. 50

Materials 51

Wechsler Adult Intelligence Scale- III. 52

Clinician Rater's Scoring Booklet. 52

Video of Interview. 53

Crawford Index. 53

Barona Formula. 55

Oklahoma Premorbid Intelligence Estimate. 58

Oklahoma Premorbid Intelligence Estimate-3. 59

Demographic Scoring Booklet. 61

Procedure 62

WAIS-III Data Group. 62

Clinician Rater's Group. 63

Demographic Information Clinician Rater Group. 64

Analysis of Clinician Judgement 65

Results 66

Analytic Plan 66

Preliminary Analysis 66

Hypotheses Testing – Part One Clinician Judgement 67

Crawford Equation Compared With Clinical Judgement. 67

Regression Equation Methods Compared With WAIS-III Data Group Obtained IQ. 70

Relationship Between Experience and Accuracy. 71

Relationship Between Clinician Confidence, Accuracy and Demographic Variables of The WAIS-III Data Group. 74

Clinician Confidence. 76

Inter-Rater Reliability. 78

Demographic Variables. 78

Relationship Between Amount of Information and Judgement Accuracy. 79

Relationship Between Clinician Confidence, Accuracy and WAIS-III Data Group Demographic Characteristics When Provided With Demographic Information Only. 81

Relationship Between Amount of Information and Clinician Confidence. 82

Discussion 84

Clinician Judgement 85

Actuarial Methods Compared With Clinical Judgement. 85

Relationship Between Experience and Accuracy. 89

Clinician Confidence. 91

Inter-Rater Reliability. 94

Demographic Variables. 96

Relationship Between Amount of Information and Judgement Accuracy. 98

Relationship Between Amount of Information, and Clinician Confidence. 99

Relationship Between Amount of Information, Demographic Variables, Accuracy and Clinician Confidence. 100

Study Two 102

Method 102

Participants 102

Australian Experienced Neuropsychologists. 102

International Experienced Neuropsychologists. 102

Australian Novice Neuropsychologists. 103

Materials 105

Neuropsychologist Questionnaire. 105

Procedure 105

Australian Experienced Neuropsychologists. 105

International Experienced Neuropsychologists. 106

Novice Australian Neuropsychologists. 107

Results 109

Hypotheses Testing –Clinicians' Practice Patterns in Assessing Premorbid IQ 109

Frequency of Premorbid IQ Estimations. 109

Circumstances of Premorbid IQ Estimations. 111

Methods Used to Estimate Premorbid IQ. 112

Clinician Training in the Area of Premorbid IQ Estimation. 115

Use of Demographic Variables in Premorbid IQ Estimation. 118

Premorbid Estimation of Other Areas of Functioning. 120

Discussion 123

Methods of Premorbid IQ Estimation 123

Methods Used To Predict Premorbid IQ. 123

Clinician Training Within The Area of IQ Estimation. 128

Clinician Use Of Demographic Variables. 130

Premorbid Estimates In Areas Other Than IQ. 131

General Discussion 133

Limitations 138

Future Research 139

Conclusion 141

REFERENCES 143

APPENDIX A 154
Informed Consent Form and Study Information
Sheet for Volunteers

APPENDIX B 157
Informed Consent Form and Study Information
Sheet for Novice and Experienced Clinicians

APPENDIX C 160
Clinician Rater Scoring Booklet

APPENDIX D 177
Clinician Rater Scoring Booklet With Demographic Information Only

APPENDIX E 182
Methods of IQ Estimation Questionnaire

LIST OF TABLES

Table 1 48
Demographic Characteristics of the WAIS-III Data Group.

Table 2 50
Demographic Characteristics of the Novice and Experienced Neuropsychologist Rater Groups.

Table 3 51
Demographic Characteristics of the Demographic Information Clinician Rater Group.

Table 4 57
Barona Formula Variable Weights.

Table 5 59
OPIE Variable Weights.

Table 6 61
OPIE-3 Variable Weights.

Table 7 68
Summary Statistics for the 15 Obtained WAIS-III IQs and Corresponding IQs Estimated By The Crawford Index, With and Without Adjustment For Use With The WAIS-III, and The 30 Clinician Raters.

Table 8 69
Comparison of Descriptive Ranges for Obtained IQ, Mean Clinician Estimates and the Mean Adjusted Crawford Index Estimates for the 15 WAIS-III Data Cases.

Table 9 70
Summary Statistics for the 15 Obtained WAIS-III IQs and Corresponding IQs Estimated By The Crawford Index, Barona Formula, OPIE And OPIE-3 with Adjustments for use with the WAIS-III, and the 30 Clinician Raters.

Table 10 73
Descriptive Statistics for Estimated IQ by the Novice and Experienced Clinician Raters.

Table 11 74
Mean Accuracy Scores for Each WAIS-III Data Case Sorted By Accuracy and Demographic Characteristics for Each Case.

Table 12 75
Pearsons Product Moment Correlations Among Clinician Accuracy in Estimating IQ, Clinician Confidence, Obtained IQ and Demographic Variables For The 15 Cases in the WAIS-III Data Group.

Table 13 77
Pearson Product Moment Correlations Between Measures of Confidence and Accuracy.

Table 14 79
Pearson Product Moment Correlations Among Clinician Estimates of IQ, Regression Equation Estimates of IQ and Obtained IQ and Gender, Social Class(Occupation), Age and Education.

Table 15 80
Descriptive Statistics for Accuracy Difference Scores.

Table 16 82
Pearsons Product Moment Correlations Among Clinician Accuracy in Estimating IQ, Clinician Confidence, Obtained IQ and Demographic Characteristics For The 15 Cases in the WAIS-III Data Group Based on Demographic Variables Only.

Table 17 104
Demographic Characteristics of the Australian Experienced, International Experienced and Novice Australian Neuropsychologists Groups.

Table 18 109
Percentage of Times Clinicians Estimate Premorbid IQ.

Table 19 110
Percentage of Time Premorbid IQ Estimate is Made Explicit.

Table 20 112
Methods Used For Estimating Premorbid IQ.

Table 21 114
Percentage Weight Given to Each Method in Determining Premorbid IQ.

Table 22 116
Percentage of Neuropsychologists Trained In Methods Of IQ Estimation.

Table 23 117
Percentage of Neuropsychologists who Believe Training In Premorbid IQ Estimation is Adequate.

Table 24 118
Demographic Variables Used To Estimate Premorbid IQ.

Table 25 120
Percentage Weight Given To Each Variable In Determining Premorbid IQ.

Table 26 121
Percentage of Time Premorbid Estimates Are Made In Areas Other Than IQ.

Table 27 122
Premorbid Abilities Routinely Estimated.

Health professionals are often required to make clinical decisions which may rely on the use of clinical judgement. Franzen, Burgess and Smith-Seemiller (1997) describe clinical judgement as "the complex organization of disparate sources of information and a concatenation of multiple assumptions" (p.712). Put simply, Holt (1958) described clinical judgement as "the problem solving or decision-reaching behavior of a person who tries to reach conclusions on the basis of facts and theories already available to him by thinking them over" (p.2).

Clinical judgement has been the subject of multiple investigations and much theoretical speculation over the past few decades. In particular, much research attention has been dedicated to clinical judgement within the fields of medicine (e.g., Einhorn, 1972; Gilpin et al., 1990) and clinical psychology (e.g., Jones, 1959; Grebstein, 1963). In more recent times, clinical judgement within the field of neuropsychology has begun to be investigated (e.g., Wedding, 1983; Kareken & Williams, 1994).

The judgement research to date has predominantly looked at clinical judgement in a wide variety of circumstances, ranging from predicting survival time after diagnosis of a disease and laterality of brain impairment, to risk of falling in the elderly and disorder diagnosis (e.g., Leli & Filskov, 1984; Ruchinskas, 2003). In particular, researchers have investigated the role experience plays in clinical judgement, clinician confidence in judgements, and inter-rater reliability between clinicians on judgement tasks, and the accuracy of clinical judgement in comparison to other methods of prediction such as actuarial formulas.

Despite the numerous studies that have been conducted within the area of clinical judgement, few straightforward answers have emerged. This will be highlighted in the following literature review, which summarizes the theory relating to human judgement and the key findings from research in the area of clinical judgement.

Limitations to Human Judgement

Clinical judgement is carried out by most clinicians working in health care settings. Recent research into cognitive functioning has begun to uncover some of the factors that make clinical judgement such a difficult task. Humans are susceptible to many errors in clinical judgement (Kahnemann, Slovic & Tversky, 1982; Garb, 1989). Flaws and biases inherent in cognition prevent humans from combining and integrating information accurately (Tversky & Kahnermann, 1974). These flaws and biases include ignoring base rates, assigning non-optimal weights to cues, and not acknowledging regression towards the mean (Tversky & Kahnermann, 1974). Heuristics such as representativeness (Kareken, 1997), also limits clinician judgment accuracy and these and other factors that contribute to judgement error will be further discussed.

Reliability.

One of the most salient of all judgement biases is the consistency and reliability with which humans think. Overall, humans are inconsistent in combining information (Kareken, 1997). Inconsistency in human judgement can occur as a result of many factors such as fatigue, distraction, time pressure or other extraneous variables (Kareken, 1997). This inconsistency in turn compromises the validity of diagnosis (Kareken, 1997). There are however other nonrandom factors that contribute to the unreliability of human judgement (Tversky & Kahnemann, 1974; Nisbett & Ross, 1980). Such factors include the availability and representative heuristics and illusionary correlation.

Representative and Availability Heuristics.

According to Kareken (1997), “under the effects of representativeness, people make predictions about an event by relying too heavily on the similarity (representativeness) of evidence to that event” (p. 703). For example, Kahnemann and Tversky (1973) gave subjects a description of a man who was very organized, not good at social interactions, was obsessive and had little interest in others and asked them to guess his area of study. The subjects guessed that this man was in the computer science field as this description fitted the stereotype of this profession, despite this being a relatively uncommon course at the time. In doing this, clinicians ignore base rates and information about how likely an event is to occur (Kahnemann &

Tversky, 1973). Availability heuristics refer to judgements that are based upon the ease with which information is retrieved from memory, such that the memory of a particularly salient case may influence the diagnosis of a subsequent patient (Kahnemann & Tversky, 1973).

Illusionary Correlation.

Illusionary correlation refers to the notion that the availability of information can create the perception that a relationship between variables exists when one does not (Chapman & Chapman, 1967). For example, a diagnostic sign paired with a pathognomic sign is likely to attract a clinician's attention, such as confabulation with Korsakoff amnesia, whereas a disease without a marker is not. This results in an inference that is based only on the conspicuousness of information and not on the actual degree of association (Jenkins & Ward, 1965). As such clinicians tend to pay attention to conspicuous details and let less salient features fade into the background (Einhorn, 1988).

According to Wedding and Faust (1989), "clinicians who work mainly with skewed or abnormal samples and attend mainly to pairings between signs and disorder will inevitably make false associations" (p. 244). They proposed that these confirming instances of personal experience override scientific evidence to the contrary. It has also been shown that the illusionary correlation increases as the amount of information available increases and is in part due to the limited working memory capacity of the human brain (Leuger & Petzel, 1979). It has been suggested that this is one reason why the illusionary correlation is robust and difficult to eliminate from clinical judgement, even when clinicians are made aware of it (Kurtz & Garfield, 1978).

Labelling.

Another limitation to clinical judgement proposed by Faust and colleagues is the tendency for more experienced clinicians to over-diagnose impairment (Faust, Guilmette, Hart, Arkes, Fishburne & Davey, 1988). Kareken (1997) proposes that a bias to find pathology may stem from a belief that on the basis of being a 'patient', a pathological condition is possessed until proven otherwise.

Anchoring.

It has also been noted that clinicians tend to make judgements subsequent to an initial value that is 'anchored' to their original estimate (Tversky & Kahnemann, 1974). For example, Friedlander and Stockman (1983) found that after initial exposure to a patient presenting as asymptomatic, clinicians disregarded pathognomic signs in the same patient that were introduced at a later time. It is also proposed that anchoring may be exacerbated by the human tendency to seek evidence that confirms a hypothesis (Watson, 1960).

Confirmatory Bias.

Clinicians form hypotheses quickly and tend to discredit or reinterpret information that is contrary to the hypotheses they hold (Arkes, 1981). Snyder (1981) found that people selectively seek evidence that confirm the hypothesis currently under consideration. According to Arkes (1981) this kind of biased information-gathering will lend support to the hypothesis.

An example of confirmatory bias within neuropsychology was provided by Wedding and Faust (1989), who stated that clinicians often posit the presence of a memory disorder and then search for supportive data to 'test' this hypothesis. They further proposed that almost everyone has some complaints of memory, and thus such a confirmation is likely to succeed even when appraising normal individuals.

Over-reliance on Salient Data.

According to Wedding and Faust (1989), "clinicians tend to overemphasize concrete and salient data in comparison to data which is more informative but 'bland'. For example a 'boring' history of normal employment may be overshadowed by a single, dramatic 'pathognomonic sign'" (p. 242). In addition Reitan (1986) noted that clinicians often mistakenly identify normal error for pathognomonic signs. Wedding and Faust (1989) proposed clinicians' preference for this kind of data explains why judgement accuracy actually decreases when more information is provided such as interview data in conjunction with test results.

Information Overload and Increased Confidence.

Confidence in one's decisions tends to rise in relation to increasing information (Wedding & Faust, 1989). Clinicians can only process a limited amount of information, and any additional information makes no difference in judgement performance (Wedding & Faust, 1989). This consistent finding is related to the limits of human cognitive capacity, whereby judgement accuracy increases until approximately seven pieces of information are presented (Miller, 1956). Further, clinicians have difficulty assigning optimal weights to information, so judgements which are based on valid data become altered by additional, but less valid data (Wedding & Faust, 1989).

The overconfidence clinicians have in their diagnoses is another factor that contributes to flawed human judgement (Arkes, 1981). For example, Oskamp (1965) found that once provided with more information, clinicians confidence increases but not their judgement accuracy. Further, Holsopple and Phelan (1954) found that the most confident diagnosticians tended to be the least accurate. Part of the reason why clinicians tend to have inflated confidence may be related to the fact that they rarely receive feedback with regards to judgement accuracy.

Lack of Feedback.

Within the field of psychology, clinicians often receive little or no feedback about their diagnoses and predictions (Dawes, Faust & Meehl, 1989). For example, clinicians asked to predict violence might never learn whether their predictions were correct. Without clear information about judgement accuracy, the validity of the variables used to make clinical judgements is questionable (Dawes, Faust & Meehl, 1989).

Under-utilization of Base Rates.

In addition to rarely receiving feedback with regards to judgement accuracy, clinicians often fail to evaluate their own accuracy relative to base rates for various disorders. For example, neuropsychologists may devote considerable time to determining a differential diagnosis between Pick's Disease and Dementia of the Alzheimer's type without recognizing that Pick's is much less likely on the basis of its lower base rate (Wedding & Faust, 1989). Wedding and Faust (1989) argued that this problem is contributed to by the fact that that many texts do not address the high frequency disorders such as dementia whilst dedicating considerably more attention to rare disorders.

Limited Insight.

According to Arkes (1981) "no matter what bias a diagnostician demonstrates, attempts to eliminate it would be fostered by awareness of one's own clinical judgement process" (p. 325). This notion is supported by

Brehmer, Kuylenstierna and Lilijergren (1974), who found humans have negligible awareness of the factors that influence their judgement.

Appraisal of one's owns cognitive processing poses a major challenge, as much processing occurs at an automatic level that is devoid of awareness and thus is not available for self analysis (Wedding & Faust, 1989). Studies that have investigated this notion have found that despite individuals believing that they have carefully weighed and considered a large number of variables in reaching a judgement, they have in fact only considered a couple of variables (Wedding & Faust, 1989).

Hindsight Bias.

Another factor that contributes to impaired human judgement is hindsight bias. Hindsight bias describes how clinicians react when given feedback or outcome information (Garb, 1989). According to Fischhoff (1975), "finding out that an outcome has occurred increases its perceived likelihood. Judges are however, unaware of the effect that outcome knowledge has on their perceptions. Thus, judges believe that this relative inevitability was largely apparent in foresight" (p. 297).

Thus, people with hindsight knowledge about an event such as a diagnosis believe that they would have been able to predict the event with considerable accuracy (Wedding & Faust, 1989). An example of hindsight bias in neuropsychology is when a clinician evaluates the test performance of a person post head injury and finds impairment "consistent with a closed head injury" (Wedding & Faust, 1989). The authors noted that this can also occur after perusing neuroimaging findings, which can lead to the establishment of 'cognitive sets' that can affect clinical judgement, resulting in the findings that the test results are 'consistent with' the neurodiagnostic findings. This notion is supported by Arkes (1981), who stated that "there is always enough evidence in a rich source of data to nurture all but the most outlandish diagnosis". However, it has been found that individuals do not seem to recognize the operation of hindsight bias and education about this phenomena does not seem to eliminate its effects (Fischhoff, 1977).

Thus many factors contribute to the fallibility of human judgement. Further, simply being aware of and recognizing these limitations is not enough to remedy many of these limitations in human reasoning. It should be noted however, that most studies investigating the limitations of human judgement have been within the field of clinical psychology. Few studies have been conducted within the area of neuropsychology, hence further research is needed in order to clarify how cognitive biases arise in the context of neuropsychological assessment.

The Role of Clinical Experience in Clinical Judgement.

Clinicians and the professional bodies to which they belong often assume that judgement accuracy improves with experience (Faust et al., 1988). This assumption has been investigated empirically, and has received many reviews in the judgement research literature (e.g., Wedding, 1983;

Garb & Schramke, 1996). Overall, research investigating the role experience plays in the accuracy of clinical judgement has yielded mixed results.

Studies Not Supporting The Role of Experience in Clinical Judgement.

Many studies that have investigated the role of experience in clinical judgement have shown that a clinician's level of training and years of experience are not related to judgement accuracy (see Garb & Schramke, 1996). In order to investigate the role experience plays in the accuracy of clinical judgement within the field of psychology, a task force was formed by the American Psychological Association in the 1980's to investigate education, training and services (American Psychological Association, 1982). This task force found no evidence that professional training and experience was related to professional competence. The finding that judgement accuracy is not related to experience is not peculiar to psychology, but has also been found to be true within the field of medicine. For example, in a study conducted by Kendell (1973), the ability of psychiatrists to make psychodiagnoses from an observed interview with psychiatric patients was not related to years of experience in psychiatry. In another study, Levenberg (1975) investigated the accuracy of psychologists compared with an expert judge in rating 'normal or disturbed' on the basis of Kinetic Family Drawing Protocols. It was found that the PhD psychologists were accurate on 72% of the cases compared with the expert judge, who was only accurate on 47% of the cases. Similar results were found in a study conducted by Werner, Rose and Yesavage (1983). In this study, 30 psychologists and psychiatrists were asked to predict assaultiveness from a Brief Psychiatric Rating Scale. It was found that the validity of predictions was not significantly related to years of clinical experience. It is possible however that these two findings reflect either inadequacy with the test, or improper use of the tests to diagnose complex behavior or predict diagnoses.

Garb (1989) conducted a more recent review of studies investigating training, experience and clinical judgement within the mental health field. He found that in relation to personality assessment, experienced clinicians were generally no more accurate than less experienced clinicians when it came to clinical judgement. He concluded that studies in this area generally fail to support the value of experience in mental health. However, the results provided some limited support for the value of training.

There are many reasons why experienced clinicians may make less valid judgements than their less experienced colleagues. Garb (1989) proposes that outdated training, failure to acquire good assessment information, lack of feedback, or flawed feedback and inadequate cognitive processes may be some factors which contribute to the poor judgements made by experienced clinicians. Overall, these studies suggest training and years of clinical experience are not related to the accuracy of clinical judgements within the fields of psychiatry and psychology.

Studies Supporting The Role of Experience in Clinical Judgements.

In contrast to the findings that clinician training and expertise play no role in the accuracy of clinical judgements, several studies have provided support for the role of training and experience in improving the accuracy of clinical judgments. For example, in an early study, Grigg (1958) found experienced clinicians to be more accurate in predicting patients' answers on personality test items after being given recordings of patient interviews than did undergraduates.

Goldberg (1959) also found training to be related to judgement accuracy in a study which assessed an expert and PhD level psychologists' diagnosis of organic brain damage using the Bender Gestalt Test. It was found that the expert judge was more accurate than the PhD level psychologists.

Jones (1959) investigated the accuracy of clinical psychologists with PhDs compared with undergraduates in rating the severity of schizophrenic pathology. Judges were given WAIS vocabulary and comprehension scores. It was found that the clinical psychologists made more valid ratings than the undergraduate students did. Another early study which showed an expert to be more accurate than a novice clinician was conducted by Grebstein (1963). In this study judges used Rorschach to predict IQ. It was found that clinical psychologists were more accurate than graduate students who had just completed a Rorschach course, but not more accurate than those students who had just done the course and had had some professional experience.

Waxer (1976) found training and experience had a positive effect on judgement accuracy. In this study, clinical psychology graduate students' judgement accuracy was compared to the accuracy of undergraduate students by providing the judges with videotapes of interviews with psychiatric patients. It was found that the graduate students made more valid ratings of depression severity than did the undergraduates.

Overall, despite evidence that experience and training plays little role in judgement accuracy, there is also some evidence to support the value of training and experience within the mental health field. In addition to investigating the role experience plays in clinical judgement within mental health in general, recent research has also focused on the effects of training on clinical judgement accuracy within the field of neuropsychology. However, this research has also yielded mixed results.

The Role of Experience in Clinical Judgement in Neuropsychology.

Several studies have compared the accuracy of experienced clinicians to novice clinicians in decision-making within neuropsychology. Wedding (1983) compared the diagnostic accuracy of one expert neuropsychologist, three graduate clinical psychologists and ten PhD level psychologists. The judges were asked to classify subjects into one of five categories using the Halstead Reitan Battery: schizophrenia, left hemisphere impairment, right hemisphere impairment, diffuse brain damage, and normal. Judgements made by the experienced neuropsychologist were found to be more accurate than the judgements made by the less experienced judges. However, experience was not strongly associated with accuracy, and for all judges combined, accuracy was not significantly correlated with years of experience using the Halstead Reitan Battery. However, as only one expert neuropsychologist was used in this study, one can question if they were representative of other experts in the field.

Faust et al. (1988) assessed the judgement accuracy of more and less experienced neuropsychologists. In this study, the clinicians were asked to record localization, etiology, and overall appraisal of cases presented to them in the form of test scores such as the WAIS-R and The Halstead Reitan Battery. It was found that there were no relationships between training, experience, and judgement accuracy.

In a later literature review on clinical judgement in neuropsychology, Wedding and Faust (1989) argued there was virtually no data to suggest that judgement accuracy is related to experience, or to confidence in accuracy of one's predictions. Garb and Schramke (1996), who conducted a review on judgement research in neuropsychology, found similar results. They found similar results for the effects of training and experience in neuropsychology as in personality assessment, which provided little support for the role of training in judgement accuracy. Garb and Schramke (1996) concluded that level of clinical experience is unrelated to validity of judgements made. The authors pointed out, however, that a pitfall of research within this area is that most studies inadequately describe the clinicians' total number years of training, highest qualifications obtained, and levels of experience the clinicians have had, such as years working as a qualified neuropsychologist.

Studies which have shown that experience plays no role in clinician judgement have received much criticism, and the findings have, in part, been attributed to poor methodological designs. For example, Saling (1994) criticized Faust et al (1988) for using decision-making tasks that ranged from "mundane to impossible", with little in between to conclude that experience played no role in clinical judgement. According to Saling (1994), "Despite relatively competent judgements, not one of the neuropsychologists identified the etiology of the disorder. The reason for this has nothing to do with expertise or training. The cystic nature of the lesion is simply not reflected in the test scores. Guessing is no measure of the effects of experience" (p.8).

Therefore, within the field of neuropsychology, studies investigating the role of experience and training on clinician judgement have produced results that do not provide strong support for the value of training in relation to judgement accuracy. However, poor research designs, lack of explicit information regarding the clinicians' level of training and years of clinical experience, and other methodological issues may have rendered these results ambiguous overall. There is therefore a need for more research with better methodologies in order to clarify the role experience plays in neuropsychologists' judgement accuracy.

Research in the Area of Clinician Confidence.

Clinicians' confidence in their judgements is another area that has been studied extensively in the judgement literature. In order to investigate clinicians' confidence, clinicians are usually asked to make ratings to describe how much confidence they have in the judgements that they make. The accuracy of their judgements are then compared to the ratings that they gave for each judgement, in order to determine the appropriateness of their confidence levels in relation to their judgements (e.g. Karaken & Williams, 1994).

Many empirical studies have shown neuropsychologists to be frequently overconfident about their judgements (Garb & Schramke, 1996). For example, in a study conducted by Faust, Hart and Guilmette (1988), in which neuropsychologists were asked to detect malingering after receiving psychometric information and a history, it was found that clinicians were overconfident. They had a 100% error rate yet stated that they were moderately confident in their judgements.

In contrast, Garb's (1989) review of the role of experience in clinical judgement concluded that there is some evidence that experienced clinicians make better confidence ratings than do inexperienced judges, as they are better at knowing which of their judgements are likely to be correct and which ones are likely to be wrong. In another study, Gaudette (1992) asked 6 neuropsychologists to diagnose the presence or absence of brain impairment and to state whether it was diffuse, or lateralized to the left or the right hemisphere. It was found that neuropsychologists were overconfident in their judgements, estimating that they were correct in 77.5% of the cases when in fact they were only accurate in 62% of the cases.

Kareken and Williams (1994) required clinicians to express degrees of confidence about IQ estimates, which were compared to those gained using an actuarial formula. In this study, the neuropsychologists were required to set 95% confidence intervals around their estimates. It was found that the clinicians' confidence estimates were appropriate for VIQ but too small for PIQ. Overall, although the clinicians' IQ estimates were close to those of the actuarial formula, their confidence was considerably higher.

Garb and Schramke (1996) conducted a review and meta-analysis which pooled the data from two studies investigating judgements made by neuropsychologists (Wedding, 1983; Gaudette, 1992). Both studies reported the average level of accuracy and confidence rating for each clinician. The correlation between accuracy and confidence was .29. The authors concluded that confident judges make ratings that are only slightly more valid than ratings made by less confident judges and that "unlike validity, there was a trend for the appropriateness of confidence ratings to be positively related to experience and presumed expertise" (p.147).

Dawes, Faust and Meehl (1989) proposed that judgement errors arise because clinicians are far more likely to evaluate individuals with significant problems than with those without them, and this skewed exposure hinders attempts to make all of the needed comparisons. They also suggested that these same factors promote the overconfidence frequently reported in the literature on judgement research.

In summary, these findings indicate that neuropsychologists may be frequently overconfident. However, it should be noted that relatively few studies within the area of neuropsychology have investigated the appropriateness of confidence ratings. Therefore, further research in this area is necessary, and the conditions under which neuropsychologists are likely to be overconfident need to be identified.

Research into Inter-rater Reliability

Another area that has also received much research attention within the judgement literature is inter-rater reliability. Unlike the research focussing on clinician confidence or the role of training and experience, some studies within this area have revealed that overall, the inter-rater reliability of estimates made by clinicians ranged from good to excellent.

For example, Brown, Spicer, Robertson, Baird and Malik (1989) investigated inter-rater reliability for a judgement task that involved neuropsychologists rating the likelihood of lateralised brain dysfunction. They obtained an intra-class correlation coefficient (ICC) of .84, indicating good rater agreement. Similar results were obtained by Garb and Schramke (1996), who conducted a review of studies investigating judgement within the field of neuropsychological assessment. The authors calculated Kappa coefficients for two studies (Grant, Mohns, Miller & Reitan, 1976; Grant, Heaton, McSweeny, Adams, & Timms, 1982) in which clinicians were required to make judgements on the basis of test scores. It was found that reliability was good (.65 and .70 respectively) for the two studies.

However, not all studies have reported good inter-rater reliabilities on judgement tasks. Poor inter-rater reliabilities have been found within the area of IQ estimation. For example, Kareken and Williams (1994) investigated the inter-rater reliability of neuropsychologists in a judgement task that involved predicting Verbal IQ and Performance IQ for fabricated subjects. The authors concluded that "inter-rater agreement was less than ideal and estimates varied between clinicians" (p.90). This study however had methodological limitations as the use of fabricated subjects meant that the actual clinical practices used by the clinicians were not being assessed.

Thus, like the other areas of research that have investigated clinician judgement, investigations into inter-rater reliability have produced mixed results. Overall, there have been relatively few studies conducted within the field of neuropsychological assessment that have specifically investigated the inter-rater reliability of clinical neuropsychologists.

Judgement Research – Clinical versus Actuarial Methods

Despite the many studies that have focused on the role of experience, clinician confidence, and inter-rater reliability in the accuracy of clinical judgement, the debate within the area of judgement research to date has predominantly centered on the accuracy of statistical or actuarial methods of prediction compared to clinical judgement. Wedding and Faust (1989) provide an eloquent definition of these two methods. "In the clinical method, a judge combines or processes information in his or her head. In the actuarial method, the human judge is eliminated and 'conclusions' are based strictly on empirically established relations between data and criterion" (p.234).

Since Meehl's seminal study in 1954 and Holt's subsequent reply in 1958, the debate over clinical versus actuarial methods of prediction has raged, not only in psychology, but also in medicine and business. According to Holt (1958), "..by posing the question of clinical versus statistical prediction, he [Meehl] has encouraged two warring camps to form" (p.1). This general view of the debate is reiterated by Einhorn (1986) who stated that "the clinical versus statistical prediction controversy is enduring and general" (p.387).

As a result, clinical versus statistical prediction is an area that has been subjected to extensive empirical investigation and much theoretical speculation. However, like the other areas of clinical judgement that have been investigated in the judgement research, few straightforward answers have emerged.

Overall, this research has shown that human judgement is fallible, but that clinical judgement can, in certain circumstances, be as accurate or more accurate than the other judgement methods available. As such, the research in this area can be divided into studies that show superiority of the actuarial method, and those that show superiority of clinical judgement.

Studies Which Support The Use of Actuarial Formulas.

Many studies have shown the superiority of statistical methods of prediction. In one of the earliest empirical studies to test the accuracy of clinical judgement compared with statistical prediction, Sarbin (1943) found that clinical judgement was no more accurate than an actuarial formula. Subjects were five clinical psychologists who were asked to predict academic success of 162 students based on an interview and students' records. This was then compared to an actuarial formula developed to measure academic success. Sarbin concluded "in short, a competent statistical clerk can make predictions as well as a highly trained clinical worker" (p.600).

Similar superiority of statistical methods of prediction was provided by Einhorn (1972), who investigated pathologists' predictions of survival time following diagnosis of Hodgkins disease. An actuarial formula was developed to predict survival time based on the pathologists' ratings and actual survival time of the first 100 cases. It was found that the actuarial formula was superior at predicting survival time to the pathologists' judgements.

Other empirical studies in the field of medicine and psychology have yielded similar results. Leli and Filskov (1984) studied the diagnosis of progressive brain dysfunction. Groups of experienced and inexperienced clinicians were asked to identify cases based on intelligence testing. They were then also given a decision rule that had been shown to predict cases. It was found that neither group matched the decision rule's 83% accuracy.

Similarly, Dawes, Faust and Meehl (1989) found the actuarial method to be superior in their review of actuarial and clinical methods of prediction. They concluded that when actuarial formulas are developed and used properly, they can provide benefits such as saving time and money even if they are merely equal to clinical judgement.

Several meta-analyses have been conducted to compare clinical and actuarial methods. In a meta-analysis of 136 judgement studies within the fields of medicine and psychology, Grove, Zald, Lebow, Snitz and Nelson (2000) found that on average, actuarial prediction techniques were about 10% more accurate than clinical predictions. It was further found that depending on the specific analysis, mechanical prediction substantially outperformed clinical prediction in 33% - 47% of the studies examined. The authors noted however, that although clinical predictions were as accurate as mechanical predictions, in only a few studies were they substantially more accurate. Only 8 of the 136 studies notably favored clinical prediction. It was also found that medically trained judges did not differ from psychologists in how inferior they were to mechanical prediction, and that the amount of data given to judges relative to the mechanical formula made little difference in the relative superiority of mechanical prediction.

Thus, outside of the field of neuropsychology, there is much empirical evidence to support the superiority of statistical methods of prediction when compared with clinical judgement. However, it is also worth reviewing the evidence that supports the use of clinical judgement methods.

Studies Which Support The Use of Clinical Judgement.

Although not an empirical study, Holt's (1958) review of Meehl's (1954) seminal study was one of the first publications to argue for the use of clinical judgement over actuarial methods of prediction. In this article, Meehl's citation of studies was criticized as none of them cross-validated clinical predictions. Holt stated that this alone was a major reason to expect superior performance from actuarial predictions.

Holt suggested that by comparing actuarial methods to "naïve" clinical methods instead of "sophisticated" clinical methods (those employed in practice) the results of these findings were therefore not accurate. "With an inadequate sample of information about a person, no matter how sophisticated a technique of prediction, there is a low ceiling on the predictive validity that can be attained" (p.8). The underlying argument of this article was that when clinicians use methods with which they are familiar and predict a performance about which they know something, as well as have access to a rich body of data, clinical prediction can achieve success. However, few studies have investigated this empirically.

Similar arguments have been put forward by other researchers in this field to support the role of clinical judgement. For example, studies that have shown a superiority of actuarial methods over clinical judgement are frequently criticized for using methodologies that place the clinician at a disadvantage. For example, Adams and Putman (1994) criticize the methodologies employed in studies that have found actuarial methods to be more accurate than clinical judgement. They argued that unsystematic sampling, inappropriate 'clinical judgement' tasks, and research designs that instruct incomplete efforts on tests are not proof that actuarial methods are superior. However, a review of the literature revealed a few studies within this area that demonstrated that clinical judgement was at least as accurate as the actuarial method (e.g. Grebstein, 1963; Goldstein, Deysach & Kleinknecht,1973).

In an early empirical study, Grebstein (1963) investigated the accuracy of actuarial prediction compared to clinician judgement. In this study, 15 clinicians (student and PhD level clinical psychologists), were asked to predict Wechsler-Bellevue IQ scores from Rorschach psychograms which were then compared to the IQs derived from a regression equation. It was found that there was no significant difference between the degree of accuracy of the judgements made by the equation or the clinicians. The predictions made by the equation had a .56 correlation with the actual IQ and the sophisticated group of clinicians had a correlation of .68 with the actual IQ scores.

Similarly, Goldstein, Deysach and Kleinknecht (1973) found superiority of clinical judgement over a statistical formula. In this study clinicians were found to be more accurate at predicting cerebral impairment than the statistical formula. Heaton, Grant, Anthony, and Lehman (1981) also conducted a study which compared the diagnostic accuracy of two clinicians with an actuarial system for interpreting the Halstead Reitan Test Battery. Clinicians' ratings were more accurate than the automated system in predicting the presence and laterality of structural cerebral lesions.

Similarly, comparative research in medicine has shown a slight clinical advantage in some studies that involve judgements that rest on firm theoretical grounds (Dawes, Faust, & Meehl, 1989). For example, Gilpin et al. (1990), conducted a study which investigated the prediction of cardiac death within 1 year of being discharged from hospital following acute myocardial infarction. It was found that there was no difference between the cardiologists' predictions and those made by the statistically derived decision rule.

Thus, outside the field of neuropsychology, there is evidence to suggest that mechanical prediction can be as accurate or more accurate than clinical prediction. However, as Grove, Zald, Lebow, Snitz, and Nelson (2000) point out in their meta-analysis of studies investigating clinical versus mechanical prediction, this superiority is not completely uniform and in certain instances, such as predicting cerebral impairment or interpreting test scores, the clinical method is notably more accurate.

Studies Which Support The Use of Actuarial Formulas in Neuropsychology.

There have been a limited number of studies conducted within neuropsychology which have investigating the superiority of either statistical prediction or clinical judgement. A review of the literature revealed only a few neuropsychological studies that favored the actuarial method of prediction. For example, Wedding (1983) conducted a study which compared PhD level neuropsychologists' judgements to that of an actuarial formula in diagnosing etiology and chronicity of hypothetical cases. It was found that the more information the judges received, the more inaccurate they were. It was also found that the actuarial approach was superior to the clinical judgement method.

Wedding and Faust (1989) conducted a review on clinical and actuarial prediction within the fields of clinical psychology and neuropsychology. The review revealed that current research on judgement in neuropsychology is consistent with research in other areas that documents the limitations of humans as judges and argues for the increased utilization of actuarial formulas. The authors concluded that "...greater overall accuracy is achieved when one adheres 'mindlessly' to actuarial procedures and never makes exceptions based on clinical judgement" (p.237).

Studies Which Support The Use of Clinical Judgement in Neuropsychology.

A review of the neuropsychology literature did not revealed any studies that supported the use of clinical judgement over an actuarial formula. However, there has been a very limited amount of research conducted within the area of clinical judgement in neuropsychological assessment, and very few studies conducted within the area of neuropsychological assessment that have sampled clinical neuropsychologists, as most have used clinical psychologists in these neuropsychologically-based judgement tasks.

In defense of neuropsychologists' clinical judgement is Garb and Schramke's (1996) meta-analysis which investigated reliability and validity judgements, appropriateness of confidence intervals, value of training, and cognitive processes of neuropsychologists. The authors questioned the repeated findings that actuarial formulas are more accurate than clinical judgement because in most studies the clinician is generally only presented with psychometric and demographic information.

Garb and Schramke suggested that the statistical formula may not be as accurate as the clinical judgement if the clinician had access to all the information that is usually available to them in clinical practice. Overall, the results of their analysis revealed that neuropsychologists are able to make moderately valid judgements and that their judgements are consistently more valid than chance level and base rate levels of accuracy.

Therefore, within the field of neuropsychology, relatively few studies have investigated the accuracy of clinical judgement compared with statistical prediction. Overall, despite the relative superiority of statistical prediction in most studies within the judgement research literature, these findings need to be interpreted with caution, as several methodological flaws have been identified, thereby rendering these findings ambiguous with regards to actual clinical practice. In particular, more studies are needed within the field of neuropsychology in order to determine the accuracy of clinical judgement compared with statistical prediction methods.

Estimation of Premorbid IQ.

Recent research attention has focussed on the accuracy of clinical judgement of premorbid intellectual function (e.g. Crawford, Millar, & Milne, 2001; Karaken & Williams, 1994). According to Franzen, Burgess, and Smith-Seemiller (1997) "clinical judgement is often used in premorbid IQ estimation in clinical neuropsychological assessment, when there is a need for some baseline against which to compare current performance.

Actual premorbid test scores antedating cognitive decline are seldom available, yet diagnosis often requires that some degree of decline be demonstrated" (p711). As with most clinical decisions, the determination of this change usually involves subjective clinical judgement in order to determine if the present test scores are likely to have occurred in the absence of the disorder or injury. However, several methods aside from clinical judgement also exist to help clinicians come to an estimate of premorbid ability.

Several methods exist to determine premorbid intelligence. These range from using a test of current functioning that is resistant to the effects of neurological and psychiatric disorders, specifically designed premorbid IQ tests, actuarial and regression based statistical formulas, and clinical judgement. The most widely used (Spreen & Strauss, 1998) methods are outlined below.

Best Performance Method.

The best performance method consists of identifying the highest test score or the highest level of functioning in everyday tasks. This is then used as the standard against which to compare other aspects of the individual's current performances (Lezak, 1993). However, this method has been criticised as it does not take into account the normal variability among tests and has been shown to overestimate premorbid IQ (e.g. Matarazzo & Prifitera, 1989; Mortensen, Gade, & Reinisch, 1991).

Mortensen, Gade, and Reinisch (1991) note that although a general 'g' factor accounts for some of the variance in intellectual performance, it cannot account for all the variance, and that scatter is often seen in the test scores of non-neurologically impaired people. Another criticism of the Best Performance Method is related to the psychometric aspects of tests and test measurement (Franzen et al., 1997). According to Matarazzo and Prifitera (1989), the magnitude of the standard error of measurement as well as less than perfect test reliabilities contribute to the scatter seen in cognitive test scores.

Clinical Judgement.

An individual's premorbid ability level can also be estimated from behavioural observation and historical facts. These estimates can be drawn from interview, reports from family and friends, prior academic achievement and employment, or an intellectual product such as an invention (Lezak, 1995). In a study investigating clinician practice patterns in estimating IQ, Smith-Seemiller et al. (1997) found that clinicians tend to use information such as demographic variables such as years of education, occupation as well as interview and school reports to make clinical judgements in this area.

Hold – Don't Hold Method.

The Hold-Don't Hold strategy estimates premorbid ability based on the individual's current performance on a measure that is considered to be relatively resistant to neurological impairment (Lezak, 1995). The most commonly used tests within this area are the Vocabulary and Information subtests of the Wechsler Scales, as these are believed to be the best indicators of premorbid intelligence (Spreen & Strauss, 1998).

However, although the Vocabulary subtest is the most resistant to brain damage, this ability can become markedly impaired in some clinical conditions, leading to an underestimation of premorbid IQ in those instances (Lezak, 1995). The Information subtest reflects crystallised intelligence in the form of general knowledge and is also relatively resistant to the effects of brain damage (Sattler, 2001). However, performance on this subtest is heavily reliant on educational opportunities and as such has been criticised as being misleading in individuals with a limited educational history (Spreen & Strauss, 1998).

Reading Tests.

Reading tests are frequently used to measure premorbid IQ (Smith-Seemiller et al., 1997). Tests of over-learned skills such as reading have been found to be highly correlated with intelligence level in the general population (Spreen & Strauss, 1998). Reading of irregular words is thought to be more resistant to cognitive decline than is reading of regular words as it has been found to tap previous knowledge while minimizing the demands on current cognitive capacity (Franzen, Burgess & Smith-Seemiller, 1997). But there is evidence that even reading declines with some conditions such as Alzheimers.

One of the earliest reading tests used to estimate premorbid intellectual functioning was the Schonell Graded Word Reading Test (GRWT) (Schonell, 1942). This test however, was originally designed for use in children and therefore a ceiling effect occurs when it is used in adults of above average intelligence (Schonell, 1942).

The National Adult Reading Test (NART; Nelson & McKenna, 1975) is another reading test that is frequently used. The NART is based on the principle that words with irregular grapheme-phoneme correspondence are better indicators of premorbid IQ than those which follow regular pronunciation rules. It consists of 50 irregularly spelled words such as ache, psalm and gouge. The conceptual basis of the use of the NART is however, open to criticism. The underlying assumption of the NART is that if an individual can correctly pronounce a word then that word was previously in that person's vocabulary. If not, it is seen as limitation of that person's vocabulary (Franzen et al., 1997). This test also has drawbacks as it cannot be used in individuals who are aphasic, have visual acuity problems, or who are dyslexic (Crawford, 1989, 1992). It has also been found to underestimate premorbid IQ at the higher end and overestimate IQ at the lower end of abilities (Ryan & Paolo, 1992).

Other variants of the NART are also available such as the National Adult Reading Test – Revised (NART-R; Blair & Spreen, 1989), which is a version of the of the NART modified specifically for use on a North

American population as opposed to a British population for which the NART was originally formulated (Franzen et al., 1997). This version however eventually became referred to as the North American Adult Reading Test (NAART; Spreen & Strauss, 1991). A third version was also developed by Grober and Sliwinski (1991) known as the AMNART which is also based on a North American population. Caution needs to be taken in using these tests outside of the populations for which they were developed as pronunciation rules can differ between countries (Franzen et al., 1997).

Another commonly used reading test to estimate premorbid IQ is the reading component of the Wide Range Achievement Test – Revised (WRAT-R), or some other edition of this test (Wilkinson, 1993). The WRAT-R is different from the NART in that it includes both regularly and irregularly spelled words (Franzen et al., 1997). Although initial studies indicate that the WRAT-R may be less likely to overpredict IQ in the lower ranges than the NART, it too has been found to have limitations, as it likely to underestimate IQ in the higher ranges (Wiens, Bryan, & Crossen, 1993). In addition, both these reading-based methods (the NART and the WRAT) have been criticized as being likely to underestimate the premorbid ability in individuals who have pre-existing verbal deficits (Spreen & Strauss, 1998).

The Wechsler Test of Adult Reading (WTAR; The Psychological Corporation, 2001) is one of the few methods available to estimate WAIS-III IQ scores. The WTAR is based on the same approach as the NART. A score is obtained by adding the number of correctly pronounced words and estimated IQ scores are obtained via three methods; demographic variables only, current reading ability, or a combination of both current reading ability and demographic variables. The combined WTAR and demographic method reportedly yields the most closely approximated IQ score to that of the distribution of a healthy adult population (The Psychological Corporation, 2001). One reported drawback of this test, along with all reading based estimates, is that it cannot be used on individuals who have a history of reading difficulties.

Actuarial and Demographic Based Methods.

Demographic variables are another measure frequently used to estimate premorbid IQ, as demographics such as education and occupation are related to IQ. As such, they may provide an estimate of premorbid level of function (Crawford, 1992). This method is particularly useful when assessing patients with neurological impairment, as patients' performance on tests of ability can be affected by acquired brain injury, but demographic background is unaffected (Franzen, et al., 1997).

Demographic variables such as education and occupation have been shown to have a high correlation with intelligence (correlation with WAIS-R Full Scale IQ = .53 and .36 respectively; Barona et al., 1984). However, it has been noted that there are few published methods using these achievement variables on their own (Schinker & Vanderploeg, 2000). Another drawback to using demographic variables is that it has been found that clinicians tend to give more weight to them than actually warranted. For example, in a study conducted by Kareken and Williams (1994), clinicians believed that the correlation between education and WAIS-R VIQ was .85 when it was actually .56. This kind of overestimation will ultimately lead to an unreliable estimate of premorbid abilities.

Actuarial methods that use demographic data such as gender, years of education, race, and type of occupation, have also been used to estimate premorbid IQ (Spreen & Strauss, 1998). The use of actuarial formulas based on demographic information is frequently cited as superior to clinical judgement as they have a higher rate of interrater reliability due to an absence of subjective judgement (Barona et al., 1984). However, in a review of methods to estimate premorbid function by Franzen and colleagues (1997), it was recommended that more empirical work was needed regarding the accuracy of demographic prediction models. A number of regression equations have been devised to calculate premorbid IQ (e.g. Wilson et al., 1978; Reynolds & Gutkin, 1979; Barona, Reynolds, & Chastain, 1984; Crawford & Allan, 1997).

The Barona Index (Barona et al., 1984) is based on the demographic variables of age, gender, race, education, occupation region and area of residence. It was thought to be promising in estimating premorbid IQ, however the standard errors of estimate for the regression equations are rather large (Spreen & Strauss, 1998). It was noted that the use of regression formulas to calculate premorbid IQ should be considered an estimate of functioning rather than an exact prediction, limiting the utility of the procedure for individual cases (Barona et al., 1984).

The Crawford Index, designed by Crawford and Allan (1997) is based on the demographic variables of occupation, education and age. It has been found that in general, prediction of Verbal IQ and Full Scale IQ is better than that for Performance IQ (Spreen & Strauss, 1998). It should be noted however, that this equation is based on predictions for the WAIS-R which has since been replaced with the WAIS-III. To date, only one demographically-based regression equations for the WAIS-III has been developed.

Another formula used to predict premorbid IQ is the Oklahoma Premorbid Intelligence Estimate (OPIE; Krull, Scott & Sherer, 1995; OPIE-3; Schoenberg, Scott, Duff & Adams, 2002). This is a linear prediction algorithm designed to combine current performance measures with demographic information to predict premorbid IQ (Spreen & Strauss, 1998). The OPIE equations provided reasonable estimates of IQ (Spreen & Strauss, 1998). As the OPIE-3 is relatively new, comparisons between this and other methods of premorbid IQ estimations such as the WTAR are not available to comment upon.

Thus, numerous methods exist to estimate premorbid intellectual function. However, none of these methods satisfy the need for a reasonably accurate estimate of premorbid ability. Several studies have been conducted within the field of IQ estimation, comparing these various methods in order to determine which is superior. In particular, some studies have investigated the relative merits of clinician judgement in relation to actuarial methods of IQ estimation.

Studies Investigating the Accuracy of IQ Estimations.

A large body of research suggests that estimation of premorbid IQ is precisely the circumstance in which clinicians render themselves susceptible to judgement error (Dawes, Faust & Meehl, 1989; Garb, 1989). However, limited literature exists in relation to the accuracy of actuarial predictions of IQ in comparison with clinicians' estimations of IQ. A review of the literature revealed only two studies that have compared an actuarial formula to clinician estimates of an individual's IQ.

Moreover, a review of the literature revealed only one study within this area utilizing neuropsychologists. Kareken and Williams (1994) compared expert neuropsychologists' estimations of IQ to that of an actuarial formula. Clinicians made their judgements based on selected demographic information of hypothetical patients. Their estimates were then compared to those obtained based on the same demographic information using the Barona formula. It was found that neuropsychologists' estimates were as a group, very similar to the Barona estimates of the same individuals. In the final analysis however, although the clinicians' mean estimates were similar to the Barona Formula's, inter-rater agreement was less than favourable and estimates varied between clinicians. The authors concluded that "In this context, actuarial formulae anchor clinicians to one estimate and control idiosyncratic bias" (p.90). One limitation of this study was that no objective IQ data was available.

Crawford, Millar, and Milne (2001) recently investigated the accuracy with which clinicians estimate IQ from demographic variables and compared this to a regression equation (The Crawford Index) which used the same variables. This study utilized clinical psychologists rather than neuropsychologists and 60 non-neurologically impaired (healthy) subjects. The clinicians were only given the demographic variables of each subject and no other clinical information. It was found that the correlations between IQ estimates and obtained IQ (using WAIS-R) were highly significant for both the regression equation and the clinical judgement.

The regression equation, however, outperformed the clinicians' estimates in that the estimated IQs derived from it were significantly more correlated with the obtained IQs than were the clinician-derived estimates. Crawford et al., (2001) suggest that it is appropriate to use a regression equation to provide an initial estimate of IQ rather than relying on a clinical estimate that is ultimately derived from the same information. It is rare however, in clinical practice, that demographic variables are the only information used to derive an IQ estimate (Smith-Seemiller et al., 1997), and thus the methodology behind these findings does not fully represent real-life practice.

Overall, there has been a limited amount of research conducted within the area of IQ estimation. Research investigating judgement has provided limited support for the use of clinical judgement in IQ estimation. However, despite the limited support for clinician judgement in IQ estimation, this is still a widely used method (Smith–Seemiller et al., 1997)

Current Practices Within The Area of Judgement.

At present, there is little standardization or general agreement about which method of judgement is best (Karaken, 1994). This lack of agreement is reflected in current practices in the field, and is the basis of the current debate within both the areas of general judgement research and research within the area of IQ estimation.

In relation to the current practices used by clinicians to assess premorbid IQ, it has been found that clinicians in this area also tend to use clinical judgement over actuarial formulas. For example, in a study conducted by Smith-Seemiler et al., (1997), neuropsychologists' most frequently used method for assessing premorbid IQ was found to be the clinical interview. The authors also found that there was little use of methods specifically designed to assess premorbid function, such as the NART or Barona Formula. Further, they concluded, on the basis of their investigations, that most clinicians did not perceive the need to use actuarial or regression based measures for estimating premorbid ability.

The use of clinical judgement over decision aides is not peculiar to the field of IQ estimation. Berg (1997) conducted a study examining the decision-making procedures in medicine. He found that in nearly all of these studies when a comparison was made between a statistical method and a human judge, the statistical model was always better. He noted that although statistical methods have existed for over 20 years they have never caught on in practice.

Several reasons have been proposed to explain why research on statistical versus clinical judgement has had little impact on everyday decision making. Dawes, Faust and Meehl (1989) suggest several reasons why statistical methods have not caught on in practice. These include lack of familiarity with the evidence, the belief that group statistics do not apply to single individuals or events, and subjective appraisal which may lead to inflated confidence in the accuracy of clinical judgement.

Thus, despite numerous studies indicating a superiority of actuarial methods of prediction, clinical judgement is the most widely used tool in clinical decision making. However, methodological flaws have plagued many studies within this area. As such, most findings to date need to be interpreted with caution.

Past Methodologies in This Area.

A variety of methodologies have been employed in the past to compare the accuracy of clinical judgement to actuarial formula predictions. One of the early pioneers in this area, Meehl (1954), specified conditions for a fair comparison of actuarial to clinical methods. He proposed that both methods should base judgements on the same data, and conditions must be avoided which artificially inflate the accuracy of the actuarial method. A variety of study designs have been employed within the area of clinician judgement.

Most studies which have compared clinician judgement to an actuarial formula have utilized hypothetical cases (Wedding, 1983), neurologically impaired cases (Goldstein, Deysach & Kleinknecht, 1973), 'normal' cases (Sarbin, 1943) or terminally ill cases (Einhorn, 1972) and provided variable amounts of information to the judges.

The only two studies that have been conducted comparing clinical judgement to actuarial formula estimations of IQ have utilized hypothetical cases (Kareken & Williams, 1994) and non neurologically impaired cases (Crawford, Millar & Milne, 2001). In both these studies only demographic information was given to the judges from which to make their estimates, and in the first of these studies, no objective IQ data was available due to the hypothetical case methodology employed.

Thus the methodology that has been most widely used within the field of neuropsychology is the utilization of non-neurologically impaired cases to study clinician judgement of IQ (e.g. Crawford, Millar & Milne, 2001). It will be appreciated that any attempt to evaluate the accuracy of estimates of premorbid IQ, must utilise non neurologically impaired participants as the actual premorbid IQ scores are not available in individuals suffering from neurologic or psychiatric disorders.

Drawbacks of Current Research in This Area.

Over the past six decades, much research attention has been devoted to the field of clinical judgement. The clinical judgement versus statistical prediction debate has been controversial and has yielded mixed results. As a result, clinicians working in the field have had no firm guidelines as to what procedures to use. The majority of studies that have compared clinical and statistical predictions in psychology have involved behavioral inferences about personality measures and used clinical psychologists.

The methodologies employed in these studies have also often been the subject of much criticism. Few studies have investigated the accuracy of judgements in neuropsychology and only one to date has investigated the accuracy of neuropsychological judgments in IQ estimation (Karaken & Williams, 1994)

In addition to the actuarial versus clinical investigations, judgement research has also focussed on the role experience plays in clinical judgements, the confidence ratings made by clinicians, and inter-rater reliability. Of the few studies that have been conducted, mixed results have emerged. Studies specifically investigating the judgement accuracy of clinical neuropsychologists are scarce and research investigating neuropsychologist's judgement accuracy within the area of IQ estimation is almost nonexistent.

Aims and Objectives

Based on a review of the literature many questions still remain unanswered within the field of judgement research. In particular very few studies have investigated judgement accuracy within the field of neuropsychological assessment utilizing clinical neuropsychologists. It was therefore the intention of this study to further investigate the accuracy of clinical judgement in IQ estimation utilizing neuropsychologists and objective IQ data, to investigate the role of neuropsychologist's training and experience in IQ estimation and to investigate the inter-rater reliability of neuropsychologist's estimates of IQ.

Further, it was the intention of this study to investigate the appropriateness of confidence ratings made by neuropsychologists, the correlation between demographic variables and clinician estimates of IQ and the accuracy of the Crawford Index in predicting IQ compared with clinical judgement. It was also the aim of this study to investigate the accuracy of the regression based methods of premorbid IQ prediction currently available, and to investigate the relationship between clinician confidence, accuracy and amount of information. Finally, it was the aim of this study to investigate the methods used by neuropsychologists to estimate premorbid IQ.

Hypotheses

Based on a review of the literature, it was hypothesized that, levels of training and experience would not be associated with the neuropsychologists' judgement accuracy in estimating IQ, and that the neuropsychologists would be over confident about the IQ estimations that they make. It was further predicted that the neuropsychologists would have good to excellent inter-rater reliabilities for their IQ estimates, and actuarial methods of IQ estimation would produce a more accurate estimation of IQ than the clinical judgement method utilized by the neuropsychologists. Further it was predicted that the neuropsychologists would overestimate the contribution of demographic variables to the estimation of IQ compared with the regression equation, and that they would be more confident in their judgements when increased amounts of information are provided. Finally, it was predicted that neuropsychologists would not use methods that are specifically designed for estimating premorbid IQ in clinical practice.

Study One

Method

Participants

There were three main groups in study one. The WAIS-III Data Group included 15 adults, the Clinician Rater Group was comprised of 15 novice and 15 expert clinical neuropsychologists while the Demographic Information Clinician Rater Group was made up of 11 neuropsychologists from the original Clinician Rater Group (7 novice, 4 experienced clinical neuropsychologists).

WAIS-III Data Group.

The WAIS-III Data Group consisted of 8 females and 7 males free of any known sensory, neurological or psychiatric disturbances. In total, 19 individuals were invited to participate in this study. Four declined to participate. The participants in this group were recruited by word of mouth and were friends, family and work colleagues of the author. All participants were recruited on a voluntary basis and gave informed consent (See Appendix A for informed consent form and description of study given to these participants). None of the participants had been previously administered the WAIS-III. The demographic characteristics of the WAIS-III Data Group are presented in Table 1.

Table 1
Demographic Characteristics of the WAIS-III Data Group

	Males (n = 7)	Females (n = 8)	Total (N = 15)
Age			
Range	26-53	21-49	21-53
Mean	37.7	30.4	33.8
SD	12.51	9.36	11.19
Years of Education			
Range	10-16	12-22	10-22
Mean	12	15.37	13
SD	1.91	3.02	3.02
Type of Employment			
Professional	1	1	2
Intermediate	2	5	7
Skilled	3	1	4
Semi-Skilled	1	1	2
Unskilled	0	0	0
IQ Distribution			
Range	98-137	100-143	98-143
Mean	115.14	114.25	114.67
SD	13.34	13.5	12.96

Clinician Rater Groups.

Fifteen participants comprised the Novice Rater Group. They were recruited by word of mouth from professional clinical neuropsychology postgraduate courses at Victoria University, Monash University and LaTrobe University. All participants were in the process of completing masters or doctoral qualifications in neuropsychology at the time of recruitment. All raters participated on a voluntary basis and provided informed consent (see Appendix B for informed consent form and information provided to Novice Raters).

Fifteen participants comprised the Experienced Rater Group. To be classed as experienced, raters had to have had at least two years' experience working as a qualified neuropsychologist, after graduation. They were recruited by word of mouth from a wide variety of occupational settings. All participants in this group were registered as psychologists and were eligible for membership of the College of Clinical Neuropsychologists (CCN). All raters participated in this study on a voluntary basis and provided informed consent (see Appendix B for consent form and information provided to Expert Raters). The demographic characteristics of the Clinician Rater Groups are shown in Table 2.

Table 2
Demographic Characteristics of the Novice and Experienced Neuropsychologist Rater Groups

	Novice (N=15)	Expert (N=15)
Age		
Range	22-53	28-61
Mean	29	41.35
SD	8.36	12.11
Highest Qualification		
Graduate Diploma	4	0
Honours	11	0
Masters	0	10
Doctorate	0	2
PhD	0	3
Amount of time WAIS-III used in Assessments		
Almost Never	1	0
Once Per Month	1	6
Once Per Week	4	3
2-5 Times Per Week	6	2
>5 Times Per Week	2	1
Number of Years Practicing as a Neuropsychologist		
1-5 years	0	4
5-10 years	0	7
>10 years	0	4
Number of Hours spent in clinical Assessment per Week		
5-10 hours	0	4
10-20 hours	0	8
>20 hours	0	2

Demographic Information Clinician Rater Group.

The Demographic Information Clinician Rater Group was comprised of eleven participants who were a subgroup of the Clinician Raters Group and were asked in person to participate in the second part of study one (see above for information regarding recruitment of these participants). The demographic characteristics of the Demographic Information Clinician Rater Group is shown in Table 3.

Table 3
Demographic Characteristics of the Demographic Information Clinician Rater Group

	Novice (N=7)	Expert (N=4)
Age		
Range	24-45	37-62
Mean	28.28	49.50
SD	7.47	17.67
Highest Qualification		
Graduate Diploma	1	0
Honours	6	0
Masters	0	3
Doctorate	0	0
PhD	0	1
Number of Years Practicing as a Neuropsychologist		
0 years	7	0
1-5 years	0	1
>10 years	0	1
>20 years	0	2

Materials

Materials for part one of study one included: 1) The Weschler Adult Intelligence Scale-3rd Edition (WAIS III: Weschler, 1997), 2) A clinician rater's scoring booklet, 3) Video of an interview, 4) The Crawford Index (Crawford & Allan, 1997), 5) The Barona Formula (Barona et al., 1984), 6) The Oklahoma Premorbid Intelligence Estimate (OPIE: Krull et al., 1995), 7) The Oklahoma

Premorbid Intelligence Estimate 3rd Edition (OPIE-3; Schoenberg et al., 2002) and 8) A second Clinician Raters scoring booklet containing only demographic information about the 15 WAIS-III Data Group cases was used for part two of study one. These materials are described below.

Wechsler Adult Intelligence Scale-III.

All participants in the WAIS–III Data Group were assessed by the experimenter with the Wechsler Adult Intelligence Scale-III (Wechsler, 1997). The standard published instructions and method of administration was used. The WAIS-III uses deviation IQ (M = 100, SD = 15) for the Verbal Scale, Performance Scale and Full Scale IQ's. An IQ is computed by comparing the examinees scores with the scores earned by a representative sample of his or her age group. The range of WAIS-III Full Scale IQ's for all age groups is 45 to 155.

Clinician Rater's Scoring Booklet.

Each clinician rater received a booklet that contained demographic questions about age, gender, education, education level reached, number of years practicing as a neuropsychologist and number of hours spent in clinical assessment per week. The booklet also contained a scoring sheet to record IQ estimations and confidence levels for each estimation made for the 15 WAIS-III Data Group participants in the video of the interview.

Demographic information (age, gender, occupation and level of education reached) for each of the WAIS-III Data Group volunteers appeared on the top of the scoring section for each individual (see Appendix C for details of booklet).

Video of Interview.

The clinician raters were issued with a video consisting of 15, three-minute interviews for each of the WAIS-III Data Group participants. In the video, the participants talked about their current and previous areas of employment and some personal aspects about themselves such as travel and interests.

Crawford Index.

In addition to providing the clinicians with information upon which to base their estimates, the demographic variables of employment class, years of education and age for each volunteer were used to provide an actuarial estimate of IQ using the Crawford Index (1997). The Crawford Index is the following regression equation based on demographic details and WAIS-R (Wechsler, 1981) data from the United Kingdom:

Predicted WAIS-R FSIQ = 87.14 – (5.21 x social class) + (1.78 x years educated) + (0.18 x age).

According to Crawford and Allan (1997), participants are credited with 0.5 years for every year spent in part-time education. Social class is coded from their occupation, using the United Kingdom Office of Population Censuses and

Surveys (1980) Classification of Occupations. Crawford and Allan (1997) use five broad categories for social class 1 (professional), 2 (intermediate), 3 (skilled), 4 (semi-skilled), and 5 (unskilled). The classification aims to bring together people with jobs of a similar social and economic status.

According to this classification, individuals classified as professional would be engaged in work that required a university qualification such as a doctor, accountant, pharmacist, solicitor or architect. Those individuals classified as intermediate and below are engaged in occupations that do not require qualifications of a university degree standard. For example, those individuals engaging in occupations such as a secretary, advertising, property managers or actors would be placed in the intermediate category. Individuals with occupations such as photographers, sportsmen, butchers, restaurateurs or cashiers would be placed in the skilled category. Individuals engaging in occupations such as security guards, gardeners, roofers, painters or miners would be placed in the semi-skilled category. Individuals engaging in occupations such as cleaners, laborers, kitchen hands, car park attendants, or refuse collectors would be placed in the unskilled category.

The Crawford Index is easily adaptable to Australian populations as it is based only on age, years of education and employment class, with no reference to region of residence, which occurs with the Barona Formula and the OPIE. However, before analysis was conducted the Crawford Index was transformed to make it suitable for prediction of IQs obtained using the WAIS-III, as it was originally designed to estimate IQs obtained using the WAIS-R. To do this, 2.9 points were added to each full scale IQ estimated by the equation, as it has been reported that there is a 2.9 point difference between IQs obtained using the WAIS-R and the WAIS-III (Wechsler, 1997).

Barona Formula.

The demographic details for each member of the WAIS-III Data Group were also used in an actuarial estimate of IQ using the Barona Formula (Barona et al., 1984). The Barona Formula is the following regression equation based on demographic details and the WAIS-R (Wechsler, 1981) data from the United States of America:

Predicted WAIS-R FSIQ = 54.96 + 0.47 (age) + 1.76 (sex) + 4.71 (race) + 5.02 (education) + 1.89 (occupation) + .59 (region).

The values for each of the variables given by Barona et al., (1984) are shown in table 4. As the regions used in this equation are based on regions of the US and not Australia, the region of North Eastern was used as this region was estimated to most closely approximate Victoria, from where the participants in the WAIS-III Data Group resided. Before analysis was conducted the Barona was also transformed to make it suitable for prediction of WAIS-III IQs, as this equation was also originally designed to estimate IQs obtained using the WAIS-

R. This was again done by adding 2.9 points to each full scale IQ estimated by the equation.

Table 4
Barona Formula Variable Weights

Variable		Value
Sex	Female	1
	Male	2
Race	White	3
	Black	2
	Other	1
Occupation	Professional / Technical	6
	Managerial / Official / Clerical / Sales	5
	Craftsman / Foremen	4
	Not In Labor Force	3
	Operatives / Service Workers / Farmers	2
	Farm Laborers / Laborers	1
Region (U.S)	Southern	1
	Western	3
	North Central	2
	North Eastern	4
Residence	Rural	1
	Urban	2
Age	16-17	1
	18-19	2
	20-24	3
	25-34	4
	35-44	5
	45-54	6
	55-64	7
	65-69	8
	70-74	9
	75-79	10
	80-84	11
	85-89	12
	90-94	13
	95-99	14
	>100	15
Education	0-7 years	1
	8	2
	9-11	3
	12	4
	13-15	5
	16+	6

Oklahoma Premorbid Intelligence Estimate.

The demographic details for each member of the WAIS-III Data Group were also used in an actuarial estimate of IQ using the Oklahoma Premorbid Intelligence Estimate (OPIE: Krull et al., 1995). The OPIE is the following regression equation based on demographic details and current performance on subtests of the WAIS-R (Wechsler, 1981):

Predicted WAIS-R FSIQ = 53.80 + 0.10 (age) + .64 (education) - 1.73 (race) - .51 (occupation) + .57 (WAIS-R Vocabulary raw score) + 1.33 (WAIS-R Picture Completion raw score).

The values for each of the variables given by Krull et al., (1995) are shown in table 5. Before analysis was conducted the OPIE was also transformed to make it suitable for prediction of WAIS-III IQs. This was again done by adding 2.9 points to each full scale IQ estimated by the equation.

Table 5
OPIE Variable Weights

Variable	Value
Race	
Caucasian	1
Non Caucasian	2
Occupation	
Professional / Technical	1
Managerial / Official / Clerical / Sales	2
Craftsman / Foremen	3
Operatives / Service Workers / Farmers	4
Laborers	5
Unemployed	6
Age	
Calculate in years	-
Education	
0-7 years	1
8	2
9-11	3
12	4
13-15	5
16+	6

Oklahoma Premorbid Intelligence Estimate-3.

The demographic details for each member of the WAIS-III Data Group were also used in an actuarial estimate of IQ using the Oklahoma Premorbid Intelligence Estimate-3 (OPIE-3: Schoenberg et al., 2002). There are several equations that can be used to estimate premorbid IQ The first is the OPIE-3 (4ST) which combines current scores on WAIS-III Vocabulary, Information, Matrix Reasoning and Picture Completion (Schoenberg et al., 2002).

However, it was reported in the literature that this equation underestimated IQ (Schoenberg et al., 2002), and that the OPIE-3 (2ST) and the OPIE-3 MR performed better. Thus these equations were used for estimating WAIS-III IQs, as the requirements of these two equations were met by the WAIS-III Data Group. It is recommended by Schoenberg et al., (2002) that the OPIE-3 (2ST) be used when Vocabulary and Matrix Reasoning scaled scores are the same, and OPIE MR when the Matrix Reasoning scaled score is one or more points greater than Vocabulary. In total two participants in the WAIS-III Data Group met the requirements for use of the OPIE-3 (2ST), whilst the other 13 met the requirements for use of the OPIE MR. The two regression equations are:

OPIE-3 (2ST) Predicted WAIS-III FSIQ = 45.979 + 0.652 (Vocabulary raw score) + 1.287 (Matrix Reasoning raw score) + .157 (age in years) + 1.034 (education) + .652 (ethnicity) - 1.015 (gender).

OPIE-3 MR Predicted WAIS-III FSIQ = 43.678 + 1.943 (Matrix Reasoning raw score) + .297 (age in years) + 3.564 (education) + 1.541 (ethnicity) + .543 (region of country) - 1.137 (gender).

The values for each of the variables given by Schoenberg et al. (2002) are shown in table 6. Again, the North East was used as the region most comparable to Victoria.

Table 6
OPIE-3 Variable Weights

Variable	Value
Ethnicity	
African American	1
Hispanic	2
Other	3
Caucasian	4
Gender	
Male	1
Female	2
Region of Country	
South	1
North Central	2
North East	3
West	4
Age	
Calculate in years	-
Education	
0-8 years	1
9-11	2
12	3
13-15	4
16+	5

Demographic Scoring Booklet.

Approximately 18 months after the initial data was collected, a subgroup of the initial Clinician Raters received a booklet that contained demographic questions about their age, gender, education, education level reached, number of years practicing as a neuropsychologist and number of hours spent in clinical assessment per week. The booklet also contained a scoring sheet to record IQ estimations and confidence levels for estimations of IQ for the 15 participants from the WAIS-III Data Group, based on only the following demographic information for each case: age, gender, occupation and level of education reached. The demographics for some of the cases had been altered slightly to decrease the chance that the clinicians would recognize them, (eg. job title of 'consultant in emergency medicine' was changed to 'doctor'), however the demographics mostly stayed the same (see Appendix D for details of booklet). Raters were also given a reply paid envelope to return their booklets.

Procedure

WAIS-III Data Group.

The WAIS-III was administered in a standardised form in order to determine the IQ for each of the volunteers. The subtests that were administered included Picture Completion, Vocabulary, Digit-Symbol Coding, Similarities, Block Design, Arithmetic, Matrix Reasoning, Digit Span, Information, Picture Arrangement and Comprehension.

Each volunteer was then given a brief interview in which they were asked to talk about their work and education history for approximately 5 minutes. This interview was video-recorded for each of the volunteers and then compiled into a video that was sent to the clinicians for viewing to aid in their determination of the IQ estimates. The participants from the WAIS-III Data Group were given oral feedback as to their performance on the WAIS-III and provided with the range under which their Full Scale IQ fell. It should be noted that the final distribution of IQ scores was not normally distributed for the WAIS-III Data Group (refer to Table 1 for IQ distribution). This intended to ensure that the clinician raters did not normally distribute their answers, thus eliminating the possibility of accuracy by chance.

Clinician Raters Group.

Clinician raters were approached either in person, via the telephone or by email and invited to participate in the study. Once they had agreed to be a participant, a pack was sent to them that included an explanatory letter about the study, a consent form, a scoring booklet and videotape. Each rater was also given two reply paid envelopes. In total, 64 scoring packs were distributed over a 12-month period to the Clinician Rater's Group. There was a 46.8% response rate, with 30 scoring sheets returned within the allocated time period.

Raters were instructed to use one envelope to return their signed consent form and the other to return the scoring sheet and video once they had completed it. This was done in order to ensure identifying information (the consent form) was kept separate for each clinician, thus enabling each scoring sheet to remain anonymous. The only identifying information on each scoring sheet was the rater's number of years working as a neuropsychologist, nature and setting of work (eg. hospital, private practice, rehabilitation), highest qualification, gender and age. This was necessary to enable the novice and expert groups to be identified for the latter part of the study.

The raters were required to watch the video and estimate the IQ of each member of the WAIS-III Data Group based on the interview and demographic variables provided. Clinicians were asked to provide both a descriptive range (Impaired, Borderline, Low Average, Average, High Average, Superior, Very Superior) and an exact number for their IQ estimate, as well as a confidence rating for their estimations, consisting of 1 (low), 2 (moderate), 3 (high) or 4 (very high). A blank section at the bottom of each page of the scoring sheet was provided for comments and note taking whilst watching the video.

Demographic Information Clinician Rater Group.

In total, 15 Clinician Raters were approached to participate in this second investigation, with 15 booklets distributed over a 2-week period. There was a 73.3% response rate, with 11 booklets returned within the allocated time period.

The only identifying information on each scoring sheet was the rater's number of years working as a neuropsychologist, nature and setting of work (eg. hospital, private practice, rehabilitation), highest qualification, gender and age. The raters were instructed to estimate the IQs for the 15 de-identified cases in the booklet. The raters were not told that these cases were based on the original WAIS-III Data Group as it was the purpose of this part of the study to compare estimates based on video and demographic information to just demographic information. Clinicians were asked to provide both a descriptive range (Impaired, Borderline, Low Average, Average, High Average, Superior, Very

Superior) and an exact number for their IQ estimate, as well as a confidence rating for their estimations, consisting of 1 (low), 2 (moderate), 3 (high) or 4 (very high).

Analysis of Clinician Judgement

Analysis of the data for this study involved comparing the clinician raters' IQ estimates for the WAIS-III data group to the IQs obtained using the Crawford Index. This was done to ascertain the level of accuracy of clinical judgement to that of an actuarial formula designed to measure the same thing. The next analysis involved comparing the actuarial IQs obtained for each participant in the WAIS-III Data Group to the various estimates, i.e., the clinician raters', the Crawford Index, Barona Formula, OPIE and OPIE-3 regression estimates. This was done to ascertain the level of accuracy of clinician judgements and the available regression estimates compared to objective IQ data.

In order to ascertain the effect of experience on judgement accuracy, the IQ estimates of novice and expert clinicians were compared. The confidence estimates given by each clinician rater was then compared to the accuracy of their estimates for each of the WAIS-III Data Group IQs. This was done to determine the relationship between accuracy and confidence of clinician judgements.

Results

Analytic Plan

Preliminary analyses were conducted to test statistical assumptions and screen data. All variables were then examined for departures from normality. Secondly the hypotheses were tested. Pearsons Product-Moment Correlations were used to estimate bivariate relationships between variables. Mixed Between-Within Subject Analyses of Variance, T-tests and One Way Analyses of Variance were used to determine if the differences between variables were statistically significant.

All analyses were conducted using Statistical Packages for the Social Sciences (SPSS) Version 11.0 for Windows. Alpha levels were set at 0.5 unless otherwise specified.

Preliminary Analysis

Firstly, tests for outliers and normality were conducted in order to ensure that the data met the assumptions of the statistical tests that were to be used in the analysis. After consideration of the differences between the actual mean scores for each of the variables and the 5% trimmed means, the one outlier that was identified in the WAIS-III adjusted regression equation predicted IQ was kept in the data set as it was not extreme. Kolmogorov-Smirnov was significant ($p>.05$) for the mean obtained IQ, mean Clinician Raters' estimated IQ, and the adjusted regression equation IQ estimate, indicating a normal distribution of scores and therefore meeting the assumption of normality. In accordance with Pallant (2001) the variables did not violate the assumptions of the statistical techniques that were to be used in later analyses. The scores revealed a normal distribution with skewness and kurtosis scores falling between -2 and +2.

Hypotheses Testing – Part One Clinician Judgement

Crawford Equation compared with Clinical Judgement.

In order to determine the accuracy of the clinician raters compared with the Crawford equation in estimating IQ, the mean estimates for these two methods were compared with the mean obtained IQ for the WAIS-III sample. Table 7 shows summary statistics for each of the variables. It can be seen that there was a close correspondence between the mean clinician-estimated IQs and the mean IQs estimated by the adjusted Crawford equation. In contrast, the mean obtained IQ from the WAIS-III sample was greater than both the clinician-based and regression-based estimates. Both the mean clinician and adjusted Crawford equation underestimated IQ by almost half a standard deviation.

Table 7
Summary statistics for the 15 obtained WAIS-III IQs and corresponding IQs estimated by the Crawford Index, adjusted for use with the WAIS-III, and the 30 clinician raters.

	M	SD
Obtained IQ	114.66	12.96
Adjusted Crawford	108.18	9.64
Clinician Raters	109.27	10.24

An Analysis of Variance (ANOVA) was conducted to explore the correspondence between the mean obtained IQ and the mean IQ estimated by the clinician raters and the adjusted Crawford equation. There was a statistically significant main effect for IQ, $F(2, 13) = 4.35, p < .05$, indicating that there was a significant difference between the obtained, clinician estimated and adjusted Crawford equation mean IQ scores. The effect size was large according to the Cohen (1988) classification ($\eta^2 = .40$).

Planned contrasts using the simple contrast method, showed a significant difference between the adjusted Crawford equation mean IQ scores and the mean obtained IQ scores for the WAIS-III sample $F(1, 14) = 9.38$, $p < .05$, with a large effect size ($\eta^2 = .40$). The mean obtained IQ was also significantly different from the clinician raters' mean IQ estimate $F(1, 14) = 6.32$, $p < .05$, with a large effect size ($\eta^2 = .31$), The clinician raters' mean IQ did not significantly differ from the Crawford regression equation IQ estimate $F(1, 14) = 0.67$, $p > .05$, the effect size was small ($\eta^2 = .04$).

In order to investigate the accuracy of clinical judgement and the Crawford Index further, the 95% confidence intervals for the obtained IQs, and the descriptive ranges under which both the clinician estimates and the Crawford Index estimates fell were compared. Table 8 shows the obtained IQs for the 15 WAIS-III Data Group cases, the 95% confidence intervals for these cases, and descriptive ranges for the obtained IQs, mean clinician estimates and the mean adjusted Crawford estimates.

Table 8
Comparison of Descriptive Ranges for Obtained IQ, Mean Clinician Estimates and the Mean Adjusted Crawford Index Estimates for the 15 WAIS-III Data Cases

Case	Obtained IQ	95% C.I.	Descriptive Range Obtained IQ	Descriptive Range Clinicians	Descriptive Range Crawford
1	105	101 - 109	Average	Average	Average
2	113	109 – 117	High Average	Average	High Average
3	103	99 – 107	Average	High Average	High Average
4	119	115 – 123	High Average	Average	High Average
5	98	94 – 102	Average	Average	Average
6	122	117 – 126	Superior	High Average	Average
7	104	100 – 108	Average	Average	Average
8	137	132 – 140	Very Superior	Superior	Superior
9	118	114 - 122	High Average	High Average	High Average
10	100	96 – 104	Average	Average	High Average
11	124	119 – 128	Superior	Average	Average
12	143	138 – 146	Very Superior	Very Superior	Very Superior
13	109	105 – 113	Average	Average	Average
14	112	108 – 116	High Average	Average	Average
15	113	109 – 117	High Average	High Average	High Average

These results suggest that although both the clinicians and the Crawford equation tend to underestimate IQ, a total of 3 estimations still fell within the 95%

confidence intervals for the obtained IQ, whilst only one of the Crawford estimates fell within these ranges. Further, both the Crawford equation and the clinician estimates underestimated obtained IQ by as much as two descriptive IQ ranges for several cases.

Regression Equation Methods compared with WAIS-III Data Group Obtained IQ.

Table 9 shows the summary statistics for each of the adjusted regression equation methods (Crawford Index, Barona Formula, OPIE, and OPIE-3). It can be seen that there was a close correspondence between the mean obtained IQs and the mean IQs estimated by the adjusted Barona, OPIE and OPIE-3 equations.

Table 9
Summary statistics for the 15 obtained WAIS-III IQs and corresponding IQs estimated by the Crawford Index, Barona Formula, OPIE and OPIE-3 with adjustment for use with the WAIS-III, and the 30 clinician raters.

Method	M	SD
Obtained IQ	114.66	12.96
Adjusted Crawford	108.18	9.64
Clinician Raters	109.27	10.24
Adjusted Barona	110.31	5.51
Adjusted OPIE	116.44	6.24
OPIE-3	113.92	6.66

An ANOVA was conducted to explore the correspondence between the mean obtained IQ and the mean IQ estimated by the clinician raters and the regression equations (Crawford Index, Barona Formula, OPIE and the OPIE-3). There was one dependent variable (obtained IQ) and one independent variable, method, with five levels (clinician estimated IQ, adjusted Barona, Crawford, OPIE, and the OPIE-3). There was a large and statistically significant main effect for method, $F(5, 10) = 8.5, p < .05$, ($\eta^2 = .81$).

Simple contrasts showed no statistically significant difference between the mean obtained IQ and the adjusted Barona Formula , the adjusted OPIE, or the OPIE-3 (all *p*'s > .05). The effect size for the difference between the adjusted Barona IQ and the mean IQ was moderate, however ($\eta^2 = .12$), while the effect sizes for the OPIE and the OPIE-3 versus the obtained IQ contrasts were small ($\eta^2 = .04$ and .01, respectively).

Relationship between Experience and Accuracy.

In order to determine the relationship between years of experience and judgement accuracy, and if there was an effect of case on clinician accuracy, a repeated measures ANOVA was conducted to compare the accuracy of the novice and experienced clinician raters. Accuracy was defined as obtained IQ minus the clinician estimated IQ. The dependent variable was video accuracy, based on the raters' video estimates, and the independent variable was group, with two levels (novice versus experienced). The within subjects factor was the

WAIS Data cases with 15 levels. There was no significant difference in the accuracy of the IQ estimates given by the novice (M = 109.41, SD = 4.41), and the expert clinician raters (M = 108.94, SD = 3.17), $F(14, 11) = 0.52$, $p > .05$, with a large effect size ($\eta^2 = .39$). This large effect size however, suggests that there may have been a significant difference in the accuracy of the two groups if there had been a larger sample size. Table 10 shows means and standard deviations for each of the 15 WAIS-III patients and the two groups of clinician raters. The mean difference score refers to experienced versus novice estimates.

Table 10
Descriptive Statistics for Estimated IQ by the Novice and Experienced Clinician Raters

	Obtained IQ	Experienced Clinicians (n = 15)		Novice Clinicians (n = 15)		Mean Difference Score
Case		M	SD	M	SD	M
1	105	104.20	8.48	104.13	5.75	0.07
2	113	109.86	6.77	107.40	4.77	2.46
3	103	114.26	6.74	111.66	7.83	2.60
4	119	103.40	6.71	105.86	7.76	-2.46
5	98	94.73	6.26	94.53	8.49	0.20
6	122	119.61	7.73	117.13	5.81	2.48
7	104	97.86	2.14	98.13	7.21	-0.27
8	137	123.00	9.81	120.00	5.47	3.00
9	118	114.86	9.62	115.73	3.19	-0.87
10	100	108.66	4.86	108.46	5.09	0.20
11	124	109.26	9.39	106.14	10.92	3.12
12	143	132.53	8.68	131.28	7.14	0.65
13	109	95.73	9.58	99.60	7.37	-3.87
14	112	100.06	9.57	98.13	8.63	1.93
15	113	115.60	9.59	116.66	4.65	-1.06

The within subjects factor of WAIS-III Data case was significant $F\ (14, 11) = 46.9, p < .05$, with a large effect size ($\eta^2 = .98$). This indicated that there was a significant difference in the raters' accuracy in predicting the 15 WAIS-III Data cases. The deviation contrast method, was used to compare each of the video based mean clinician estimates for the 15 WAIS-III Data cases with Case 1, as that was the most accurate (refer to Table 11). Table 11 shows the mean video accuracy estimates (in order of accuracy) for each of the 15 WAIS-III cases along with demographic characteristics for each of the cases in order to determine if these was a trend between the characteristics of each case and judgment accuracy .

Table 11

Mean Accuracy Scores for Each WAIS-III Data Case Sorted By Accuracy (IQ-estimated IQ) and Demographic Characteristics of Each Case

Case	Mean Accuracy	Education	Occupation Code	Age	Obtained IQ	Mean Estimate IQ
1	0.46	12	4	21	105	104
9	2.04	13	2	49	118	115
15	-3.08	16	2	26	113	116
5	3.81	12	3	26	98	94
6	3.85	12	2	28	122	118
2	4.07	14	2	35	113	109
7	6.61	11	3	29	104	98
10	-8.73	15	3	27	100	108
3	-10.30	16	2	23	103	113
12	11.00	22	1	37	143	132
13	12.07	10	3	53	109	98
14	14.15	12	4	28	112	99
4	14.46	15	2	25	119	105
8	15.04	16	1	49	137	122
11	17.35	11	2	51	124	108

Note: IQ Ranges are as follows; Average (90-109), High Average (110-119), Superior (120-129), Very Superior (>130).

Occupation Codes based on Crawford & Allan (1997) classification of occupations for use with the Crawford Index. The codes are as follows; 1 (Professional), 2 (intermediate), 3 (skilled), 4 (semi-skilled), 5 (unskilled). See method for further information regarding coding system.

Relationship Between Clinician Confidence, Accuracy and Demographic Variables of the WAIS-III Data Group.

Peasons Product Moment Correlations were used to investigate the bivariate relationship between clinician judgement accuracy, confidence and demographic characteristics of the WAIS-III Data Group. Table 12 shows the correlations between obtained IQ, mean clinician accuracy, demographic variables and mean clinician confidence for the 15 WAIS-III Data cases.

Table 12
Pearson Product Moment Correlations Among Clinician Accuracy In Estimating IQ, Clinician Confidence, Obtained IQ and Demographic Variables For The 15 Cases in The WAIS-III Data Group

	1	2	3	4	5	6	7
(1) Accuracy	1.00						
(2) Confidence	.17	1.00					
(3) Obtained IQ	.61 *	.73 **	1.00				
(4) Gender	.52 *	-.14	.04	1.00			
(5) Age	.55 *	.14	.49	.34	1.00		
(6) Education	-.12	.69 **	.55 *	-.58 *	-.13	1.00	
(7) Occupation	-.18	-.66 **	-.74 **	.18	-.39	-.64 **	1.00

Note: ** $p < .01$, * $p < .05$

There was a strong and statistically significant positive correlation between clinician accuracy and obtained IQ, gender and age. There was also a strong statistically significant correlation between clinician confidence and obtained IQ and education. There was a strong statistically significant negative correlation between clinician confidence and occupation. Education was positively correlated with obtained IQ, while IQ was negatively correlated with occupation. There was a strong significant correlation between education and gender, and a strong statistically significant correlation between occupation and education.

Clinician Confidence.

Pearsons Product Moment Correlations were used to further investigate the bivariate relationship between clinician confidence and accuracy for each of the 15 WAIS-III Data Cases. There was a weak negative correlation between the two variables, but this was not statistically significant $r(29) = -.10, p > .05$, showing that confidence was not associated with judgement accuracy. Table 13 shows the bivariate correlations between clinician confidence and judgement accuracy for the 30 clinicians. The table reveals only one statistically significant confidence-accuracy correlation for the 15 cases. This case was a medical specialist with 22 years of education. Obtained IQ was 143, with the mean estimated IQ for this case was approximately 130.

Table 13
Pearson Product-Moment Correlations Between Measures of Confidence and Accuracy

Case Accuracy	Confidence
1	-1.70
2	-2.90
3	-1.20
4	0.25
5	-0.18
6	-1.30
7	-0.22
8	-0.04
9	-0.20
10	0.25
11	0.18
12	-0.49 **
13	0.22
14	-0.05
15	-0.25

$^{**}p < .01$

The relationship between the number of years working as a clinician and reported confidence in judgement accuracy was also investigated using Pearson Product-Moment Correlation coefficient. There was a medium negative correlation between the two variables $r(29) = -.40$, $p < .05$, with more years of experience associated with higher confidence ratings.

Inter-rater Reliability.

Inter-rater reliability was obtained by comparing each of the 30 clinician's IQ estimates for the 15 WAIS-III cases, resulting in a Chronbach's alpha of 0.82, indicating good inter-rater reliability for the IQ estimates. Inter-rater reliabilities were also calculated for the novice and experienced using Chronbach's alpha. Inter-rater reliability was 0.75 for the novice group and 0.86 for the experienced group. This indicates good inter-rater reliabilities for both groups (Pallant, 2001).

Demographic Variables.

Pearsons Product Moment Correlations were used to investigate the bivariate relationship between the demographic variables of gender, social class (occupation), age and education with estimated IQ. Table 14 shows the correlations between gender, social class (occupation), age and education mean obtained IQ, mean adjusted Crawford equation IQ and the mean clinician-estimated IQ.

Table 14
Pearsons Product-Moment Correlations among Clinician Estimates of IQ, Regression Equation Estimates of IQ, Obtained IQ and Gender, Social Class (Occupation), Age and Education

IQ Measure	Gender	Occupation	Age	Education
Clinician Estimated IQ	-.37	-.79 **	.18	.78 **
Adjusted Crawford Estimate	-.34	-.93 **	.32	.84 **
Obtained IQ	.03	.74 **	.49	.54 *

Note: N=15; **$p < .01$, * $p < .05$

There were no statistically significant correlations between gender or age and the IQ estimates or observed scores. There were however, statistically significant correlations between occupation and education and the estimated and observed IQ scores (i.e. increased years of education and higher occupational attainment was associated with increased IQ).

Relationship Between Amount of Information and Judgement Accuracy.

The judgement accuracy of the 11 clinicians who agreed to participate in the second part of study one was analyzed as the dependent variable in a multivariate repeated measures ANOVA. Amount of information was the between subjects factor and WAIS-III data cases was the within subjects factor (with 15 levels). Accuracy was again defined as obtained IQ minus the clinician estimated IQ. There was no significant differences in the accuracy of clinicians IQ estimates using either the video and demographic information (M = 107.95, *SD* = 2.85) or demographic information only (*M* = 107.27, *SD* = 5.11), $F\ (1,8) = 0.22$, $p > .05$, with a small to medium effect size ($\eta^2 = .03$).

The within subjects factor WAIS-III Data case was significant $F\ (14, 112) = 4.65$, $p < .05$, with a large effect size ($\eta^2 = .37$). This indicated greater accuracy of prediction for some cases compared to others. Table 15 shows the descriptive statistics for the difference scores obtained by subtracting the video accuracy scores from the demographic accuracy scores for each of the 15 WAIS-III Data cases.

Table 15
Descriptive Statistics for Accuracy Difference Scores

Case	M	SD
1	-3.33	11.19
2	-7.67	9.26
3	4.44	5.94
4	9.33	8.23
5	11.33	5.94
6	-7.00	12.04
7	-3.33	6.59
8	-3.22	8.61
9	-2.89	14.35
10	3.00	6.46
11	-6.00	9.79
12	-3.67	5.41
13	-2.33	14.4
14	-.11	9.37
15	-1.56	4.79

When contrasted against case 14, (the case for which the clinicians were the most accurate) there were large and significant effects for cases two $F\ (1, 8)$

$= 8.50, p < .05, \eta^2 = .51$, three $F(1, 8) = 5.67, p < .05, \eta^2 = .41$, four $F(1, 8) = 15.89, p < .05, \eta^2 = .66$ and five $F(1, 8) = 39.28, p < .05, \eta^2 = .83$. This indicated that there was a significant difference in the raters' accuracy for these four cases.

Relationship Between Clinician Confidence, Accuracy and WAIS-III Data Group Demographic Characteristics When Provided With Demographic Information Only.

Peasons Product Moment Correlations were used to investigate the relationship between clinician judgement accuracy, confidence and demographic characteristics of the WAIS-III Data Group for raters using only demographic information upon which to base their estimates. Table 16 shows the correlations between obtained IQ, mean clinician accuracy, demographic variables and mean clinician confidence for the 15 WAIS-III Data cases.

Table 16
Pearson Product Moment Correlations Among Clinician Accuracy In Estimating IQ, Clinician Confidence, Obtained IQ and Demographic Characteristics For The 15 Cases in The WAIS-III Data Group Based On Demographic Variables Only

Measures	1	2	3	4	5	6	7
(1)Accuracy	1.00						
(2)Confidence	.33	1.00					
(3)Obtained IQ	.74 **	.56 *	1.00				
(4)Gender	.43	-.34	.04	1.00			
(5)Age	.58 *	.04	.49	.34	1.00		
(6)Education	-.07	.58 *	.55 *	-.58 *	-.13	1.00	
(7)Occupation	-.27	-.44	-.74 **	.18	-.39	-.64 **	1.00

Note: ** $p < .01$, * $p < .05$

There were statistically significant correlations between obtained IQ, age and clinician accuracy. There were also statistically significant correlations between obtained IQ, education and clinician confidence. Education and occupation were again correlated with IQ, whilst education was found to be correlated with gender. Finally, a statistically significant correlation was again found between the demographic variables of occupation and education. There were no significant correlations between clinician confidence and clinician judgement accuracy using only demographic variables to estimate IQ.

Relationship Between Amount of Information and Clinician Confidence.

Another repeated measures ANOVA was conducted to investigate the relationship between amount of information (video and demographic information versus demographic information only) and clinician confidence for the 11 clinicians who agreed to participate in the second part of study one. There was a large and statistically significant difference in the confidence of clinicians IQ estimates using the video and demographic information (M = 2.06, SD =*0 .22*) compared with demographic information only (M = 1.64, SD = 0.45), $F(1,6) = 10.41, p < .05, \eta^2 = .63$. This indicated that the raters were more confident in estimating IQ when they had the video as well as demographic information upon which to base their estimates.

Discussion

Study One investigated clinician judgement and the role of experience in the area of premorbid IQ estimation. It also investigated the relationship between clinician confidence and judgement accuracy, inter-rater reliability within the area of premorbid IQ estimation, and the relationship between various methods of IQ estimation and actual IQ.

The accuracy of neuropsychologists' clinical judgement was investigated by comparing clinician rater's estimates of IQ, estimates obtained using a regression equations specifically designed to estimate premorbid IQ, and other methods of IQ estimation with actual IQ scores. Contrary to the prediction that the adjusted Crawford equation would be more accurate at predicting WAIS-III IQ than the clinician raters, it was found that there was no difference in the two methods of IQ estimation. As predicted, clinician IQ judgement accuracy was not related to years of experience, nor was clinician confidence related to their judgement accuracy and clinician confidence was increased when provided with more information. Inter-rater reliability was found to be good, as predicted. A further discussion of these findings and their implications for clinical practice are considered in the following sections.

Clinician Judgement

Actuarial Methods Compared with Clinical Judgement.

This study investigated the accuracy of the Crawford Index, a regression based equation and clinician judgement in estimating premorbid IQ for a non neurologically impaired sample, thus providing objective IQ data with which to compare the estimates with. The results indicated that neither the clinician raters, nor the equation were accurate in predicting IQ.

The mean IQs estimated by the clinicians and by the Crawford regression equation both significantly underestimated the mean IQ obtained from the WAIS-III sample. In addition, there was no significant difference between the mean IQ estimated by the clinicians and that of the regression equation. However, further analysis revealed that several of the estimates given by the clinicians actually fell within the 95% confidence interval for the obtained IQs, but there were still several estimates that underestimated obtained IQ. Overall, this suggests that both clinician judgement and the equation did not yield accurate IQ estimates.

These findings are consistent with previous studies which have found actuarial methods of prediction to be as accurate as clinical judgement (e.g., Sarbin, 1943; Grebstein, 1963; Gilpin, Olshen & Chatterjee et al.,1990). However, this finding should not be taken to argue for the use of actuarial prediction (Meehl, 1954), as both methods were found to significantly underestimate the IQ of the sample used in this study.

Moreover, these results are contradictory to other findings which have shown a superiority of actuarial methods of prediction over clinician judgement. For example, Wedding (1983), who conducted a study which compared neuropsychologists' judgements to that of an actuarial formula, found that the actuarial approach was superior to that of the clinical judgement method. Thus it would seem that suggestions such as those made by Wedding and Faust (1989) who stated that "...greater overall accuracy is achieved when one adheres 'mindlessly' to actuarial procedures and never makes exceptions based on clinical judgement" (p.237), are not accurate within the area of IQ estimation.

These findings, in relation to IQ estimation parallel other studies within the area of premorbid IQ estimation. For example, Kareken and Williams (1994), also found that neuropsychologists' estimates were very similar to an actuarial formula. However, this study did not utilize objective IQ data and thus the estimates could not be compared to any obtained IQ scores.

Crawford, Millar and Milne (2001) did however utilize objective IQ data in their comparison of IQ estimation methods. In their study, the regression equation (Crawford Index) outperformed the clinician's estimates in that the estimated IQ's derived from it were significantly more correlated with the obtained IQ's than were the clinician-derived estimates. One possible explanation for the discrepancy between the findings of this study and the present study is that the Crawford index was specifically designed for use with the WAIS-R, whereas the current study utilized the WAIS-III. Also, the equation is based on a United Kingdom population and not an Australian one, as was utilized in this study. Furthermore, neuropsychologists (as opposed to clinical psychologists), who are more likely to estimate premorbid IQ's were used in this study, which could possibly account for the better performance of the clinicians in the present study.

This study also investigated the accuracy of the other available regression based methods of premorbid IQ estimation, The Barona Formula (Barona et al., 1984), The OPIE (Krull et al., 1995) and the OPIE-3 (Schoenberg et al., 2002) compared with the obtained IQs for the WAIS-III Data Group.

Before analysis began, the OPIE and the Barona, like the Crawford equation were adjusted by 2.9 points to make them suitable for prediction of WAIS-III IQs as they were originally designed for use with the WAIS-R. The OPIE-3 did not require adjustment as it is one of the only currently available regression based equations that has been designed for use with the WAIS-III.

The results indicated that the mean adjusted Barona and OPIE estimates and the estimates of the OPIE-3 were not significantly different to the mean obtained IQs for the WAIS-III Data Group. This suggests that all three methods yielded accurate IQ estimates. However, caution should be taken when interpreting these results as the OPIE estimates are based on current performance on WAIS-R/III subtests, (Vocabulary scores and Matrix Reasoning or Picture Completion) and therefore it would be expected that these scores would reflect current IQ in a non-neurologically impaired sample.

The Barona does not utilize current performance measures, and the finding that this yielded accurate IQ estimates is promising for use with Australian samples. One drawback of using the Barona and the OPIE equations is that they are based on American regions, and as such, guidelines in the region chosen (e.g North East = Victoria) would be required if this is to be used routinely with Australian samples.

The results from the present study expand on previous research by not only comparing clinician judgement to the Crawford equation in the estimation of IQ, but by also comparing the Barona, OPIE and OPIE-3 to objective IQ data. This is the only study conducted within the area of premorbid IQ estimation to compare all currently available regression equations to objective WAIS-II IQ data using an Australian sample.

Ideally the Wechsler Test of Adult Reading (WTAR; The Psychological Corporation, 2001) would have also been administered to the WAIS-III Data Group, however this test was not available at the time of data collection for this part of the study.

Clinician judgement is therefore as accurate as the Crawford Index in estimating IQ in an Australian population. However, both methods of prediction underestimate IQ as obtained using the WAIS-III. The Barona Formula, the OPIE and the OPIE-3 however provide reasonable estimates of IQ as obtained using the WAIS-III, but the Barona is the only one that doesn't rely on current functioning. This suggests that clinicians who need to estimate premorbid IQ should be aware of these findings, and seek a method of IQ estimation that provides the most accurate estimate of premorbid abilities (see Vanderploeg, 2000).

Further, as most of the regression equations were originally designed for use with the WAIS-R, clinicians are reminded to adjust these equations by 2.9 points in order to make them suitable for use with the WAIS-III. These findings highlight the fact that there is clearly a need for a regression equation to be designed that is specifically designed for use with the WAIS-III in Australian populations.

Relationship Between Experience and Accuracy.

To date, studies that have investigated the role of experience in clinical judgement have yielded mixed results. Some studies have shown that a clinician's level of training and years of experience are not related to judgement accuracy (e.g., Kendell, 1973; Werner, Rose & Yesavage, 1983), while several studies have provided support for the role of training and experience in improving the quality of clinical judgments (e.g., Grigg, 1958; Waxer, 1976). Of the few studies of judgement accuracy that have been conducted in neuropsychology, there seems little evidence to support the notion that experience or training is related to judgment accuracy (see Wedding & Faust, 1989; Garb & Schramke, 1996).

The results of this study indicted that there was no difference between the IQ estimates given by the novice and experienced clinician raters. These results are consistent with the findings from other neuropsychological studies exploring the relationship between experience, training, and judgement accuracy (see Garb, 1989). For example, Faust et al.(1988) found no relationship between training, experience and judgement accuracy in a judgement task that utilized neuropsychologists.

While being consistent with past research, the findings of the current study add to previous research in examining the relationship between years of experience and judgement accuracy within the field of IQ estimation, by utilizing neuropsychologists estimates of IQ rather than a diagnostic task such as diagnosisng laterality of brain impairment or diagnosing neuropsychological disorders from test scores, as most previous studies have done (e.g. Heaton et al., 1981; Leli & Filskov, 1984). The current findings add to the previous research by investigating the judgement accuracy of novice versus experienced clinical neuropsychologists, as opposed to comparing general psychology graduate students or clinical psychologists as other studies within this area have done (e.g. Jones, 1959; Waxer, 1976).

Training and experience has, therefore, been identified as playing little role in the judgement accuracy of clinical neuropsychologists estimates of IQ. This finding has implications for clinical neuropsychologists in clinical practice, and suggests they should be cautious when making clinical judgements within the area of IQ estimation as experience does not equate with judgement accuracy within this area. The tendency for clinician confidence to increase with years of experience ($r = 0.40$) is a particular concern as this is not reflected in increased judgement accuracy.

Clinician Confidence.

In addition to studies within the area of neuropsychology showing a limited relationship experience, training, and judgement accuracy, many studies have also shown that neuropsychologists are frequently overconfident about the judgements that they make (see Garb & Schramke, 1996). The results of this study indicate that overall, clinician confidence and accuracy were not related.

There was however one case for which clinician confidence and judgement accuracy were related. This case for an individual who had an IQ within the very superior range and educational and occupational attainment to match (refer to case 12, Table 11), suggesting that clinician confidence may be related to accuracy within the area of premorbid IQ estimation within the extreme end of the IQ spectrum. However, the mean IQ estimate was still approximately 10 points below the observed IQ for this case. This finding has implications for estimating decline after traumatic brain injury in high functioning individuals, as the tendency for clinicians to underestimate IQ in this range suggests that many patients are possibly being under compensated on the basis of percentage loss.

The findings of the present study parallel previous studies within the area of judgement research which explored the relationship between confidence and judgement accuracy (e.g., Faust et al., 1988; Kareken & Williams, 1994). For example, studies have shown a weak link between judgement accuracy and confidence ratings. In a study conducted by Gaudette (1992), in which neuropsychologists were asked to diagnose the presence or absence of brain impairment on the basis of test scores, it was found that the neuropsychologists were overconfident in their judgements. In the present study, confidence was not related to judgement accuracy in premorbid IQ prediction.

The finding that years of experience was related to confidence is also consistent with previous studies within this area. For example, experienced clinicians give higher confidence ratings than do inexperienced judges (Garb, 1989). Garb reviewed the role experience plays in clinical judgement and found a trend for the appropriateness of confidence ratings to be positively related to experience and presumed expertise.

The lack of an association between judgement accuracy and confidence ratings within the area of IQ estimation is consistent with Dawes et al.,(1989) proposal that judgement error occurs because clinicians are far more likely to evaluate individuals with significant problems than with those without them. In turn, this skewed exposure hinders attempts to make all of the needed comparisons and thus promotes overconfidence. This overconfidence can in turn impede judgement accuracy. This latter contention was not directly tested in the current study, however, the universal tendency of both novice and experienced clinicians to underestimate IQ and the tendency for confidence to increase with years of experience is consistent with previous findings that confidence is not related to accuracy.

The results of the present study also indicated that there were no differences between the IQ estimates given by the novice and experienced raters, but there was an effect of case for the volunteer WAIS-III subjects. This suggests that the clinician's accuracy differed across the 15 cases. In order to determine if this was due to the varying ability and background characteristics of the cases they were predicting and to try to determine why clinicians were more accurate at predicting some cases rather others, the relationship between demographic characteristics of the cases clinician confidence and clinician accuracy were explored.

The findings indicated that there was a positive correlation between gender, age and obtained IQ and clinician accuracy, and a positive correlation between clinician confidence, Obtained IQ and education. There was also a negative correlation between confidence and occupation (i.e. clinicians were more confident making estimates for individuals with more skilled occupations). This finding in conjunction with the results from Table 11, suggest that clinicians were more accurate at estimating IQ for cases who were younger females and that they were more confident in estimating cases with higher years of education and higher occupational attainment.

These findings add to previous research by examining the relationship between clinician confidence and accuracy in IQ estimation and the demographic characteristics of the cases for which they are predicting. The results of the present study indicate that clinicians who have to estimate premorbid IQ should be aware of the relationship between demographic variables, confidence and accuracy. Clinicians would be advised to seek alternate methods of IQ estimation other than clinical judgement since simply relying on the patients educational and occupational history does not seem to result in accurate IQ estimations.

Inter-rater Reliability.

Inter-rater reliability within the area of premorbid IQ estimation was another aspect of clinician judgement that was investigated. Unlike the research focussing on clinician confidence or the role of training and experience, some studies within this area have revealed that overall, reliability ratings made by clinicians ranged from good to excellent (see Garb & Schramke,1996).

The findings of the present study indicated good inter-rater reliability for the premorbid IQ estimations (α = 0.82). Further, when split into groups of novice and experienced clinicians, the inter-rater reliabilities were still good for both groups, with the experienced clinician group displaying a slightly better inter-rater reliability for IQ estimation than the novice group.

This finding is consistent with previous studies within the area of neuropsychological assessment. For example, Brown, Spicer, Robertson, Baird and Malik (1989), found good rater agreement in a task which required neuropsychologists to lateralise brain dysfunction (left hemisphere versus right hemisphere). However, this finding is inconsistent with the only other study which has investigated inter-rater reliability of neuropsychologists in a premorbid IQ estimation task. In this study, conducted by Kareken and Williams (1994), it was found that inter-rater agreement was less than ideal and estimates varied between clinicians. However, there is less scope for error in a task such as determining lateralization compared with estimating IQ.

The results from the present study expand on previous research by comparing the inter-rater reliabilities of novice and experienced clinicians within the area of IQ estimation. The good inter-rater reliabilities between the novice and experienced clinicians lends further support to the previous finding relating to judgement accuracy and the role of training and experience, as both the novice and experienced psychologists gave similar IQ estimates for each case.

These findings have implications for clinical practice as they highlight the fact that neuropsychologists tend to give similar estimates of premorbid IQ. As a result, two neuropsychologists may give the same individual a similar premorbid IQ estimate, however this estimate may not be correct. As such clinicians should use methods other than discussing their premorbid estimates with other clinicians as a means of verifying the accuracy of their premorbid IQ estimations.

Demographic Variables.

The relationship between demographic variables and IQ has already been established, whereby certain demographic variables have been shown to be correlated with IQ (e.g. Crawford, 1992). In an empirical study conducted by Barona et al., (1994), the demographic variables of education and occupation were shown to have a high correlation with intelligence. However, the demographic variables of age and gender were not associated with IQ.

In this study it was found that social class and education were associated with the obtained IQs and the clinician estimates of IQ. Age and gender were not associated with the obtained IQs or the clinician estimates of IQ. These findings are consistent with other findings in the literature within this area.

The weight clinicians give to demographic variables in determining premorbid IQ has also been investigated. For example, Kareken and Williams (1994) found that clinicians tend to give more weight to the demographic variable of education than it actually had, overestimating the amount this demographic variable contributes to IQ. That is, the clinicians made judgments as if the correlation was much higher than that established by empirical studies. It was suggested by the Karaken and Williams that this kind of overestimation would ultimately lead to an unreliable estimate of premorbid abilities.

The results of the present study parallel previous findings in which the weight of education in predicting IQ was overestimated. In the current study, the clinicians' estimates of IQ had a .78 correlation with education whereas the obtained IQ only had a .54 correlation with education. This finding suggests that clinicians overestimate the contribution of education in predicting premorbid IQ. Overall, this finding supports the notion proposed by Crawford et al., (1997) that clinical estimates of IQ may be based on vague or distorted impressions of IQ-demographic relationships.

However, the findings of the current study are not consistent with previous research which has investigated the association between age and IQ. For example, Crawford et al., (1997) found clinicians' estimates of IQ to be negatively correlated with age. It was proposed that clinicians may have overestimated the level of decline in cognitive abilities that occurs with age. In the present study, clinicians' estimates of IQ were not found to be associated with age. This finding suggests that 'ageism' was not a contributing factor to the clinical judgments made in the present study. It should also be noted however, that the sample used to obtain the WAIS-III IQs in the present study were relatively young (no participants over the age of 60 years). Perhaps a similar negative correlation between age and IQ may have been found if the sample had included individuals in an older age bracket.

Relationship Between Amount of Information and Judgement Accuracy.

The relationship between amount of information and judgement accuracy was also investigated in this study. Clinician estimates based solely upon demographic information about the 15 WAIS-III Data cases was compared with the same clinicians estimates based on a short video-taped interview in conjunction with the demographic information. The results indicated that there was no difference in the accuracy of the clinician estimates of IQ with or without the provision of more information about the cases. However the effect size of this difference was small to moderate with low standard deviations for the video and demographic information group suggesting that more information leads to more consistent responses across clinicians.

This finding is consistent with past research in this area. For example, Wedding (1983) conducted a study investigating neuropsychologists' judgements and found the amount of information was not associated with judgement accuracy. Wedding and Faust (1989) propose that judgement accuracy actually decreases when more information is provided (such as interview data in conjunction with test results) as a result of clinicians' preference for salient data. In doing this clinicians tend to overemphasize salient data and overlook other bland but informative data. Further, clinicians can only process a limited amount of information, and any additional information makes no difference in judgement performance (Wedding & Faust, 1989). This consistent finding is related to the limits of human cognitive capacity, whereby judgement accuracy increases until approximately seven pieces of information are presented (Miller, 1956).

This study also investigated the effect of case in clinician judgement using only the demographic information. Again, there was a significant effect for case suggesting that despite being given less information, there was still some cases that the clinicians were more accurate at predicting than others. This suggests that clinicians are not uniformly accurate (or not accurate) in predicating cases, and that some cases are easier to estimate an IQ for than others such as those with congruent educational, and occupational attainment. Individual clinician differences in judgement style may also account for some of these differences.

Relationship Between Amount of Information and Clinician Confidence.

The relationship between amount of information and clinician confidence was also investigated. The results indicated that clinicians were more confident in estimating IQ when they had both the video and demographic information upon which to base their estimates, compared to when they had only demographic information.

This finding parallels previous research within the area of clinician confidence which have shown that confidence in one's decisions tends to rise in relation to increasing information (Wedding & Faust, 1989). It has been proposed that the overconfidence clinicians have in their diagnoses is a factor that contributes to flawed human judgement (Arkes, 1981). For example, Oskamp (1965) found that once provided with more information, clinicians confidence increases but not their judgement accuracy. Further, Holsopple and Phelan (1954) found that the most confident diagnosticians tended to be the least accurate.

Relationship Between Amount of Information, Demographic Variables, Accuracy and Clinician Confidence.

Finally, the relationship between demographic characteristics of the WAIS-III Data Group, clinician confidence and clinician accuracy was investigated when clinicians were provided with only demographic information about the 15 cases. It was found that accuracy was related to age and obtained IQ, whereas confidence was related to education and obtained IQ.

This finding was different to that obtained when clinicians were presented with the video as well as the demographic information, as gender was not found to be correlated with accuracy and occupation was not correlated with confidence. This change in findings may in part be due to clinicians not having visual information upon which to base their estimates, and factors such as dress, speech, and other physical features may have influenced the clinicians judgements. Although this was not directly tested in the current study, it raises the possibility that clinical judgement is possibly influenced by a patient's appearance.

Study Two

Method

Participants

Australian Experienced Neuropsychologists.

A total of 31 participants comprised the Australian Experienced Neuropsychologists (AE) group. Subjects were recruited by two methods. First, those clinicians in Study 1 who had agreed to be interviewed at a later date were approached via telephone or in person to participate in the second part of this study. Secondly, an invitation to clinicians to participate in the survey was posted on the Australian Psychological Society College of Clinical Neuropsychologists' (CCN) Listserve. A copy of the questionnaire was sent to those clinicians who requested one. All participants in this group were registered as psychologists and were eligible for membership with CCN. All neuropsychologists participated on a voluntary basis. The demographic characteristics of the Australian Experienced neuropsychologists questionnaire group are shown in Table 17.

International Experienced Neuropsychologists.

The International Experienced Neuropsychologist (IE) Group consisted of 28 participants. Subjects were recruited through the internet and via word of mouth. A letter inviting clinicians to participate was posted on the npsych Listserve (www.npsych.com is a neuropsychologist discussion group with > 600 members). As before, a copy of the questionnaire was sent to those clinicians who requested one. All participants in this group were registered as psychologists, and participated on a voluntary basis. The demographic characteristics of the International Experienced Neuropsychologists' Group are shown in Table 17. Countries of origin included UK, Canada, and USA.

Australian Novice Neuropsychologists.

Fourteen participants comprised the Australian Novice Neuropsychologist (AN) Group. Subjects were recruited by three methods. First, those students Study One who had agreed to be interviewed at a later date were approached via telephone, or in person, to participate in the Study 2. Secondly, a copy of the questionnaire and letter inviting students to participate was emailed to the directors of the postgraduate training programs in neuropsychology at various universities in Australia. Thirdly, student neuropsychologist trainees at Victoria University who had not previously been approached in the first part of this study were invited to participate by the experimenter after a class. All participants in this group were undertaking training in an APS accredited postgraduate psychology course (Masters, Doctorate or PhD) and upon completion would be eligible to apply for membership with CCN. All student neuropsychologists participated on a voluntary basis. The demographic characteristics of the Australian Novice Group are also shown in Table 17.

Table 17
Demographic Characteristics of the Australian Experienced, International Experienced and Novice Neuropsychologists Groups

	AE (N=31)	IE (N=28)	AN (N=14)
Years Worked as a neuropsychologist			
0 (student)	0	0	14
0-5	6	5	0
5-10	10	6	0
>10	9	8	0
>20	6	9	0
Highest Qualification			
Honours	2	2	6
Graduate Diploma	0	0	4
Masters	17	3	3
Doctorate	6	2	1
PhD	6	21	0
Amount of time WAIS-III used in Assessments			
Never	4	6	0
1-5 times per week	25	19	14
5-10 times per week	2	3	0
Predominant area of clinical work			
Student	0	0	14
Rehabilitation	10	4	0
Acute	4	5	0
Paediatric	3	4	0
Geriatric	10	1	0
Medicolegal	4	10	0
Subscribe to neurospsychological journals			
Yes	19	27	2
No	12	1	12

Materials

Neuropsychologist Questionnaire.

For the second part to this study, clinicians were issued with a short, three page, structured questionnaire to complete (see Appendix E). The questionnaire contained demographic questions including age, gender, education level reached, number of years practicing as a neuropsychologist and number of hours spent in clinical assessment per week.

The questionnaire also contained questions relating to the types of methods used by the neuropsychologists in estimating premorbid IQ in clinical practice. For example, the neuropsychologists were asked to indicate the type of method they used to estimate premorbid IQ. Options included the NART, The Best Three Method, Interview, Demographic Variables, and School Reports. Other questions related to the amount of time each clinician assessed premorbid IQ and whether or not they believed that their formal education had provided them with sufficient training within this area.

Procedure

Australian Experienced Neuropsychologists.

Clinicians who agreed to participate were given a questionnaire and an explanatory letter about the study. This was sent to them via email, posted to them as a hard copy, or given to them in person. The clinicians for whom the questionnaires were not emailed were given a reply paid envelope and they were instructed to return the completed questionnaire via mail. The clinicians who were completing the questionnaire via email were instructed to click on the boxes in order to give their answers, and to then return the questionnaire via email to the researcher.

Identifying information about the clinicians was deleted and only the questionnaires were kept. For all questionnaires, the only identifying information on each was the clinicians' number of years working as a neuropsychologist, nature and setting of work, highest qualification, gender and age. This was necessary in order to establish the demographics of the clinicians used in this part of the study.

All participants participated on a voluntary basis. In total, 30 questionnaires were distributed to individuals who were approached in person to participate, in addition to the invitations to participate that were sent out to all neuropsychologists whose names appeared on the Australian Psychological Society CCN Listserve. Overall, 70 questionnaires were distributed over a 3-month period. There was a 44.3% response rate, with 31 questionnaires returned within the allocated time period.

International Experienced Neuropsychologists.

Clinicians who expressed interest in participating were sent a copy of the questionnaire and a further explanatory letter about the second part of the study. Clinicians were instructed to complete the questionnaire by clicking on the appropriate boxes to record their answers, and return the completed questionnaire via email to the researcher.

Participation was on a voluntary basis. Of the 600 plus neuropsychologists whose names appeared on the npsych Listserve as well as the 22 neuropsychologists who were approached in person, a total of 60 questionnaires were distributed over a 2-month period. There was a 46.6% response rate, with 28 questionnaires returned within the allocated time period.

Novice Australian Neuropsychologists.

Students who agreed to participate were given a questionnaire and an explanatory letter about the second part of the study. This was sent to them via email, posted to them as a hard copy, or given to them in person. They were asked to complete the questionnaire and return to the researcher in the reply paid envelope provided or via email.

An invitation to participate in the study was also emailed to the directors of the postgraduate training programs in neuropsychology at the universities of Western Australia, New South Wales, Queensland, Latrobe University and Melbourne University. The directors were asked to forward this on to their students. The invitation asked students who were interested in participating to contact the researcher in order for a questionnaire to be distributed to them. They were then asked to complete the questionnaire and return it via email to the researcher. A total of 27 questionnaires were distributed over a 2-month time period. There was a 50.0% response rate, with 14 questionnaires returned within the allocated time period.

Results

Hypothesis Testing –Clinicians' Practice Patterns in Assessing Premorbid IQ

Frequency of Premorbid IQ Estimations.

Table 18 shows the percentage of neuropsychologists who routinely estimate premorbid IQ. Overall, the majority of clinicians in the AE, IE and AN groups indicated that they always estimated a premorbid IQ.

Table 18
Percentage of Time Clinicians Estimate Premorbid IQ

		% Of Responses By Group		
Code	Description	AE (n = 31)	IE (n = 28)	AN (n = 14)
1	Always	54.8	46.4	42.9
2	Often	38.7	25.0	35.7
3	Sometimes	0	14.3	14.3
4	Occasionally	0	3.6	0
5	Rarely	6.5	7.1	0
6	Never	0	3.6	7.1

Note. Always = 100%; Often = 80%; Sometimes = 60%; Occasionally = 40%; Rarely = 20%; Never = 0%. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

A chi-square test for independence was conducted to investigate whether there was a difference in the frequency of IQ estimations for the three groups. Overall, the three groups did not significantly differ in the frequency of their premorbid IQ estimates, χ^2 (10, N = 73) = 10.15, $p > .05$.

Table 19 shows the percentage of neuropsychologists who made this estimate explicit in their reports. Overall, the majority of clinicians in the Australian Experienced and International Experienced groups indicated that they always made their premorbid IQ estimates explicit. The students in the Novice Australian group indicated that they often made this premorbid IQ estimation explicit.

Table 19
Percentage of Time Premorbid IQ Estimate is Made Explicit

		% Of Responses By Group		
Code	Description	AE (n = 31)	IE (n = 28)	AN (n = 14)
1	Always	54.8	50.0	28.6
2	Often	38.7	32.1	57.1
3	Sometimes	6.5	3.6	7.1
4	Occasionally	0	3.6	0
5	Rarely	0	3.6	7.1
6	Never	0	7.1	0

Note. Always = 100%; Often = 80%; Sometimes = 60%; Occasionally = 40%; Rarely = 20%; Never = 0%. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

A chi-square test of independence revealed no statistically significant differences between the three groups with regards to how often premorbid IQ estimations were made explicit in reports, χ^2 (10, N = 73) = 9.95, p > .05.

Circumstances of Premorbid IQ Estimations.

In relation to circumstances in which clinicians make a premorbid IQ estimation, 6.5% of respondents in the Australian Experienced group reported that they only made a premorbid IQ estimate in medico-legal cases, 32.3% reported that they only made a premorbid IQ estimate when there was a known neurological impairment, and 58.1% of respondents reported that they made a premorbid estimate for every assessment.

In the International Experienced group, 14.3% of respondents reported that they only made a premorbid IQ estimate in medico-legal cases, 39.3% reported making a premorbid IQ estimate when there was a known neurological impairment, 42.9% reported estimating premorbid abilities for every assessment, and 3.6% stated that they never made a premorbid IQ estimate (these clinicians worked with paediatric populations). For the Novice Australian group, 57.1% reported that they only made a premorbid IQ when there was a known neurological impairment and 42.9% of respondents reported that they made a premorbid estimate for every assessment.

There were no statistically significant differences in the circumstances in which the three groups estimated premorbid IQ, χ^2 (6, N = 72) = 6.41, p > *.05*.

Methods Used To Estimate Premorbid IQ.

Table 20 shows the methods for estimating premorbid IQ that were included in the questionnaire, as well as the percentage of respondents who indicated that they routinely use these particular methods for each group. Overall, the most frequently used methods for estimating premorbid IQ was vocational information and interview for the Australian Experienced group, the Wechsler Test of Adult Reading, vocational information and school records for the International Experienced group and the National Adult Reading Test and vocational information for the Australian Novice group.

Table 20
Method Used For Estimating Premorbid IQ

Method	% of Clinicians Routinely Using Method		
	AE (n = 31)	IE (n = 28)	AN (n = 14)
NART	64.5	32.1	92.9
Best 3 Performance	22.6	3.6	28.6
Barona Estimate	12.9	14.3	7.2
Crawford Index	0	7.1	0
Hold Subtests of the WAIS	51.6	32.1	64.3
Interview	74.2	42.9	78.6
Wide Range Achievement Test	12.9	25.0	14.3
Vocational Information	77.4	50.0	85.7
School Records	54.8	50.0	50.0
Wechsler Test of Adult Reading	38.7	60.7	7.1
Other	9.7	10.7	0

Note. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

A chi-squared test of independence revealed statistically significant differences between the groups in relation to their use of clinical interview, χ^2 (2, $N = 73$) = 7.99, $p < .05$, vocational information, χ^2 (2, $N = 73$) = 7.51, $p < .05$, and the WTAR, χ^2 (2, $N = 73$) = 11.19, $p < .05$. There were no statistically significant differences between the groups in their use of any other measures of premorbid IQ.

Qualitative analysis was conducted to determine the importance that clinicians give to each method of IQ estimation. Table 21 shows the percentage weight given to each method by the neuropsychologists. Overall, the Experienced Australian group gave the most weight to vocational information. The International Experienced group gave the most weight to the Wechsler Test of Adult Reading, vocational information and school records and the Australian Novice group gave the most weight to vocational information and the National Adult Reading Test, vocational information, and clinical interview.

Table 21
Percentage Weight Given to Each Method in Determining Premorbid IQ

		% of Clinicians		
Method	Percentage Weight	AE (n = 31)	IE (n = 28)	AN (n = 14)
NART				
	0%-50%	38.8	14.3	21.4
	50%-100%	38.7	21.5	78.6
School Records				
	0%-50%	22.6	10.7	42.8
	50%-100%	42.0	46.4	35.7
Best 3 Performance				
	0%-50%	6.4	10.7	28.5
	50%-100%	22.6	0	7.1
Vocational Information				
	0%-50%	0	10.7	0
	50%-100%	80.7	42.9	100.00
Barona Formula				
	0%-50%	6.5	14.3	0
	50%-100%	9.7	14.3	7.1
Interview				
	0%-50%	16.2	17.8	21.4
	50%-100%	54.8	32.2	71.4
Crawford Index				
	0%-50%	3.2	7.1	7.1
	50%-100%	0	3.6	0
WRAT				
	0%-50%	16.1	21.5	35.7
	50%-100%	6.4	17.9	14.2
Hold-Don't Hold				
	0%-50%	13.0	14.3	28.6
	50%-100%	41.9	17.9	42.9
WTAR				
	0%-50%	13.0	14.2	14.3
	50%-100%	25.9	46.4	7.1

Note. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

Clinician Training in the Area of Premorbid IQ Estimation.

The neuropsychologists were asked to indicate the methods of premorbid IQ estimation in which they had received training. Table 22 shows the percentage of clinicians who have received training in the various methods of premorbid IQ estimation by group. The method of premorbid IQ estimation that the most clinicians in the Experienced Australian neuropsychologist group had received training in was the National Adult Reading Test. The method of premorbid IQ estimation that the most clinicians in the Experienced International neuropsychologist group had received training in was the Hold Subtests of the Wechsler Adult Intelligence Scale. The Novice Australian neuropsychologist group reported receiving the most training in the use of the National Adult Reading Test and demographic variables. Overall, both the experienced groups reported little training in the use of Crawford Index to estimate premorbid IQ.

Table 22

Percentage of Neuropsychologists Trained in Methods of IQ Estimation

Method	% of Respondents		
	AE (n = 31)	IE (n = 28)	AN (n = 14)
NART	80.6	64.3	92.9
Best Performance	29.0	14.3	14.3
Hold Tests	77.4	67.9	64.3
Demographic Variables	64.5	57.1	92.9
Barona Formula	29.0	42.9	7.1
Crawford Index	3.2	3.6	21.4
Lezak	29.0	25.0	21.4
WRAT	22.6	42.9	57.1
WTAR	12.9	25.0	14.3

Note. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

A chi-square test of independence revealed no significant difference between the 3 clinician groups in the methods of premorbid IQ estimation in which they had been trained (all *p's* > *.05).*

Clinicians were also asked to indicate whether or not they felt that their training had been adequate within this area. Table 23 shows the percentage of clinicians in each group who either agreed or disagreed that they have been provided with sufficient levels of training in the area of premorbid IQ estimation.

Table 23
Percentage of Neuropsychologists who Believe Training In Premorbid IQ Estimation is Adequate

Beliefs About Training	% of Respondents		
	AE (n = 31)	IE (n = 28)	AN (n = 14)
Strongly Agree	9.7	10.7	7.1
Agree	48.4	28.6	35.7
Neutral	19.4	10.7	35.7
Disagree	22.6	25.0	21.4
Strongly Disagree	0.0	21.4	0.0

Note. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

In the Australian Experienced group, over 50% of respondents indicated that they believed that their training had provided them with sufficient instruction in methods of estimating premorbid IQ. Of the remaining clinicians, almost 25% of the sample disagreed that their level of training in the area of premorbid IQ estimation was sufficient.

In the International Experienced group, 39.3% of respondents reported that they believed that their training had provided them with sufficient instruction in methods of estimating premorbid IQ. Of the remaining clinicians, 46.4% disagreed that their level of training in the area of premorbid IQ estimation was sufficient. In the Australian Novice group, 42.8% agreed that their training in methods of premorbid IQ estimation was sufficient. As there were only second and third year students in the sample, the responses indicating that they felt there had been insufficient training was not a result of not having completed premorbid IQ estimation in their curriculum at the time of data collection.

The three groups did not significantly differ in their views about training, χ^2 (8, N = 72) = 14.47, p > *.05.*

Use of Demographic Variables in Premorbid IQ Estimation.

Table 24 shows the percentage of clinicians in each group that used each type of demographic variable when determining premorbid IQ. Overall, the most frequently used demographic variable appeared to be level of education for all three groups.

Table 24
Demographic Variables Used to Estimate Premorbid IQ

Demographic	% Clinicians Who Routinely Use		
	AE (n = 31)	IE (n = 28)	AN (n = 14)
Age	32.3	53.6	50.0
Gender	16.1	28.6	21.4
Level of Education	87.1	85.7	100.00
Type of Occupation	83.9	75.0	92.9
None	9.7	3.6	0

Note. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

A chi-squared test of independence was conducted to compare the demographic variables used by the Australian Experienced, International Experienced and Australian Novice Groups. There was no statistically significant differences in variables used between the three groups (all *p's* $>.05$).

The clinicians were also asked to report the percentage weight they would give to each of these demographic variables in estimating premorbid IQ (Table 25). Qualitatively, the Australian and International Experienced neuropsychologists gave the most weight to the demographic variable of education when estimating premorbid IQ. The Novice Australian neuropsychologists gave the most weight to type of occupation.

Table 25
Percentage Weight Given To Each Variable in Determining Premorbid IQ

Demographic	% of Clinicians		
	AE (n = 31)	IE (n = 28)	NA (n = 14)
Age			
0%-50%	22.6	25.0	35.7
50%-100%	16.1	10.7	14.2
Gender			
0%-50%	29.0	17.9	21.4
50%-100%	6.5	3.6	14.3
Level of Education			
0%-50%	16.1	17.9	21.4
50%-100%	58.0	42.9	64.3
Type of Occupation			
0%-50%	22.6	35.7	21.4
50%-100%	54.8	35.7	71.4

Note. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

Premorbid Estimation of Other Areas of Functioning.

The neuropsychologists were asked to indicate how often they estimated premorbid abilities for areas other than IQ (eg memory, executive functioning). Table 26 shows the percentage of neuropsychologists who indicated that they routinely estimated premorbid abilities in other areas. More than 40% of the Australian Experienced and International Experienced Neuropsychologists reported always or often estimating premorbid abilities in other areas. In contrast, only 28.6% of the Novice Australian group reported estimating premorbid abilities in other areas often or always and greater than 60% indicated estimating other premorbid abilities occasionally or less often.

Table 26
Percentage of Time Premorbid Estimates are Made In Areas Other Than IQ

Description	% Of Responses By Group		
	AE (n = 31)	IE (n = 28)	AN (n = 14)
Always	25.8	10.7	14.3
Often	16.1	32.1	14.3
Sometimes	22.6	10.7	7.1
Occasionally	9.7	10.7	28.6
Rarely	19.4	17.9	21.4
Never	3.2	17.9	14.3

Note. Always = 100%; Often = 80%; Sometimes = 60%; Occasionally = 40%; Rarely = 20%; Never = 0%. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

However, there was no statistically significant differences between the three groups in relation to the frequency of their estimations of premorbid functioning, χ^2 (10, N = 72) = 12.17, $p > .05$.

The neuropsychologists were also asked to indicate the types of other abilities for which they routinely estimated a premorbid estimate. Table 27 shows the percentage of clinicians who estimated other premorbid abilities in each group. All three groups reported making estimates of premorbid memory function the most often. However, a chi-square test of independence revealed that the three groups did not significantly differ in the areas of premorbid function routinely estimated (all *p's* > .05).

Table 27
Premorbid Abilities Routinely Estimated

Ability	% Clinicians Who Routinely Estimate		
	AE (n = 31)	IE (n = 28)	AN (n = 14)
Memory	67.7	64.3	64.3
Visuospatial	41.9	25.0	21.4
Motor Skills	19.4	17.9	21.4
Executive Functioning	51.6	42.9	50.0
Other	12.9	14.3	7.1
None	19.4	25.0	21.4

Note. AE = Australian Experienced neuropsychologists; IE = International Experienced neuropsychologists; AN= Australian Novice neuropsychologists.

Discussion

Methods of Premorbid IQ Estimation

Methods Used To Predict Premorbid IQ.

At present, there is little standardization or general agreement about which method of IQ estimation is best (Karaken, 1994). Most methods of IQ estimation have their shortcomings. This lack of general agreement is reflected in current practices in the field of IQ estimation and was the basis for the second part of this investigation that explored the methods used by neuropsychologists' to estimate premorbid IQ. Smith-Seemiller et al. (1997), conducted a similar investigation into the practice patterns utilized by American neuropsychologists in estimating premorbid IQ, finding relatively little use of strategies specifically designed to assess premorbid ability.

The present study not only investigated the practice patterns of neuropsychologists in Australia, but also aimed to replicate and further the study conducted by Smith-Seemiller et al., (1997) by investigating the practice patterns of neuropsychologists overseas in America, Canada and the UK. Further, the practice patterns of novice Australian neuropsychologists were also investigated. Overall, the findings indicated that the three groups used different methods for estimating premorbid IQ.

The results of this study indicated that in general, the Australian Experienced and Australian Novice neuropsychologists tended not to use methods specifically designed for the estimation of premorbid IQ, whereas the International group tended to mainly use only one of the many specifically designed premorbid IQ estimates available.

Consistent with the findings of the Smith-Seemiller et al. (1997) study, in the Australian Experienced group, the most frequently used methods for estimating premorbid IQ were clinical interview and vocational information. Clinicians were not specifically asked what they did with either the interview information or the vocational information, however it is assumed they utilized this information in making a clinical judgement regarding premorbid IQ. In contrast, the International Experienced neuropsychologists reported that the most frequently used tools for estimating premorbid IQ were the Wechsler Test of Adult Reading, followed by school records and vocational information.

This finding seems to suggest that the International neuropsychologists are more likely to use tests that are specifically designed to estimate premorbid IQ than the Australian neuropsychologists. However, it may be that due to its recent release, or co-norming with the WAIS-III and WMS-III , the WTAR is the method of choice at the moment. Further, the finding that school records and vocational information were the next most frequently used methods in the International sample (and not other methods specifically designed to estimate premorbid IQ such as the NART or the Crawford Index) suggest that overall the overseas clinicians tended not to use methods specifically designed to estimate premorbid IQ. It is possible however, that the international neuropsychologists have taken on board the findings from the Smith-Seemiller et al., (1997) study and changed their methods to reflect best practice.

The current study's findings are contrary to Smith-Seemiller et al's. finding that the NART was one of the least used methods of premorbid IQ prediction, and that the WRAT was used more routinely (1997). In the present study the NART was reportedly the third most frequently used method of IQ estimation in the Australian neuropsychologists group and the fifth most common in the International group, whereas the WRAT was one of the least commonly used methods in both groups.

Another routinely used method for estimating IQ reported by both the Australian and International neuropsychologist groups were school records. However, although Smith-Seemiller et al. (1997) also found that this was a frequently used method of IQ estimation, they noted that little is known about the extent that the information contained in school records is directly related to the tests utilized by neuropsychologists. Clinicians were not asked to state whether or not they actually obtained the school records (which may prove difficult in many cases), or whether they only obtained a verbal report of school performance. If the latter were the case, the accuracy of the individuals self-reports must be questioned when using this method (eg. the individual may state that they were an average student but may not have been).

The Barona Formula and the Crawford Index were two of the least frequently used methods of IQ estimation as reported by all three neuropsychologist groups. This is surprising, as these are two methods that have been specifically designed to estimate premorbid IQ, and have been the focus of many empirical studies and extensive reviews. Despite having shortcomings as premorbid IQ estimators, several studies have shown the superiority of these methods over clinical judgement (eg Kareken and Williams, 1994). This finding is also consistent with the view of Siegert (1999), who stated that despite many studies showing the superiority of statistical methods of prediction "most clinicians carry on their work unconcerned by such findings, probably unconvinced by them too, or just simply unaware of the studies" (p.39).

The Barona and the Crawford Indices are both designed for use with the WAIS-R and not the WAIS-III, which is currently used by most clinicians, which may be why they are not used. Also, the Crawford Index is relatively new, having only been developed in 1997. Further, as it is based on UK populations, it may not be well known outside of the UK.

In general, the findings of the present study indicated that methods relying on the use of clinical judgement were most routinely used by clinicians when making estimates of premorbid IQ. This has implications for practice as clinician judgement has been shown to have many flaws (e.g. Kareken, 1997). For example, humans have flaws and biases that prevent them from combining and integrating information accurately (Tversky & Kahnermann, 1974). Premorbid IQ estimation utilizing clinical judgement requires the combining of multiple sources of information as well as the appropriate weighting of that information in order to reach a judgement. However, this is exactly the type of instance under which human judgement is fallible (Miller, 1956; Wedding & Faust, 1989).

It may be the case that clinicians feel that they are aware of these reported flaws in human judgement and as such make these premorbid IQ estimations with this theory in mind. It should be noted however, that just being aware of these limitations in human reasoning is not enough to control their effects (Fischhoff, 1977). If clinicians are made aware of the common problems in decision-making they might be less likely to be influenced by heuristics and less biased in their judgement. However, many errors in reasoning are caused by the limited capacity of human memory and the human brain's inability to perform complex arithmetical procedures. Thus, although the effects of these limitations can be lessened by being aware of them, simply making more conservative judgements and trying harder cannot totally remedy them (Fischhoff, 1977).

Clinician Training within The Area Of IQ Estimation.

The findings of the current study indicated that there was no significant differences in the types of methods clinicians in the three groups had received training in. The majority of clinicians in all three groups reported that they had received training in the use of the NART, Hold-Don't Hold strategies and demographic variables to predict premorbid IQ. Very few clinicians stated that they had received training in the use of regression based estimates of premorbid IQ such as the Barona Formula or the Crawford Index. This finding may possibly account for the fact that relatively few clinicians reported actually using these latter methods in clinical practice to estimate premorbid IQ. It also highlights the need for clinicians who routinely estimate premorbid IQ as part of their clinical work, to obtain further training in the use of these methods.

However, only a minority of clinicians reported that they believed that their level of training within the field of premorbid IQ estimation was insufficient. The majority of clinicians reported that they were either neutral with regards to this issue or agreed that their training had been sufficient. Thus despite reporting that they had received little to no training in the methods specifically designed to estimate IQ (ie the Barona Formula or the Crawford Index), the clinicians in general, believed that their training within this area of premorbid IQ estimation had been adequate. Overall, the beliefs of the clinicians in all three groups with regards to the training they had received in this area did not significantly differ.

Although not actually investigated, perhaps this finding suggests that clinicians do not see the need to use regression estimates in order to predict IQ as found in the study conducted by Smith-Seemiller et al (1997). Alternatively, it could possibly reflect an inflated confidence in clinical ability. The fact that the majority of clinicians in each of the three groups indicated that they made a premorbid estimate for over 80% of their cases, and the majority of clinicians indicated making this estimate explicit in reports, suggests that they have confidence in the estimates that they make.

Further, many methods of premorbid IQ estimation may not have been available at the time many of the experienced clinicians were in training. As such, clinicians may have felt that their training was sufficient, given the availability of materials and the established empirical relationships at that time. For example, over 90% of the Novice Australian group reported that they had received training in the use of demographic variables whereas less that 65% of clinicians in the experienced group reported that they had received training in this method. Perhaps this reflects the better knowledge we now have with regards to the relationship between these variables and IQ.

Further, the Australian Novice group would have had training in most methods currently available as they are still completing their Postrgraduate courses. However, clinicians in the experienced groups may have completed their formal prior to the development of many methods currently available (e.g. WTAR, Crawford Index, OPIE-3) and as such never received training in their use.

These findings have implications for clinical practice, as it highlights the fact that many clinicians are not trained in methods designed to estimate premorbid IQ. Postgraduate neuropsychology courses could address this issue by including regression-based estimates as part of their curriculum. Further, clinicians should be aware of gaps in training and seek to fill these gaps in knowledge through appropriate channels such as training courses or continuing education.

Clinicians' Use of Demographic Variables.

As previous research has indicated, the demographic variables of education and occupation are related to IQ (Crawford, 1992; Barona et al., 1994). The extent to which clinicians use demographic variables to estimate premorbid IQ was investigated in the current study. The results indicated that the demographic variables used by all three groups did not differ and the most frequently used demographic variable used to estimate IQ was education followed closely by occupation. Both these demographic variables were also given the most weight as information used to determine premorbid IQ, compared with the other demographic variables. However, the demographic variables of gender and age were also reportedly used by all three groups, albeit to a lesser extent. This is an unfortunate finding as it is well established in the literature that age and gender are not associated with IQ, and should therefore not be used to estimate premorbid IQ (Barona et al., 1994).

These findings suggest that clinicians are not fully aware of the relationship between demographic variables and IQ. It also lends even further support to the notion proposed by Crawford et al. (1997), that clinical estimates of IQ are based on vague or distorted impressions of IQ-demographic relationships. As such clinicians who routinely use demographic variables as a means of estimating premorbid IQ should be aware of the exact relationship between demographics and IQ in order to ensure that they do not attribute a relationship where there is none.

Premorbid Estimates of Areas Other Than IQ.

Despite the existence of numerous methods to estimate premorbid IQ, there are virtually no methods available to estimate premorbid abilities in other areas. The current study investigated the practice patterns of Experience Australian, Experienced International and Novive Australian neuropsychologists in estimating premorbid abilities in areas other than IQ. The findings indicated no significant differences in the practice patterns of clinicians in the three groups with regards to premorbid estimation in areas other than IQ, and that all clinicians in the current study estimated premorbid abilities outside of the area of

IQ. The majority of clinicians however, did this less than 40% of the time. The most commonly estimated ability was memory, followed by executive functioning.

These results suggest that clinicians use clinical judgement to estimate premorbid levels of functioning within these areas, as actual tests to measure premorbid ability within the areas of memory and executive functioning do not exist. Given the reported findings in relation to limitations on clinical judgement (eg. Wedding & Faust, 1989), clinicians should be cautious in making estimates based on clinical judgement alone.

General Discussion

The judgement literature to date suggests that overall there is a limited relationship between training, experience and judgement accuracy, that clinicians are overconfident in their judgements, that inter-rater reliability is generally good for most judgement tasks, and that actuarial formulas are more accurate than clinical judgement. With respect to neuropsychologist's judgement within the area of IQ estimation, too few studies exist to draw any firm conclusions. Of the studies that do exist, objective IQ data was not available and judges other than neuropsychologists were used.

It was therefore the intention of this study to further investigate the role of experience, clinical judgement and actuarial methods of prediction within the area of premorbid IQ estimation. This study differed from previous research in this area and contributes to the literature by utilizing objective IQ data, clinical neuropsychologists, and an Australian sample.

This study consisted of two parts. Study One investigated clinician judgement of intellectual function whilst Study Two investigated clinicians practice patterns within the area of premorbid IQ estimation. The following discussion will draw together key findings from each study and make some general inferences with regards to how the results from this study tie in with previous research and relate to clinical practice.

The findings from Study One suggest that clinician judgement is no more accurate than the Crawford index in estimating IQ in an Australian sample, and that both methods are not accurate predictors of IQ. However, the adjusted Barona Formula, the adjusted OPIE and the OPIE-3 were accurate predictors of IQ in this same sample. However, given the current performance component in the OPIE and OPIE-3, their use with neurologically impaired samples may not lead to the accurate results seen in this study. Thus, the Barona Formula seems to be the most promising of all regression based methods currently available for use with Australian samples as there are no regression equations designed for specific use with Australian populations.

However, findings from Study Two suggest that clinical judgement is one of the most widely used methods of premorbid IQ estimation, both here and overseas. Further, consistent with predictions, there was very little use of methods specifically designed to measure premorbid IQ utilized by neuropsychologists' in clinical practice. In addition, the majority of clinicians believed that their training in this area had been sufficient despite not having received training in many methods of IQ estimation. These findings seem to suggest that most clinicians are content with their current methods of premorbid IQ estimation and do not perceive the need to incorporate other methods, despite the well-documented biases in human decision making upon which clinical judgement is based (e.g.

Faust et al, 1988; Kareken, 1997). Alternatively clinicians may not actually be aware of the flaws in human judgement and a further study could explore this issue further.

The findings that clinician confidence became greater with the amount of information provided, but judgement accuracy did not, further highlights the flaws inherent in human cognition. This finding is not peculiar to this study, and has been well documented in the judgement literature (Arkes, 1981; Wedding & Faust, 1989). In fact, as neuropsychologists, clinicians should be aware that human information processing capacity is approximately plus or minus seven pieces of information, yet the clinicians still reported greater confidence in their judgements in the task which provided them with more information upon which to base those judgements.

Another example of flawed human reasoning was the finding that clinicians overestimated the relationship between IQ and education in Study One. Further, in Study Two, clinicians reported basing IQ estimates on demographic variables such as age and gender, despite no empirically established relationships between these variables. Relationships have only been shown to exist between education, occupation and IQ (Barona et al., 1984).

These findings suggest that in estimating premorbid IQ, clinicians should be aware of their tendency to overestimate the relationship between education and IQ, and endeavor to ensure that they stick to the empirically determined relationships between these variables such as regression equations. In doing this, clinicians will ensure that the role of education is not inflated and ultimately lead to more accurate prediction of premorbid IQ. Clinicians should also be aware of the effects of their own implicit biases and ensure that they do not attribute a relationship between variables such as age and IQ or gender and IQ when there is none.

The findings in relation to the role of experience in judgement accuracy suggest that experienced clinicians are no more accurate than novice clinicians at estimating IQ. This finding was consistent with previous research in this area (Faust et al., 1988, Garb, 1989). Further, it was found that Clinician confidence was not related to judgement accuracy, but that years of experience were related to higher confidence ratings.

This suggests that neuropsychologists conducting assessments which require them to make premorbid IQ estimates should be aware that their confidence in their predictions may not be related to the accuracy of their judgement, and alternative steps should be taken to ensure their judgements are accurate. Further, experienced neuropsychologists need to be aware of the association between years of experience and increased confidence and ensure they question the accuracy and validity of the judgements they make.

Overall, these findings raise the question as to why clinicians continue to use clinical judgement as a means of premorbid IQ estimation despite the limitations inherent in this method. Garb (1989) proposes outdated training, failure to acquire good assessment information, lack of feedback or flawed feedback and inadequate cognitive processes as some factors that contribute to the poor judgements made by clinicians. A lack of knowledge about the inaccuracy of clinical judgements versus actuarial formulae may also contribute to increased confidence.

Siegert (1999) proposed that despite many studies showing the superiority of statistical methods of prediction, "most clinicians carry on their work unconcerned by such findings, probably unconvinced by them too, or just simply unaware of the studies" (p.39). Although not directly investigated in the present study, it would seem, given the reported practice patterns in IQ estimation, that Australian clinicians are either unaware of the limitations in human judgement or unconcerned by them, or unaware of other methods of IQ estimation.

According to Tversky and Kahnermann (1974), flaws and biases inherent in cognition prevent humans from combining and integrating information accurately. Perhaps it is these flaws and biases that prevent clinicians from recognizing their own limitations, ultimately leading them to continue to use their own clinical judgement over more accurate methods. It may also be these biases that lead clinicians to assume years of experience equate with clinical judgement accuracy.

Taken together, the findings from the present study suggest that neuropsychologists conducting assessments which require premorbid IQ estimation for diagnostic decisions should endeavor to update training within the area of premorbid IQ estimation, obtain accurate feedback regarding their predictions, and acknowledge limitations to cognition which can influence decision-making accuracy. On the basis of the current findings, clinicians working with Australian samples would be advised to supplement current practices with accurate regression-based methods such as the Barona Formula, until such time as empirically tested premorbid IQ estimate methods specifically designed for use with Australian populations become available.

Limitations

A principle limitation of this study was that is was cross sectional. A longitudinal study would have been better suited to determining the relationship between training and judgement accuracy as well as the practice patterns of experienced and novice neuropsychologists.

A second limitation to this study was that recruitment was not random, and as such results may reflect a response bias. That is, only clinicians who were interested in the area of premorbid IQ estimation may have agreed to participate in the study and therefore a representation of all neuropsychologists practices was unable to be obtained.

A third limitation of this study was that the majority of the respondents from the International Experienced group were acquaintances of the research supervisor and may have held similar views with regards to evidence-based practice. This may have also skewed results towards this direction and not provided an accurate reflection of the current practices employed by neuropsychologists overseas.

Another limitation of this study was that reading-based methods were not compared to the regression based estimates in order to get a more comprehensive evaluation of methods currently available. Ideally, the WTAR would have been also administered (as this has been specifically designed for use with the WAIS-III), however this test was not available at the time of data collection.

A final limitation of this study was the relatively small sample size of each group. Ideally, larger samples would have been used in order to gain a better picture of current practices within the area of premorbid IQ estimation.

Future Research

As one of only three studies that have investigated clinician judgement within the area of IQ estimation (Kareken & Williams, 1994; Crawford et al., 2001), and one of only two studies that have investigated clinical neuropsychologists' judgement within the area of premorbid IQ estimation (Kareken & Williams, 1994), there is clearly a need for more research in this area. Future studies may wish to replicate this study using a larger and more representative sample.

Further, this is the only study that has investigated Australian neuropsychologists' practice patterns in assessing premorbid IQ and the only study to compare the practice patterns of both novice and expert clinicians. Thus there is a clear need for more research in this area in order to establish trends and best practice within the area of premorbid IQ estimation. Future research may wish to replicate and extend these preliminary findings.

There is also a need to establish a regression based equation based on Australian demographic data and the WAIS-III, as current regression estimates are based on the demographic data of the United Kingdom or the USA and are designed for use with the WAIS-R.

Similarly, there is a need for future research to look at the prediction of cognitive variables other than IQ. The question of how we determine premorbid estimates of motor, executive and memory functioning needs to be addressed. Perhaps regression formulas can also be developed for these functions.

Finally, the current study investigated clinical judgement within the area of adult premorbid IQ estimation. To date relatively few studies have investigated premorbid IQ estimation on pediatric populations, and this could be an interesting area for future research.

Conclusion

The estimation of premorbid IQ is a complicated task. Numerous methods exist, however none are without their flaws. Results of the present study indicate that the Crawford Index is no more accurate than clinical judgement in predicting premorbid IQ in an Australian sample. In addition, judgement accuracy was not related to years of experience, clinician confidence was not related to judgement accuracy, and inter–rater reliability within the area of premorbid IQ estimation was found to be good. Further clinician confidence was shown to rise with increasing information, but accuracy was not found to differ.

In relation to the current practices used by clinicians to assess premorbid IQ, it was found that clinicians tend to use clinical judgement over actuarial formulas. It was also found that most clinicians had received little training in methods specifically designed to estimate premorbid IQ and that all clinicians endorsed estimating premorbid abilities in areas other than IQ.

While the pattern of the current findings indeed suggest that neither clinical judgement nor the Crawford index are accurate at predicting premorbid IQ in an Australian Sample, it emphasizes the need for clinicians to be wary in applying these methods. Further, the finding that clinician judgement was one of the most widely used methods of premorbid IQ estimation despite the documented biases in human decision making also highlights this point. However, as there is no universally accepted premorbid estimate method, clinicians working with Australian samples would be best advised to supplement their current practices with several other methods of premorbid estimates and base estimates upon coinciding data.

References

Adams, K. M., & Putman, S. H. (1994). Coping With Professional Skeptics: Reply to Faust. *Psychological Assessment*, 6, 5-7.

American Psychological Association. (1982). *Report of The Task Force On The Evaluation of Education, Training, and Service in Psychology.* Washington, DC: Author.

Arkes, H. (1981). Impediments to Actual Clinical Judgement and Possible Ways to Minimize Their Impact. *Journal of Consulting and Clinical Psychology, 49,* 323-330.

Barona, A., Reynolds, C. R., & Chastain, R. (1984). A Demographically Based Index of Premorbid Intelligence For The WAIS-R. *Journal of Consulting and Clinical Psychology*, 52, 885-887.

Berg, M. (1997). *Rationalising Medical Work: Decision Support Techniques and Medical Practices*. Cambridge, MA: MIT Press.

Blair, J., & Spreen, O. (1989). Predicting Premorbid IQ: A Revision of The National Adult Reading Test. *The Clinical Neuropsychologist*, 3, 129-136.

Brehmer, B., Kuylenstierna, J., & Lilijergen, J. (1974). Effects of function form and cue validity on the subjects' hypotheses in probabilistic inference tasks. *Organizational and Human Behaviour,* 11, 338-354.

Brown, G. G., Spicer, K. B., Robertson, W. M., Baird, A. D., & Malik, G. (1989). Neuropsychological Signs of Lateralised Arteriovenous Malformations: Comparisons With Ischemic Stroke. *The Clinical Neuropsychologist,* 3, 340-352.

Chapman, L., & Chapman, J. (1967). Genesis of Popular but Erroneous Psychodiagnostic Observations. *Journal of Abnormal Psychology, 72,* 193-204.

Crawford, J. R. (1989). Estimation of Premorbid Intelligence: A review of recent developments. In J. R. Crawford & D. M. Parker (Eds), *Developments in_Clinical and Experimental Neuropsychology_* (p.55-74). New York: Plenum Press.

Crawford, J. R. (1992). Current and Premorbid Intelligence Measures In Neuropsychological Assessment. In J. R. Crawford and D. M Parker (Eds), *Developments In Clinical and Experimental Neuropsychology* (p. 21-49). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Crawford, J. R., & Allan, K. M. (1997). Estimating Premorbid IQ With Demographic Variables: Regression Equations Derived From a UK Sample. *The Clinical Neuropsychologist*, 11, 192-197.

Crawford, J. R., Millar, J., & Milne, A. B. (2001). Estimating Premorbid IQ From Demographic Variables: A Comparison of a Regression Equation vs. Clinical Judgement. *British Journal of Clinical Psychology*, 40(1), 95-105.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus Actuarial Judgement. *Science*, 243, 1668-1674.

Einhorn, H. J. (1972). Expert Measurement and Mechanical Combination. *Organizational Behaviour and Human Performance,* 7, 86-106.

Einhorn, H. J. (1986). Accepting Error to Make Less Error. *Journal of Personality and Assessment*, 50, 387-395.

Einhorn, H. J. (1988). Diagnosis and Causality in Clinical and Statistical Prediction. In Turk, D., & Salovey, P. (eds.), *Reasoning and Inference in Clinical Psychology*. New York: The Free Press.

Faust, D., Guilmette, T. J., Hart, K., Arkes, H. R., Fishburne, F. J., & Davey, L. (1988). Neuropsychologists' Training, Experience, and Judgement Accuracy. *Archives of Clinical Neuropsychology*, 3, 145-163.

Fischhoff, B. (1975). Hindsight = Foresight: The Effect of Outcome Knowledge on Judgement Under Uncertainty. *Journal of Experimental Psychology: Human Perception and performance, 1,* 288-299.

Fischhoff, B. (1977). Perceived Informativeness of Facts. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 349-358.

Flynn, J. R. (1984). The Mean IQ of Americans: Massive Gains 1932-1978. *Psychological Bulletin,* 95, 29-51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ Tests Really Measure. *Psychological Bulletin,* 101, 171-191.

Flynn, J. R. (1998). WAIS-III and WISC-III IQ Gains in the United States from 1972 to 1995: How to Compensate for Obsolete Norms. *Perceptual and Motor Skills*, 86, 1231-1239.

Franzen, M. D., Burgess, E. J., & Smith-Seemiller, L. (1997). Methods of Estimating Premorbid Functioning. *Archives of Clinical Neuropsychology,* 12, 711-738.

Friedlander, M., & Stockman, J. (1983). Anchoring and Publicity Effects in Clinical Judgement. *Journal of Clinical Psychology, 39,* 637-643.

Garb, H. N. (1989). Clinical Judgement, Clinical Training, and Professional Experience. *Psychological Bulletin,* 105, 387-396.

Garb, H, N., & Schramke, C. J. (1996). Judgement Research and Neuropsychological Assessment: A Narrative review and Meta-Analyses. *Psychological Bulletin*, 120(1), 140-153.

Gilpin, E. A., Olshen, R., Chatterjee, K., Kjekshus, J., Moss, A., Henning, H., Engler, R., Blacky, R., Dittrich, H., & Ross, J. (1990). Predicting 1 Year outcome Following Acute Myocardial Infarction: Physicians Versus Computers. *Computers and Biomedical Research,* 23, 46-63.

Goldberg, L. R. (1959). The Effectiveness of Clinician's Judgements: The Diagnosis of Organic Brain Damage From The Bender Gestalt Test. *Journal of Consulting Psychology*, 23, 25-33.

Goldstein, S. G., Deysach, R. E., & Kleinknecht, R. A. (1973). Effect of Experience and Amount of Information On Identification Of Cerebral Impairment. *Journal of Consulting and Clinical Psychology,* 41, 30-34.

Grant, I., Heaton, R. K., McSweeny, J., Adams, K. M., & Timms, R. M. (1982). Neuropsychologic Findings in Hypoxemic Chronic Obstructive Pulmonary Disease. *Archives of Internal Medicine,* 142, 1470-1476.

Grant, I., Mohns, L., Miller, M., & Reitan, R. (1976). A Neuropsychological Study Of Poly Drug Users. *Archives of General Psychiatry,* 33, 973-978.

Grebstein, L. (1963). Relative Accuracy of Actuarial Prediction, Experienced Clinicians, and Graduate Students in a Clinical Judgement Task. *Journal of Consulting Psychology,* 37, 127-132.

Grigg, A. E. (1958). Experience Of Clinicians, and Speech Characteristics and Statements of Clients as Variables In Clinical Judgement. *Journal of Consulting Psychology,* 22, 315-319.

Grober, E., & Sliwinski, M. (1991). Development and Validation of a Model For Estimating Premorbid Verbal Intelligence In The Elderly. *Journal of Clinical and Experimental Neuropsychology*, 13, 933-949.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus Mechanical Prediction: A Meta Analysis. *Psychological Assessment,* 12 (1), 19-30.

Guadette, M. D. (1992). Clinical Decision Making In Neuropsychology:Bootstrapping the Neuropsychologist Utilizing Brunswik's Lens Model. (Doctoral Dissertation, Indiana University of Pennsylvania, 1992). *Dissertation Abstracts International,* 53, 2059B.

Heaton, R. K., Grant, I., Anthony, W. Z., & Lehman, R. A. (1981). A Comparison Of Clinical and Automated Interpretation of The Halstead Reitan Battery. *Journal of Clinical Neuropsychology,* 3, 121-141.

Holsopple, J., & Phelan, J. (1954). The Skills of Clinicians in Analysis of Projective Tests. *Journal of Clinical Psychology, 10,* 307-320.

Holt, R. R. (1958). Clinical and Statistical Prediction. A Reformulation and Some New Data. *Journal of Abnormal and social Psychology,* 56, 1-12.

Jenkins, H., & Ward, W. (1965). The Judgement Contingency Between Responses and Outcomes. *Psychological Monographs, 79,* (Whole No. 594).

Jones, N. F. (1959). The Validity of Clinical Judgements in Schizophrenic Pathology Based on Verbal Responses to Intelligence Test Items. *Journal of Clinical Psychology,* 15, 396-400.

Karaken, D. A. (1997). Judgement Pitfalls in Estimating Premorbid Intellectual Function. *Archives of Clinical Neuropsychology,* 12, 701-709.

Karaken, D. A., & Williams, J. M. (1994). Human Judgement and Estimation of Premorbid intellectual Function. *Psychological Assessment,* 6, 83-91.

Kahnemann, D., & Tversky, K. (1973). On the Psychology of Prediction. *Psychological Review, 80,* 237-251.

Kendell, R. E. (1973). Psychiatric Diagnosis: A Study of How They Are Made. *British Journal of Psychiatry*, 122, 437-445.

Krull, K. R., Scott, J. G., & Sherer, M. (1995). Estimation of Premorbid Intelligence From Combined Performance and Demographic Variables. *The Clinical Neuropsychologist*, 9, 83-88.

Kurtz, R., & Garfield, S. (1978). Illusory Correlation: A Further Exploration of Chapman's Paradigm. *Journal of Consulting and Clinical Psychology, 5,* 1009-1015.

Leli, D. A., & Filskov, S. B. (1981). Clinical-Actuarial Detection and Description of Brain Impairment With The WB Form I. *Journal of Clinical Psychology,* 37, 623-629.

Leuger, R., & Petzel, T. (1979). Illusory Correlation in Clinical Judgement, Effects of Amount of Information To Be Processed. *Journal of Consulting and Clinical Psychology, 47,* 1120-1121.

Levenberg, S. B. (1975). Professional Training, Psychodiagnostic Skill, and Kinetic Family Drawings. *Journal of Personality Assessment,* 39, 389-393.

Lezak, M. D. (1983). *Neuropsychological Assessment (2nd Ed).* New York: Oxford University Press.

Lezak, M. D. (1995). *Neuropsychological Assessment (3rd Ed).* New York/Oxford: Oxford University Press.

Meehl, P. E. (1954). *Clinical vs Statistical Prediction: A Theoretical Analysis and a Review Of The Evidence.* Mineapolis: University of Minnesota Press.

Nelson, H. & McKenna, P. (1975). The Use of Current Reading Ability In The Assessment of Dementia. *British Journal of Social and Clinical Psychology,* 14, 259-267.

Nisbett, R., & Ross, L. (1980). *Human Inference: Strategies and Shortcommings of Social Judgement.* Englewood Cliffs: Prentice Hall.

Office of Population, Censuses and Surveys. (1980). *Classification of Occupations.* London: Her Majesty's Stationary Office.

Oskamp, S. (1965). Overconfidence in Case-Study Judgements. *Journal of Consulting Psychology, 29,* 261-265.

Pallant, J. (2001*). SPSS Survival Manual.* New South Wales: Allen &Unwin.

Reitan, R. (1986). Theoretical and Methodological Bases of The Halstead Reitan Neuropsychological Battery. In Grant , G., & Adams, K. (eds), *Neuropsychological Assessment of Neuropsychiatric Disorders.* New York: Oxford University Press.

Reynolds, C. R., & Gutkin, T. B. (1979). Predicting The Premorbid Intellectual Status Of Children Using Demographic Data. *Clinical Neuropsychology,* 1, 36-38.

Ruchinskas, R. (2003). Clinical Prediction of Falls in The Elderly. *American Journal of Physical Medicine and Rehabilitation*, 82, 273-278.

Ryan, J. J., & Paolo, A. M. (1992). A Screening Procedure for Estimating Premorbid Intelligence in The Elderly. *The Clinical Neuropsychologist,* 6, 53-62.

Saling, M. (1994). Neuropsychology Beyond 2001: The Future of Diagnostic expertise. In S. Touyz, D. Byrne & A. Gilandas (1994). *Neuropsychology in Clinical Practice.* Sydney: Harcourt Brace & Co.

Sarbin, T. R. (1943). A Contribution to the Study of Actuarial and individual Methods or Prediction. *American Journal of Sociology,* 48, 593-602.

Sattler, J. (2001). *Assessment of Children: Cognitive Applications.* USA: Sattler Publisher, Inc.

Schinka, J., & Vanderploeg, R. (2000). Estimating Premorbid Level of Functioning. In Vanderploeg, R. (ed). *Clinicians Guide to Neuropsychological Assessment.* London: Lawrence Erblaum Associates.

Schoenberg, M., Scott, J., Duff, K., & Adams, R. (2002). Estimation of WAIS-III Intelligence from Combined Performance and Demographic Variables. Development of the OPIE-3. *The Clinical Neuropsychologist, 16,* 426-438.

Schonell, F. (1942) *Backwardness in The Basic Subjects.* Edinburgh: Oliver & Boyd.

Siegert, R. J. (1999). Some Thoughts About Reasoning In Clinical Neuropsychology. *Behavior Change,* 16(1), 37-48.

Smith-Seemiler, L., Franzen, M. D., Burgess, E. J., & Prieto, L. R. (1997). Neuropsychologists' Practice Patterns in Assessing Premorbid Intelligence. *Archives of Clinical Neuropsychology,* 12(8), 739-744.

Snyder, M. (1981). Seek and Ye Shall Find. In Higgins, E., Herman, C., & Zanna, M. (eds), *Social Cognition: The Ontario Symposium on Personality and Social Psychology.* Hillsdale: Erlbaum.

Spreen, O., & Strauss, E. (1998). *A compendium of Neuropsychological Tests.* New York: Oxford University Press.

The Psychological Corporation. (2001). *Manual For the Wechsler Test of Adult Reading (WTAR).* San Antonio, TX: Author.

Tversky, A., & Kahnemann, D. (1973). Availabilty: A Heuristic For Judging Frequency and probability. *Cognitive Psychology, 4,* 207-232.

Tversky, A., & Kahnemann, D. (1974). Judgement Under Uncertainty: Heuristics and Biases. *Science, 185,* 1124-1131.

Watson, P. (1960). On Failure to Eliminate Hypotheses in a Conceptual Task. *Quarterly Journal of Experimental Psychology, 12,* 129-140.

Waxer, P. (1976). Nonverbal Cues For Depth of Depression: Set Versus No Set. *Journal of Consulting and Clinical Psychology,* 44, 493.

Wechsler, D. (1997). *Manual for the Wechsler Adult Intelligence Scale (3rd ed.).* San Antonio, TX: The Psychological Corporation.

Wedding , D., & Faust, D. (1989). Clinical Judgement and Decision Making In Neuropsychology. *Archives of Clinical Neuropsychology*, 4, 233-265.

Wedding, D. (1983). Clinical and Statistical Prediction in Neuropsychology. *Clinical Neuropsychology*, 5, 679-684.

Werner, P. D., Rose, T. L., & Yesavage, J. A. (1983). Reliability, Accuracy, and Decision Making Strategy in Clinical Predictions of Imminent Dangerousness. *Journal of Consulting and Clinical Psychology*, 51, 815-825.

Wiens, A. N., Bryan, J. E., & Crossen, J. R. (1993). Estimating WAIS-R FSIQ From The National Adult Reading Test – Revised in Normal Subjects. *The Clinical Neuropsychologist,* 8, 70-84.

Wilkinson, G. (1993). *Wide Range Achievement Test – 3 Manual.* Wilmington, DE: Wide Range Inc.

Wilson, R. S., Rosenbaum, G., Brown, G., Rourke, D., Whitman, D., & Grisell, J. (1978). An Index of Premorbid Intelligence. *Journal of Consulting and Clinical Psychology,* 46, 1554-1555.

APPENDIX A
Study Information Sheet for Volunteers

VICTORIA UNIVERSITY
DEPARTMENT OF PSYCHOLOGY
INVITATION TO PARTICIPATE IN A RESEARCH STUDY - VOLUNTEERS

Dear

My name is Michelle Lindsay, I am a Clinical Neuropsychology Doctorate Student, supervised by Izabela Walters, within the Department of Psychology at Victoria University St Albans Campus.

I am undertaking research on clinician judgements of intellectual function. Neuropsychologists are Frequently asked to conduct intelligence tests on patients in order to assess the loss of functioning that a patient has suffered following a neurological injury. In order to assess how much loss of functioning has occurred, test results pre and post injury are compared. This poses a problem for neuropsychologists as pre injury test results rarely exist. As a result, neuropsychologists must make a clinical judgement, based on information that the patient provides such as age, education level reached, occupation and other information about themselves as to what the patients IQ was prior to the injury. I am interested to see how accurate these judgements are.

What I require from you is approximately 2 hours of your time. During this time I will administer the Wechsler Adult Intelligence Scale. This scale comprises of several tests that measure functioning in different areas. All you will be required to do is answer questions that I will ask you. You can be informed of your IQ if you wish following the administration of the test. This will be given in the form of the range under which your IQ falls. Once the test has been administered you will be asked some questions regarding your demographic variables. In addition, a short interview will take place that will be video taped with your consent. In this video you will be asked a series of questions. The questions are designed to give more of an insight into who you are and your experiences. These will not be personal questions, rather questions surrounding your work, and accomplishments you have made in your life. The interview will take approximately 5 minutes.

The video will be shown to neuropsychologists participating in the study. You will not have to state yourname at any time during the video. This interview in conjuntion with your demographic variables will be the only information given to the neuropsychologists in order for them to 'judge' your IQ. Your actual IQ score will not be made available to the neuropsychologists at any time. The investigators (myself and Izabela Walters) will be the only people who will know your actual IQ. No identifying details will be included in the final report either.

If you are willing to participate, please complete the consent forms and return them via the reply paid envelopes. Once I receive your consent to participate in the study, I will contact you to arrange a time and place that is convenient to administer the test and interview.

Should you have any concerns regarding the manner in which this research project is conducted, please do not hesitate to inform the researchers directly or the Victoria University Human Research Ethics Committee (ph. 9688 4710).Results will be available at the end of the project from the Department of Psychology. If you have any queries you can contact myself or Izabela Walters on (03) 9365 2778.

Thank you

Yours Sincerely,
Michelle Lindsay

APPENDIX A
Consent Form For Volunteers

Victoria University of Technology
Consent Form for Participants Involved in Research

INFORMATION TO VOLUNTEERS:

We would like to invite you to be a part of a study into clinician judgement of intellectual function. The aims of the study are to investigate the accuracy of clinical judgement compared with a statistical equation that has been designed to measure IQ based on a persons demographic variables, the role experience plays in clinical judgement, and the processes involved in clinical judgement.You will be administered the Wechsler Adult Intelligence Scale in order for your IQ to be determined. A short interview, which will be video taped, will follow. This videotape along with your demographic variables will be shown to the neuropsychologists who are participating in this study. They will be asked to 'guess' your IQ based on the interview and your demographic variables. At no stage will the neuropsychologists be given your actual IQ score. By knowing your actual IQ we can assess how accurate the clinicians judgements have been. In addition, by knowing your demographic variables, we can enter those into an equation that will provide a statistical (actuarial) estimate of your IQ.

The final report will compare two groups of neuropsychologists, expert and novice and clinical judgement to an actuarial formula. As such no individuals or individual results will be identifyable or published.

CERTIFICATION BY PARTICIPANT

I, ____

of ____

certify that I am at least 17 years old and that I am voluntarily giving my consent to participate in the study entitled: Clinician Judgement of Intellectual Function being conducted at Victoria University of Technology by Ms Michelle Lindsay and Ms Izabella Walters

I certify that the objectives of the study, together with any risks to me associated with the procedures listed below to be carried out in the study, have been fully explained to me by Michelle Lindsay,and that I freely consent to participation involving the use of these procedures.

Procedures:
Administration of the Wechsler Adult Intelligence Scale (WAIS-III) and a short clinical interview which will consist of questions regarding demographic background, work and educational history.

I certify that I have had the opportunity to have any questions answered and that I understand that I can withdraw from this study at any time and that this withdrawal will not jeopardise me in any way.
I have been informed that the information I provide will be kept confidential.

Signed: .... }

Witness other than the experimenter (as appropriate) } **Date: ....**

....}

Any queries about your participation in this project may be directed to the researcher (Ms Michelle Lindsay and Ms Izabella Walters on 9365 2778). If you have any queries or complaints about the way you have been treated, you may contact the Secretary, University Human Research Ethics Committee, Victoria University of Technology, PO Box 14428 MCMC, Melbourne, 8001 (telephone no: 03-9688 4710).

APPENDIX B
Study Information Sheet for Clinicians – Part One

VICTORIA UNIVERSITY
DEPARTMENT OF PSYCHOLOGY
INVITATION TO PARTICIPATE IN A RESEARCH STUDY - CLINICIANS

Dear

My name is Michelle Lindsay, I am a Clinical Neuropsychology Doctoral Student, supervised by Izabella Walters, within the Department of Psychology at Victoria University St Albans Campus.

I am undertaking research on clinician judgements of intellectual function. Clinical judgement is often used in the area of clinical neuropsychological assessment in relation to premorbid IQ estimation, as there is frequently a need for some baseline against which to compare current performance. Actual premorbid test scores antedating cognitive decline are rarely available, yet diagnosis requires that some decline be demonstrated. Within the field of IQ estimation, several methods exist to determine premorbid IQ. Although clinical judgement is frequently cited as one of the least effective measures, it is the only one that can be used in the case of migrant populations as reading tests and the use of demographic predictors are generally rendered ineligible. There is therefore a need for the investigation of how accuarte neuropsychologist's clinical judgements are in relation to IQ estimation.

Should you wish to take part in this study, I firstly require you to complete and sign the enclosed consent form. If you are available for an interview at a later stage, please also complete the form entitled "I am available for an interview". Next send all this information back to me in the reply paid envelope provided. I will then send you out a scoring sheet and videotape which contains a short clinical interview with several volunteers. You will be required to watch the video and estimate the IQ of each of the cases. The demographic variables of each case will be provided on the scoring sheets as well.

The video runs for approximately half an hour. I recognise that many demands are made on your time and energy. A response from all clinicians will enable me to gain an overall picture of the accuracy of neuropsychologists' judgement of IQ. Information given will be treated as confidential by the researchers; scoring sheets will be identified only by the years of experience each clinician has had, and only group results will be published.

Should you have any concerns regarding the manner in which this research project is conducted, please do not hesitate to inform the researchers directly or the Victoria University Human Research Ethics Committee (ph. 9688 4710).

Results will be available at the end of the project from the Department of Psychology. If you have any queries you can contact myself or Izabella Walters on (03) 9365 2778.

Thank you
Yours Sincerely,

Michelle Lindsay

APPPENDIX B
Consent Form For Clinicians – Part One

Victoria University of Technology
Consent Form for Participants Involved in Research

INFORMATION TO CLINICIANS:

We would like to invite you to be a part of a study into clinician judgement of intellectual function. The aims of the study are to investigate the accuracy of clinical judgement using demographic data and a clinical interview alone. It is important to know the accuracy of clinical judgement in relation to IQ estimation, as clinical judgement is the only method available for premorbid IQ estimation when working with non-english speaking populations. At these times, clinicians can only rely on demographic information provided and their clinical judgement following an interview.

If you agree to participate, you will be sent a video, which contains a short clinical interview of 15 volunteers, and a scoring sheet, which contains each cases demographic variables. You will be required to make an IQ estimate for each of the cases based on the information provided and your clinical judgement. The cases have been administered the WAIS-III, so their actual IQ scores are known by the investigators. Your estimates will be compared to the actual scores to see how accurate they were.

The final report will compare expert and novice neuropsychologists judgements, clinical judgement to an actuarial formula, and clinical judgement and the results of an actuarial formula to the obtained IQ scores. As such no individuals or individual results will be identifyable and only group data will be published.

A second part to this study involves an interview with you regarding the processes involved in clinical judgement. Should you wish to take part in this interview, (and you have agreed to by filling in the form "I am available for an interview") you will be contacted at a later stage. You do not have to agree to participate in the second part of the study if you do not wish to do so. If you agree to participate in the interview, no identifying details from the interview will appear in the report

CERTIFICATION BY PARTICIPANT

I, ______________________________________

of ____

certify that I am at least 17 years old and that I am voluntarily giving my consent to participate in the study entitled: Clinician Judgement of Intellectual Function being conducted at Victoria University of Technology by Ms Michelle Lindsay and Ms Izabella Walters

I certify that the objectives of the study, together with any risks to me associated with the procedures listed below to be carried out in the study, have been fully explained to me by Michelle Lindsay,and that I freely consent to participation involving the use of these procedures.

Procedures:

Watch a video and estimate the IQ of each case.

If involved in part two of this study - participate in a short interview.

I certify that I have had the opportunity to have any questions answered and that I understand that I can withdraw from this study at any time and that this withdrawal will not jeopardise me in any way.

I have been informed that the information I provide will be kept confidential.

Signed: .... }

Witness other than the experimenter (as appropriate) } **Date: ....**

....}

Any queries about your participation in this project may be directed to the researcher (Ms Michelle Lindsay and Ms Izabela Walters on 9365 2778). If you have any queries or complaints about the way you have been treated, you may contact the Secretary, University Human Research Ethics Committee, Victoria University of Technology, PO Box 14428 MCMC, Melbourne, 8001 (telephone no: 03-9688 4710).

APPENDIX C
Clinician Rater's Scoring Booklet

CLINICIAN JUDGEMENT OF INTELLECTUAL FUNCTION

VICTORIA UNIVERSITY OF TECHNOLOGY
ST ALBANS CAMPUS
MELBOURNE

CLINICIAN SCORING SHEET

Please note that this sample is not representative of a normal population

Victoria University of Technology
Neuropsychologist Demographic Information

Please Complete the following information:

Age __________
Highest Qualification ____
Number of Years Practicing as a Neuropsychologist Full Time ______________
Number of Years Working Part Time
(Please also indicate as a percentage the amount of time spent ie if worked 5 years part time 50% was 2 days per week, 20% was 4 days per week and 30% was 1 day per week)

Days Per Week	Number of Years
______________________	______________________
______________________	______________________
______________________	______________________
______________________	______________________

Total number of years in which you worked part time________________

Predominant Nature of Neuropsychological Work (Please Circle)

Rehabilitation	Children
Diagnostic	Adults
Elderly	Student Neuropsychologist

Predominant Setting in Which Neuropsychological Work Undertaken (e.g. Hospital, Rehabilitation Centre)____
____

No. Of Hours Spent in Clinical Assessment Per Week __________________

Amount of time WAIS used in practice (please circle)

almost never
once per month
once per week
2 to 5 times per week
more than 5 times per week

Scoring Sheet

1. Case 1 Demographic Information

Age	21 years old
Gender	Female
Occupation	Food & Beveridge Attendant
Level of Education	***Yr 12 and was accepted into an Arts Undergraduate degree, however deferred and has never undertook it.***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average

High Average Superior Very Superior

IQ Estimation Exact Number ______________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 2 Demographic Information

Age	35 years old
Gender	Female
Occupation	Administration Manager
Level of Education	***Year 12 and completed 2 years of an undergraduate degree***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle)   Impaired   Borderline   Low Average   Average

High Average   Superior   Very Superior

IQ Estimation Exact Number ________________________

Confidence Rating for Estimate: (Please Circle)   Low   Moderate   High   Very High

Comments:____

1. Case 3 Demographic Information

Age	23 years old
Gender	Female
Occupation	Textile Designer
Level of Education	***Honours Degree RMIT in Textile Design***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average

High Average Superior Very Superior

IQ Estimation Exact Number ______________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments: ____

1. Case 4 Demographic Information

Age	25 years old
Gender	Female
Occupation	Retail Assistant
Level of Education	***Undergraduate Degree in Textile Design at RMIT***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

1. Case 5 Demographic Information

Age	26 years old
Gender	Male
Occupation	Plumber
Level of Education	***Yr 12 and Plumbing Apprenticeship***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average

High Average Superior Very Superior

IQ Estimation Exact Number ________________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 6 Demographic Information

Age	28 years old
Gender	Male
Occupation	Field Gaming Specialist
Level of Education	***Yr 12 and Boat Building Apprenticeship***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average

High Average Superior Very Superior

IQ Estimation Exact Number ______________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 7 Demographic Information

Age	29 years old
Gender	Male
Occupation	Electrician
Level of Education	***Yr 11 and a Trade Apprenticeship***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average

High Average Superior Very Superior

IQ Estimation Exact Number ________________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 8 Demographic Information

Age	49 years
Gender	Male
Occupation	Computer EDI Manager
Level of Education	***Honours Degree***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average

High Average Superior Very Superior

IQ Estimation Exact Number ________________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 9 Demographic Information

Age	49 years old
Gender	Female
Occupation	Data Manager
Level of Education	***Yr 12 and Applied Microbiology Diploma***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average

High Average Superior Very Superior

IQ Estimation Exact Number ________________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 10 Demographic Information

Age	27 years old
Gender	Female
Occupation	Assistant Manager Hospitality
Level of Education	***Yr12 and Completed a 3 year Marketing Diploma at Monash University***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 11 Demographic Information

Age	51 years old
Gender	Male
Occupation	Self –Employed (Car Importer)
Level of Education	***Yr 11 and Fitter and Turner Apprenticeship***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 12 Demographic Information

Age	37 years old
Gender	Female
Occupation	Consultant in Emergency Medicine
Level of Education	***Postgraduate Medical Fellowships x 2***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ______________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 13 Demographic Information

Age	53 years old
Gender	Male
Occupation	Technical Support
Level of Education	***Yr 10 and Carpentry Apprenticeship***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ________________________

Confidence Rating for Estimate: (Please Circle) — Low Moderate High Very High

Comments:____

1. Case 14 Demographic Information

Age	28 years old
Gender	Male
Occupation	Plastics Injection Operator
Level of Education	***Yr 12 and Hospitality Certificate (Not Completed)***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ________________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

1. Case 15 Demographic Information

Age	26 years old
Gender	Female
Occupation	Automotive Designer
Level of Education	***Completed undergraduate degree in Textile Design at RMIT. Currently undertaking degree in Multimedia and Interactive design part time (2 year degree) whilst working full time***

2. After viewing the interview and noting the demographic information provided, please record your IQ estimation (both range and an exact estimation).

IQ Descriptive Range (Please Circle) Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ______________________________

Confidence Rating for Estimate: (Please Circle) Low Moderate High Very High

Comments:____

Thank you for your participation in this study. Please return the videotape, along with your completed scoring sheet to the investigators in the reply paid envelope provided.

APPENDIX D

CLINICIAN JUDGEMENT OF INTELLECTUAL FUNCTION

VICTORIA UNIVERSITY OF TECHNOLOGY
ST ALBANS CAMPUS
MELBOURNE

Please Complete the following information:

Age ________
Highest Qualification ____________

Number of Years Practicing as a Neuropsychologist ____________

Predominant Nature of Neuropsychological Work (Please Circle)

Rehabilitation	Children
Diagnostic	Adults
Elderly	Student Neuropsychologist

On the basis of the following demographic information, please estimate an IQ for each of the following hypothetical cases:

Case 1
Age: **21 years old**
Gender: **Female**
Occupation: **Hospitality/ Bar/ Waitress**
Level of Education: ***Yr 12***

IQ Descriptive Range (Please Circle)
Impaired Borderline Low Average Average
High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: Low Moderate High Very High

Case 2
Age **35 years old**
Gender **Female**
Occupation **Secretarial**
Level of Education **Year 12**

IQ Descriptive Range (Please Circle)
Impaired Borderline Low Average Average
High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: Low Moderate High Very High

Case 3

Age	**23 years old**
Gender	**Female**
Occupation	**Designer**
Level of Education	**Honours Degree Arts**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: *Low* *Moderate* *High* *Very High*

Case 4

Age	**25 years old**
Gender	**Female**
Occupation	**Retail**
Level of Education	**Undergraduate Degree Arts**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: *Low* *Moderate* *High* *Very High*

Case 5

Age	**26 years old**
Gender	**Male**
Occupation	**Trade**
Level of Education	**Yr 12 and Apprenticeship**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: *Low* *Moderate* *High* *Very High*

Case 6

Age	**28 years old**
Gender	**Male**
Occupation	**Hospitality Management**
Level of Education	**Yr 12 and Apprenticeship**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: Low Moderate High Very High

Case 7

Age	**29 years old**
Gender	**Male**
Occupation	**Trade**
Level of Education	**Yr 11 and Apprenticeship**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ______________________

Confidence Rating for Estimate: *Low* *Moderate* *High* *Very High*

Case 8

Age	**49 years**
Gender	**Male**
Occupation	**I.T.**
Level of Education	Honours Degree

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ______________________

Confidence Rating for Estimate: *Low* *Moderate* *High* *Very High*

Case 9

Age	**49 years old**
Gender	**Female**
Occupation	**Administration Manager / Project Officer**
Level of Education	**Yr 12 and Science Graduate Diploma**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ______________________

Confidence Rating for Estimate: *Low* *Moderate* *High* *Very High*

Case 10

Age	**27 years old**
Gender	**Female**
Occupation	**Assistant Manager Hospitality**
Level of Education	**Yr12 and Marketing Degree**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ______________________

Confidence Rating for Estimate: *Low* *Moderate* *High* *Very High*

Case 11

Age	**51 years old**
Gender	**Male**
Occupation	**Sales**
Level of Education	**Yr 11 and Apprenticeship**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: Low Moderate High Very High

Case 12

Age	**37 years old**
Gender	**Female**
Occupation	**Doctor**
Level of Education	**Medical Fellowship**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: Low Moderate High Very High

Case 13

Age	**53 years old**
Gender	**Male**
Occupation	**Trade**
Level of Education	**Yr 10 and Apprenticeship**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: Low Moderate High Very High

Case 14

Age:	**28 years old**
Gender:	**Male**
Occupation:	**Factory Worker**
Level of Education:	**Yr 12**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ____________________

Confidence Rating for Estimate: Low Moderate High Very High

<u>**Case 15**</u>

Age	**26 years old**
Gender	**Female**
Occupation	**Designer**
Level of Education	**Arts Degree**

IQ Descriptive Range (Please Circle) — Impaired Borderline Low Average Average High Average Superior Very Superior

IQ Estimation Exact Number ________________________

Confidence Rating for Estimate: Low Moderate High Very High

THANK-YOU FOR YOUR PARTICIPATION

APPENDIX E

Methods Of IQ Estimation Questionnaire

Clinician Demographic Information

(If filling in by email, for all items that require you to tick a box, please place an X in the box by clicking on the box. ***Please SAVE the file after completion so that you can return via email*****)**

1.Sex **M** ☐ **F** ☐

2.Years Working as a Qualified Neuropsychologist **0** ☐ **1-5** ☐ **5-10** ☐ **>10** ☐ **>20** ☐

3. Highest Qualification (Please Circle) Graduate Diploma ☐ Honours ☐ Masters ☐ Doctorate ☐ PhD ☐

4. Area of Work Student ☐ Rehabilitation ☐ Acute ☐ Pediatric ☐ Elderly ☐ Medicolegal ☐

5. Number of Times WAIS used in 1 week: A. Never ☐ B. 1-5 ☐ C. 5-10 ☐ D. >10 ☐

6. Do subscribe to Neuropsychological Journals A. Yes ☐ B. No ☐

PREMORBID ESTIMATION QUESTIONAIRE

1. How often do you estimate a premorbid IQ for an individual you are assessing?

A. Always (100%) ☐ B. Often (80%) ☐ C. Sometimes (60%) ☐
D. Occasionally (40%) ☐ E. Rarely (20%) ☐ F. Never (0%) ☐

2. How often is this estimate made explicit in your reports?

A. Always (100%) ☐ B. Often (80%) ☐ C. Sometimes (60%) ☐
D. Occasionally (40%) ☐ E. Rarely (20%) ☐ F. Never (0%) ☐

3. In what circumstances are you likely to estimate a premorbid IQ?

A. For medicolegal reports only or ones in which it is likely to go to court ☐
B. For assessments when there is definitely a known neurological impairment (ie after stroke or head injury) ☐
C. Every assessment ☐
D. Never ☐

4. What methods do you use routinely to estimate premorbid IQ? (You can choose more than 1)

A. NART ☐
B. Best 3 performance ☐
C. Barona Estimate ☐
D. Crawford Index ☐
F. Interview ☐
G. Wide Range Achievement Test ☐
H. Vocational information ☐
I. School records ☐

E. Hold subtests of the WAIS ☐

5. Which methods of premorbid IQ estimation have you received training in (ie as part of graduate training)

A. NART / Schonell / Reading Test ☐
B. Best 3 performance ☐
C. Barona Estimate ☐
D. Lezak Best Performance Method ☐
E. Other (please specify) ____________
E. Hold subtests of the WAIS ☐
F. Wide Range Achievement Test ☐
G. Demographic variables ☐
H. Crawford Index ☐

6. If you use demographic variables to estimate IQ, which ones do you use?

A. Age ☐
B. Gender ☐
C. Level of Education ☐
D. Type of Occupation ☐
E. Other (please specify)____________________
F. None of the above (ie I don't rely on demographics) ☐

8. Please indicate whether or not you believe that your training as a neuropsychologist provided sufficient instruction in methods to estimate premorbid IQ.

A. Strongly Agree ☐ **B. Agree** ☐ **C. Neutral** ☐ **D. Disagree** ☐ **E. Strongly Disagree** ☐

9. Do you routinely estimate other premorbid areas of cognitive functioning other than IQ (ie memory, motor skills)

A. Always (100%) ☐ B. Often (80%) ☐ C. Sometimes (60%) ☐
D. Occasionally (40%) ☐ E. Rarely (20%) ☐ F. Never (0%) ☐

10. Which other areas do you routinely estimate premorbid abilities in?

A. Memory ☐
B. Visuospatial ☐
C. Motor Skills ☐
D. Executive Function ☐
E. Other ☐
F. None ☐

11. Please indicate the percentage of times you would use each method to assess premorbid IQ.

A. NART	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
B. School records	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
C. Best 3 performance	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
D. Vocational info	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
E. Barona Estimate	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
F. Interview	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
G. Crawford Index	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
H. WRAT	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
I. Hold subtests WAIS	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐

12. Please specify the percentage weight given to each demographic variable in determining premorbid IQ.

A. Age	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
B. Occupation	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
C. Gender	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
D. Other	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
E. Level of Education	10% ☐	20% ☐	30% ☐	40% ☐	50% ☐	60% ☐	70% ☐	80% ☐	90% ☐	100% ☐
F. None of above	☐									

(don't rely on demographics)

Please return this completed Questionnaire to the researcher via email at: mlindsay_1@hotmail.com **or by mail to 63 Humphries Rd Frankston South 3199 as soon as possible. If you would like to add any other comments, please feel free to do so.**

THANK YOU FOR YOUR PARTICIPATION.