University of Rhode Island

DigitalCommons@URI

Open Access Master's Theses

2007

Impact of Sociodemographic Adjustments on Test Score Patterns

David C. Ahern

University of Rhode Island

Follow this and additional works at: <https://digitalcommons.uri.edu/theses>

Recommended Citation

Ahern, David C., "Impact of Sociodemographic Adjustments on Test Score Patterns" (2007). *Open Access Master's Theses*. Paper 944.

<https://digitalcommons.uri.edu/theses/944>

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons-group@uri.edu.

**IMPACT OF SOCIODEMOGRAPHIC
ADJUSTMENTS ON TEST SCORE PATTERNS**

BY

DAVID C. AHERN

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ARTS
IN
PSYCHOLOGY**

UNIVERSITY OF RHODE ISLAND

2007

APPROVAL PAGE

MASTER OF ARTS THESIS

OF

DAVID C. AHERN

APPROVED:

Thesis Committee:

Major Professor David Gant

U. Hill

[Signature]

Harold Stolt

DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2007

Abstract

Sociodemographic adjustment of neuropsychological test scores is widely embraced. However, limited research has examined the frequency with which such adjustments alter score profiles, psychologists' decision-making, or diagnostic accuracy. Of particular interest, in an unknown number of cases test scores that are and are not adjusted may yield conflicting interpretations, both of which cannot be correct. Working with the *WAIS-III* and *WMS-III*, this study examined how often, and the extent to which, sociodemographic adjustment altered test score patterns among normal and abnormal groups. Potential impact on classification and diagnostic accuracy was also examined by determining the frequency with which sociodemographically adjusted profiles best matched their original diagnostic prototypes as opposed to other prototypes, and with which abnormal and normal profiles shifted to more closely resemble the opposing classification. Analysis showed that sociodemographic adjustment frequently produced substantial change in profiles for cases based on the prototype for traumatic brain injury, with these same findings replicated for cases of alcohol abuse, as well as for various normal profiles. Further, matches with original prototypes were often altered such that in a large percentage of cases, profiles from specific pathologic conditions formed better matches with different conditions. Both normal and abnormal profiles additionally switched classes with modest frequency. This research demonstrates that sociodemographic adjustment often has a considerable affect on test score patterns, and hence research on the ultimate impact of such adjustments on clinical judgment and diagnostic accuracy is urgently needed.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Tables	iv
List of Figures	v
Introduction	1
Problems with Interpreting Test Data by Means of Pattern Analysis	2
Hypotheses	12
Instruments	15
Pilot Study	18
Main Study: Overview	23
Part I: Are Abnormal Patterns Altered by Sociodemographic Adjustment?	24
Part II: Are Normal Patterns Altered by Sociodemographic Adjustment?	32
Part III: Do Adjustments Alter Matches Within Abnormal Prototypes and Across Normal and Abnormal Prototypes?	37
Summary of Results	42
Discussion	44
Appendix A: Examples of <i>WAIS-III</i> Pattern Reversals	68
Bibliography	78

List of Tables

Table 1: <i>Features of the WAIS-III – WMS-III – WIAT-II Writer</i>	50
Table 2: <i>Rules for Classifying Pattern and Elevation</i>	51
Table 3: <i>Characteristics and Test Scores of Disorder Prototypes</i>	52
Table 4: <i>Sociodemographic Data for Alcohol Abuse Profiles</i>	53
Table 5: <i>Sociodemographic Data for TBI Profiles</i>	54
Table 6: <i>Generated Normal Profiles</i>	55
Table 7: <i>Rate of Profile Match Across Multiple Prototypes</i>	56

List of Figures

Figure 1: <i>Chronic Alcohol Abuse – Modal Profile: WAIS-III</i>	57
Figure 2: <i>Chronic Alcohol Abuse – Modal Profile: WMS-III</i>	58
Figure 3: <i>Classifications of Adjusted Chronic Alcohol Abuse Profiles: WAIS-III</i>	59
Figure 4: <i>Classifications of Adjusted Chronic Alcohol Abuse Profiles: WMS-III</i>	60
Figure 5: <i>TBI – Modal Profile: WAIS-III</i>	61
Figure 6: <i>TBI – Modal Profile: WMS-III</i>	62
Figure 7: <i>Classifications of Adjusted TBI Profiles: WAIS-III</i>	63
Figure 8: <i>Classifications of Adjusted TBI Profiles: WMS-III</i>	64
Figure 9: <i>Classification of Normal Profiles</i>	65
Figure 10: <i>Normal Prototype A</i>	66
Figure 11: <i>Normal Prototype B</i>	67
Figure 12: <i>Pattern Alteration: Example A</i>	68
Figure 13: <i>Pattern Alteration: Example B</i>	69
Figure 14: <i>Pattern Alteration: Example C</i>	70
Figure 15: <i>Pattern Alteration: Example D</i>	71
Figure 16: <i>Pattern Alteration: Example E</i>	72
Figure 17: <i>Pattern Alteration: Example F</i>	73
Figure 19: <i>Pattern Alteration: Example G</i>	74
Figure 19: <i>Pattern Alteration: Example H</i>	75
Figure 20: <i>Pattern Alteration: Example I</i>	76
Figure 21: <i>Pattern Alteration: Example J</i>	77

Introduction

Although sociodemographic adjustment of psychological and neuropsychological test scores is widely embraced, limited research has examined the frequency with which such adjustments alter test score profiles or patterns and, consequently, psychologists' diagnostic decision-making. When sociodemographic adjustment is performed, the magnitude of score alteration may vary both across demographic groups and within the individual. As such, certain measures or scores may change more than others. With particular tests or subtests changing more or considerably more than others in the same diagnostic assessment, alteration in test score pattern is inevitable. One potential result is to distort normal performances to appear clinically abnormal, and vice versa. Put another way, profiles may be made to appear spuriously normal or spuriously abnormal by virtue of the sociodemographic adjustments alone. When the same set of test scores produce conflicting results dependant on the use of – or type of – sociodemographic adjustment, one outcome will by necessity be incorrect.

This research explored the frequency, magnitude, and potential impact of sociodemographic adjustment on normal and abnormal test profiles in certain clinical syndromes using data from the *Wechsler Adult Intelligence Scale, Third Edition (WAIS-III; Wechsler, 1997a)*, and the *Wechsler Memory Scale, Third Edition (WMS-III; Wechsler, 1997b)*. Adjusted profiles were classified to determine the degree of change, thereby offering some insight into the potential impact such score-profile alteration may have on clinical decision making.

Problems with Interpreting Test Data by Means of Pattern Analysis

Pattern analysis, also interchangeably referred to as profile analysis, is a process for interpreting multiple scores from a test battery or the various indices of a single test by examining the relative strengths and weaknesses in performance across the various measures or scores. This process furthermore involves comparing the score configuration to the prototypic or modal performance of other individuals with the cognitive disorders in question. A central concern in pattern analysis is variability or scatter among scores, which takes two forms: intertest scatter, or variability among a set of tests, and intratest scatter, or variability between the set of subscores or indices within a single measure (Lezak, Howieson, Loring, Hannay, & Fischer, 2004). Beyond a comparison of relative highs and lows, the process of profile analysis is concerned with the interrelations or configurations among the various scores. This method of test interpretation is often described as a central tenant of psychological assessment, and is espoused particularly as a tool for differential diagnosis. Pattern analysis has long been promoted (e.g., Gaier & Lee, 1953) and advocates are present across various psychological domains, from personality assessment (e.g., Exner & Erdberg, 2005; Graham, 2006) to neuropsychological evaluation (e.g., Lezak et al., 2004, Zakzanis, Kaplan, & Leach, 1999).

Neuropsychological literature often sets forth test profiles and patterns purportedly associated with a myriad of diagnoses. Even a brief literature search yielded descriptions of patterns espoused to characterize such diverse conditions as Alzheimer's Disease (Reed et al., 2007), AIDS (Cysique, Maruff, & Brew, 2006), vascular cognitive impairment (Nyenhuis et al., 2004; Sachdev et al., 2004), schizotypal personality disorder

(Voglmaier et al., 2005), bipolar disorder (Torres, Boudreau, & Yatham, 2007; Martínez-Arán et al., 2004), ADHD (Hervey, Epstein, & Curry, 2004), major depression (Porter, Gallagher, Thompson, & Young, 2003), borderline personality disorder (Dinn et al., 2004), mercury exposure (Rohling & Demakis, 2006), recurrent headaches (O'Bryant, Marcus, Rains, & Penzien, 2006), Williams syndrome (Bellugi, Lichtenberger, Jones, & Lai, 2006), and traumatic brain injury (Salmond & Sahakian, 2005). It should be evident that pattern or profile analysis is a diagnostic tool that clinicians regularly implement or describe. Lezak et al. (2004) indicated that appraisal of test score patterns is perhaps the most common technique for the psychological assessment of brain disorders, and they suggested that a professional practitioner can integrate a multitude of test scores with proficiency and effectiveness.

There is, however, a strong analytic and scientific basis to question the reliability and validity of complex configural analysis (e.g., Faust, 2003). Although simple comparisons between small numbers of variables may be relatively straightforward, limits in human information processing capacities may thwart or stymie more complex subjective analyses. Humans are susceptible to a number of problematic decision-making practices and limitations, including a failure to properly utilize base rates, suboptimal ability to assign weight or importance among competing variables, and a number of cognitive biases. Such biases include, for example, confirmatory bias (i.e., a tendency to search for information confirming preconceived beliefs and to disregard contradictory evidence), hindsight bias (i.e., the tendency to believe events are more predictable in retrospect than they in fact were prior to occurrence), and the availability heuristic (i.e., the propensity to be overly influenced in judgments of frequency by the ease with which

instances of occurrence can be brought to mind). For an overview of these and other cognitive biases see Kahneman, Slovic, & Tversky (1982).

Furthermore, extensive literature shows that across a wide variety of situations, statistical prediction almost always equals or exceeds the accuracy of clinician judgment, largely by virtue of increased consistency and a reduced vulnerability to a variety of the pitfalls to which clinicians are susceptible (e.g., Dawes, Faust, & Meehl, 1989; Grove, Zald, Lebow, Snitz, & Nelson, 2000; Egisdóttir et al., 2006). These findings are supported within the neuropsychological domain as well (e.g., Wedding & Faust, 1989). Such findings, coupled with the realization that scores on a test or battery of tests may often be reproduced and interpreted as simple linear composites, makes suspect the rationale for purportedly complex clinician interpretation of patterns.

Profile analysis is also susceptible to increased error by virtue of being a joint product of multiple measures, each with a distinct error rate. All else being equal, the reliability of a test score pattern is lower or much lower than the respective isolated error rates of the component measures from which the pattern is comprised. When interpretation depends on interrelations between multiple test scores, the reliability of those interrelations or patterns is far lower than the reliabilities of the component parts because of interdependence (Anastasi & Urbina, 1996). Each subtest within a larger battery likewise has a unique level of associated error or specificity, the degree of which may be inconsistent among subtests. A common hazard of clinical interpretation is the excessive belief in the ability of the clinician to reduce error through complex configural analysis. However, overly complex strategies, in part by depending on results with inadequate reliabilities, may often *increase* error in comparison to reliance on actuarial

methods, even relatively simple actuarial methods that accept some level of error in order to make less error (Einhorn, 1986).

In part because of this increased error, as the number of tests included in a battery rises so does the appearance of artifactual scatter, or relative highs and lows in test data due primarily to inconsistent error rates and low inter-measure correlations. Research further indicates that, all else being equal, as the number of tests in a battery increases so too does the frequency with which a profile is judged to be abnormal (Mitrushina, Boone, Razani, & D'Elia, 2005). Whereas this could merely reflect the increased number of false-positive results that are almost inevitable as more tests are used, this finding may also be attributable to human misperception of an acceptable level of deviant findings. It is normal to be abnormal to a point, with some research (Heaton, Miller, Taylor, & Grant, 2004) indicating about 87% of normal individuals demonstrate at least one abnormal score on a lengthy battery, with an average rate of 15% abnormal performances when cutoffs are set at common levels.

A number of additional factors likewise introduce error into a pattern or profile, potentially decreasing accuracy and confounding interpretation (Faust, Bridges, & Ahern, in preparation). For example, unequal Flynn effects across measures, that is, non-uniformity in the rates at which the norms across tests and subtests become obsolete, can distort the pattern of test scores (Strauss, Sherman, & Spreen, 2006). More specifically, if scores across or within tests become obsolete at different rates, differences in obtained test scores may reflect varying distortions in normative standards for test performance versus differential ability levels (Hiscock, 2007). A test battery may contain measures normed up to decades apart, for instance. Ironically, although rarely mentioned, rate of

obsolescence also can vary within components of a test, thereby also producing artificial scatter. For example, the Digit Symbol – Coding subtest of the Wechsler intelligence scales becomes obsolete much more rapidly than a subtest such as Vocabulary (Faust, Bridges, & Ahern, in preparation).

Along similar lines, the comparison of tests with dramatically different normative groups may impact profile interpretation to a great degree (Mitrushina et al., 2005). It is not atypical for comparisons to be drawn between measures with norms of inconsistent sample size, mean, variability, and demographic constituency, (e.g., convenience samples with very high ability levels versus other samples of average abilities). Psychological instruments may also have inconsistent floor or ceiling effects, truncated ranges, differing reliabilities, or mixed evidence of validity, all of which have the potential to distort test score patterns (Faust, Ziskin, & Hiers, 1991).

Sociodemographic Adjustments and Pattern Analysis

To the extent obtained patterns are artifactual rather than true reflections of a patient's status, clinical interpretation that emphasizes pattern can be hindered or enfeebled. Among the various factors that can alter or distort pattern analysis and that deserve thorough exploration, such as those described above, one that appears to be of particular importance is the increasingly common practice of sociodemographic adjustment.

Determining the pattern of test scores that various psychological conditions produce often depends, in principle, on some knowledge of an individual's baseline or pre-event functioning. For example, a low average score of 88 on the *WAIS-III* Processing Speed Index (a composite of subtests measuring the ability to rapidly manage

information) may represent a sharp and troublesome decline in an individual previously performing in the high average range, and it may be indicative of a certain disease process. This score may represent no change at all, however, if prior functioning was within the same low average range of ability. Thus, even if various conditions or diseases have similar impact on cognitive functioning, given individual differences in overall level of functioning and in relative level of functioning across areas, the presence of such conditions often cannot be properly assessed without reasonable knowledge of an individual's baseline.

Attempts to take pre-existing functioning or baseline into account has been a major motivation in creating increasingly complex systems of sociodemographic adjustment in test scores. These processes are often intended to serve as a surrogate baseline, or rough approximation of prior level of performance, as direct or true baseline information is often unavailable. Take, for example, scores on a measure of grip strength. Scores in the extreme high or low ranges may be interpretable regardless, but what of a more ambiguous performance falling closer to the middle of the bell curve? Research has indicated a reasonably strong relationship between this variable and certain demographic characteristics such as age and gender. Therefore, a performance that may be well within normal limits and carry no pathological implications for a 50-year-old female may indicate an abnormality, such as the early onset of a muscular degenerative disorder, for an otherwise healthy 25-year-old male.

In this context, adjustment in relation to certain variables with known associations to the function in question narrows the normative group to one more likely representative of the patient's true baseline. For example, if males typically obtain considerably higher

scores on measures of grip strength, the patient is male, and decline in functioning is at issue, then comparing his performance to other males – versus a combined group of males and females – will probably serve as a better indicator of deviance from his true baseline than would otherwise be the case. In this instance, where the question is one of decline, the clinician does not want to know how the patient compares to others but rather how he or she compares to him- or herself in a disease or disorder-free state. There are many situations where decline is the key issue in neuropsychological assessment, as might be the case, for example, when evaluating for the presence of dementia or the sequelae of a traumatic brain injury.

Sociodemographic adjustment of psychological test scores by multiple variables as a substitute for knowledge of baseline functioning is a relatively recent phenomenon, but one that has become nearly ubiquitous in neuropsychology. Certain tests, such as those designed to assess academic ability, intelligence, or memory, have long been adjusted in relation to the age of the examinee. Pediatric psychology and neuropsychology aside¹, the rationale for age adjustment in initial measures was primarily to prevent older individuals from being misclassified as having abnormal test performances on dimensions greatly impacted by age (e.g., processing speed) when compared to the mean scores of much younger adults (Heaton, Taylor, & Manly, 2003).

Performance on cognitive testing is often associated with multiple demographic characteristics, including level of education, gender, and ethnicity. A significant amount of variability in test scores may be accounted for by even one demographic variable, exceeding 40% in some cases (Heaton, Grant, & Matthews, 1992). Furthermore, several

¹ Note that pediatric assessment and related sociodemographic adjustment carries a unique set of problems that are beyond the scope of this study, and thus will not be addressed here despite the considerable importance of these issues.

demographic variables in combination may account for the majority of test score variance. For example, Heaton et al. (2003) found significant relations between age, education, sex, and ethnicity on intelligence and memory testing. Heaton et al. (1992) developed one of the first comprehensive systems for simultaneous sociodemographic adjustment among multiple variables for a large test battery. Heaton et al. (1992; 2003; see also 1986) likewise demonstrated that diagnostic accuracy of neuropsychological test scores can be enhanced when a normative database is subdivided by sociodemographic variables. These demonstrations did not include complex configural analysis but rather consideration of scores singly or in one-dimensional combinations, such as the cumulative frequency of false-positive error with or without sociodemographic adjustment.

Despite the broad-based practice of adjusting neuropsychological test scores by sociodemographic variables and isolated demonstrations of improved accuracy for certain narrow interpretive practices, there is a paucity of research examining the impact of such adjustments on other common interpretations tasks and diagnostic activities. For example, there has been minimal study on the impact of sociodemographic adjustment on the type of complex pattern that is pervasively advised in the neuropsychological literature, nor on differential diagnosis. Note that differential diagnosis in neuropsychology often is critical in directing treatment efforts, and that intervention choices may influence outcome considerably. Furthermore, sociodemographic adjustment may be performed, or available, on some tests but not on others. Thus we have the odd combination of circumstances in which sociodemographic adjustment may

or may not be performed, but in which there is almost no research to guide practice, leading to inconsistent clinical procedures.

The standard scoring for various tests and test batteries, in a case like the traditional Halstead-Reitan Battery (Reitan & Wolfson, 1993), does not adjust for any sociodemographic variables or, in the case of the *WAIS-III* and *WMS-III*, adjusts for age alone using relatively broad groupings (e.g., 25 years – 34 years). For these and many other such tests, the clinician has the option of performing additional or finer adjustments, such as using narrower age bands and also adding adjustments for education and gender. The adjustments may come in packaged combination such that multiple variables are incorporated without the option for examining fewer or isolated variables, such as adjusting for education alone. Thus, within certain restraints, the practicing neuropsychologist often has an array of options that could dramatically alter outcome but have been minimally studied for impact on diagnostic judgments or accuracy.

It is a virtual certainty that sociodemographic adjustment impacts not only the level of individual scores, but also generates change in the relative level of scores in relation to one another due to unequal use of adjustments across all measures or subtests and unequal impact on individual test scores. For example, Test A may allow for adjustment by age alone, Test B for adjustments by age and education, and Test C for adjustments by age and gender. As a further example, if one is adjusting for a variable such as gender, scores on a measure of grip strength may change a great deal whereas scores on a measure of vocabulary may change minimally, if at all, thereby frequently altering the relative standing or “pattern” between the two scores. A patient who is assessed twice and who has not changed whatsoever may consequently obtain substantial

inconsistency in test scores due predominantly to contrasting clinical practices in the use of sociodemographic adjustment.

The very real possibility that decisions to perform or not perform sociodemographic adjustment can markedly alter profiles and clinical decisions is a very serious one. Many clinical decisions involve dichotomous or mutually exclusive choices. For example, one may need to decide if the patient is brain damaged or not, or has either Condition A or Condition B. With dichotomous or mutually exclusive decisions, competing or inconsistent answers cannot both be correct, and sometimes both will be wrong. The diametric nature of some decisions preclude combining the outcomes when they differ: an applicant cannot be both hired and rejected, and a tumor cannot be malignant and benign. As there are many potential instances in which results are opposed, cannot be integrated, and a decision is necessary, it is important to determine how often adjustment of a score pattern impacts clinician judgment. Thus, in a larger context of concerns regarding the theoretical rationale, feasibility, and success of complex pattern analysis in clinical assessment, this study was designed to explore one such pressing issue.

Hypotheses

Considerable research has been conducted on the impact of sociodemographic variables on clinical judgment. For example, Perlick and Atkins (1984) found that clinicians were more likely to diagnose depression for younger individuals and dementia for older individuals, even when subjects obtained the same underlying score profile (which could well be rational given the differences in base rates for the different age groups). This present study differs substantially, however, from such types of studies, and in some sense involves the impact and success of taking sociodemographic variables into account in a more objective, quantitative manner. This research examines the degree of profile change based on adjustment for multiple sociodemographic variables, a topic which, to the author's knowledge, has not been previously investigated.

This study explored one primary research question: Do sociodemographic adjustments result in a substantial alteration in test score patterns with sufficient frequency to be of potential clinical relevance? It is evident that sociodemographic adjustments can and do alter and potentially distort test score patterns, but the rate with which such changes occur remains unclear. Three basic sources of information were used to formulate specific research hypotheses. These sources included, first and foremost, consideration of the pertinent literature already discussed. Second, by virtue of work in a consulting practice reviewing a diversity of psychological and neuropsychological evaluations, it has been possible to examine the impact of sociodemographic adjustments on a fairly wide range of tests and test batteries. Third, exploratory analysis was performed on programs for sociodemographic adjustments, in particular those created for the *WAIS-III* and *WMS-III*. This analysis, which included

examinations of multiple hypothetical test protocols, strongly suggested that sociodemographic adjustments often have considerable impact.

Based on these sources of information and analyses, three hypotheses were formulated: two primary and one exploratory, which can be stated here and further explicated below. These include:

H1: A nontrivial² percentage of abnormal profiles will change considerably when sociodemographic adjustment is performed; that is, abnormal profiles will alter considerably from their respective prototypes by virtue of such adjustments alone.

H2: Likewise, a nontrivial percentage of normal profiles will change considerably from their respective prototypes when sociodemographic adjustment is performed.

H3 (exploratory): Profiles from one disorder or condition will come to equally or more closely resemble those of other disorders or conditions in a nontrivial percentage of cases when sociodemographic adjustments are performed. As such, a certain number of normal profiles will be made to appear abnormal and *vice versa*.

Currently the extent to which resultant alteration in test scores and test score patterns influences clinical judgment and decision making is uncertain. If clinical judgments and decisions are altered by the adjustments, then the nature of these alterations and the impact on judgmental accuracy are of great interest. However, it makes no sense to study judgment unless it is determined that significant alteration of the pattern is a non-trivial occurrence. Significant alteration in sociodemographically

² Although a term like "nontrivial" is ambiguous, it was anticipated the level and frequency of change would produce an overwhelming or non-ambiguous outcome.

adjusted profiles could be anticipated to impact clinician judgment significantly, including shifts in judgment regarding the presence of abnormality, and regarding the level of match with modal normal and pathological profiles. As such, the occurrence and degree of profile alteration is an important research question as a first step in determining the impact on clinician judgment.

The materials that follow are divided into three sections, each of which corresponds with a main or exploratory hypothesis. The methodology and results of each part are detailed as subsections under the appropriate major heading. The rationale for, and composition of, the sample for each part is detailed in the respective methods section.

Instruments

Psychometric Properties of Measures to be Utilized

The *WAIS-III* and *WMS-III* were employed as these instruments are excellent measures with which to explore the present research questions due to their wide prevalence of use. According to Hogan (2005), the *WAIS-III* is the most commonly used psychological measure in the clinical and forensic domain, second in neuropsychology, and within the top five most frequently utilized instruments in counseling and school psychology. Further, the *WMS-III* is the third most commonly utilized measure in the forensic and neuropsychological domains, and the ninth most commonly utilized in the clinical realm. It is the most commonly used memory assessment tool across all domains.

These instruments also exemplify a number of positive psychometric properties. The standardization samples of the *WAIS-III* and *WMS-III* were large (*WAIS-III* $N = 2,450$; *WMS-III* $N = 1,032$) and stratified to be proportionally representative of several sociodemographic variables from 1995 census data. The measures were stratified by 13 age bands, and within each age band by: sex; race/ethnicity; education level; and geographic location. Average split-half reliability coefficients for the three *WAIS-III* IQ scores range from $r = 0.94$ to 0.98 , and from $r = 0.88$ to 0.96 for the four composite indices. Average split-half reliability coefficients range from $r = 0.82$ to 0.93 across the eight *WMS-III* composite scores (The Psychological Corporation, 2002a). All corrected IQ score and Index test-retest reliabilities on the *WAIS-III* are well within the desired $r \geq 0.80$ range espoused by Anastasi and Urbina (1996), ranging from a low of $r = 0.83$ to a high of $r = 0.97$. Corrected test-retest reliabilities for the *WMS-III* are also largely within the desired range. An abundance of research examining convergent and criterion-based

validity supports both measures as valid tools for cognitive assessment of normal individuals and individuals suffering from numerous and varied developmental, psychological, and neuropsychological difficulties (The Psychological Corporation, 2002a).

Additional Reasons for Selection of Measures

The *WAIS-III* and *WMS-III* further lend themselves to the study of sociodemographic adjustments to score profiles for two additional reasons. First, both tests utilize multiple subtests and present test users with multiple composite scores and indices from which a pattern or profile may be interpreted. The *WAIS-III* provides three IQ scores: Full Scale IQ (FSIQ, an overall composite representing global cognitive functioning), Verbal IQ (VIQ, a composite of six acquired knowledge and verbal reasoning tasks), and Performance IQ (PIQ, a composite of five visual-motor, fluid reasoning, and spatial orientation tasks). In addition, four indices are obtained: Verbal Comprehension (VCI, verbal reasoning and acquired knowledge), Processing Speed (PSI, rapid processing of visual information), Working Memory (WMI, attending to and processing information in memory), and Perceptual Organization (POI, fluid reasoning and visual-motor integration). The *WMS-III* provides eight indices each exploring a different aspect of memory: Auditory Immediate, Auditory Delayed, Visual Immediate, Visual Delayed, Immediate Memory (a compilation of Auditory Immediate and Visual Immediate memory), Auditory Recognition Delayed (memory for auditory information with some cues or reminders), General Memory (retention over a delay), and Working Memory. The IQ scores and indices across both tests use the same standard scoring ($M = 100$, $SD = 15$), thereby facilitating comparisons between results and profile analysis.

Second, sociodemographic adjustments are readily available for the *WAIS-III* and *WMS-III*. Heaton et al. (2004) provide a means by which to adjust *WAIS-III* and *WMS-III* scores for age, education, gender, and ethnicity as a component of a comprehensive neuropsychological battery. For this study, the *WAIS-III – WMS-III – WIAT-II Writer* program developed and published by The Psychological Corporation (2002b) was utilized. This program allowed generation of adjusted test profiles along the same dimensions as those of the Heaton et al. system. Raw subtest scores are entered into the program and then adjusted, with output of graphs or tables in standardized T-score ($M = 50$, $SD = 10$) format. Unfortunately, little information is available from The Psychological Corporation about the method for developing the normative database for this program or its psychometric qualities. Personal communication with an employee of The Psychological Corporation indicated that this system is based on data developed by Heaton and his colleagues. However, this assertion could not be independently verified and no published information from The Psychological Corporation that attributes these data to Heaton et al. could be located.

Pilot Study

The rationale behind pilot work was manifold. The feasibility of the design needed to be ascertained, as did the sufficiency of the selected modal profiles and the interpretability of graphical output. The investigator needed to acquire thorough knowledge about the *Writer* program and its operation (much of which was not readily available), including proper data entry, demographic adjustment, and the generation of graphical output such as charts and tables. The dimensions available for adjustment were identified, and the range of coding on each dimension was explored. Table 1 summarizes the findings as are detailed below, in particular the match or mismatch between the options for data entry on sociodemographic variables and the impact, if any, of alteration in standing on these variables.

Characteristics of the Writer Program

As can be seen in Table 1, the *Writer* program provides six options for ethnicity. Although each of these ethnicities can be selected from the program's pull-down menu, normative information exists only for "White not Hispanic Origin" and "African/African-American." The selection of other listed ethnicities, including "American Indian/Alaskan Native," "Hispanic," "Asian/Asian American," and "Pacific Islander," produced an error message and would not permit the program to run. This, of course, forces the clinician to select a non-matching sociodemographic group for various ethnicities in order to run the program and generate a demographically adjusted profile, as the program will not allow adjustment without all variables entered. Therefore, it is not possible to adjust by other distinct variables, such as gender or education.

Exploration work also demonstrated that for the education variable, the range of coding does not coincide with the divisions for sociodemographic adjustment. The program allows the selection of 1 to 20 years of education in one-year increments, but it was revealed through manipulation that some age ranges are demographic brackets. For example, 17 to 20 years of education appears to be one category, as variation within this range does not appear to further alter profiles, although this was not tested for all possible interaction effects.

Age is not selected from a menu but rather calculated by the program based on date of birth. The age groupings do not appear to coincide with those age bands listed in the *WAIS-III* and *WMS-III* administration manuals. That is, the age divisions found in the standard manuals are apparently not adhered to in the *Writer* program, as variation was noted from changing ages within brackets specified by the test manuals.

Gender appears to alter profiles along the processing speed domain, with the program generating higher scores for females than males given comparable original (unadjusted) scores. The impact of gender on other domains is unclear. The variable for handedness, while selectable, does not appear to produce any adjustments in test scores. Of interest, the system provides dimensional options that all create the appearance of influencing the score profile, when not all variables or ranges of variables impact outcome. No information from the publisher could be located as to what dimensions or levels influence the adjusted profiles, or about possible interactions among variables. It is possible, if not likely, that some users will falsely believe that adjustments are being made for some variables when this is not the case, or that certain subdivisions alter output (e.g., differences of one or two years in education) when no such adjustments are made.

Also, at present, one cannot determine if the absence of adjustment reflects research showing a lack of effect, or rather a lack of sufficient data to measure or adjust for effects.

Methods

The *WAIS-III – WMS-III Technical Manual* (The Psychological Corporation, 2002a) is a text designed to supplement the administration manuals for both measures. This manual describes the rationale for revision from each previous version, provides psychometric information on both tests, details the norming process, elaborates on uses of the measures and interpretive strategies, and relates modal profile information for multiple cognitive disorders. These disorders and conditions include mental retardation, hearing impairment, attention-deficit/hyperactivity disorder (ADHD), learning disabilities, chronic alcohol abuse, Alzheimer's disease, Huntington's disease, Parkinson's disease, traumatic brain injury (TBI), multiple sclerosis, temporal lobe epilepsy, and schizophrenia.

Preliminary analysis with the *Writer* program utilized the modal profile for individuals with chronic alcohol abuse³ as reported in the *Technical Manual* due to the relatively distinct pattern of test scores in this affliction as compared to normal profiles and the pattern of other cognitive impairments. Additionally, the difference in the chronic alcohol abuse pattern from a normal profile was not as extreme in comparison to other disorders, such as mental retardation. This disorder also represents a frequent and

³ While the *WAIS-III WMS-III Technical Manual* refers to the condition as "chronic alcohol abuse" [emphasis added], the description of the disorder in the text more closely conforms to the DSM-IV category of 'alcohol dependence.' Due to this definitional ambiguity, as well as problems in the literature such as confusion of terms and amalgamation of data from both conditions into one category, this study uses data from both dependence and abuse of alcohol. This appears to be a viable approach as the chronic alcohol abuse profile functions here as an exemplar for studying pattern analysis rather than being the focus of the research.

important differential in clinical assessment, as alcoholism often overlaps with other conditions or presents comorbidly with other psychological disorders (Petrakis, Gonzalez, Rosenheck, & Krystal, 2002).

Data from the modal alcohol abuse profile is provided in the first column of Table 2. Figure 1 presents the *WAIS-III* modal alcohol profile plotted in standardized T-score units ($M = 50$, $SD = 10$), as the *Writer* program produces sociodemographically adjusted profiles in T-score format. Note that the mean FSIQ of the chronic alcohol abuse group was 106, exceeding the exam mean score of 100, and therefore the mean T-score across the profile exceeds $T = 50$.

A test protocol of raw scores was generated to match the IQ scores and Indices provided by the publisher, and data were entered into the *Writer* program. Random sampling with replacement via an online random number generator (Haahr, n.d.) was used to produce a sample ($N = 20$) stratified to roughly match the demographic characteristics of individuals with a history of alcohol abuse per the *DSM-IV-TR* (American Psychiatric Association, 2000) and census data regarding age and education (United States Census Bureau, 2007). The sample was 80% male and 85% White. Age for males ranged from 25 to 70 in increments of five, with two profiles each for ages 25 to 50, and one each for ages in the range of 55 to 70. One female profile was assigned to each increment of age 30 to age 60, in denominations of 10 years. For the males one profile was generated with 6 years of education, two with 8 years, four with 10 years, four with 12 years, three with 14 years, and two with 16 years. One female profile was generated for 8 years of education, two for 12 years, and one for 16 years.

The 20 sociodemographically adjusted profiles were generated as graphs, as is routinely done with the program. Two raters (the primary investigator and his major professor) then sorted the resultant graphs into one of three categories in relation to the modal chronic alcohol abuse profile: “similar or highly similar,” “neither similar or dissimilar,” and “dissimilar or highly dissimilar.” The profile categorization was conducted blind to the underlying characteristics of the sociodemographic adjustment.

Results

The investigator rated 80% of the profiles as “dissimilar” or as “neither similar nor dissimilar.” The investigator’s major professor rated 60% of the profiles as falling in one of these two categories. Inter-rater agreement was $r = 0.73$. The most frequent disagreement between raters involved the “neither similar nor dissimilar” category. The basis for the difference in ratings quickly became evident in that the investigator was using more stringent criteria than the major professor for classifying profiles as “similar or highly similar.” (This difference seemed correctable and was subsequently addressed by creating a set of classification rules as described below.)

This preliminary analysis provided a crude estimate of the frequency with which significant profile change occurs subsequent to sociodemographic adjustment, or as a way of determining if more extensive investigation seemed warranted. Even with such a small sample, the obtained rate of substantial pattern change $\geq 60\%$ strongly suggested alteration is not uncommon and as such may have a meaningful influence on diagnostic decision making. This pilot work helped to clarify procedural and feasibility issues and suggested that the matters under consideration merited further research efforts.

Main Study: Overview

Given the results of pilot analyses, the main study seemed feasible and warranted. This research involved three basic parts, the first two of which were designed to address the two primary research hypotheses. The third part was designed to address the exploratory hypothesis. In the first part of the study, chronic alcohol abuse profiles were evaluated to determine rate of change from their respective prototype following sociodemographic adjustment. The process was then repeated with a smaller sample from a separate disorder, traumatic brain injury. In the second part of the study, multiple normal prototypes were generated, with sociodemographically adjusted profiles compared to their respective prototype to assess rate of alteration. In the third or exploratory part of the study, normal and abnormal profiles were compared to several prototypes to provide a rough indication of the potential misclassification rate.

Part I: Are Abnormal Patterns Altered by Sociodemographic Adjustment?

Further Analysis of the Chronic Alcohol Pattern

Method. A simple set of decision rules were generated to improve inter-rater reliability, which are detailed in Table 3. It was decided to rate pattern and elevation respectively, as pilot work revealed instances in which the shape of the pattern was nearly identical and yet the elevation of the test scores changed markedly. That is, although pattern and elevation are often interrelated, it was sensible to create a separate rating for each as instances can occur in which changes in one are more or less independent of changes in the other. Patterns were classified into the “similar or highly similar” category if no substantial shifts in direction or slope were present⁴, although trivial alterations that did not change the ratings were noted. Profiles were classified as “neither similar nor dissimilar” if a minor directional shift was present, or if there were no changes in direction but dramatic change in slope. Finally, a “dissimilar” label was assigned to those profiles with either: a) two or more minor directional changes; b) at least one dramatic directional shift; or c) a minor directional shift *and* a dramatic change in slope. Change in elevation was classified as “none or minimal,” “> minimal but < moderate,” “modest or moderate,” “marked,” and “extreme.” Obviously all of these classification guidelines are somewhat arbitrary but served as a means for generating a rough estimate of occurrence.

⁴ Here, “direction” refers broadly to whether a segment of the graph (as interpreted from left to right) moved downward, upward, or plateaued. “Slope” refers broadly to the distance or scatter between the endpoints of a line segment; that is, whether the line segment was very long (and thus very steep), or was very short (and thus very shallow or flat). It is possible for slope to change independent of direction. In other words, all line segments may conform to the prototypic pattern of ups and downs, but, due to change in the level of scatter, the slopes of certain line segments may be dramatically altered.

A total sample of $N = 102$ profiles was generated for the alcohol abuse condition in the manner described in the “pilot study” section. This sample size was chosen for ease of subdivision over various stratified demographic categories (see below). Random sampling with replacement via an online random number generator (Haahr, n.d.) was utilized to select demographics of each profile. The sample was stratified to crudely match available base rate criteria for the disorder, in order to roughly approximate the clinical presentation of such a disorder. As can be seen in Table 4, which details sociodemographic composition, the sample therefore included over twice as many males as females (Anthony, Warner, & Kessler, 1994; International Consortium on Psychiatric Epidemiology, 2007), a somewhat increased number of White profiles than African American profiles (National Institute on Alcohol Abuse and Alcoholism, 2004; 2005a; 2005b; Kandel, Chen, Warner, Kessler, & Grant, 1997; Grant et al., 2004), and a bias in age towards the lower (25 – 40) and middle (40 – 60) age brackets both to match the age of the prototype sample ($M = 53$; The Psychological Corporation, 2002a) and data indicating higher rates of alcohol abuse and dependence in lower (National Institute on Alcohol Abuse and Alcoholism, 2005a; 2005b; Grant et al., 2004) and middle-age ranges (International Consortium on Psychiatric Epidemiology, 2007). Education level was stratified to roughly match census data (United States Census Bureau, 2007) with a slight bias towards lower and middle levels of education (Crum, Helzer, & Anthony, 1993; Kandel et al., 1997; Droomers, Schrijvers, & Mackenbach, 2004).

Ages below 25 years were not be sampled as *chronic* abuse may be defined as several years of continued heavy drinking. Additionally, some research (Nelson, Heath, & Kessler, 1998) indicates the median age of onset for many alcohol abuse and

dependence criteria is 20 years of age or greater. Furthermore, lower age ranges (those below 20 years old) may create a sampling problem in that a random profile could be drawn in which the assigned level of education in years is greater than is feasible given the age of the individual.

In order to establish the feasibility and reliability of the classification rules a limited sample of $n = 20$ profiles was initially generated. The profiles were classified blindly by the investigator and his major professor as previously described. As the classification system proved sufficiently reliable for these initial 20 profiles, an additional $n = 82$ profiles were generated and classified.

Results. Inter-rater reliability was found to be acceptably high for all domains. Agreement for rating change in *WAIS-III* profiles was $r = 0.85$ for pattern and $r = 0.75$ for elevation. Concordance was higher for *WMS-III* ratings, with inter-rater reliability of $r = 0.89$ for pattern and $r = 0.88$ for elevation. Discrepancies in ratings $> \pm 1$ point were examined by the investigators to identify any instances in which one rater made an objective error (which did occur in two instances and was corrected). Where differences were impressionistic they were discussed with the aim of increasing uniformity so that, at some point, all profiles did not have to be evaluated by both raters.

With satisfactory agreement established, analysis was next directed at profile shifts. Although the initial intent was to analyze shifts by descriptive and inferential statistics, the dramatic results abridged this process. Figures 3 and 4 provide overviews of the results. As can be seen in Figure 3, for pattern on the *WAIS-III* the primary investigator classified 25 profiles as essentially unchanged from the prototype, 48 profiles as neither similar nor dissimilar to the prototype, and 29 profiles as markedly altered.

The investigator's major professor categorized 20 profiles as unchanged, 46 as neither similar nor dissimilar, and 36 as markedly different. In total, between 75% - 80% of the profiles demonstrated more than minor alteration in pattern or shape from the prototype, with 28% - 35% altered markedly enough to classify as distinctly different.

Approximately 20% - 25% of the profiles showed no change in structure.

As can be seen in Figure 4, the *WMS-III* proved slightly more robust to alteration in pattern, with the principle investigator rating 43 profiles as unaltered, 29 profiles as neither similar nor dissimilar to the prototype, and 30 profiles as markedly different. The investigator's major professor classified profiles into a nearly equal distribution, rating 35 profiles as unchanged, 33 as neither similar nor dissimilar, and 34 as markedly changed. Approximately 59% - 66% of the profiles demonstrated more than minor alteration, with about one-third of the profiles (30% - 33%) showing substantial alteration. No change in the pattern was detected for approximately 34% - 42% of the profiles.

Rate of change observed in *WAIS-III* patterns was significantly different from rate of change observed in patterns of *WMS-III* profiles, $\chi^2(1, N = 204) = 7.15, p < 0.01$.

Unevenness in the frequency of change raises an additional concern because a commonly recommended interpretation strategy is to compare *WAIS-III* results to *WMS-III* results, and if both do not change in tandem the interrelation between the tests will be altered as well. The potential implications of these differential shifts will be further delineated and explained in the "Discussion" section.

In regards to elevation, as can be seen in Figure 3, the principle investigator rated approximately 51% of the *WAIS-III* profiles as demonstrating less than moderate change

in elevation, and approximately 50%⁵ of the profiles as demonstrating a moderate or greater degree of elevation shift. Roughly 14% of the profiles demonstrated marked change in elevation, and 8% displayed extreme shifts, such as differences in average elevation of over 15 T-score points (1.5 *SD*'s). The investigator's major professor rated 39% of the *WAIS-III* profiles as exhibiting a moderate or greater degree of elevation shift, with roughly 6% demonstrating marked change, and 5% displaying extreme shifts. As shown in Figure 4, according to the ratings of the principle investigator, approximately 64% of the *WMS-III* profiles demonstrated less than moderate elevation shifts, with 37% demonstrating moderate or greater alteration. About 6% of the profiles demonstrated marked change in elevation, and 7% displayed extreme shifts. According to the ratings of the investigator's major professor, approximately 33% of *WMS-III* profiles demonstrated moderate or greater change in elevation, with approximately 7% showing marked shifts and 3% extreme shifts. No significant differences were observed for rate of elevation change between the *WAIS-III* and *WMS-III*, $\chi^2(4, N = 204) = 5.40, p = 0.25$.

Analysis of Profiles from an Additional Cognitive Disorder: Traumatic Brain Injury

Method. It seems highly unlikely that change in pattern with sociodemographic adjustment would be unique to a single disorder. However, for the sake of thoroughness, and also to obtain some data on generalizability, the process of generating and sorting profiles was replicated with another form of cognitive impairment, traumatic brain injury (TBI). Based on the profound degree of pattern alteration noted in the chronic alcohol abuse sample, a smaller sample size ($N = 50$) seemed sufficient for this analysis. The TBI modal profile as presented in the *Technical Manual* (The Psychological Corporation, 2002a) was utilized. Test scores for this profile can be found in Table 2. Figures 5 and 6

⁵ Totals may not equal 100% due to rounding.

display the respective line graph plots of the WAIS-III and WMS-III profiles, drawn from those test scores.

Random sampling with replacement via an online random number generator (Haahr, n.d.) was again utilized to select the demographics for each of the 50 profiles. The sample was stratified to crudely match available incidence data for TBI, in order to roughly approximate the clinical presentation of this disorder. The sample therefore included approximately twice as many males as females (National Center for Injury Prevention and Control, 2006; Centers for Disease Control and Prevention, 2007). White Non-Hispanic profiles and African/African-American profiles were represented equally, despite some evidence of higher incidence among African-Americans, due to confounds in the literature of ethnicity with socioeconomic status (Bruns & Hauser, 2003). Younger ages were over-sampled both to match the age of the prototype sample ($M = 27$; The Psychological Corporation, 2002a) and data indicating younger individuals are most at risk for sustaining a TBI (National Center for Injury Prevention and Control, 2006; Centers for Disease Control and Prevention, 2007; Bruns & Hauser, 2003). Unlike the alcohol abuse sample in which ages ranged from 25 to 70 years, ages as low as 20 years and as high as 75 years were represented in the TBI group to capture the increased risk of brain injury among younger and elderly individuals (Bruns & Hauser, 2003). Education level was stratified to crudely approximate census data (United States Census Bureau, 2007). The sociodemographic composition of the sample is presented in Table 5.

Classification of profiles was conducted blindly by the principal investigator and his major professor in the same manner as with the alcohol abuse profiles, utilizing the

categories delineated in Table 3. Level of agreement between the two raters was calculated, and the results analyzed by means of descriptive statistics.

Results. Most likely as a result of previous discussion regarding discrepancies in ratings, inter-rater reliability improved for this series of classifications and was high for all domains. Agreement for rating change in *WAIS-III* profiles was $r = 0.94$ for pattern and $r = 0.90$ for elevation. For *WMS-III* ratings, inter-rater reliability was $r = 0.90$ for pattern and $r = 0.91$ for elevation. No discrepancies in ratings were $> \pm 1$ point.

Figure 7 summarizes the *WAIS-III* classifications and Figure 8 the *WMS-III* classifications. For pattern on the *WAIS-III*, according to both raters 16% of the profiles remained essentially unchanged and 84% demonstrated some degree of alteration. Further, across the raters, 70% - 72% of the profiles showed great enough alteration to be classified as dissimilar. The *WMS-III* once again proved more robust to alteration in pattern, with approximately 40% - 42% of the profiles remaining essentially unchanged and 58% - 60% demonstrating some degree of alteration in pattern or shape from the prototype. Alteration sufficient to be classified as a distinct change occurred in 10% of the profiles. A significant difference in the observed rate of change was again obtained when comparing the *WAIS-III* to the *WMS-III*, $\chi^2(1, N = 100) = 7.14, p < 0.01$.

WAIS-III profiles exhibited less than moderate change in elevation in 72% - 78% of the cases, whereas about 22% - 28% exhibited moderate or greater elevation shift. Only 2% - 6% demonstrated marked or extreme change in elevation. *WMS-III* profiles yielded similar results, with 66% - 86% demonstrating less than moderate elevation shifts. Moderate or greater alteration was obtained in 14% - 34%, with 6% demonstrating marked change in elevation and 0% - 4% displaying extreme shifts. Again no significant

differences were obtained between the *WAIS-III* and *WMS-III* for degree of change in elevation, $\chi^2(1, N = 100) = 4.03, p = 0.40$.

Overall, as with the chronic alcohol abuse profiles, many *WAIS-III* and *WMS-III* profiles showed shifts in pattern as a result of sociodemographic adjustments. Likewise, rates of change in pattern differed significantly between the two measures. A nontrivial percentage likewise demonstrated changes in elevation, although as before these changes occurred less frequently than changes in pattern and no significant differences in the rate of elevation change were noted between these two tests.

Part II: Are Normal Patterns Altered by Sociodemographic Adjustment?

It is conceivable, although unlikely, that normal *WAIS-III* and *WMS-III* profiles do not show the same types of changes as profiles associated with pathologic categories. To appraise this possibility, parallel investigation of normal profiles was undertaken.

Method

Six normal *WAIS-III* profiles were constructed to match the overall level of intellectual functioning, (i.e., Full Scale IQ) of the chronic alcohol abuse modal profile (FSIQ = 106). To ensure a normal distribution and range of scatter, the intercorrelations among IQ scores and indices as presented in the *Technical Manual* (The Psychological Corporation, 2002a) were examined. Based on the 25 years – 29 years age group, all possible pairings of IQ scores and indices were listed in rank order of highest to lowest correlations. Although these values are not identical across age brackets, they tend to be similar and no composite listing is available. As such, the listing from this age range seemed as suitable for the present purpose as any other age bracket. The list was roughly divided into three categories of correlation strength: strongest ($r \approx 0.90$'s – 0.80 's), moderate ($r \approx 0.70$'s – 0.60 's), and weakest ($r \leq 0.60$). Degree of scatter was assigned by randomly choosing pairs with stratification, so that those pairs with the weakest correlations – and thus the greatest likelihood of deviating considerably from one another – were more likely to be assigned the largest score discrepancy. Likewise, score pairs with high correlations were more likely to be assigned minimal scatter or discrepancy. The high and low score was determined randomly, with each score having an equal likelihood of deviating in either direction.

After ascertaining which IQ score or indices would be the most discrepant, the level or value of scatter across the scores was determined. The *WAIS-III Administration and Scoring Manual* (Wechsler, 1997a) provides the mean difference between VIQ and PIQ, as well as between all indices. Averaging across these means gives an overall value of $M = 10.3$ ($SD = 8.1$). Given the seven indices to be plotted and the total number of possible comparisons among scores, levels of scatter were randomly assigned to vary within approximately ± 1.5 SD from the mean to replicate the modal level of variability across indices in a typical examinee. The values of the six normal profiles are presented in Table 6.

It was decided not to construct normal *WMS-III* profiles for multiple reasons. First, due to the more robust nature of the *WMS-III* in relation to profile change, possibly due to low correlations between *WMS-III* variables and sociodemographic variables, the *WAIS-III* is a better instrument for detecting change in pattern. Second, the absence of an overall composite score on the *WMS-III* makes it more difficult to match normal profiles to the characteristics of the previously generated abnormal profiles. Third, the presence of some negative correlations between indices on the *WMS-III* complicates the generation of a normal profile, particularly as correlations values may vary from negative to positive by age bracket.

What became apparent in this exercise is the vast number of possible variations in normal profiles, and, further, the extent of problems involved in determining representativeness. For example, there are literally thousands of possible score combinations falling within parameters for normal range and variation across IQ and index values. Further, it is not straightforward to determine how to group these

variations, and whether there are groupings of scores that might be considered genuinely prototypical or whether such categorization is merely arbitrary. Consequently, when there are considerable uncertainties about the defining features and parameters of normal profiles, how does one determine how to sample from this diverse pool of potential profiles, and how can one evaluate generalization to other normal profiles?

It was obviously beyond the scope of this thesis to solve these still largely unresolved questions in the field. However, a few matters seemed clear. First, given the impressive frequency and magnitude of change observed with the two abnormal prototypes, should similar results occur with normal profiles, which seemed likely, only relatively small n 's would be necessary to detect them. Further, given questions about representativeness, it seemed preferable to examine a smaller number of cases across a larger number of normal protocols than a large number of cases across one or two normal protocols. Additionally, even given a modest number of normal protocols, consistency of results across the protocols, if present, would also provide helpful information about potential generalization. Of course, with greater resources, it would be most preferable to study a large number of cases across a large number of normal protocols, but for purposes of this study expanding the number of normal protocols and studying a limited number of cases per protocol seemed the most judicious and realistic compromise.

As such, 10 adjusted profiles were generated for each of the six respective normal *WAIS-III* profiles, for a total sample of $N = 60$. Three profiles (A, C, E) were randomly assigned to be held at age 53 years to match the mean of the chronic alcohol abuse group and facilitate later comparison between the profiles. Likewise, Profiles B, D, and F were held to age 27 years to match the mean of the TBI group. Ethnicity, gender, and

education levels were set to match the 30 alcohol abuse profiles and 30 TBI profiles closest to the mean age of each respective group. That is, the demographic stratification of the 30 previously generated alcohol abuse profiles closest in age to 53 years, and the demographic stratification of the 30 previously generated TBI profiles closest in age to 27 years, were assigned to the normal profiles in order to create matched demographic samples. These matched normal profiles were then sorted in the manner described previously, as compared to each profile's respective original non-adjusted pattern.

Results

Inter-rater reliability was again high, with agreement of $r = 0.95$ for rating change in pattern, and $r = 0.92$ for rating elevation. No discrepancies in ratings were $> \pm 1$ point.

The amalgamated classifications for all six profiles are presented in Figure 9. For pattern, 43% of the profiles changed minimally, with 57% showing greater change. Among the 57%, 20% of the profiles changed moderately and 37% became dissimilar. Obviously, this is a high rate of change and mirrors the fundamental results for the pathological prototypes. Combining the categories of moderate change and greater change (to the point of dissimilarity), all six of the normal prototypes showed a minimal frequency of 30%, thus suggesting that alteration is probably not an isolated phenomenon across normal profiles but a more general one. Further, although differences have to be interpreted very cautiously given the low n 's per normal prototype, the frequency of change to dissimilar patterns ranged from 10% to 100%. This differential rate allows the very tentative conjecture, which is almost a given *a priori* anyway, that rate and level of change may vary considerably across normal profiles.

As can also be seen in Figure 9, approximately 68% - 70% of normal profiles exhibited minimal or less than moderate change in elevation, with 30% - 32% of the profiles demonstrating a moderate or greater degree of elevation shift. Only 8% - 10% demonstrated marked change in elevation, and no profile exhibited an extreme elevation change. These results suggest that, while not trivial, the changes in elevation are less pronounced than pattern shifts, which emulates previous findings with abnormal profiles. However, because the metric for evaluating level of change across these two dimensions of pattern and elevation was not uniform, such comparisons are tentative and it is not even clear how a common metric would be created.

Part III: Do Adjustments Alter Matches Within Abnormal Prototypes and Across Normal and Abnormal Prototypes?

The analysis to this point provided strong support for the primary hypotheses. Both abnormal and normal prototypes were found to change often and substantially when sociodemographic adjustments were performed. It remained to examine the exploratory hypothesis, which addressed whether such changes might result in altered matches among different pathological prototypes and also in switches across normal and abnormal categories. For example, might a result that resembles the prototype for alcoholism come to more closely resemble the prototype for, say, TBI? Alternatively, might a prototype that appears normal come to more closely resemble some abnormal prototype and *vice versa*? Certain flexibility was called for in addressing this last hypothesis, both because of its exploratory nature and because guidance and direction depended in part on the results of the prior analyses.

Best Fit Among Represented Disorders

Method. A sample of normal and abnormal profiles was selected for comparison of best fit to multiple pattern prototypes. Those alcohol abuse profiles ($n = 30$) closest in age to 53 years and those TBI profiles ($n = 30$) closest in age to 27 years were selected for their comparability to the mean ages of the respective disordered groups as presented in the *Technical Manual*. Additionally, 40 of the 60 original normal profiles were randomly selected, resulting in a total sample of $N = 100$.

Each of the 100 protocols was examined for match with four prototypes: TBI, chronic alcohol abuse, and two normal prototypes (A and B from the prior analysis). In choosing the normal prototypes, the researcher selected, from among the six previously

generated normal prototypes, the one that most closely resembled the TBI prototype and the one that most closely resembled the chronic alcohol abuse prototype. Although it might initially seem that this procedure biased outcome to produce poor test performance, it was not as if the closest matches were selected from among hundreds of generated normal profiles. Rather, only six were generated and two (or one-third of the normal prototypes) selected. If anywhere approaching one in three normal protocols match any one of two pathological prototypes to some extent, the implications are clear. Thus, it could be argued that, if anything, the procedure created conditions less demanding than those often found in clinical application. The normal prototypes appear in Figure 10 and Figure 11, respectively.

All protocols were sequentially compared to the four prototypes (alcohol abuse, TBI, normal A and normal B), and sorted into one of four categories: “same,” “same and other,” “none,” and “other.” A rating of “same” indicated that the profile was closest in match to its respective prototype, and was not miscategorized. “Same and other” indicates a reasonable degree of match between the profile and both its respective template and another distinct condition. Ratings of “none” were assigned to those profiles not demonstrating a reasonable match to any prototype, and ratings of “other” were assigned to those profiles that best matched a prototype from a separate condition but not the prototype from which the profile was created.

Profiles were rated in random sequence to control for potential order effects. Given the extremely high levels of agreement among raters as the project progressed, it only seemed necessary to obtain one set of ratings. Also, although it was originally planned to perform the ratings blindly, judgment in almost all cases seemed very

straightforward and the decision was made to forego the time and effort needed to prepare the protocols for blind evaluation.

Results. Collectively, in 26% of the cases, profiles exhibited no or minimal change in pattern and matched the proper prototype exclusively. Profiles demonstrated reasonable match to the proper prototype as well as another prototype in 11% of the cases, and matched a different prototype in 4% of the cases. Profiles were found to not match any prototype in more than half of the instances (59%).

When examining the chronic alcohol abuse profiles alone, 33% matched the appropriate alcohol abuse prototype exclusively, 10% matched that and another prototype, 7% better matched a distinct prototype exclusively, and 50% did not match any prototype. TBI profiles exhibited similar trends, with 30% matching the correct TBI prototype exclusively, 20% matching both the TBI prototype and another prototype, 7% exhibiting better fit to a different prototype only, and 43% not matching any prototype. To state the obvious, the low frequency of exclusive match with the respective prototype raises serious questions or concerns.

In a small but notable percentage, normal profiles were made to appear abnormal and *vice versa*. For alcohol abuse profiles, 10%, demonstrated match to both a normal prototype and another abnormal condition while another 7% best matched a normal prototype exclusively. In the TBI profiles, 20% matched both the correct prototype and the TBI prototype with a reasonable degree of match, while about 7% matched a normal prototype exclusively. About 3% of normal profiles showed best match to both the respective normal prototype and an abnormal condition.

Best Fit Among Represented Disorders and Additional Cognitive Disorders

Method. To further explore the issue of miscategorization or misdiagnosis, the process was repeated including additional prototypes from other pathological conditions. In this manner a rough estimate could be generated of the rate with which sociodemographic adjustment potentially leads to misdiagnosis or confusion among differing etiologies in questions of differential diagnosis.

The other cognitive disorder prototypes were selected from data available in the *Technical Manual*, excluding: a) those profiles so aberrant from the normal profile or the chronic alcohol abuse profile as to virtually preclude any possibility of misclassification (e.g., mental retardation), and b) those disorders for which misclassification is unlikely based on clinical presentation (e.g., hearing disorders). Of those remaining disorders, three were selected based on frequency of occurrence and relevance in a differential diagnosis: Alzheimer's disease, Parkinson's disease, and schizophrenia. Note, for example, in cases of alcohol abuse, excessive drinking can lead to diminished cognitive functioning, altered sensory experience, or movement disorders that could be confused with these other pathological conditions. Data regarding the modal presentation for these conditions are available in Table 3. Graphical templates were created and comparison was conducted as before.

Results. When additional cognitive disorders were included as categories, correct classification rates or matches with respective prototypes decreased even further. The total number of profiles classified as solely matching the correct prototype dropped from 26% to 7%. For the subset of alcohol abuse profiles, correct classification dropped from 33% to 0%, with no profiles exhibiting best match solely to the correct prototype. Best match to another prototype alone was observed in 13 profiles, while 9 profiles matched

both the correct prototype and another abnormal prototype equally well and 2 profiles did not match any prototype. TBI proved more robust, with a decrease in classification accuracy from 30% to 23%. One profile matched the TBI prototype and other pathological prototypes equally well, 7 profiles matched both the TBI prototype and a normal prototype equally well, 2 profiles matched normal prototypes exclusively, and 13 profiles did not match any prototype. The results with the normal prototypes were also very poor, with the original rate of 10% also dropping to 0%.

Brief Exploratory Analysis using Normal Profiles with Increased and Decreased Scatter

Method. Through the above illustrations it became clear to the investigator that profiles with normal ranges of scatter often appear more abnormal or discrepant than authentically abnormal prototypes. To further assess the impact of normal scatter on pattern interpretation, a brief exploratory analysis was conducted. Ten additional normal profiles were constructed in the manner described previously, five with decreased or lower levels of scatter, and five with increased levels of scatter. High scatter profiles were assigned random levels of scatter up to approximately ± 2 SD's, while low scatter profiles were randomly assigned varying levels < 1 SD. These profiles were then compared to the various cognitive disorder templates.

Results. Based on impressionistic judgment of the primary investigator, misclassification seemed possible for 2 of the 5 low scatter profiles, 1 demonstrating reasonable match to the schizophrenia template, and 1 exhibiting reasonable match to both Parkinson's disease and Alzheimer's disease. Of the 5 high scatter profiles, 1 could be potentially misclassified, showing a reasonable degree of match to both the Parkinson's and Alzheimer's conditions.

Summary of Results

Both primary hypotheses were supported. Frequent and substantial pattern change was detected across all parts of this study for both normal and abnormal profiles. The *WAIS-III* was most vulnerable to alteration in pattern, with more than three out of four profiles changing at least moderately for the alcohol condition, and about a third of the profiles changing enough to create a distinct pattern. The effect was even more pronounced for the TBI group, with nearly three of four profiles changing substantially enough to create a new pattern. About three in five normal profiles exhibited at least moderate change and about two in five substantial change. Although level and rate of alteration varied across the six normal prototypes, all six showed moderate or greater changes on at least 30% of the profiles, thereby suggesting that such change is not uncommon across normal individuals. Across the abnormal and normal cases, high levels of alteration and partial to complete pattern reversals were not uncommon. For examples of this type of dramatic change in profiles secondary to sociodemographic adjustment, see Appendix A.

The *WMS-III* showed less extreme change than the *WAIS-III*, differing significantly in both alcohol and TBI conditions, but still demonstrated a clinically noteworthy level of alteration. Three in five profiles demonstrated some level of change for both alcohol abuse and TBI. In about one third and one tenth of the alcohol abuse and TBI profiles, respectively, dissimilar profiles were generated.

Overall profile elevation change was likewise impacted by sociodemographic adjustment, with moderate or smaller changes occurring most often. Marked or extreme changes for the *WAIS-III* were noted in 5% - 14% of the alcohol abuse profiles and 6% or

less of the TBI profiles. About 10% of the normal profiles were likewise impacted. For the *WMS-III*, about one in ten alcohol abuse profiles showed marked or extreme elevation change, while very few to none of the TBI profiles showed variation to such a degree.

Proper matching of a profile to its respective prototype occurred at approximately chance level in the first round of sorting (involving two abnormal and two normal prototypes), with only one in four profiles properly matching its respective prototype. Roughly 10% of the profiles matched another prototype to the same degree as they matched the appropriate one, and for about 4% a misclassification was the best match. Almost two fifths of the profiles could not be properly matched to any category. On the subgroup level alcohol and TBI profiles were each correctly classified in about one in three instances. When additional cognitive disorders that could arise as potential differential diagnoses in a clinical setting were added to the sorting process, rates of proper classification fell further overall, with about one in six profiles being correctly categorized. On the subgroup level, no alcohol profiles were correctly classified, while TBI profiles proved more robust yet still dropped to chance level, with less than one in four categorized appropriately. A brief exploration of additional normal profiles with varying degrees of scatter likewise illustrated potential misclassification to be a common problem.

Discussion

Theoretical Implications and Future Directions

The results of this study strongly suggest that sociodemographic adjustment often impacts test score profiles. A considerable percentage of profiles changed at least moderately from the original prototype and commonly resembled other conditions or states as well as, or more closely, than the prototypic pattern. Likewise, in a fairly sizable percentage of cases, normal patterns looked abnormal and vice versa following sociodemographic adjustment.

Although intended to serve as a simple illustration of one potential confound to pattern analysis, the results of this study should generalize to clinical practice and apply in particular to attempts to reach a differential diagnoses via complex configural analysis. Alcohol abuse and brain injury profiles were both altered substantially from their respective original patterns, and proper 'diagnosis' or correct classification was reduced to below chance level. These conditions were selected (at least as per their prototypes presented in the official test manual) due to their frequent relevance in clinical settings and the likelihood that both would be included as part of the differential diagnosis in many cases. As differential diagnosis of such conditions are additionally hampered by factors such as unknown prior levels of cognitive functioning and high levels of comorbidity with other cognitive disorders (e.g., various dementias), the startling misclassification rates found in the present study certainly raise concerns.

Certain implications of these findings for diagnosis by means of pattern analysis are obvious. The level of scatter and the configuration of highs and lows in a test battery may be determined in considerable part by the use or misuse of sociodemographic

adjustments. First, as was demonstrated in this study, such adjustment may differentially impact subtests and indices within a test, thus changing the relationships of such scores with one another. Second, a larger battery may exhibit similar uneven effects across tests, magnifying the problem by producing both intra-test and inter-test alterations. Furthermore, if one uses a battery in which adjustments are applied inconsistently, such as by drawing a comparison between Test A (which adjusts for age and education) and Test B (which accounts for only gender), the potential distorting impact on test score pattern and elevation is greatly exacerbated. This inconsistent alteration in scores compounds other hindrances to pattern analysis, such as inconsistent Flynn effects, varying error rates, and discrepant normative groups to weaken – if not stymie – the process of complex configural analysis as it is commonly practiced.

It is often questionable how much, even under relatively good conditions (e.g., limited measurement error), attempts at complex pattern analysis contributes to the accuracy of decision making in psychology and neuropsychology. Patterns, to begin with, are often highly susceptible to error. Further, because such factors as uneven sociodemographic adjustment or insufficient research on the meaning and impact of such adjustments creates additional ambiguity, simplified approaches that are less error-prone may often, at present, prove better alternatives. For example, other methods may prove unhindered by – or at least more robust to – the issues discussed here. Linear composites with equal unit weights, for example, tend to be much more reliable than test patterns, especially when the latter is appraised by subjective judgment (e.g., Dawes, 1979; Dawes & Corrigan, 1974; Faust, 2007). Further research as to valid actuarial diagnostic

processes that could supercede complex pattern analysis would be both prudent and beneficial to the field of neuropsychology.

Furthermore, the logical progression for this line of research would be to explore the impact of such pattern alteration on clinician judgment. If such alterations do not impact clinician interpretation of the profile, their occurrence then is a moot point. It borders on absurd, however, to assume such drastic alteration in pattern as marked changes or complete reversals would not impact subjective interpretation. The classification system in this study was designed to mirror clinician judgment regarding stability or judgment in pattern, supporting the prediction that clinical judgment will demonstrate similar effects.

Another aspect of this study that warrants further attention is the interpretation of normal profiles that are not sociodemographically adjusted. Generating normal profiles proved to be a difficult and complicated undertaking. The process was hindered by the enormous range of theoretical normal profiles to select from, in addition to impediments imposed by both varying test relationships and inconsistent normal estimates of scatter.

It was interesting to note that the generated normal profiles often appeared more "abnormal" or deviant than the prototypic abnormal profiles used in this research. Prototypes, which average across individuals, tend to reduce the level and range of scatter, which of course does not occur with single normal profiles, yet higher levels of scatter are often presumed to reflect pathology. Whether or not greater scatter is associated with pathology, an involved matter that resists a "yes or no" answer, in clinical cases individual results are routinely compared against prototypes that usually induce scatter. This gives cause to question the rate with which a profile demonstrating normal

scatter and inter-index variability would be clinically judged as abnormal or indicative of some disease state. Future research should examine the degree to which normality in test score variations is viewed as abnormality. For an example of such a study in a population of children, see Glutting et al. (1997).

Limitations

A number of potential limitations arose in the course of conducting this research and should be addressed here along with some potential criticisms. For one, the validity of the test score patterns used in this study can be questioned. One can ask, for example, whether the TBI prototype presented in the *Technical Manual* is a good representation of legitimate TBI patients, or whether the alcohol abuse pattern is an accurate representation of individuals with chronic alcohol abuse. Although the concern itself is legitimate, it would not seem to bear on the viability of the current study. The focus of this research was not on the true state of any particular pattern but rather on the impact of sociodemographic adjustment on test patterns, and, consequently, the potential impact on clinical judgment. To this end the employed prototypes served in the capacity of demonstratives only. Again, the focus is mainly on the consistency of the profile or pattern of the adjusted scores. If the same raw scores produce markedly discrepant patterns depending whether sociodemographic adjustment is preformed or not, those differences are likely to produce contrasting conclusions, and, as must be true in many cases, one or both of those conclusions must be wrong. This study strongly suggests there is such an issue that merits attention.

A second potential criticism involves the nearly exclusive use of descriptive statistics for analysis and interpretation. Although a number of statistical analyses were

contemplated, the obvious and overwhelming magnitude of results generally rendered such analyses unnecessary. There is no doubt that the frequency of change was “significantly” and substantially above zero, and showing that one after another result achieved statistical significance would do little to convey the power of the findings. Furthermore, inferential statistical analysis is stymied by a lack of consensus in the literature regarding the interpretation and assessment of change in pattern. As significant variation in pattern (or lack thereof) may not account for shape or distribution of scatter, the inferential interpretation of configural analysis remains an unresolved issue in the field.

Along similar lines, it remains unclear how to calculate an effect size for the degree of profile change due to sociodemographic adjustment. Whereas a correlation coefficient may partially account for level of change, problems with this method arise. First, the prototype used for comparison is a constant and thus invariable. Variation within the prototype is not necessarily reflective of the degree of variations between the prototype and an adjusted profile. Likewise, any change detected by the correlation coefficient would not account for direction of pattern shift and may be negated by the respective highs and lows of the profile.

Finally, one notable confound of this analysis is the use of one set of raw scores for the generation of chronic alcohol abuse and TBI profiles. Raw scores were calculated from the available prototypes and all adjusted profiles for these two groups were derived from this set of raw scores. An interaction with the intrinsic age-correction of the *WAIS-III* and *WMS-III* is thus likely, exacerbating deviation in the profiles as age diverges further from the mean of each group. Despite this confound the study remains

interpretable on two grounds. Foremost, as the samples are segmented by age, notable rates of pattern distortion are found for those age levels similar to the mean of the prototypic group as well as for the outlying ages. Further, the later portions of this study focusing on misclassification controlled for this confound by including only those profiles from comparable ages to the mean of the prototype. A brief analysis of subtests most likely to vary by age revealed the alteration in the selected profiles to be negligible. This remains a limitation meriting later follow-up, which is planned.

Table 1

Features of the WAIS-III – WMS-III – WIAT-II Writer

Variable	Available Options	True Functions
Age	Calculated from birth date, giving appearance each year is distinct	Ages are banded, but not in manner coinciding with WAIS-III manual age bands
Education	Yearly, from 1 to 20 years	Appears to be banded, such as higher levels of education
Ethnicity	White not Hispanic Origin Hispanic African/African American American Indian/Alaskan Native Asian/Asian American Pacific Islander	White not Hispanic Origin African/African American
Gender	Male Female	Male Female
Handedness	Right Left Ambidextrous	No effect on scores

Table 2

Characteristics and Test Scores of Disorder Prototypes

	Disorder:				
	Chronic Alcohol Abuse	TBI	Alzheimer's	Parkinson's	Schizophrenia
Sample Size (N =)	28	22	35	10	42
	<u>M (SD)</u>	<u>M (SD)</u>	<u>M (SD)</u>	<u>M (SD)</u>	<u>M (SD)</u>
Age	53.3 (10.2)	26.9 (5.9)	72.2 (7.8)	71.2 (11.5)	38.2 (6.1)
WAIS-III					
Verbal IQ	108.6 (12.7)	89.6 (12.4)	92.2 (13.1)	94.6 (8.3)	88.9 (15.2)
Performance IQ	101.2 (14.5)	84.5 (13.8)	81.7 (13.2)	82.3 (12.3)	86.2 (12.9)
Full Scale IQ	106.1 (13.5)	86.5 (10.9)	86.6 (13.1)	88.2 (10.1)	86.8 (14.8)
Verbal Comp. Index	109.0 (11.4)	89.6 (12.7)	93.0 (12.0)	96.9 (7.3)	93.3 (16.4)
Perceptual Org. Index	102.0 (14.0)	92.1 (15.0)	84.8 (12.4)	84.7 (12.1)	89.6 (13.9)
Working Memory Index	104.6 (12.3)	89.8 (13.1)	87.2 (17.3)	89.6 (12.1)	85.0 (15.1)
Processing Speed Index	97.7 (12.5)	73.4 (10.7)	79.6 (14.4)	81.7 (10.3)	83.4 (11.8)
WMS-III					
Auditory Immediate	108.0 (15.4)	89.3 (19.3)	68.7 (11.0)	86.0 (17.4)	83.3 (15.6)
Visual Immediate	96.0 (14.5)	74.9 (13.9)	70.6 (10.9)	84.7 (15.1)	82.3 (14.3)
Immediate Memory	102.5 (16.5)	78.9 (17.7)	62.9 (11.4)	82.3 (17.8)	79.1 (15.7)
Auditory Delayed	107.3 (15.5)	89.6 (21.8)	66.1 (9.6)	85.8 (17.2)	84.4 (15.9)
Visual Delayed	97.9 (13.5)	74.3 (13.9)	67.5 (8.1)	80.9 (13.1)	79.3 (14.8)
Auditory Recog. Delayed	109.6 (13.0)	93.6 (16.6)	65.6 (8.6)	90.0 (14.1)	86.1 (14.9)
General Memory	105.2 (14.2)	81.9 (16.5)	60.4 (8.9)	81.8 (15.1)	79.7 (15.8)
Working Memory	98.0 (9.3)	91.9 (11.9)	80.4 (16.9)	85.6 (13.0)	85.6 (13.9)

Note: All data from The Psychological Corporation, 2002a.

Table 3

Rules for Classifying Pattern and Elevation

Pattern Classifications	
similar or highly similar	no shifts
neither similar nor dissimilar	no dramatic shifts no shifts, but dramatic change in slope
dissimilar	no dramatic shifts, but dramatic change in slope one or more dramatic shifts two or more shifts that are > trivial but < dramatic
Elevation Classifications	
1	none or minimal change
2	more than minimal, less than moderate change
3	modest or moderate change
4	marked change
5	extreme change

Table 4

Sociodemographic Data for Alcohol Abuse Profiles

Category	n =
Gender	
Male	71
Female	31
Race	
White Non-Hispanic	55
African/African-American	47
Age	
25y	13
30y	13
35y	13
40y	13
45y	8
50y	8
55y	9
60y	8
65y	7
70y	10
Education	
6y	6
8y	7
10y	16
12y	25
14y	23
16y	15
18y	10
TOTAL SAMPLE	N = 102

Table 5

Sociodemographic Data for TBI Profiles

Category	n =
Gender	
Male	33
Female	17
Race	
White Non-Hispanic	25
African/African-American	25
Age	
20y	9
25y	8
30y	6
35y	4
40y	4
45y	4
50y	3
55y	3
60y	3
65y	2
70y	2
75y	2
Education	
8y	3
10y	7
12y	20
14y	10
16y	5
18y	5
TOTAL SAMPLE	N = 50

Table 6

Generated Normal Profiles

WAIS-III Score	Profile A	Profile B	Profile C	Profile D	Profile E	Profile F
Verbal IQ	108	110	107	103	112	107
Performance IQ	102	100	104	109	98	104
Full Scale IQ	106	106	106	106	106	106
Verbal Comp. Index	105	114	105	114	116	105
Perceptual Org. Index	118	95	109	111	99	109
Working Memory Index	113	111	94	99	109	94
Processing Speed Index	96	103	117	96	103	117

Table 7

Rate of Profile Match Across Multiple Prototypes

<u>Prototype</u>	Classification			
	<u>Same</u>	<u>Same & Other</u>	<u>None</u>	<u>Other</u>
Normal	7	2	31	0
Alcohol Abuse	10	3	15	2
TBI	9	6	13	2
TOTAL	26	11	59	4

Total N = 100

Figure 1

Chronic Alcohol Abuse – Modal Profile: WAIS-III

Figure 2

Chronic Alcohol Abuse – Modal Profile: WMS-III

Figure 3

Classifications of Adjusted Chronic Alcohol Abuse Profiles: WAIS-III

Note: may not total to 100% due to rounding; pattern 1 = similar or highly similar; pattern 2 = neither similar nor dissimilar; pattern 3 = dissimilar; elevation 1 = none or minimal change, elevation 2 = >minimal, <modest change; elevation 3 = modest or moderate change; elevation 4 = marked change; elevation 5 = extreme change

Figure 4

Classifications of Adjusted Chronic Alcohol Abuse Profiles: WMS-III

Note: may not total to 100% due to rounding; pattern 1 = similar or highly similar; pattern 2 = neither similar nor dissimilar; pattern 3 = dissimilar; elevation 1 = none or minimal change, elevation 2 = >minimal, <modest change; elevation 3 = modest or moderate change; elevation 4 = marked change; elevation 5 = extreme change

Figure 5

TBI – Modal Profile: WAIS-III

Figure 6

TBI – Modal Profile: WMS-III

Figure 7

Classifications of Adjusted TBI Profiles: WAIS-III

Note: may not total to 100% due to rounding; pattern 1 = similar or highly similar; pattern 2 = neither similar nor dissimilar; pattern 3 = dissimilar; elevation 1 = none or minimal change, elevation 2 = >minimal, <modest change; elevation 3 = modest or moderate change; elevation 4 = marked change; elevation 5 = extreme change

Figure 8

Classifications of Adjusted TBI Profiles: WMS-III

Note: may not total to 100% due to rounding; pattern 1 = similar or highly similar; pattern 2 = neither similar nor dissimilar; pattern 3 = dissimilar; elevation 1 = none or minimal change, elevation 2 = >minimal, <modest change; elevation 3 = modest or moderate change; elevation 4 = marked change; elevation 5 = extreme change

Figure 9

Classifications of Normal Profiles

Note: may not total to 100% due to rounding; pattern 1 = similar or highly similar; pattern 2 = neither similar nor dissimilar; pattern 3 = dissimilar; elevation 1 = none or minimal change, elevation 2 = >minimal, <modest change; elevation 3 = modest or moderate change; elevation 4 = marked change; elevation 5 = extreme change

Figure 10

Normal Prototype A

Figure 11

Normal Prototype B

Appendix A: Examples of *WAIS-III* Pattern Reversals

Figure 12

Pattern Alteration: Example A

Note: 70 year old Black female with 12 years of education. Compare to Figure 1.

Figure 13

Pattern Alteration: Example B

Note: 55 year old Black female with 8 years of education. Compare to Figure 1.

Figure 14

Pattern Alteration: Example C

Note: 70 year old Black male with 12 years of education. Compare to Figure 1.

Figure 15

Pattern Alteration: Example D

Note: 70 year old White male with 14 years of education. Compare to Figure 1.

Figure 16

Pattern Alteration: Example E

Note: 55 year old Black female with 8 years of education. Compare to Figure 1.

Figure 17

Pattern Alteration: Example F

Note: 40 year old White male with 16 years of education. Compare to Figure 5.

Figure 18

Pattern Alteration: Example G

Note: 75 year old White male with 14 years of education. Compare to Figure 5.

Figure 19

Pattern Alteration: Example H

Note: 30 year old White female with 16 years of education. Compare to Figure 5.

Figure 20

Pattern Alteration: Example I

Note: 20 year old White male with 12 years of education. Compare to Figure 5.

Figure 21

Pattern Alteration: Example J

Note: 27 year old Black male with 8 years of education. Compare to Figure 11.

Bibliography

- Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34*, 341 – 382.
- American Psychiatric Association. (2000). *The diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.
- Anthony, J. C., Warner, L. A., & Kessler, R. C. (1994). Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: Basic findings from the national comorbidity survey. *Experimental and Clinical Psychopharmacology, 2*, 244 – 268.
- Anastasi, A., & Urbina, S. (1996). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Bellugi, U., Lichtenberger, L., Jones, W., & Lai, Z. (2000). I. The neurocognitive profiles of Williams syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience, 12 suppl. 1*, 7 – 29.
- Bruns, J. Jr., & Hauser, W. A. (2003). The epidemiology of traumatic brain injury: A review. *Epilepsia, 44*, 2 – 10.
- Centers for Disease Control and Prevention. (2007, May 30). *Incidence of traumatic brain injury in the United States, 2002 and 2003 updated data tables*. Retrieved from http://www.cdc.gov/ncipc/tbi/TBI_in_US_04/TBI_Incidents.htm
- Crum, R. M., Helzer, J. E., & Anthony, J. C. (1993). Level of education and alcohol

- abuse and dependence in adulthood: A further inquiry. *American Journal of Public Health*, 83, 830 – 837.
- Cysique, L. A. J., Maruff, P., & Brew, B. J. (2006). The neuropsychological profile of symptomatic AIDS and ADC patients in the pre-HAART era: A meta-analysis. *Journal of the International Neuropsychological Society*, 12, 368 – 382.
- Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34, 571 – 582.
- Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81, 95 – 106.
- Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668 – 1674.
- Dinn, W. M., Haris, C. L., Aycicegi, A., Greene, P. B., Kirkley, S. M., & Reilly, C. (2004). Neurocognitive function in borderline personality disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 28, 329 – 341.
- Droomers, M., Schrijvers, C. T. M., & Mackenbach, J. P. (2004). Educational differences in starting excessive alcohol consumption: Explanations from the longitudinal GLOBE study. *Social Science & Medicine*, 58, 2023 – 2033.
- Einhorn, H. J. (1986). Accepting error to make less error. *Journal of Personality Assessment*, 50, 387 – 395.
- Exner, J. E. Jr., & Erdberg, P. (2005). *The Rorschach: A comprehensive system – Volume 2 advanced interpretation* (3rd ed.). Hoboken, NJ: John Wiley & Sons.
- Faust, D. (2003). Holistic thinking is not the whole story: Alternative or adjunct approaches for increasing the accuracy of legal evaluations. *Assessment*, 10, 428 –

- Faust, D. (2007). Decision research can increase the accuracy of clinical judgment and thereby improve patient care. In S. O. Lilienfeld & W. T. O'Donohue (eds.), *The great ideas of clinical science: 17 principles that every mental health professional should understand* (pp. 49 – 76). New York: Routledge.
- Faust, D., Bridges, A., & Ahern, D. C. (in preparation). Reasons complex pattern analysis cannot be successful in neuropsychology and superior alternatives.
- Faust, D., Ziskin, J., & Hiers, J. B. (1991). *Brain damage claims: Coping with neuropsychological evidence*. Los Angeles: Law & Psychology Press.
- Gaier, E. L., & Lee, M. C. (1953). Pattern analysis: The configural approach to predictive measurement. *Psychological Bulletin*, 50, 140 – 148.
- Glutting, J. J., McDermott, P. A., Watkins, M. M., Kush, J. C., & Konold, T. R. (1997). The base rate problem and its consequences for interpreting children's ability profiles. *School Psychology Review*, 26, 176 – 188.
- Graham, J. R. (2006). *MMPI-2: Assessing personality and psychopathology* (4th ed.). New York: Oxford University Press.
- Grant, B. F., Dawson, D. A., Stinson, F. S., Chou, S. P., Dufour, M. C., & Pickering, R. P. (2004). The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991 – 1992 and 2001 – 2002. *Drug and Alcohol Dependence*, 74, 223 – 234.
- Grove, W. H., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12, 19 – 30.

- Haahr, M. (n.d.) *Random integer generator*. Retrieved May 1, 2007 from <http://www.random.org/integers/>
- Heaton, R. K., Grant, I., & Matthews, C.G. (1986). Differences in neuropsychological test performance associated with age, education and sex. In I. Grant & K. M. Adams (eds.), *Neuropsychological assessment in neuropsychiatric disorders: Clinical methods and empirical findings* (pp. 100 – 120). New York: Oxford University Press.
- Heaton, R. K., Grant, I., & Matthews, C. G. (1992). *Comprehensive norms for an expanded Halstead-Reitan battery: Demographic corrections, research findings, and clinical applications*. Odessa, FL: Psychological Assessment Resources.
- Heaton, R. K., Miller, W., Taylor, M. J., & Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological norms for African-American and Caucasian adults – Professional Manual*. Lutz, FL: Psychological Assessment Resources.
- Heaton, R. K., Taylor, M. J., & Manly, J. (2003). Demographic effects and use of demographically corrected norms with the WAIS-III and WMS-III. In D. S. Tulsky et al. (eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 181 – 210). New York: Academic Press.
- Hervey, A. S., Epstein, J. N., & Curry, J. F. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology*, 18, 485 – 503.
- Hiscock, M. (2007). The Flynn effect and its relevance to neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, 29, 514 – 529.
- Hogan, T. P. (2005). 50 widely used psychological tests. In G. P. Koocher, J. C.

- Norcross, & S. S. Hill III (eds.), *Psychologists' desk reference* (2nd ed., pp. 101 – 104). New York: Oxford University Press.
- International Consortium on Psychiatric Epidemiology. (2007, July 19). *Table 1. Lifetime prevalence of DSM-IV/WHM-CIDI disorders by sex and cohort*. Retrieved from http://www.hcp.med.harvard.edu/ncs/ftpd/dir/table_ncsr_LTPrevgenderxage.pdf
- Kahneman, D., Slovic, P., & Tversky, A. (eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Kandel, D., Chen, K., Warner, L. A., Kessler, R. C., & Grant, B. (1997). Prevalence and demographic correlates of symptoms of last year dependence on alcohol, nicotine, marijuana and cocaine in the U.S. population. *Drug and Alcohol Dependence*, 44, 11 – 29.
- Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, H. J., & Fischer, J. S. (2004). *Neuropsychological assessment* (4th ed.). New York: Oxford University Press.
- Martínez-Arán, A., Vieta, E., Reinares, M., Colom, F., Torrent, C., Sánchez-Moreno, J., Benabarre, A., Goikolea, J. M., Comes, M., & Salamero, M. (2004). Cognitive function across manic or hypomanic, depressed, and euthymic states in bipolar disorder. *American Journal of Psychiatry*, 161, 262 – 270.
- Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment* (2nd ed.). New York: Oxford University Press.
- National Center for Injury Prevention and Control. (2006, September 19). *TBI overview*. Retrieved from <http://www.cdc.gov/ncipc/tbi/Overview.htm>
- National Institute on Alcohol Abuse and Alcoholism. (2004, June 10). *Alcohol abuse*

increases, dependence declines across decade: Young adult minorities emerge as high-risk subgroups. Retrieved from

<http://www.niaaa.nih.gov/NewsEvents/NewsReleases/NESARCNews.html>

National Institute on Alcohol Abuse and Alcoholism. (2005a, January). *Twelve-month prevalence and population estimates of DSM-IV alcohol dependence by age, sex, and race-ethnicity: Unites States, 2001 – 2002 (NESARC).* Retrieved September 25, 2007 from <http://www.niaaa.nih.gov/Resources/DatabaseResources/QuickFacts/AlcoholDependence/abusdep2.htm>

National Institute on Alcohol Abuse and Alcoholism. (2005b, January). *Twelve-month prevalence and population estimates of DSM-IV alcohol abuse by age, sex, and race-ethnicity: Unites States, 2001 – 2002 (NESARC).* Retrieved September 25, 2007 from <http://www.niaaa.nih.gov/Resources/DatabaseResources/QuickFacts/AlcoholDependence/abusdep1.htm>

Nelson, C. B., Heath, A. C., & Kessler, R. C. (1998). Temporal progression of alcohol dependence symptoms in the U.S. household population: Results from the national comorbidity survey. *Journal of Consulting and Clinical Psychology, 3*, 474 – 483.

Nyenhuis, D. L., Gorelick, P. B., Geenen, E. J., Smith, C. A., Gencheva, E., Freels, S., & de Toledo-Morrell, L. (2004). The pattern of neuropsychological deficits in vascular cognitive impairment – no dementia (vascular CIND). *The Clinical Neuropsychologist, 18*, 41 – 49.

Perlick, D., & Atkins, A. (1984). Variations in the reported age of a patient: A source of bias in the diagnosis of depression and dementia. *Journal of Consulting and Clinical Psychology, 52*, 812 – 820.

- Petrakis, I. L., Gonzalez, G., Rosenheck, R., & Krystal, J. H. (2002). Comorbidity of alcoholism and psychiatric disorders: An overview. *Alcohol Research & Health*, 26, 81 – 89.
- Porter, R. J., Gallagher, P., Thompson, J. M., & Young, A. H. (2003). Neurocognitive impairment in drug-free patients with major depression disorder. *The British Journal of Psychiatry*, 182, 214 – 220.
- Reed, B. R., Mungas, D. M., Kramer, J. H., Ellis, W., Vinters, H. V., Zarow, C., Jagust, W. J., Chui, H. C. (2007). Profiles of neuropsychological impairment in autopsy-defined Alzheimer's disease and cerebrovascular disease. *Brain*, 130, 731 – 739.
- Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan neuropsychological test battery* (2nd ed.). Tucson, AZ: Neuropsychology Press
- Rohling, M. L., & Demakis, G. J. (2006). A meta-analysis of the neuropsychological effects of occupational exposure to mercury. *The Clinical Neuropsychologist*, 20, 108 – 132.
- Sachdev, P. S., Brodaty, H., Valenzuela, M. J., Lorentz, L., Looi, J. C., Wen, W., & Zagami, A. S. (2004). The neuropsychological profile of vascular cognitive impairment in stroke and TIA patients. *Neurology*, 23, 912 – 919.
- Salmond, C. H., & Sahakian, B. J. (2005). Cognitive outcome in traumatic brain injury survivors. *Current Opinion in Critical Care*, 11, 111 – 116.
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford University Press.
- The Psychological Corporation. (2002a). *WAIS-III – WMS-III technical manual*. San

Antonio, TX: Author.

The Psychological Corporation. (2002b). *WAIS-III – WMS-III – WIAT-II Writer*. San Antonio, TX: Author.

Torres, I. J., Boudreau, V. G., & Yatham, L. N. (2007). Neuropsychological functioning in euthymic bipolar disorder: A meta-analysis. *Acta Psychiatrica Scandinavica*, 116, 17 – 26.

United States Census Bureau. (2007, March 15). *Table 1. Educational attainment of the population 15 years and over, by age, sex, race, and Hispanic origin: 2006*. Retrieved from <http://www.census.gov/population/www/socdemo/education/cps2006.html>

Voglmaier, M. M., Seidman, L. J., Niznikiewicz, M. A., Dickey, C. C., Shenton, M. E., & McCarley, R. W. (2005). A comparative profile analysis of neuropsychological function in men and women with schizotypal personality disorder. *Schizophrenia Research*, 74, 43 – 49.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale – Third Edition: Administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale – Third Edition: Administration and scoring manual*. San Antonio, TX: The Psychological Corporation.

Wedding, D., & Faust, D. (1989). Clinical judgment and decision making in neuropsychology. *Archives of Clinical Neuropsychology*, 4, 233 – 265.

Zakzanis, K. K., Kaplan, E. F., & Leach, L. (1999). *Neuropsychological differential diagnosis*. Lisse, the Netherlands: Swets & Zeitlinger.