Development of the O*NET™ Paper-and-Pencil Work Importance Locator

Rod McCloy, Gordon Waugh, Gina Medsker
Human Resources Research Organization

Janet Wall
ACT

David Rivkin and Phil Lewis
National Center for O*NET Development

National Center for O*NET Development
Employment Security Commission
Post Office Box 27625
Raleigh, North Carolina 27611
e-mail: onet@ncmail.net

1999

Addendum

April, 2001

This technical development report contains information based on O*NET 98, whose occupational classification system contains 1,122 occupational units (OUs) based on the Occupational Employment Statistics classification system.

Since the writing of this report, the O*NET 3.0 database has been developed. The major difference between this database and the O*NET 98 database is its compatibility with the 1998 Standard Occupational Classification (SOC) System.[1] By making O*NET 3.0 compatible with the SOC system, the O*NET 3.0 database contains 974 occupations. Please note that the U.S. Office of Management and Budget has mandated that all federal agencies' occupational classification systems be compatible with the 1998 SOC system.

All O*NET 98 data have been converted to O*NET 3.0 data and verified.

O*NET 3.0 and O*NET OnLine, a Web-based application that allows users to view and use the O*NET 3.0 database, can be accessed via the National Center for O*NET Development's Website, **www.onetcenter.org** .

[1] United States Department of Labor, Bureau of Labor Statistics. (1999). *Revising the Standard Occupational Classification System*. Washington, DC: Author.

Table of Contents

Table of Contents (continued)

Appendices

List of Tables

Table of Contents (continued)

Acknowledgments

The O*NET™ Work Importance Locator was produced and funded by the O*NET project of the U.S. Department of Labor, Employment and Training Administration, Office of Policy and Research (OPR) under the direction of Gerard F. Fiala, Administrator. The O*NET project is directed by Jim Woods, Office of Policy and Research, and Donna Dye, Office of Workforce Security.

The instrument was created through Department of Labor grants with the National O*NET Consortium, National Center for O*NET Development; the North Carolina Employment Security Commission, Southern Assessment Research and Development Center; the Utah Department of Employment Security, Western Assessment Research and Development Center; the New York Department of Labor, Eastern Assessment Research and Development Center; and the Michigan Employment Security Commission, Northern Assessment Research and Development Center.

The development of the O*NET Work Importance Locator was completed under the direction of David Rivkin and Phil Lewis. The Human Resources Research Organization (HumRRO) was the primary contractor in developing the instrument. Their team was led by Dr. Rodney McCloy. Dr. Gary Carter, John Boyle, and Dan Connell initiated and participated in many of the developmental research efforts.

Numerous research phases were conducted in developing the O*NET Work Importance Locator. Each phase required multiple steps and the participation of a wide variety of staff, contractors, and pilot sites. Grateful acknowledgment for their prominent roles in this process is made to Jeannie Stroupe, Helen Tannenbaum, Ray Dobrin, Enid Israelson, Ann Kump, Diana Williams, Dr. Steve Atkins, Ronald Boese, Brenda Dunn, and Jerry Pickett.

A special thanks is given to Dr. René V. Dawis, who served as the overall technical advisor for the project and whose previous research inspired the work.

Executive Summary

The U.S. Department of Labor's (USDOL's) Office of Policy and Research has developed the Occupational Information Network (O*NET), a comprehensive system for collecting, organizing, describing, and disseminating data on occupational characteristics and worker attributes (see *O*NET Final Technical Report*, 1998). O*NET is the replacement for the *Dictionary of Occupational Titles* (DOT; U.S. Department of Labor, 1991). O*NET includes the Content Model, a skills-based structure that serves as the framework for organizing the information describing the world of work presented within O*NET (see *Development of Prototype Occupational Information Network (O*NET) Content Model*, 1995). The Office of Policy and Research initiated several projects aimed at producing valid and reliable data covering a majority of the variables described in the O*NET Content Model. This report focuses on the effort to generate work-related values information included in the Worker Characteristics domain of the model (i.e., identifying features of employment which O*NET users may value or view as personally important). Inclusion of such work importance (i.e., work values) information within O*NET provides an important data set for career guidance and research.

It is important to note that USDOL's Office of Policy and Research has developed career exploration and development tools in an effort to create more complete, flexible services. The career exploration tools link directly to O*NET. For example, the Work Importance instruments USDOL has developed will enable users to link their results directly to the work values information provided in O*NET. These materials will allow individuals to use a variety of assessment information about themselves (e.g., vocational interests, skills and abilities, education, experience, as well as work values) to explore careers either individually, with a career counselor, or in a group.

The U.S. Department of Labor's (USDOL's) National O*NET Consortium contracted with the Human Resources Research Organization (HumRRO) to complete a project aimed at developing two measures of work values. The work values measures would be a part of a set of measures to complement the USDOL Occupational Information Network (O*NET), a new computerized database of occupational information that will replace the *Dictionary of Occupational Titles* (DOT; U.S. Department of Labor, 1965, 1977, 1991).

The Work Values project had three parts:

a) the design and evaluation of a computerized measure of work values,
b) the design and evaluation of a paper-and-pencil measure of work values, and
c) the determination of the work values score profiles (i.e., occupational reinforcer patterns) of the 1,122 Occupational Units (OUs) of the O*NET.

The purpose of this report is to describe and document the development and evaluation of the paper-and-pencil work values measure (WIL-P&P). Two separate reports describe the development of the computerized work values measure (WIP-C; please see McCloy, Waugh, Medsker, Wall, Rivkin, and Lewis, 1999a) and the occupational reinforcer patterns (ORPs; please see McCloy, Waugh, Medsker, Wall, Rivkin, and Lewis, 1999b).

The WIL-P&P was modeled on the Minnesota Importance Questionnaire (MIQ; Rounds, Henly, Dawis, Lofquist, and Weiss, 1981). The measures were designed for self-assessment, thus affording

individual job seekers the chance to receive immediate feedback about the characteristics of occupations they find most important (i.e., the needs that the client wishes the occupation to reinforce—presumably enhancing job satisfaction). Some items on the WIL-P&P were modified from the wording on the MIQ so that the statements could more closely match the descriptors used in O*NET.

WIL-P&P development involved three studies:

- Pre-Pilot Study – 21 employment center clients completed a draft version of the WIL-P&P. Results were used to improve the profiler.
- Pilot Study – 48 employment center participants completed the improved WIL-P&P. Test administrators took notes while observing the participants taking the measure, and the participants provided ratings of and comments about the measure on reaction forms. This information was used to make further improvements to the WIL-P&P.
- Main Study – the WIL-P&P, WIP-C, and MIQ were administered to employment center clients and junior college students at 23 sites. Respondents took two of the three instruments so that information on the same respondents taking different measures would be available. Some respondents also took the same instruments at two points in time, so that test-retest reliabilities could be computed. Respondents providing useable data totaled 1,199 for the WIL-P&P, 941 for the WIP-C, and 550 for the MIQ.

Adequate numbers of male, female, white, Hispanic, and African American respondents provided useable data to allow separate statistical analyses for these subgroups. Results indicated that the WIL-P&P required less time to complete than did the WIP-C (less by 12-14 minutes) and MIQ (less by 6-7 minutes).

Analyses from these studies revealed the following information:

- Level of education was related to differences on the values of Achievement, Autonomy, and Safety. Those with more education tended to have higher expressed values for Autonomy and Achievement and lower expressed values for Safety.
- Some racial and ethnic differences were found. African Americans tended to express higher value for Status than did whites.
- Females tended to express higher value for Safety and Altruism than did males.
- Because the WIL-P&P requires respondents to rank order the 20 needs, it uses ipsative scoring. Ipsative scoring limits item intercorrelations and, hence, lowers internal consistency reliability estimates.
- As expected, coefficient alpha internal consistency reliabilities were quite low for the WIL-P&P. Adjustments to the item scores aimed at reducing the effects of ipsatization did raise internal consistencies for the six value scales.
- Adjustments to reduce the effects of ipsatization also improved correlations among the six values on the WIL-P&P and WIP-C. These correlations (which represent alternate forms reliability) are reasonably high given the properties of these instruments. Test-retest reliability results showed moderate score consistency over time.

- Respondents had the same top work value 62 percent of the time, and 80 percent of the time the highest-ranked value at Time 1 was ranked either first or second at Time 2. Profiles of need and value ranks between Time 1 and Time 2 demonstrated very respectable correspondence.
- Validity of the WIL-P&P scores was investigated by comparing its scores to those from the MIQ for respondents who had completed both measures. Correlations between the need items and values for the WIL-P&P and MIQ were low to moderate. For those respondents who completed both the WIL-P&P and MIQ, 57 percent had the same top value on both measures. The top value on the WIL-P&P was the first or second value on the MIQ 79 percent of the time. This lends support to the use of the top scoring value of the WIL-P&P in career exploration.
- WIL-P&P items with the lowest relationship to those on the MIQ were those items that underwent wording changes.
- For more than 79 percent of respondents, the top-ranked value on the WIL-P&P was one of the top two ranked values on the MIQ. These analyses suggest reasonable links to previously demonstrated validity evidence associated with the MIQ.
- The profile correlations between the WIL-P&P and MIQ were higher for the needs than for the work values. The profile correlation for the needs was .67, and for the values it was .57. This indicates that the needs themselves may be useful in understanding a person's work values, and techniques might be developed for using needs in the exploration process.
- Relationships of the WIL-P&P and the MIQ are sufficiently strong to support the use of the WIL-P&P as a measure of work values.
- Taken together and considering the difficulties interpreting traditional statistical analyses on measures using ipsative scoring, the WIL-P&P appears to offer very useful information at both the work values and needs levels to individuals exploring potential careers or occupations.

The WIL-P&P shows promise as an engaging exercise for helping individuals identify their work values and needs for purposes of career exploration. Low reliabilities, due largely to problems related to ipsative scoring and moderate-to-good decision consistency, would suggest that individuals may be enlightened about themselves through the use of this instrument. As with any measure with these statistical qualities, the user should be cautioned to use additional instruments for gaining information for career exploration. Critical career decisions justify using a "whole-person" approach, (i.e., using diverse information about oneself to explore careers and the world of work).

Chapter 1. Introduction

Overview and Purposes of Project

The U.S. Department of Labor's (USDOL's) Office of Policy and Research has developed the Occupational Information Network (O*NET), a comprehensive system for collecting, organizing, describing, and disseminating data on occupational characteristics and worker attributes (see *O*NET Final Technical Report*, 1998). O*NET is the replacement for the *Dictionary of Occupational Titles* (DOT; U.S. Department of Labor, 1991). O*NET includes the Content Model, a skills-based structure that serves as the framework for organizing the information describing the world of work presented within O*NET (see *Development of Prototype Occupational Information Network (O*NET) Content Model*, 1995). The Office of Policy and Research initiated several projects aimed at producing valid and reliable data covering a majority of the variables described in the O*NET Content Model. This report focuses on the effort to generate work-related values information included in the Worker Characteristics domain of the model (i.e., identifying features of employment which O*NET users may value or view as personally important). Inclusion of such work importance (i.e., work values) information within O*NET provides an important data set for career guidance and research.

It is important to note that USDOL's Office of Policy and Research has developed career exploration and development tools in an effort to create more complete, flexible services. The career exploration tools link directly to O*NET. For example, the Work Importance instruments USDOL has developed will enable users to link their results directly to the work values information provided in O*NET. These materials will allow individuals to use a variety of assessment information about themselves (e.g., vocational interests, skills and abilities, education, experience, as well as work values) to explore careers either individually, with a career counselor, or in a group. Examples of career exploration materials USDOL has developed include:

1. The **O*NET Interest Profiler**, which measures six broad vocational interest areas that coincide with the RIASEC model (Holland, 1997).
2. The **O*NET Ability Profiler**, which measures nine different abilities directly linked to job performance.
3. The **O*NET Computerized Work Importance Profiler**, which allows individuals to identify work values that are important to them (e.g., features of employment they personally value or find to be important).

Most of these tools will be available in both automated and paper formats to meet the needs of a variety of users. The U.S. Department of Labor's National O*NET Consortium contracted with the Human Resources Research Organization (HumRRO) to assist in the completion of a project aimed at developing two measures of work values. The work values measures would be part of a set of measures to complement the USDOL's Occupational Information Network (O*NET), a new computerized database of occupational information that will replace the *Dictionary of Occupational Titles* (DOT; U.S. Department of Labor, 1965, 1977, 1991).

The O*NET Work Importance Locator project had three distinct parts:

- Part I - to design and evaluate a new computerized measure of work values
- Part II - to design and evaluate a similar paper-and-pencil measure of work values
- Part III - to determine the work values score profiles, or Occupational Reinforcer Patterns (ORPs), of the 1,122 Occupational Units (OUs - groups of occupations from the DOT) contained in O*NET.

This report describes the development and evaluation process of the paper-and-pencil measure of work values (WIL-P&P). Two additional reports describe the development and evaluation of the computerized work values measure (WIP-C; please see McCloy, Waugh, Medsker, Wall, Rivkin, and Lewis, 1999a) and the development of the Occupational Reinforcer Patterns (ORPs; please see McCloy, Waugh, Medsker, Wall, Rivkin, and Lewis, 1999b) of the OUs in O*NET.

The O*NET Work Importance Locator (WIL-P&P) was designed to be available as a stand-alone measure of work values and as a piece of USDOL's collection of career exploration tools. At present the collection includes: a) a set of career assessment tools, b) occupational classification structures that are accessed via assessment data, and c) occupational information from O*NET. Assessment tools to be made available are the Work Importance instruments, an Ability Profiler, and two Interest Profilers (computerized and paper-and-pencil formats). Individuals requesting career guidance or wishing to engage in career exploration will be guided toward occupational groupings based on their scores on the various assessment tools.

The work values instruments developed in this project measure individuals' work needs and values and are based on a previously developed measure of work values, the Minnesota Importance Questionnaire (MIQ; Rounds, Henly, Dawis, Lofquist, and Weiss, 1981). All items on the WIP-C and WIL-P&P were based on items from the MIQ (with some changes in wording to be consistent with the wording developed for O*NET). The MIQ was not used because it required complex scoring, resulting in an answer sheet that must be sent to the publisher to be scored. The first purpose of the work values project was to develop new work values measures that could be scored immediately. The work values instruments were specifically designed to make scores immediately available to the respondent. In the WIP-C, the computer calculates the scores and displays them on the screen; in the WIL-P&P, the respondent calculates the scores by hand after completing the instrument. Additional materials provide interpretive information to the user—including groups of occupations that "match" (or are most likely to satisfy) the user's work values.

This report focuses on the development of the WIL-P&P, which involved three studies: a Pre-Pilot Study, Pilot Study, and Main Study. Data for these studies were collected at four points in time (referred to as "Time 1," "Time 2," "Time 3," and "Time 4"). The purpose of each data collection is noted below:

- Time 1, the Pre-Pilot Study: conducted to examine the initial development of the WIL-P&P based on the MIQ.
- Time 2, the Pilot Study: conducted to further refine the new measure to prepare it for use in the Main Study.
- Times 3 and 4, the Main Study: used to examine the test-retest reliability of the new measure.

Data from Time 3 also were used to evaluate other psychometric characteristics of the measure, including:

a) descriptive statistics on respondents, items, and value scales;

b) comparative statistics for the three work values measures (WIP-C, WIL-P&P, and MIQ);

c) response consistencies and errors;

d) completion times;

e) relationships of values to education level, race/ethnic group, and gender;

f) consistency of needs and values expressed on the three need/value measures;

g) internal consistency within measures; and

h) factor analyses of how need items group together into values.

The Pre-Pilot, Pilot, and Main Study are discussed in several chapters of this report that immediately follow this first chapter.

The following sections describe historical and theoretical background on the measurement of work needs and values and ORPs. This information should help explain the foundation on which the newly developed WIL-P&P is based.

Historical Background on Work Values: The Theory of Work Adjustment

The studies on work adjustment began in 1957 by the Work Adjustment Project at the University of Minnesota under the direction of René Dawis and Lloyd Lofquist. The impetus of their research was to explore aspects of an individual's work adjustment and to develop assessment tools that measure and predict an individual's adjustment to work. The Theory of Work Adjustment (TWA) was first conceptualized in 1964 (Dawis, Lofquist, and Weiss, 1968; Weiss, Dawis, England, and Lofquist, 1964) and was given more comprehensive treatment in the book *Adjustment to Work* (Lofquist and Dawis, 1969). Early work on the theory was supported by the Rehabilitation Services Administration, Social and Rehabilitation Service, U.S. Department of Health, Education, and Welfare (Dawis and Lofquist, 1984).

The Theory of Work Adjustment is a comprehensive model of vocational adjustment based on the concept of correspondence between individual and environment (Dawis and Lofquist, 1984). The TWA postulates that vocational needs and abilities are instrumental elements of the work personality, while ability requirements and reinforcer systems are significant aspects of the work environment. The degree of correspondence between an individual's skills and abilities with the ability requirements of the work environment will predict *satisfactoriness*. In addition, the degree of correspondence between an individual's needs and values and the reinforcers available in the work environment will predict satisfaction with work. Dawis and Lofquist summarized the TWA as follows:

- Work is conceptualized as an interaction between an individual and a work environment.
- The work environment requires that certain tasks be performed, and the individual brings skills to perform the tasks.
- In exchange, the individual requires compensation for work performance and certain preferred conditions, such as a safe and comfortable place to work.

- The environment and the individual must continue to meet each other's requirements for the interaction to be maintained. The degree to which the requirements of both are met may be called *correspondence*.
- Work adjustment is the process of achieving and maintaining correspondence. Work adjustment is indicated by the satisfaction of the individual with the work environment and by the satisfaction of the work environment with the individual (the individual's satisfactoriness).
- Satisfaction and satisfactoriness result in tenure, the principal indicator of work adjustment. Tenure can be predicted from the correspondence of an individual's work personality with the work environment.
- Work personalities and work environments can be described in terms of structure and style variables that are measured on the same dimensions (Dawis and Lofquist, 1984, pp. 9-10).

To completely operationalize the TWA, one must measure characteristics of both the individual and the work environment to determine the amount of correspondence between the two. The TWA considers both abilities and vocational needs to be instrumental characteristics of individuals that are relevant to determine the correspondence between the individual and work environment. Instruments such as the General Aptitude Test Battery (GATB; U.S. Department of Labor, 1970), which measures work-related abilities, and the MIQ (Rounds et. al., 1981), which measures workers' needs and values, are both examples of worker assessments. Complementary to the ability and need characteristics of individuals are the ability requirements and reinforcer systems of work environments. To assess the degree of correspondence between the needs of an individual and the reinforcer systems of occupational environments, a third measurement tool, the Minnesota Job Description Questionnaire (MJDQ; Borgen, Weiss, Tinsley, Dawis, and Lofquist, 1968), can be used. The MJDQ provides a description of the work environment in need/value terms. The need-reinforcer statements included in the MJDQ are very similar to the statements included in the MIQ (Dawis and Lofquist, 1984) to enable the individual's needs/values to be matched to what the work environments provide in terms of need and value satisfaction or fulfillment. To assess the ability requirements of jobs, a fourth tool, job analysis, can be used (specifically, job analysis tools yielding job profiles compatible with worker ability profilers like the GATB).

The MIQ, (an instrument upon which the WIL-P&P is modeled) is described in the section to follow.

The Minnesota Importance Questionnaire (MIQ)

The MIQ was based on the N-Factors Questionnaire, which in turn was based on a questionnaire by Schaffer (1953). The MIQ has been through three revisions since its creation in 1964. It is designed to provide information about an individual's needs and values. Persons completing the MIQ are asked to indicate the relative importance, to them, of 21 vocationally relevant need reinforcers (e.g., receiving recognition, having steady employment). The need-reinforcer dimensions measured by the MIQ have been found to be important to job satisfaction (Gay, Weiss, Hendel, Dawis, and Lofquist, 1971). The 21 needs can be grouped into 6 value dimensions (derived through factor analysis) named Achievement, Comfort, Status, Altruism, Safety, and Autonomy (though these names were later changed - see Chapter 8 and Table 23).

The original form of the MIQ consisted of 20 scales of 5 items each. Respondents were asked to rate the importance of specific aspects of work on a 5-point Likert scale. This form produced negatively skewed distributions of scale scores and yielded high intercorrelations among scale scores (Gay et. al., 1971). Ipsative forms of the MIQ, including a paired-comparison form and a multiple ranking form, were developed to overcome these deficiencies.

The Multiple Rank Order 5 (MRO5) version of the MIQ was the basis for the automated version of the O*NET Work Importance Locator. As discussed later in Chapter 3, the paper/pencil version uses a card-sorting format that differs rather markedly from the other versions of the MIQ. The MRO5 produces scores for 21 needs. Related needs are combined into six work values scales. According to the TWA, needs and values with high scores are important to a person's satisfaction; needs and values with low scores have little or no effect upon a person's satisfaction. For example, the level of Independence inherent in a specific job will greatly affect the satisfaction of people who have high scores on the Independence need of the MIQ, but it will have little effect on people who have low scores.

The multiple ranking form (MRO5) consists of two sections: a ranked section and an absolute zero section. In the ranked section, stimuli are grouped into a *balanced incomplete block* in which each stimulus is paired with every other stimulus an equal number of times. The 21 statements are presented in 21 blocks. Each block has five statements. Within each block, respondents rank-order the statements according to the relative importance of the needs on their ideal jobs. Each need appears in five blocks and with every other need exactly once. Using this format, 210 paired comparisons can be reduced to 21 blocks of 5 stimuli each. This format produces profiles similar to those provided using the paired-comparison form and reduces administration time and the number of judgments required of respondents (Rounds, Miller, and Dawis, 1978).

The responses in this first part of the MIQ indicate the *relative* importance of the 21 needs. That is, the scoring thus far is ipsative. The scores do not show the *absolute* importance of each need. For some people, however, only a few needs are important; for others, most needs are important. Therefore, the second part of the MIQ asks the respondents to rate each need as either *important* or *not important*. This places each need on an absolute scale. Possible scores range from -4.00 to +4.00, although each *person's* scores will have a range no greater than 4.00.

Because each need appears in five different blocks, a respondent's consistency can be computed. The following example demonstrates inconsistent responding: Need A is ranked higher than Need B, Need B is ranked higher than Need C, and Need C is ranked higher than Need A. This is called a *circular triad*. Within the MIQ, there are 440 triads of needs. For the MIQ, the percentage of circular triads (PCT) is computed as the percentage of the 440 triads that are circular. The converse of this statistic is the *coefficient of consistency*, which is the proportion of triads that are *not* circular. If the coefficient of consistency is less than .50, then it is assumed that the respondent is either responding carelessly or is unsure of the importance of his or her needs. Score results for the MIQ include scores on the 21 needs, scores on the six values, and the percentage of circular triads (Rounds et. al., 1978).

The MIQ was not selected as one of the tools for direct use in the O*NET program because of several factors. First, the items on the instrument did not fit exactly with the work values information included in the O*NET model. It was a prime objective of the USDOL to provide opportunities for

clients to use their results (profiles) to explore occupations in O*NET. More importantly, an instrument that could be self-scored, self-administered, and self-interpreted was critical to the practical success of the effort. The MIQ requires complex, and thus machine, scoring. As a result, feedback was not immediately available. Still, the MIQ served as an excellent basis of the development of the O*NET Work Values assessment instruments. A paper/pencil version was needed so that people could take the instrument on their own and so that users who did not have computers available could be served as well.

Chapter 2. Overall Research Design for Developing the WIL-P&P

Introduction

The research design for developing the WIL-P&P (Paper-and-Pencil O*NET Work Importance Locator) involved three studies: a Pre-Pilot Study, a Pilot Study, and a Main Study. Data for these studies were collected at four points in time (referred to as "Time 1," "Time 2," "Time 3," and "Time 4") as shown in Table 1. Table 1 also shows the sample size (i.e., number of respondents) involved in each of the phases of the project. Each phase of the research design is discussed briefly below and in greater detail in the following chapters.

Pre-Pilot Study to Develop New Measure

Prior to the Pre-Pilot Study at Time 1, a draft version of the WIL-P&P was developed using items from the MIQ as a basis. At Time 1, in the Pre-Pilot study, 21 employment center clients in North Carolina completed the WIL-P&P. Errors made by the clients while completing the instrument and clients' comments on a Participant Reaction Questionnaire were used to determine how the instrument could be improved. Modifications were made as a result of this testing. More detailed information about the Pre-Pilot is provided in Chapter 3.

Pilot Study to Refine New Measure

The purpose of the Pilot Study (Time 2) was to get feedback and information useful for further refinement of the WIL-P&P. Forty-eight employment center participants took the WIL-P&P at an employment center in Utah. Administrators observed participants completing the instruments, and then the participants completed Participant Reaction Questionnaires. Once again, comments from the questionnaires were used to guide modifications to the WIL-P&P. More detailed information about the Pilot Study is provided in Chapter 3.

Main Study: Psychometric Properties of the New Measures

Data for the Main Study on the WIL-P&P were collected at Times 3 and 4. In this Main Study, the WIL-P&P, WIP-C, and MIQ were administered to employment center clients and junior college students. Respondents in the Main Study took two of the three instruments (WIL-P&P and MIQ, or WIP-C and MIQ, or WIL-P&P and WIP-C) so that information on the same respondents taking different measures would be available. The design attempted to balance the pairing and ordering of measures taken by respondents so that all pairs and orders would be represented.

The subsamples of respondents numbered 5 and 6 in Table 1 were junior college students who took both the WIL-P&P and WIP-C at Times 3 and 4. These data were used to examine the test-retest reliability of the profilers. For the test-retest reliability analyses, the sample size was 232 for the WIL-P&P and 213 for the WIP-C. Time 3 data from these students and Time 3 data from employment center clients (subsamples 7 through 12 in Table 1) were used to analyze other psychometric properties of the profilers, including: a) descriptive statistics on respondents, items, and value scales; b) comparative statistics for the three measures (WIP-C, WIL-P&P, and MIQ); c) response consistencies and errors; d) completion times; e) relationships of values to education,

race/ethnic group, and gender; f) consistency of needs and values expressed on the measures; g) internal consistency within measures; and h) factor analyses of how need items group together into values. Based on Time 3 data, the overall sample sizes in the Main Study for these analyses were 1,199 for the WIL-P&P, 941 for the WIP-C, and 550 for the MIQ.

Table 1. ***Sample and Subsample Sizes for the Research Project***

Study	Time 1 *n* per Subsample	Time 2 *n* per Subsample	Time 3 *n* per Subsample	Time 4 *n* per Subsample	Total *n* in Study
Pre-Pilot	1. C = 20				C+P = 31
	2. P = 21				
Pilot		3. C = 43			C+P = 91
		4. P = 48			
Main Study: Test-Retest			5. P-C=148-133	5. P-C=121-111	C=213
			6. C-P=132-142	6. C-P=1-2-111	P=232
Main Study: Other Analyses			5. P-C=148-133		C=941
			6. C-P=132-142		P=1199
			7. P-C=269-210		MIQ=550
			8. C-P=218-276		
			9. P-MIQ=174-141		
			10. MIQ-P=158-190		
			11. C-MIQ=120-121		
			12. MIQ-C=130-128		

Note: C = WIP-C, P = WIL-P&P. Unique subsamples of subjects are numbered 1-12; the numbers 5 and 6 refer to two subsamples who completed the WIL-P&P and WIP-C at both Times 3 and 4. The numbers shown after the measure labels are the number of people included in the analyses for those instruments. For example, in subject group 12, 130 MIQ cases and 128 WIP-C cases were used in the analyses; these subjects took the MIQ first, immediately followed by the WIP-C. Some completed measures were excluded from analyses due to excessive response inconsistency (>50%), missing or invalid responses, or more than one math error during self-scoring.

Chapters 4 through 8 of this report discuss the data and results from this Main Study. Chapter 4 discusses the sample of respondents included in analyses and presents descriptive statistics on the respondents. Chapter 5 presents descriptive statistics for the WIL-P&P. Chapter 6 discusses the relationships between respondents' reported values and their education level, race/ethnic group, and gender. Chapter 7 provides information on the reliability of the WIL-P&P and discusses the results of different types of reliability analyses. Chapter 8 presents results of analyses used to preliminarily assess the construct validity of the WIL-P&P.

Chapter 3. Development of the WIL-P&P in the Pre-Pilot and Pilot Studies

Introduction

The purpose of this chapter is to describe the initial stages of development of the WIL-P&P. This section briefly describes the development of the items and the scoring procedures. A description of the lessons learned from the pre-pilot and pilot studies is provided.

Development of the Content and Scoring of the Items

All items on the Work Values profiles developed during this project (i.e., the WIP-C and WIL-P&P) were based on the original 21 need statements from the MIQ. Many of the items on the new work values measures, however, are worded somewhat differently from their MIQ source items, because several of the MIQ items were reworded during the O*NET project. Table 2 provides the new wording of the need statements, the original wording of each statement from the MIQ, and the reason for the wording change. Because the WIL-P&P and WIP-C were to be used with O*NET, the O*NET rewording was adopted for the new profilers.

In the original design of the WIL-P&P, the respondent sorted 21 need statements from the MIQ into seven card piles of three cards each. This design preserved all 21 of the MIQ needs, but required more complex scoring. To compute the score for each value, the respondent had to refer to a table (which showed the weighted score for each need) and then add up three two-digit numbers. The table made it unnecessary for the respondent to do any multiplication. The numbers in the table took into account the pile number and the number of items in the value scale.

Though the seven-pile (or point) scoring procedure was given serious consideration, it was thought to be too complicated for some respondents because they would have to look up numbers in a table and add two-digit numbers. Therefore, a second version of the measure was developed that omitted the table lookup and required the addition of only one-digit numbers. This version was very similar to the final WIL-P&P.

There were two possible disadvantages of this design: a) one of the items would have to be dropped to keep the math simple, and b) it required multiplication. After a discussion with one of the MIQ's co-authors, it was concluded that the loss of one carefully chosen item (Item 16 on the need for social status) would have very little effect on the quality of the measure. Item 16 was chosen because a) the wording had changed at least moderately from the original MIQ wording, and b) its deletion did not yield another two-item value (i.e., it was from a scale with at least four items). With the rewording, the item seemed to have a different meaning from the other items meant to measure social status (see Table 2). With regard to the computation required, the most difficult multiplications that might be required with only 20 items were 2×15 and 3×10.

Table 2. ***Wording Changes from the O*NET Project for the 21 Need Statements***

WIL-P&P	Original MIQ Items	Reasons for Change Made
1. On my ideal job it is important that I make use of my abilitities.[1]	1. On my ideal job it is important that I could do something that makes use of my abilities.	MIQ MJDQ non-parallel. Follow MJDQ and O*NET
2. On my ideal job it is important that the work could give me a feeling of accomplishment.[1]	2. On my ideal job it is important that the job could give me a feeling of accomplishment.	Work consists of tasks that are done on a job. It is more clear and less redundant.
3. On my ideal job it is important that I could be busy all the time.	3. On my ideal job it is important that I could be busy all the time.	
4. On my ideal job it is important that the job provide an opportunity for advancement.	4. On my ideal job it is important that the job would provide an opportunity for advancement.	
5. On my ideal job it is important that I could give directions and instructions to others. [2]	5. On my ideal job it is important that I could tell people what to do.	Consistent with O*NET
6. On my ideal job it is important that I would be treated fairly by the company. [2]	6. On my ideal job it is important that the company would administer its policies fairly.	Consistent with O*NET
7. On my ideal job it is important that my pay would compare well with that of other workers.	7. On my ideal job it is important that my pay would compare well with that of other workers.	
8. On my ideal job it is important that my co-workers would be easy to get along with.[2]	8. On my ideal job it is important that my co-workers would be easy to make friends with.	Consistent with O*NET
9. On my ideal job it is important that I could try out my own ideas.[1]	9. On my ideal job it is important that I could try out some of my own ideas.	MIQ MJDQ items non-parallel. Follow MJDQ wording
10. On my ideal job it is important that I could work alone.[1]	10. On my ideal job it is important that I could work alone on the job.	Reduce redundancy
11. On my ideal job it is important that I would never be pressured to do things that go against my sense of right and wrong. [3]	11. On my ideal job it is important that I could do work without feeling that it is morally wrong.	Consistent with O*NET
12. On my ideal job it is important that I could receive recognition for the work I do.[1]	12. On my ideal job it is important that I could get recognition for the work I do.	O*NET Change. MIQ/MJDQ items non-parallel. Follow MJDQ wording

Table 2. ***Wording Changes from the O*NET Project for the 21 Need Statements*** *(Continued)*

WIL-P&P	Original MIQ Items	Reasons for Changes Made
13. On my ideal job it is important that I could make decisions on my own.	13. On my ideal job it is important that I could make decisions on my own.	
14. On my ideal job it is important that the job would provide for steady employment.	14. On my ideal job it is important that the job would provide for steady employment.	
15. On my ideal job it is important that I could do things for other people.	15. On my ideal job it is important that I could do things for other people.	
16. On my ideal job it is important that I would be looked up to by others in my company and my community. [2]	16. On my ideal job it is important that I could be "somebody" in the community.	Consistent with O*NET
17. On my ideal job it is important that I have supervisors who would back up their workers with management. [1]	17. On my ideal job it is important that my boss would back up the workers (with top management).	
18. On my ideal job it is important that I would have supervisors who train workers well. [1]	18. On my ideal job it is important that my boss would train their workers well.	Consistent with O*NET
19. On my ideal job it is important that I could do something different every day.	19. On my ideal job it is important that I could do something different every day.	
20. On my ideal job it is important that the job would have good working conditions.	20. On my ideal job it is important that the job would have good working conditions.	
21. On my ideal job it is important that I could plan my work with little supervision.	21. On my ideal job it is important that I could plan my work with little supervision.	

[1] Minor difference in the wording between WIL-P&P and MIQ.
[2] Moderate difference in wording between WIL-P&P and MIQ.
[3] Substantial difference in the wording between WIL-P&P and MIQ.

Initial Development of the WIL-P&P

The WIL-P&P employs a card-sorting task that is similar to the Q-sort technique (Stephenson, 1953). Respondents complete the WIL-P&P by categorizing 20 need statements in terms of their relative importance in the respondents' ideal job on a 5-point scale from very important (5) to least important (1). This decision scale allows for relatively easy decision making and scoring. The following are examples of need statements:

22. "My pay would compare well with that of other workers."
23. "I could do things for other people."
24. "I could be busy all the time."
25. "I could try out my ideas."
26. "The job would provide an opportunity for advancement."

At this stage, the instrument consisted of:

a) two pages of instructions,

b) 20 cards which each display a need statement and have a letter (A through T) printed on them,

c) a card sorter sheet used to place the cards into five groups, and

d) a scoring page on which need scores are reported and value scores are calculated.

An example of the materials is included in Appendix A.

When a respondent sorts the 20 cards of need statements into five piles (four cards per pile), the four needs that are the most important are placed in the first pile, the four needs that are next in importance are placed in the second pile, and so on. The respondent then records the pile number for each card on the scoring page. The pile number represents the score for that need (e.g., each need in the *most important* pile gets a score of 5). There are six tables on the scoring page. The scoring page is laid out so that the cards that represent the same work value are grouped together in the same table. After all the pile numbers have been recorded, the respondent computes the six value scores by adding and multiplying one-digit numbers.

Pre-Pilot Study for the WIL-P&P

A Pre-Pilot Study of the WIL-P&P was conducted to determine how easy the WIL-P&P was to use (e.g., were the instruments easy to follow?) and what improvements should be made. The measure was completed by 21 employment center clients in North Carolina. The sum or tally of the cards in each pile and the product of each pile's sum multiplied by the appropriate multiplier for the pile were taken from each individual's scoring page and entered into a computer database. A computer program checked for values that were out of the possible range and remultiplied the sum of each pile by the multiplier for each pile. The products from the computer program's remultiplication were compared to the values the clients had calculated to determine if the clients had made errors. Nine of the 21 people made at least one error while taking the measure. The number of people making each type of error is shown in Table 3.

Table 3. ***Number of People Making Errors in the WIL-P&P Pre-Pilot Study***

Type of Error	Number of People Who Made an Error
Addition	4
Multiplication	2
Wrong number of cards in a pile	5
Multiplied each pile number by four before writing it on the scoring page	1
Any type of error	9

Note: N=21. Three people made two types of errors.

The large number of pile errors was a concern. The respondents were supposed to put four cards into each of the five rectangles printed on the card sorting sheet. Apparently, some people put the wrong number of cards in some piles. Additionally, math errors were made by these respondents. In an attempt to reduce the amount and type of errors, the card sorting sheet was modified so that there were 20 rectangles printed on the sheet. Each card would be placed in its own rectangle. The math errors were addressed by improving the instructions and making minor format changes to the scoring page to make it easier to understand which numbers to add and which numbers to multiply.

The respondents also completed a Participant Reaction Questionnaire. Only one participant found the task boring or tiring, two became frustrated with the task, and one was frustrated that some important work values were not explicitly included in the measure (e.g., flexible scheduling and leave policy). Almost all of the respondents said the instructions were either *clear* or *very clear*. Only two of the six people who made pile or score transfer errors thought the instructions were *very clear*, suggesting that an improvement in the instructions would probably reduce errors. Instructions were clarified.

Pilot Study for the WIL-P&P

The modified WIL-P&P was tested in a Pilot Study at an employment center in Utah. In this study, the WIL-P&P and a Participant Reaction Questionnaire were completed by 48 clients of the center. Test administrators observed the clients completing the measure.

The test administration team made the following observations during WIL-P&P pilot testing:

- Most participants completed the task without difficulty.
- Mathematical errors in score computation were the most common problem.
- Some participants expressed frustration with having to rank some cards on the lower end of the importance scale.
- The average completion time was 14 minutes. Errors made by the respondents in the Pilot Study are detailed in Table 4. The Participant Reaction Questionnaire results for the WIL-P&P are shown in Table 5.

Table 4. ***Number of People Making Errors in the WIL-P&P Pilot Study***

Type of Error	Number of People Who Made an Error
Addition	7
Multiplication	2
Wrong number of cards in a pile	0
Multiplied the total for Work Value 2. (This total should not be multiplied.)	1
Any type of error.	8

Note: N=48. Two people made two types of errors.

Table 5. ***Participant Reaction Results from the Pilot Study***

	Percent Positive Responses
Evaluation Question	Paper Version (*N*=48)
How clear were the instructions on the survey?	90%
Did you find the rankings easy to do?	92%
Was the survey easy to score? (Paper)	95%
Did you get tired or bored at any time during the survey?	100%
Are the results of the survey consistent with how you would describe yourself?	87%

As Table 4 shows, the number of apparent errors decreased from the level in the Pre-Pilot Study, with sorting errors disappearing entirely. Unfortunately, the test administrators stated during a debriefing that they had helped many of the respondents. Therefore, the number of errors would probably have been higher if all of the respondents had completed the measure without assistance. The administrators speculated that the number of errors might have been as high as 30 percent, but it is not certain what the error rate would have been.

Final Revisions

Further revisions to the measure were made based on comments by respondents and administrators. These changes were small, but affected several parts of the measure. At this point, the Work Values Explanation sheet was developed. This sheet told the respondents how to interpret their scores.

The final version of the WIL-P&P consisted of:

- 20 need statement cards, labeled A through T;
- a 3-page booklet containing the instructions and scoring page;
- a card sorter sheet (11-inch × 17-inch), which also includes some instructions; and
- the *Work Values Explanation* sheet, which tells the respondents how to interpret their scores.

The procedure for completing the final version of the WIL-P&P was the same as for the pilot version (see Appendix A):

- After reading the initial instructions, the 20 cards containing the need statements are sorted into 5 columns on the card sorter sheet. The cards with the most important needs are placed in Column 5, the five next most important needs are placed in Column 4, and so forth. On the scoring page, there are six tables named Work Value 1 through Work Value 6. Within each table is a column of card labels (letters).
- After sorting the cards, the respondent looks at the first card in pile 1 to see what its letter label is. He or she locates that letter in one of the rectangles on the scoring page and writes the card's pile number next to the letter on the scoring page. This number is the score for that need. The respondent continues recording the pile number for each card.
- Once the score has been recorded for every card, the scores for each work value (i.e., within each table) are summed.
- Finally, the sums are multiplied by a constant for that work value. This multiplication is necessary to weight the scales because the six work values consist of different numbers of needs. For example, Work Value 1 has two needs (i.e., cards), whereas Work Value 2 has six needs. The lowest and highest possible scores for each Work Value are 6 and 30, with the exception of Work Value 2, which has a minimum score of 8 and a maximum score of 28.

With the completion of the Pre-Pilot Study and modification of the measure and accompanying materials based on the results of these studies, the revised measure was considered ready for use in the Main Study. Results of the Main Study are discussed in the next four chapters.

Summary

The WIL-P&P is a self-scored measure of work values that employs a card-sorting method and is based on items from the MIQ. One item was dropped from the original set of 21 to simplify the scoring process for respondents. The Pre-Pilot and Pilot Studies were conducted to gain information to improve the measure before administering it to a larger sample. Many improvements were made based on feedback from these two preliminary studies. With the completion of the Pre-Pilot Study and Pilot Study and subsequent modification of the measure and accompanying materials, the WIL-P&P was considered ready for use in the Main Study.

Chapter 4. Main Study: Sample Description and Data Cleaning

Introduction

This chapter describes the creation of the dataset and the characteristics of the respondents in the Main Study. Various frequencies and descriptive statistics are reported which describe the sample both before and after the data-cleaning process. The data-cleaning process is used to identify and correct problems in the data that are due to errors, incorrect coding, careless responding, or other factors that can lead to erroneous conclusions in analyses.

Administration of Instruments in the Main Study

Various versions of the work values instruments and the MIQ were administered to participants in the main study. For purposes of calculating the test-retest reliabilities, the WIL-P&P and WIP-C were administered to students in junior college classes. Specifically, Table 6 shows that at Time 3 the WIL-P&P was administered first, followed by the WIP-C. Several weeks later (Time 4), the instruments were administered again to the same individuals and in the same order as the first administration. Other students responded to the WIP-C first, followed by the WIL-P&P. Approximately 230 usable data sets resulted from this administration. This sample was also used for additional statistical analyses.

For other statistical analyses, such as alternative forms reliability, the WIL-P&P, WIP-C and MIQ were administered to junior college students and individuals at employment centers. The design that was used attempted to balance the pairing and ordering of measures taken by respondents, so that all pairs and orders would be represented. For example, some respondents took the WIL-P&P first, followed by the WIP-C or the MIQ. In other cases, the WIP-C was administered first, followed by the WIL-P&P or MIQ. Still other respondents took the MIQ first, followed by the WIL-P&P or WIP-C. These administrations led to the collection of sufficient numbers of individuals for each instrument so that analyses could be performed for gender and racial/ethnic groupings.

Construction of the Dataset

Table 6 also shows the number of subsamples of respondents who participated and provided complete, useable data at Times 3 and 4 in the Main Study. The construction of the database of information (from these respondents) was complex. The data were collected at 23 sites over a period of 3 months. MIQ forms, completed by 550 respondents, were sent together to be computer-scored by the MIQ publisher. The data from the WIP-C were copied onto floppy disks at each administration site at the end of each day. WIL-P&P form data and all other data (i.e., roster, biodata) were entered manually by National O*NET Center (O*NET) personnel; there was a separate file for each combination of site and form. The data at this stage consisted of the following data files:

- WIL-P&P files,
- WIP-C files,
- MIQ file,

- Biodata form files, and
- Roster files.

In order to perform various analyses, the data from these files had to be combined to create one large dataset. Participants' Social Security numbers (SSNs) served as the linking variable. Errors in the data were corrected before merging. Common errors included incorrect SSNs and incorrect coding of whether or not the test administration was a retest.

Some respondents' results were excluded from the dataset to prevent the responses of careless or unmotivated participants from distorting the analyses. There were a few cases where the test administrator concluded that a respondent was completely unable or unwilling to complete a measure. This form was discarded and, consequently, its data were not used. The number of times this happened was not recorded. Although the number of people who had *all* their results discarded is unknown, informal reports suggest that it was very few. It was possible, however, to identify the number of people who had *one* measure discarded. After excluding these results and merging data files by participants' SSNs, there were 1,609 people in the total dataset; of these 1,199 (74.5 percent) had useable WIL-P&P data.

Table 6. ***Sample and Subsample Sizes for the Main Study***

Study	Time 1 *n* per Subsample	Time 2 *n* per Subsample	Time 3 *n* per Subsample	Time 4 *n* per Subsample	Total *n* in Study
Main Study: Test-Retest			5. P-C=148-133	5. P-C=121-111	C=213
			6. C-P=132-142	6. C-P=102-111	P=232
Main Study: Other Analyses			5. P-C=148-133		C=941
			6. C-P=132-142		P=1199
			7. P-C=269-210		MIQ=550
			8. C-P=218-276		
			9. P-MIQ=174-141		
			10. MIQ-P=158-190		
			11. C-MIQ=120-121		
			12. MIQ-C=130-128		

Note. P is the WIL-P&P; C is the WIP-C

Characteristics of the Respondents Before Data Cleaning

Each measurement site implemented procedures to ensure adequate numbers of respondents from each of the following gender and racial/ethnic groups: male, female, white, Hispanic, and African American. Each site also took steps to ensure that there were enough respondents for each pair of measures. These subgroup sample sizes for the Main Study are shown in Table 7. The figures shown in the table and discussed below represent the people in the dataset before data cleaning.

Table 7. ***Main Study Subsample Sizes and Percentages by Gender and Race/Ethnic Group (Before Data Cleaning)***

		Gender		Race/Ethnic Group					
Measures	Row Total	F	M	Af. Amer.	Hispanic	White	Am. Indian	Asian	Other
				Frequencies					
C - P	598	260	336	294	80	199	5	8	8
C - MIQ	321	147	172	161	44	105	2	2	7
P - MIQ	389	184	204	132	71	167	5	4	9
All	1308	591	712	587	195	471	12	14	24
				Percentages					
C - P	100%	44%	56%	49%	13%	34%	1%	<1%	1%
C - MIQ	100%	46%	54%	50%	14%	33%	<1%	<1%	2%
P - MIQ	100%	47%	53%	34%	18%	43%	1%	1%	2%
All	100%	45%	55%	45%	15%	36%	1%	<1%	2%

Note. C is the WIP-C; P is the WIL-P&P. Some row values do not sum to the total because of missing values. The table includes all people scheduled for only one administration of the measures (including people whose measures were eliminated from subsequent analyses).

In the Main Study, representation by race/ethnic group and gender was well-balanced across the three pairs of measures (C-P, C-MIQ, P-MIQ). There were no meaningful differences in percent representation by gender or employment situation across the three pairs of measures. There were some differences by race/ethnic group; in the P-MIQ pair, there were relatively more whites and fewer African Americans. Overall the sample was 45 percent female, 45 percent African American,

26 percent white, and 19 percent Hispanic. The respondents' employment situation was equally represented across the three pairs of measures, as well (see Table 8). The largest group consisted of employment center customers, who represented 42 percent of the sample.

Table 8. ***Main Study Subsample Sizes and Percentages by Employment Situation (Before Data Cleaning)***

		Employment Situation							
Measures	Row Total	Employment Service	Junior College	Emp-ployed	Unemp-loyed	Community Group	High School Student	Military Service	Other
				Frequencies					
C - P	598	241	30	84	82	0	13	5	137
C - MIQ	321	126	10	36	45	1	6	1	88
P - MIQ	389	171	18	54	45	1	6	1	88
ALL	1308	538	58	174	172	2	25	7	313
				Percentages					
C - P	100%	41%	5%	14%	14%	0%	<1%	<1%	28%
C - MIQ	100%	40%	3%	12%	14%	<1%	2%	<1%	28%
P - MIQ	100%	45%	5%	14%	12%	<1%	2%	<1%	23%
ALL	100%	42%	4%	13%	13%	<1%	2%	<1%	24%

Note. C is the WIP-C; P is the WIL-P&P. Some row values do not sum to the total because of missing values. The table includes all people scheduled for only one administration of the measures (including people whose measures were eliminated from subsequent analyses).

For the dataset used in the analysis of test-retest reliabilities, representation was well balanced (see Table 9 and Table 10). There were no meaningful differences in percent among three measures in representation by gender, race/ethnic group, or employment situation. The sample was 70 percent female, 26 percent African American, 64 percent white, and 8 percent Hispanic. The majority (66 percent) listed Junior College Student as their employment situation. There were substantial differences between the test-retest sample and the sample used for the rest of the Main Study in all these percentages. This was not surprising considering that the majority of respondents in the Main Study had taken the need/value measures at employment centers, whereas the test-retest reliability analysis was based on a subsample of respondents who took the need/value measures at junior colleges.

Table 9. ***Test-Retest Analyses Subsample Sizes and Percentages by Gender and Race/Ethnic Group (Before Data Cleaning)***

Time	Row Total	Gender		Race/Ethnic Group					
		F	M	Af. Amer.	Hispanic	White	Am. Indian	Asian	Other
				Frequencies					
Time 1	301	201	98	78	23	192	2	2	4
Time 2	240	167	72	61	16	158	1	1	1
				Percentages					
Time 1	100%	67%	33%	26%	8%	64%	<1%	<1%	1%
Time 2	100%	70%	30%	25%	7%	66%	<1%	<1%	1%

Note. Some row values do not sum to the total because of missing values. The table includes all people scheduled for two administrations of the measures (including people whose results were eliminated from subsequent analyses). For these people, the two measures taken at each administration were the WIL-P&P and the WIP-C.

Table 10. ***Test-Retest Analyses Subsample Sizes and Percentages by Employment Situation (Before Data Cleaning)***

		Employment Situation							
Time	Total	Employment Service	Junior College	Emp-ployed	Unemp-loyed	Community Group	High School Student	Military Service	Other
				Frequencies					
Time 1	301	6	196	45	30	1	3	1	18
Time 2	240	6	158	36	20	1	3	0	15
				Percentages					
Time 1	100%	2%	65%	15%	10%	<1%	1%	<1%	6%
Time 2	100%	3%	66%	15%	8%	<1%	2%	0%	6%

Note. Some row values do not sum to the total because of missing values. The table includes all people scheduled for two administrations of the measures (including people whose measures were eliminated from subsequent analyses). For these people, the two measures taken at each administration were the WIL-P&P and the WIP-C.

Overall, for the Main Study, the mean age of the sample was 35 years, and the mean years of education was 13 (see Table 11). The test-retest reliability subsample was similar; the average age was 32 years, and the mean years of education was 13. There were no large differences in age and education among the design subsamples.

Table 11. ***Descriptive Statistics for Age and Education (Before Data Cleaning)***

	Age (in years)		Education (in years)	
Measure Design Group	No-Retest Group	Test-Retest Group	No-Retest Group	Test-Retest Group
C - P				
M	35.6	31.0	12.5	12.7
SD	11.3	10.7	2.1	1.9
n	303	147	300	145
P - C				
M	36.8	33.1	12.5	12.4
SD	11.0	10.5	1.9	1.7
n	293	153	290	154
C - MIQ				
M	36.7		13.0	
SD	11.0		1.8	
n	154		153	
MIQ - C				
M	35.9		12.7	
SD	12.1		1.8	
n	167		165	
P - MIQ				
M	36.7		12.9	
SD	12.8		2.1	
n	190		187	
MIQ - P				
M	36.6		12.6	
SD	11.6		2.0	
n	198		197	

Note. C is the WIP-C; P is the WIL-P&P. The *n* values represent the number of non-missing values for age or education in each cell of the design

Data Cleaning Based on Response Irregularities

The initial analyses examined response inconsistencies, missing responses, invalid responses, and errors in self-scoring (see Table 12). These will be called *response irregularities* in the remainder of the report. A total of 1,609 individuals participated in the Main Study. This does not include people who had all their results discarded by the test administrator because the test administrator concluded that these respondents were completely unable or unwilling to complete the measures, as discussed earlier. Because its administration was computerized, the WIP-C had no missing responses.

It was decided that the data analyses should be based, as much as possible, upon measures that were properly completed by motivated respondents. The following criteria were used to omit additional data from the analyses:

For the WIL-P&P
- any missing values
- any invalid values
- more than one math error during self-scoring

For the MIQ
- consistency $\leq .5$
- any missing values (the MIQ publisher does not score measures with any missing values)
- any invalid values (the MIQ publisher does not score measures with any invalid values)

For the WIP-C
- consistency $\leq .5$

The approach taken when determining the screening criteria was to ensure, as much as possible, that responses on the measures were true reflections of the respondents' attitudes. Therefore, stringent screening standards were used. The response consistency cutoff of .50 is about three standard deviations above the level for random responding (which is .16). In addition, the MIQ publisher uses the same cutoff of .50 and considers people with lower consistencies either to have poorly defined values or be responding carelessly.

It was assumed that an invalid or missing response was also evidence of carelessness. Therefore, measures with any invalid or missing responses were excluded. On the other hand, it was assumed that the presence of a *single* math error during the WIL-P&P's self-scoring process was *not* due to carelessness, but that the presence of *more* than one math error *was* due to carelessness. This assumption was supported by the finding that many respondents (14 percent) made one math error, whereas fewer respondents (2.9 percent) had a missing or invalid response - or more than one math error (5.1 percent). To prevent respondents' math errors from affecting any analyses, the work values scale scores were calculated by the computer when creating the dataset.

Several analyses showed that the presence of response irregularities in one of a respondent's measures had little relation to the presence or absence of response irregularities in his or her other measure. Therefore, data were screened out on a measure-by-measure basis. That is, the data for a respondent might be screened out for one measure, but kept for another.

Table 12. ***Frequency of Respondents in Main Study with Response Irregularities***

Type of Response Irregularity	Number of Respondents			Percent of Total		
	WIL-P&P	WIP-C	MIQ	WIL-P&P	WIP-C	MIQ
Consistency $\leq .5$		264	105		21.6	14.8
1 Missing	9		0	0.7		0.0
>1 Missing	11		5	0.9		0.7
1 Invalid	11		0	0.9		0.0
> 1 Invalid	5		44	0.4		6.2
1 Addition Error	148			11.5		
> 1 Addition Error	61			4.7		
1 Multiplication Error	110			8.5		
> 1 Multiplication Error	59			4.6		
1 Math Error	180			14.0		
> 1 Math Error	66			5.1		
Measure Discarded by Administrator	5	15	6	0.4	1.2	0.8
Total Respondents	1,288	1220	710	100	100	100
Useable Respondents	1,199	941	550	93.1	77.1	77.5

Note. Numbers in the WIL-P&P column do not sum to the total because some respondents had more than one type of response irregularity. Any miscalculation of a WIL-P&P value scale score was considered a *math error*. There were 1,609 individual respondents in Study 2.

Characteristics of the Respondents After Data Cleaning

As just described, suspect data were screened out on a measure-by-measure basis. That is, a respondent could have one measure excluded from the analyses, but the other measure is included. Therefore, the frequencies and descriptive statistics are presented by instrument because the number of subjects changes somewhat depending on the instrument.

Table 13 shows the frequencies by gender and race/ethnic group. The distributions of gender and race/ethnic group changed slightly after data cleaning. The percentage of females increased 6 percent, and the percentage of whites increased 9 percent. In addition, the percentages for whites and

African Americans reversed. Before the data cleaning, the sample was 45 percent African American and 36 percent white; after the data cleaning, the sample was 38 percent African American and 45 percent white. Although subject loss was disproportionately high for African Americans, the final percentage of 38 percent provides an adequate subsample of African Americans for statistical comparisons.

Table 13. ***Frequencies and Percentages for Gender and Race/Ethnic Group by Instrument and Time (after Data Cleanup)***

Group:		Gender		Race/Ethnic Group					
Measures	Row Total	F	M	Af. Amer.	Hispanic	White	Am. Indian	Asian	Other
Time 1				Frequencies					
WIL-P&P	1199	606	588	454	157	539	12	14	19
WIP-C	941	496	440	351	113	441	8	12	14
MIQ	550	272	277	216	83	228	7	4	12
Time 2									
WIL-P&P	232	163	68	59	16	152	1	1	3
WIP-C	213	150	62	49	14	146	1	1	2
				Percentages					
Time 1									
WIL-P&P	100%	51%	49%	38%	13%	45%	1%	1%	2%
WIP-C	100%	53%	47%	37%	12%	47%	1%	1%	1%
MIQ	100%	50%	50%	39%	15%	41%	1%	1%	2%
Time 2									
WIL-P&P	100%	71%	29%	25%	7%	66%	<1%	<1%	1%
WIP-C	100%	41%	29%	23%	7%	69%	<1%	<1%	1%

Note. Some row values do not sum to the total because of missing values in Gender or Race/Ethnic Group. Only one measure was used for some people because of excessive errors on their other measure. On the MIQ and WIP-C, measures with consistency $\leq .5$ were excluded from all analyses. On the WIP-P, measures with any missing values, any invalid entries, or more than one math error were excluded from all analyses.

Summary

Numerous data files from 23 sites were collected and combined into a single dataset for analyses. Data were merged into this single data set through the use of respondents' SSNs. After the exclusion of tests that respondents were unable or unwilling to complete, the total number of respondents in the data set was 1,609. Errors and incorrect coding were corrected before the data were merged. Some additional data were not used in analyses because response inconsistencies, missing values, invalid values, and scoring errors were found. There were 1,199 respondents with useable data on the WIL-P&P, 941 on the WIP-C, and 550 on the MIQ for analyses. Data collection sites had implemented procedures to ensure that adequate numbers of male, female, white, Hispanic, and African American respondents were included (so that there would be large enough subsamples from these groups to do statistical analyses). Although the relative percentages of the various demographic subgroups did change somewhat after data cleaning, subgroups remained sufficiently large for statistical analyses.

The next step in analyzing the data was to calculate descriptive statistics on the measures, sample, and subsamples. These statistics are presented in the next two chapters.

Chapter 5. Main Study: Descriptive Statistics for the Instruments

Introduction

This chapter presents the descriptive statistics for the WIL-P&P and the MIQ. The variables of interest from these instruments are the 20 need items and 6 value scales, instrument completion times, and error rates. Table 14 shows the means and standard deviations for each instrument. Means from the WIL-P&P and MIQ are considerably different because the WIL-P&P used ipsative scoring; hence, the metric for it and the MIQ are not the same.

Table 14. ***Main Study Need and Value Means and Standard Deviations***

				WIL-P&P		MIQ	
Value		Item	Need Name	Mean	Std. Dev.	Mean	Std.Dev.
1. Achievement		1	Ability Utilization	3.97	1.19	1.47	0.71
		2	Achievement	3.83	1.15	1.47	0.68
2. Comfort		3	Activity	2.34	1.35	0.53	0.88
		7	Compensation	3.13	1.43	0.94	0.91
		10	Independence	1.81	1.17	0.42	0.99
		14	Security	3.94	1.27	1.70	0.89
		19	Variety	2.03	1.20	0.51	0.77
		20	Working Conditions	3.41	1.27	1.25	0.74
3. Status		4	Advancement	4.12	1.12	1.56	0.75
	*	5	Authority	2.03	1.17	-0.28	0.89
		12	Recognition	3.16	1.26	1.11	0.75
	*	16	Social Status	NA	NA	0.10	1.00
4. Altruism	*	8	Co-Workers	2.73	1.26	0.41	0.79
	*	11	Moral Values	3.30	1.39	1.05	1.08
		15	Social Service	2.50	1.30	0.61	0.82
5. Safety	*	6	Company Policies	3.73	1.18	1.10	0.75
		17	Supervision: Human Relations	2.97	1.25	0.83	0.75
		18	Supervision: Technical	3.02	1.28	0.80	0.71
6. Autonomy		9	Creativity	2.62	1.20	0.80	0.75
		13	Responsibility	2.80	1.22	0.99	0.76
		21	Autonomy	2.60	1.19	0.95	0.72
1. Achievement				23.39	5.44	1.47	0.62
2. Comfort				16.65	2.91	0.89	0.56
3. Status				18.62	4.04	0.63	0.61
4. Altruism				17.05	4.68	0.69	0.64
5. Safety				19.44	4.91	0.91	0.63
6. Autonomy				16.04	4.90	0.91	0.65

Note. For each instrument, scores on the retest were excluded. $N = 1{,}119$ and 550 for the WIL-P&P and MIQ, respectively.
*Moderate or substantial differences in wording.

Analysis of Effects of Completing Two Measures During the Same Administration

When a person responds to two very similar measures back-to-back, there is always a concern that the act of taking the first measure will affect the scores on the second measure. Increased fatigue, perseveration (i.e., attempting to answer the same way on both measures), decreased motivation, and true changes in the characteristic being measured can all systematically affect examinees' scores on the second of two similar measures - and with the exception of the latter, these are primarily problematic when there is only a very *short* interval between the administrations.

To examine the seriousness of these effects, results from the first measure respondents completed were compared with results from the second measure they completed in terms of the proportion of respondents with math errors on the card sorting procedure of the WIL-P&P. Administration time also was examined to determine if there were drastic reductions in the time required for completing the second measure. The results of these analyses are given in Table 15.

Table 15. ***Errors and Completion Times by Measure and by Order of Administration***

		Mean Number of Math Errors	Mean Completion Time	
Measure Order	*n*	WIL-P&P	MIQ	WIL-P&P
WIL-P&P/MIQ	134	.08	17.6	11.3
MIQ/WIL-P&P	155	.10	17.8	10.3
		p=.66	p=.78	p=.04

Note. Measures discarded before data entry were excluded from these analyses. Under each pair of values in the table is the *p*-value representing significance test results from the comparison of the numbers to determine whether or not they are significantly different.

The order effects were minimal and indicated careful responding. There was no significant increase in the proportion of respondents making math errors on the card sort procedure of the WIL-P&P when the WIL-P&P was the second measure taken, compared with when it was the first measure taken. As expected, the time to take the WIL-P&P decreased when it was the second measure taken.

Overall, the evidence indicates that the participants took the measures seriously. The error rates and completion times remained relatively stable regardless of whether a measure was taken first or second. If participants became bored, one would expect the completion times and accuracy to decrease substantially for the second measure. The slight drop in completion times for the second measure is easily attributable to greater familiarity with the concepts and process involved in the

measure.

Summary

Results of the above analyses showed that one can confidently use data from both the first and second instruments individuals completed at a single administration. The data do not indicate that individuals took the second instrument less seriously (or showed significantly greater fatigue) based on error rates and completion times.

Chapter 6. Main Study: Relationships of Work Values with Gender, Race/Ethnic Group, and Education

Introduction

Several analyses were performed to determine if work values were related to gender, race/ethnic group, or educational attainment. One of the analyses compared mean work values of respondents in groups divided by four levels of education: two or more years of college, some college but less than two years, a high school degree, and no high school degree. A second analysis compared mean work values between males and females and among whites, African Americans, and those of Hispanic background. Results of these analyses are provided in the following sections.

Education

A multivariate analysis of variance (MANOVA) was performed on the WIL-P&P and the MIQ to determine if the level of education affected the magnitude of the scores. The dependent variables were the six value scales, and the independent variable was the level of education.

The MANOVA was significant ($p<.0001$) for each of the instruments. The pattern of the effects showed some similarities between the two instruments. The *d-scores* for the two measures are shown in Table 16. Those which are significantly different for differing levels of education are indicated. The *d-score* is the difference between the group means divided by the standard deviation to standardize the differences between means. This allows direct comparison of the instruments on this statistic. Post-hoc tests determined, for each scale, which levels of education had significantly different mean scores.

The following paired comparisons were significant for all of the instruments:

- Scores on the Achievement (Value 1) scale of subjects who had completed at least two years of college tended to be higher than the scores of subjects with only a high school education.
- Scores on the Safety (Value 5) scale of subjects who had completed two years of college tended to be lower than the scores of subjects with only a high school education.
- Scores on the Autonomy (Value 6) scale of subjects with at least two years of college tended to be higher than the scores of subjects who had not completed high school.

Table 16. ***Mean Work Values Scores by Educational Attainment***

Instrument		Mean				*d*-score					
						HS-	SC-	C2+-	SC-	C2+-	C2+
	Value	NHS	HS	SC	C2+	NHS	NHS	NHS	HS	HS	SC
P&P	1. Achievement	21.92	23.03	23.56	25.53	0.20	0.29*	0.65*	0.10	0.10	0.38*
	2. Comfort	16.95	16.78	16.77	15.85	-0.06	0.06*	-0.39*	0.00	-0.32*	-0.31*
	3. Status	19.00	18.52	18.89	18.19	-0.12	-0.03	-0.19	0.10	-0.08	-0.18*
	4. Altruism	18.18	17.00	15.99	17.31	-0.25*	-0.48*	-0.18	-0.22*	0.07	0.29*
	5. Safety	19.57	19.98	19.26	18.02	0.09	-0.06*	-0.30*	-0.15	-0.41*	-0.24*
	6. Autonomy	14.97	15.52	16.79	17.69	0.12	0.36*	0.55*	0.27*	0.46*	0.17
	Sample Size=>	190	558	247	192						
MIQ	1. Achievement	1.39	1.42	1.51	1.60	0.05	0.19	0.34	0.15	0.30*	0.14*
	2. Comfort	0.92	0.95	0.92	0.72	0.05	0.00	-0.40	-0.05	-0.42	-0.37
	3. Status	0.73	0.64	0.62	0.54	-0.15	-0.19	-0.31	-0.03	-0.16	-0.13
	4. Altruism	0.72	0.66	0.69	0.73	-0.10	-0.05	0.02	0.05	0.11*	0.06
	5. Safety	1.00	0.96	0.91	0.77	-0.07	-0.14	-0.37	-0.08	-0.31*	-0.21
	6. Autonomy	0.72	0.89	0.92	1.11	0.25	0.32*	0.68	0.05	0.34	0.33
	Sample Size=>	69	227	144	107						

Note. NHS = Did not graduate high school; HS = Graduated high school; SC = Attended some college, but did not get a degree; C2+ = Obtained a college degree (2- or 4- year).
*p < .05 using Tukey's correction for experiment wise error.

The following paired comparisons were significant for the WIL-P&P:

- Scores on Achievement (Value 1) tended to be greater for those with at least two years of college than those with no high school.
- Those with at least two years of college tended to be lower on Achievement (Value 1) than those with less than two years of college.
- Those with no high school diploma, a high school diploma, and some college tended to be lower on Comfort (Value 2) than those with a two or four year degree.

- Those with no high school tended to be lower on Altruism (Value 4) than those with a high school diploma and some college.
- Those with a two or four year college degree tended to be higher on Altruism (Value 4) than those with some college. Those with some college were lower on this value than those with a high school diploma.
- Those with a two or four year college degree tended to be lower on Safety (Value 5) than those with some college, a high school diploma, or no high school diploma.
- Those with some college or a two or four year college degree tended to be higher on Autonomy (Value 6) than those with a high school diploma or no high school diploma.

Gender and Race/Ethnic Group

A two-way MANOVA was performed on each of the instruments to determine if the levels of race/ethnic group or gender affected the magnitude of the scores. Years of education was used as a covariant to remove the potential confounding influence of education. Only whites, African Americans, and Hispanics were included in the analyses because of the small number of individuals in the other race/ethnic groups.

Significant effects related to race/ethnic group, gender, and education were detected for each of the instruments. The interaction between race/ethnic group and gender was not statistically significant ($p > .05$). The results from the univariate ANOVAs for the WIL-P&P differed somewhat from the results of the MIQ.

- Status apparently was more important to males than to females.
- Altruism apparently was more important to females than males.
- Autonomy apparently was more important to males than females.
- Achievement apparently was more important to whites than African Americans.
- Status apparently was more important to African Americans than to whites.
- Altruism apparently was more important to African Americans than to whites.

The *d-scores* are shown in Tables 17 and 18.

Table 17. ***Mean Work Value Scores by Gender***

Instrument	Value	Least-Square Mean		*d*-score Female - Male	*p*-value of significant difference
		Female	Male		
P & P	1. Achievement	23.66	23.02	0.12	
	2. Comfort	16.54	16.65	-0.04	
	3. Status	18.36	19.08	-0.18	.009
	4. Altruism	17.62	16.59	0.22	.001
	5. Safety	19.67	19.05	0.13	
	6. Autonomy	15.47	16.59	-0.23	.001
MIQ	1. Achievement	1.53	1.43	0.16	
	2. Comfort	0.88	0.85	0.05	
	3. Status	0.60	0.66	-0.10	
	4. Altruism	0.79	0.63	0.25	.011
	5. Safety	1.01	0.74	0.43	<.001
	6. Autonomy	0.89	0.94	-0.08	

Note. The least-square means are adjusted for race/ethnic group and years of education.

Table 18. ***Mean Work Value Scores by Race/Ethnic Group***

Instrument	Value	LS Mean			*d*-score Af.Amer. - White	*p*-value of signif. diff.	*d*-score Hisp - White	*p*-value of signif. diff.
		Af. Amer.	Hisp.	White				
P & P	1. Achievement	22.71	23.51	23.80	-0.20	.002	-0.06	
	2. Comfort	16.83	16.35	16.60	0.08		-0.09	
	3. Status	19.00	18.92	18.26	0.18	.004	0.17	
	4. Altruism	16.53	17.43	17.35	-0.18	.006	0.02	
	5. Safety	19.75	19.03	19.29	0.10		-0.05	
	6. Autonomy	16.05	16.07	15.97	0.02		0.02	
MIQ	1. Achievement	1.47	1.53	1.43	0.06		0.16	
	2. Comfort	0.95	0.79	0.87	0.14		-0.15	
	3. Status	0.72	0.66	0.51	0.35	<.001	0.25	
	4. Altruism	0.64	0.79	0.70	-0.09		0.14	
	5. Safety	0.98	0.76	0.88	0.16		-0.19	
	6. Autonomy	0.93	0.94	0.88	0.08		0.09	

Note. The least-square means are adjusted for gender and years of education.

Summary

Level of education was related to differences on the values of Achievement, Autonomy, and Safety. In general, those with more education tended to have higher expressed values for Autonomy and Achievement and lower expressed values for Safety. The exception was that those with two or more years of college tended to have lower expressed value for Achievement than those with less than two years of college. African Americans tended to express higher value for Status than whites, while

females tended to express higher value for Safety and Altruism than males. Differences tended to be quite small, however.

Chapter 7. Main Study: Internal Consistency and Test-Retest Reliability Evidence for the WIL-P&P

Introduction

The term *reliability* refers to the degree to which a measurement procedure is free from unsystematic errors of measurement and the degree to which a measurement instrument gives one the same values if the procedure is repeated. An individual responding to a measure is likely to have different results if measured more than once on the same instrument. Systematic differences in scores (e.g., improvement on a test taken at two different times because the individual's knowledge has increased between tests) should not be considered to contribute to the *unreliability* of an instrument. But an individual's results may change when measured more than once on the same instrument because of unsystematic effects (e.g., mismarking a response to an item, feeling tired one day but not the next, varying interpretations of the same item from one to another). Such unsystematic differences are considered unreliability. Differences could be due to changes in motivation, energy, attention, anxiety, clarity, or other such factors. Unreliability limits the ability to generalize from individuals' results on a single instrument. The higher the reliability of an instrument, the better it is for drawing conclusions based on values obtained on the instrument.

There are several ways to assess the reliability of measurement, depending on the type of consistency with which one is most concerned. *Test-retest reliability* refers to the consistency of results when the same individual is assessed on the same instrument at two points in time. This information is obtained by looking at the degree of relationship (i.e., correlation) between the examinee's scores obtained on the instrument at the different points in time. Estimates of test-retest reliability are particularly useful if the characteristic being measured is not expected to change over the time between the two measurement periods (e.g., a measure of personality characteristics of normal adults at two points in time that are a month apart, as opposed to a measure of knowledge administered before and after a course on the subject of the measure). Given that work values of adults are considered to be relatively stable characteristics, it would be expected that individuals' responses to the WIL-P&P should be stable across time. As a result, the correlation between the results of the two measurements should be reasonably high.

Another important way to assess reliability is the evaluation of responses by the same individuals on forms that have been created to be *alternative* or *parallel forms* of the same instrument. This estimate of reliability is useful here because the WIL-P&P and WIP-C were designed to be used interchangeably, depending on the computer resources of the location where measurement is taking place. These forms also were designed to closely parallel the original MIQ, although some differences can be expected due to wording changes in some items (see Table 2). Similar results for the same individuals on these different instruments are desirable and would support using the measures interchangeably.

A third type of reliability analysis, *internal consistency reliability*, is used to determine whether different items that are measuring the same subject on the same instrument have highly related

results. For example, if a test included 10 items on addition and 10 items on reading ability, one would expect to see higher interrelationships within the set of 10 addition items and within the set of 10 reading ability items than between items from the two sets. Thus, internal consistency reliability is another type of reliability analysis that can be applied to the WIL-P&P to assess the adequacy of its development. It is desirable to have high internal consistencies among items within the same scale (i.e., the needs that are used to measure each of the six work values).

The Effects of Ipsative Scoring

Before proceeding to the reliability analyses for the WIL-P&P, it is important to understand how the scoring process for the WIL-P&P can be expected to affect reliability results. The WIL-P&P is scored differently from the WIP-C and the MIQ. In all three instruments, the respondents first rank the needs. This procedure puts the needs on a rank-order scale. In the WIL-P&P, the scoring stops here: A need's rank is its score. This is called an ipsative scale. In the WIP-C and MIQ, however, there is an important second step. The respondent rates each need as being either important or not important. The scoring algorithm combines the two steps to put each need on an absolute scale.

Because the WIL-P&P yields an ipsative scale, most of the correlations among its items are negative. This is because rank ordering causes items to "compete" with each other; a high rank for one item means a low rank for another. This has ramifications for the item intercorrelations, and, therefore, for the reliability and validity analyses. Because of the low item intercorrelations, subscales tend to have low internal consistency. Subscales with more items tend to have lower internal consistencies because there is more opportunity for competition among the items (e.g., Value 2 which has 6 items). Thus, for the WIL-P&P, the values scales have low internal consistencies. As a consequence, the values also tended to have low correlations with each other and with other variables, as would be expected from the type of scale used in the measure. Note, however, that the low internal consistency reliability estimates and low correlations with other measures are a function of the scoring procedure - they do not speak to the psychometric strength or operational utility of the WIL-P&P. The WIP-C and MIQ are not restricted in this way and are able to have higher correlations with each other because they are not scored ipsatively. Internal consistency estimates are provided because they are standard in technical development reports and, thus, the information is provided. Standard and corrected reliabilities are provided in an attempt to adjust the low correlations that are likely due to ipsative scoring. Perhaps test-retest reliabilities to include decision consistency are more appropriate measures of instrument reliability.

Test-Retest Reliability

Study 1 assessed the test-retest reliability of the WIL-P&P. In this study, the work values instruments were administered in junior college classes and then re-administered in the same classes 4-8 weeks later. Each participant completed both the WIP-C and the WIL-P&P at each administration. The order of the two instruments was balanced: approximately one-half of the participants completed the WIP-C first (n = 234), while the others completed the WIL-P&P first (n = 269). At the second administration, each participant took the measures in the same order as at the first administration. Elimination of cases due to errors and missing data yielded useable data for 230 participants.

Test-retest reliability was examined in three ways. First, the correlations between the Time 1 and Time 2 results were computed for each need and value. Second, the proportion of the participants whose top values or needs were the same (or nearly the same) at Time 1 and Time 2 was also determined. Third, the correlations between the profiles for an instrument at Time 1 versus Time 2 were computed. This was done for both the needs profiles and the work values profiles.

Table 19 shows the test-retest correlations for each need and value. An attempt was made to reduce the WIL-P&P's ipsative scoring effects. Specifically, each person's paper-and-pencil need scores were adjusted by adding to each score his or her average need score from the WIP-C (after converting the WIL-P&P scores to the same metric as the WIP-C). The score adjustment was expected to increase the test-retest correlations for the work values. This correlation should be considered merely a rough approximation because the correction value is taken from a different measure. In some instances, the adjustment lowered the estimated test-retest reliability for both needs and values.

The WIL-P&P test-retest correlations for the needs were low to moderate. They ranged from .26 to .62 with a median of .42. These moderate figures may be due partly to the ipsative scoring artifact (i.e., the attempted correction may not have worked). The figures for the values scores were slightly higher, with a median of .47. Adjusted values ranged from .33 to .61 with a median of .52.

Table 19. ***Test-Retest Correlations***

Value	Item	Need Name	WIL-P&P	WIL-P&P$_{adjusted}$
1. Achievement	1[1]	Ability Utilization	32	37
	2[1]	Achievement	37	43
2. Comfort	3	Activity	62	61
	7	Compensation	55	52
	10[1]	Independence	47	51
	14	Security	41	37
	19	Variety	42	43
	20	Working Conditions	32	38
3. Status	4	Advancement	41	41
	5[2]	Authority	40	37
	12[1]	Recognition	46	44
	16[2]	Social Status		
4. Altruism	8[2]	Co-Workers	53	51
	11[3]	Moral Values	57	57
	15	Social Service	55	51
5. Safety	6[2]	Company Policies	30	36
	17[1]	Supervision: Human Relations	50	44
	18[1]	Supervision: Technical	44	41
6. Autonomy	9[1]	Creativity	34	45
	13	Responsibility	28	35
	21	Autonomy	26	33
1. Achievement			35	43
2. Comfort			42	45
3. Status			50	42
4. Altruism			53	48
5. Safety			58	50
6. Autonomy			43	53

Note. Decimal points omitted. N = 230 for WIL-P&P and N = 222 for WIL-P&P$_{adjusted}$. WIL-P&P$_{adjusted}$ = paper-and-pencil profiler with scores adjusted using importance ratings from WIP-C.

[1]Minor difference in the wording of the MIQ *vs*. other versions for this item.

[2]Moderate difference in the wording of the MIQ *vs*. other versions for this item.

[3]Substantial difference in the wording of the MIQ *vs*. other versions for this item.

Decision Consistency - Values

Test-retest reliability was examined also from a *decision consistency* perspective. If respondents will use the measurement results to help select an occupation, then one must consider exactly how the results will be used. It is likely that in exploring occupations, the respondent will consider only the top one or two values in the exploration process. In this case, the appropriate measure of test-retest reliability is how often people come up with the same top one or two values each time they take the measure. These matching analyses answer the question: "Are the top few needs or work values the same ones each time a person completes the instrument?"

Table 20 shows how consistent the value rankings were between Time 1 and Time 2. For the WIL-P&P, the top value was the same at Times 1 and 2 for about 62 percent of all respondents (including those who took the WIL-P&P before the WIP-C and those who took the measures in the reverse order). The top two values were the same for about 38 percent of all respondents. These figures may seem to be lower than one would like. A look at the second row in Table 20, however, shows that when the top value at Time 1 is not ranked first at Time 2, it is usually ranked second. Thus, the top value at Time 1 is ranked first or second at Time 2 for 80 percent of respondents. This finding lends support for using the top one or two values in a career exploration process.

Table 20. ***Percentage of the Time the Values Matched at Time 1 and Time 2***

Type of Match Among Scale (Value) Ranks	C-P	P-C
Top scale at Time 1: Time 2 rank = 1	60%	63%
Top scale at Time 1: Time 2 rank = 2	20%	17%
Top scale at Time 1: Time 2 rank = 3	7%	6%
Top scale at Time 1: Time 2 rank = 4	4%	6%
Top scale at Time 1: Time 2 rank = 5	7%	9%
Top scale at Time 1: Time 2 rank = 6	2%	0%
1st- and 2nd-ranked scales match (any order)	37%	39%
1st- and 2nd-ranked scales match (same order)	30%	29%
Sample size	108	118

Note. The subjects in this table took either the WIL-P&P or the WIP-C at two separate times (about 6 weeks apart). The sample sizes for the WIL-P&P and WIP-C are different because some subjects had one of the two instruments dropped due to excessive response errors. An instrument could be dropped from either one or both administrations. P-C and C-P represent the order in which the two measures were taken.

Decision Consistency - Needs

Similar analyses were performed at the need (item) level to see if the needs that were in the top 5 at Time 1 were likely to also be in the top five at Time 2. Fifty-eight percent of the time, the top five needs at Time 1 were also the top five needs at Time 2 on the WIL-P&P. This analysis showed that need scores appear to be as stable as work value scores for the purpose of using just the top five needs versus the top work value to help choose an occupation. The histograms and frequency tables in Appendix B show the test-retest matching results for both needs and work values in greater detail.

Needs and Values Profiles Consistency

The third type of test-retest reliability analysis involved calculating the correlations between WIL-P&P score profiles at Time 1 and Time 2. Correlations were obtained for the needs profiles and the work values profiles. The correlations between the profiles at Time 1 and Time 2 were moderate for the WIL-P&P (r =.64, .62) for both the need ranks and the work value ranks, respectively.

In summary, the unadjusted WIL-P&P test-retest correlations for the individual needs were moderate, with a median of .42. These moderate reliability estimates are likely due largely to the ipsative scoring artifact. The unadjusted reliability estimates for the values scores were slightly higher, with a median of .47. For the WIL-P&P, the top value was the same at Times 1 and 2 for about 62 percent of respondents, and the top value was ranked one or two for 80 percent of respondents. Fifty-eight percent of the time, the top five needs at Time 1 were also the top five needs at Time 2 on the WIL-P&P. Correlations between the profiles at Time 1 and Time 2 were respectable (r =.62 for values). The degree of reliability represented by these results supports the use of scores to determine the top few needs or top two work values.

Correlations Between Instruments

As mentioned previously, the two work values instruments were developed to be alternative measures to each other. Therefore, scores should correlate highly between these instruments. Table 21 shows the correlations between the instruments. For one comparison, the WIL-P&P scores were adjusted using the mean WIP-C score in an attempt to reduce the effects of ipsative scoring.

Table 21. ***Correlations Between the Instruments***

			Measures	
Value	Item	Need Name	C–P	C-$P_{adjusted}$
1. Achievement	1	Ability Utilization	42	62
	2	Achievement	40	61
2. Comfort	3	Activity	54	71
	7	Compensation	57	73
	10	Independence	48	69
	14	Security	51	69
	19	Variety	46	69
	20	Working Conditions	39	65
3. Status	4	Advancement	42	65
	5	Authority	48	70
	12	Recognition	51	71
	16	Social Status		
4. Altruism	8	Co-Workers	44	65
	11	Moral Values	58	71
	15	Social Service	57	73
5. Safety	6	Company Policies	35	61
	17	Supervision: Human Relations	43	64
	18	Supervision: Technical	46	67
6. Autonomy	9	Creativity	45	68
	13	Responsibility	40	67
	21	Autonomy	37	64
1. Achievement			47	73
2. Comfort			32	85
3. Status			38	76
4. Altruism			42	77
5. Safety			43	77
6. Autonomy			45	81

Note. Decimal points omitted. N = 668 for C-P and C-$P_{adjusted}$. C = computerized profiler, P = paper-and-pencil, profiler, $P_{adjusted}$ = paper-and-pencil profiler with scores adjusted using importance ratings from WIP-C.

The WIL-P&P did not correlate well with the WIP-C; however, the correction for ipsatization increased the correlations substantially for both the needs and the values. It appears that the low correlations using the uncorrected WIL-P&P scores were largely due to score ipsatization rather than any problems with the sorting task. The six values had correlations in the .70 and .80s, and a median correlation of .77, for the WIL-P&P and WIP-C with the correction for ipsatization used on the WIL-P&P. This indicates high agreement for values.

Another analysis examined the similarity of the respondents' WIP-C and WIL-P&P profiles. For each respondent, profile correlations were computed for both the need and value profiles (i.e., the instruments were considered to be variables and the needs/values were considered to be cases). The profile correlations were somewhat higher for the needs (r =.71) than for the values (r =.66).

In summary, the WIL-P&P showed moderate correspondence with the WIP-C when the two forms were considered as alternative or parallel forms of one another. With the correction for ipsatization, correlations between individual needs ranged from .61 to .73, and correlations among individual values ranged from .73 to .85. Profile correlations were of similar magnitude, with a value of .71 for the need profile and .66 for the values profile. These analyses provide moderate support for obtaining the same results from either the paper/pencil or computerized Work Values Profile.

Internal Consistency Reliability

The internal consistency reliabilities of the six work values scales were estimated using coefficient alpha, an index of how well the items in a scale measure the same construct. High values arise when items are highly correlated and, thus, indicate that the items are measuring the same construct; low values indicate either that the items are not measuring any clear construct, or that they are measuring two or more constructs that are not highly related.

As discussed previously, the ipsative scoring of the WIL-P&P attenuates internal consistency values because most of the inter-item correlations are necessarily negative. Scales with more items encounter greater attenuation because there is more competition among the items within the scale. Therefore, Value 1 (which has only two items) would be attenuated the least, and Value 2 (which has six items) would be attenuated the most.

The coefficient alpha statistics appear in Table 22. The effects of ipsatization were approximately removed from the WIL-P&P scores by adding each respondent's WIP-C or MIQ mean need score to his or her WIL-P&P need scores (after transforming the WIL-P&P scores to the same metric as the WIP-C). Alpha is reported for both the corrected and uncorrected WIL-P&P scores. The reliabilities for the uncorrected WIL-P&P scores are more relevant because these scores, rather than the corrected scores, will be the ones used in practice. Reliabilities were very low for the WIL-P&P (median alpha = .20). The ipsatization process had a marked adverse affect on the WIL-P&P internal consistency. The average decrement was .38.

In summary, the internal consistency reliabilities for the WIL-P&P scales are quite low largely because of the WIL-P&P's ipsative scoring. This can be expected to reduce correlations of WIL-P&P scores with scores on the WIP-C or MIQ because low reliabilities restrict the degree to which a measure can correlate with other measures.

Table 22. ***Internal Consistency Reliability Estimates (Coefficient Alpha)***

		Tests			
	Number	WIL-P&P		$WIL\text{-}P\&P_{adjusted}$	
Scale	of Items	Test	Retest	Test	Retest
1. Achievement	2	33	30	58	57
2. Comfort	6	-21	-41	55	59
3. Status[1]	3	-05	02	47	53
4. Altruism	3	06	-07	35	37
5. Safety	3	35	49	57	59
6. Autonomy	3	42	50	64	68
N		1207	234	668	228

Note. Decimal points omitted. WIL-P&P = paper-and-pencil profiler; $WIL\text{-}P\&P_{adjusted}$ = paper-and-pencil profiler with scores adjusted using importance ratings from WIP-C or MIQ. Test = instrument taken at Time 1; Retest = instrument taken at Time 2.

[1]Scale 3 has 4 items in the MIQ but only 3 items in the WIL-P&P.

Summary of the Reliability Results

The WIL-P&P displayed moderate reliability in most analyses. The reliabilities were attenuated largely due to its ipsative scoring. Test-retest results showed moderate correspondence within individuals tested at two times several weeks apart. Individuals had the same top value 62 percent of the time. They ranked the top value at Time 1 with the first or second rank at Time 2 80 percent of the time. Profiles of need and value ranks between Time 1 and Time 2 scores correlated .64 and .62 for needs and values, respectively. A method used to adjust for the effects of ipsatization did help somewhat to improve correlations of the WIL-P&P with the WIP-C. Values on the two forms correlated in the .70s and .80s, with a median of .77, after the correction for ipsatization. Internal consistencies were low, with a median value of .20, arguably due to (ipsatization).

Chapter 8. Main Study: Alternate Forms Reliability and Preliminary Validity Evidence for the WIL-P&P

Introduction

The WIL-P&P was designed to measure the same constructs as the original MIQ. The WIL-P&P closely paralleled the MIQ, but wording changes on items between the MIQ and WIL-P&P limit their treatment as parallel or alternative forms (see Table 2). The analyses described in this chapter were intended to determine the degree to which: a) the WIL-P&P and the MIQ appeared to be measuring the same constructs[2], and b) whether the constructs were those identified by the Theory of Work Adjustment (Dawis, Lofquist, & Weiss, 1968). For the alternative (computer-based) profiler, the WIP-C, exploratory and confirmatory factor analyses were conducted to see if the factor structure of the needs (i.e., items) in the instruments reflected the six hypothesized work values (i.e., scales) from the MIQ. The largely supportive outcomes of these factor analyses are presented in Appendices C and D. The WIL-P&P was not included in the factor analyses because ipsative scales do not yield correlations amenable to factor analysis. Here, the initial eigenvalues of the item correlation matrix (with mean squared correlations in the diagonal) indicated a 20-factor solution, an expected result due entirely to the WIL-P&P's ipsative scoring.

Analyses

Preliminary validity analyses for the WIL-P&P focused on the relationships between WIL-P&P and MIQ results to determine whether they appeared to represent the same constructs. Table 23 shows correlations between WIL-P&P and MIQ scores for the 20 individual needs and for the 6 work values. The correlations for the items ranged from .27 to .63, and the correlations for the values ranged from .30 to .49. Thus, correlations were low to moderate. Note that, except for Item 13, all items showing correlations less than .40 had their wording changed. This table reinforces the conclusion that the wording changes had a substantial impact on the properties of this instrument.

[2]Construct labels were changed shortly after completion of these analyses. Changes were designed to improve user interpretability. For instance, the traditional psychological construct label "autonomy" was changed to the lax term "independence." These changes are depicted in the first column of Table 23.

Table 23. ***Correlations Between the Instruments***

			Measures
Value	Item	Need Name	P–M
1. Achievement	1[1]	Ability Utilization	35
	2[1]	Achievement	27
2. Working Conditions	3	Activity	50
	7	Compensation	63
	10[1]	Independence	57
	14	Security	51
	19	Variety	46
	20	Working Conditions	37
3. Recognition	4	Advancement	44
	5[2]	Authority	39
	12[1]	Recognition	42
	16[2]	Social Status	
4. Relationships	8[2]	Co-Workers	54
	11[3]	Moral Values	48
	15	Social Service	53
5. Support	6[2]	Company Policies	38
	17[1]	Supervision: Human Relations	41
	18[1]	Supervision: Technical	45
6. Independence	9[1]	Creativity	47
	13	Responsibility	35
	21	Autonomy	44
1. Achievement			34
2. Comfort			43
3. Status			30
4. Altruism			47
5. Safety			49
6. Autonomy			45

Note. Decimal points omitted. N = 670 for P-M. M = MIQ and P = WIL-P&P.
[1]Minor difference in the wording of the MIQ vs. other versions for this item.
[2]Moderate difference in the wording of the MIQ vs. other versions for this item.
[3]Substantial difference in the wording of the MIQ vs. other versions for this item.

An additional set of analyses assessed the decision consistency between instruments. These analyses are most appropriate when the top values are used to help select an occupation. The analyses focused on the top two values for each person. Table 24 shows that the top value was usually the same for each pair of instruments (on average, 57 percent of the time), and the top value on the WIL-P&P was one of the top two values on the MIQ 79 percent of the time.

Table 24. ***Percentages of Matches for Top Values Across Work Values Measures by Order of Administration***

Measures	Top value matches	Top two values match any order	Top value on PP is one of top two on MIQ
MIQ / PP either order (n = 289)	57%	16%	79%
MIQ - PP (n = 155)	57%	18%	79%
PP - MIQ (n = 134)	58%	14%	80%

Note: C = computerized profiler; PP = paper-and-pencil profiler.

The final set of analyses examined the similarity among the respondents' MIQ and WIL-P&P profiles. For each respondent, profile correlations were computed for both the need and value profiles. That is, for each respondent, the instruments were considered to be variables and the needs/values were considered to be cases. The profile correlations were higher for the needs than for the work values. The mean profile correlation for the needs was .67, and for the values it was .57.

Summary

The WIL-P&P showed moderate correspondence with the MIQ in correlations of individual needs and individual values and on the top need. Many of the low or moderate correlations between the WIL-P&P and MIQ are the result of ipsatization on the WIL-P&P. Ipsatization prevents high internal consistencies for WIL-P&P scales and, thereby, attenuates the size of correlations between the WIL-P&P and other measures such as the MIQ. Decisions based on the WIL-P&P's most important work value could be expected to have a very respectable level of correspondence with the MIQ, based on the top value match data for the two instruments (Table 24). Correlations of profiles of needs and values also showed moderately sized correlations. Relationships of the WIL-P&P and the MIQ are sufficiently strong to support the use of the WIL-P&P as a measure of work values.

Chapter 9. Summary and Conclusions

Summary of Methods and Results

This report has detailed the development of the paper-and-pencil version of the O*NET Work Importance Locator (WIL-P&P). Two other reports describe the development of the O*NET Computerized Work Importance Profiler (WIP-C) and the Occupational Reinforcer Patterns (ORPs) that were also components of the Work Values project sponsored by USDOL.

The WIL-P&P is a self-administered measure of work values that is modeled on the Minnesota Importance Questionnaire (MIQ; Rounds et. al., 1981). Respondents complete the WIL-P&P by categorizing 20 need statements in terms of their relative importance in the respondents' ideal job. Clients can self-score the results of the assessment and obtain immediate feedback regarding their most important work values and work needs.

The development of the WIL-P&P involved three studies. In the Pre-Pilot Study, 21 employment center clients completed the draft version of the WIL-P&P. Results were used to improve the profiler (e.g., revised instructions, modified card sorter sheet, improved format). In the Pilot Study, 48 employment center participants took the improved WIL-P&P. Administrators observing the participants identified computational errors by respondents to be the biggest problem, with one out of six respondents (on average) making at least one error. Participant reactions to the measure, however, were favorable. In the Main Study, the WIL-P&P, WIP-C, and MIQ were administered to employment center clients and junior college students at 23 sites. Respondents took two of the three instruments so that information on the same respondents taking different measures would be available. Some respondents also took the same instruments at two points in time so that test-retest reliabilities could be computed. Results included the following:

a) the WIL-P&P required the least administration time;
b) few subgroup differences were found for race/ethnic group and gender groups, although several differences by educational level were apparent;
c) reliability analyses were mixed, with encouraging results from the test-retest estimates and discouraging results from the internal consistency estimates (likely due to the ipsative scoring of the measure); and
d) the ipsative scoring also attenuated correlations of the WIL-P&P with the other work values measures, but respectable consistency (in identifying the most important work value) was observed across time.

Conclusions

The WIL-P&P provides clients who are interested in career exploration with a fast, engaging measure for identifying their most important work value. The WIL-P&P can be expected to accurately identify a client's primary work value. The WIL-P&P also has great utility in providing many clients (especially those just entering the job market) with an opportunity to consider a facet

of job satisfaction that they might not have considered previously. Further, experienced clients will benefit from the WIL-P&P because the complaints of experienced employees about their previous jobs are typically couched in terms of work values that were not sufficiently reinforced (e.g., insufficient pay, lack of promotion opportunities, lack of support from upper management). As such, the WIL-P&P can be expected to have high levels of face validity for its users.

References

Borgen, F.H., Weiss, D.J., Tinsley, H.E.A., Dawis, R.V., & Lofquist, L.H. (1968). The measurement of occupational reinforcer patterns. *Minnesota Studies in Vocational Rehabilitation, 25*.

Dawis, R.V., & Lofquist, L.H. (1984). *A psychological theory of work adjustment*. Minneapolis, MN: University of Minnesota Press.

Dawis, R.V., Lofquist, L.J., & Weiss, D.J. (1968). A theory of work adjustment (a revision). *Minnesota Studies in Vocational Rehabilitation, 23*.

Gay, E.G., Weiss, D.J., Hendel, D.D., Dawis, R.V., & Lofquist, L.H. (1971). Manual for the Minnesota Importance Questionnaire. *Minnesota Studies in Vocational Rehabilitation, 28*.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Englewood Cliff, NJ: Prentice-Hall.

Lofquist, L.H., & Dawis, R.V. (1969). *Adjustment to work*. New York: Appleton-Century-Crofts.

McCloy, R., Waugh, G., Medsker, G., Wall, J., Rivkin, D., & Lewis, P. (1999a). Development of the O*NET Computerized Work Importance Profiler. Raleigh, NC: National Center for O*NET Development.

McCloy, R., Waugh, G., Medsker, G., Wall, J., Rivkin, D., & Lewis, P. (1999b). Determining the occupational reinforcer patterns for O*NET occupational units. Raleigh, NC: National Center for O*NET Development.

Peterson, N., Mumford, M. Borman, W., Jeanneret, P., & Fleishman, E. (1995). *Development of prototype Occupational Information Network (O*NET) content model volume I: Report*. Salt Lake City, UT: Utah Department of Workforce Services.

Peterson, N., Mumford, M., Borman, W., Jeanneret, P., & Fleishman, E. (1995). *Development of prototype Occupational Information Network (O*NET) content model volume II: Appendices*. Salt Lake City, UT: Utah Department of Workforce Services.

Peterson, N., Mumford, M., Borman, W., Jeanneret, P., Fleishman, E., & Levin, K.Y. (1997) *O*NET final technical report* (Vols. 1-3). Salt Lake City UT: Utah Department of Workforce Services.

Rounds, J.B., Henly, G.A., Dawis, R.V., Lofquist, L.H., & Weiss, D.J. (1981). *Manual for the Minnesota Importance Questionnaire: A measure of needs and values*. Minneapolis: University of Minnesota Department of Psychology.

Rounds, J.B., Miller, T.W., & Dawis, R.V. (1978). Comparability of multiple rank order and paired comparison methods. *Applied Psychological Measurement,* 2, 415-422.

Schaffer, R.H. (1953). Job satisfaction as related to need satisfaction at work. *Psychological Monographs, 67(*364).

Stephenson, W. (1953). *The study of behavior*. Chicago: University of Chicago Press.

Stewart, E.S., Greenstein, S.M., Holt, N.C., Henly, G.A., Engdahl, B.E., Dawis, R.V., Lofquist, L.H., & Weiss, D.J. (1986). *Occupational reinforcer patterns*. Minneapolis: Vocational Psychology Research, University of Minnesota Department of Psychology.

U.S. Department of Labor. (1965). *Dictionary of occupational titles* (3rd ed.). Washington, DC: U.S. Government Printing Office.

U.S. Department of Labor. (1977). *Dictionary of occupational titles* (4th ed.). Washington, DC: U.S. Government Printing Office.

U.S. Department of Labor. (1991). *Dictionary of occupational titles* (Rev. 4th ed.). Washington, DC: U.S. Government Printing Office.

U.S. Department of Labor. (1970). *Manual for the USES General Aptitude Test Battery*, *Section III: Development*. Washington, DC: U.S. Government Printing Office.

Weiss, D.J., Dawis, R.V., England, G.W., & Lofquist, L.H. (1964). Construct validation studies of the Minnesota Importance Questionnaire. *Minnesota Studies in Vocational Rehabilitation, 18*.

Appendix A

WVP-P Materials
(WVP-P later changed to WIL-P&P)

Name:_____
Social Security Number:_____ - _____ - _____

WORK VALUES PROFILER

Instructions

This survey is designed to help you find out what you consider important in your ***ideal job*** (that is, the kind of job you would most like to have). You can use the results of this survey to find out about the kinds of jobs you prefer.

This survey is different from others you may have taken in the past. Instead of questions, this survey has cards with statements about different kinds of jobs on them. You will sort the cards into groups based on how important the statement on each card is to you.

Please complete and score the survey yourself by following the numbered steps in order.

Before You Begin

First, write your name and social security number at the top of this page.

There are no right or wrong answers in this survey; you are just being asked how you feel. Answer as honestly and carefully as you can.

You will need the following materials to complete the survey:

- 20 work values cards;
- a work values card sorting page;
- these instructions.

Remember, do all the steps in order.

Step 1. Read the Cards

Think about how important it is for you to have a job like the one described on each card. Read the cards before going to Step 2.

Go To Step 2

(Turn the page)

Name:_____
Social Security Number:_____ - _____ - _____

Step 2. Sort the Cards

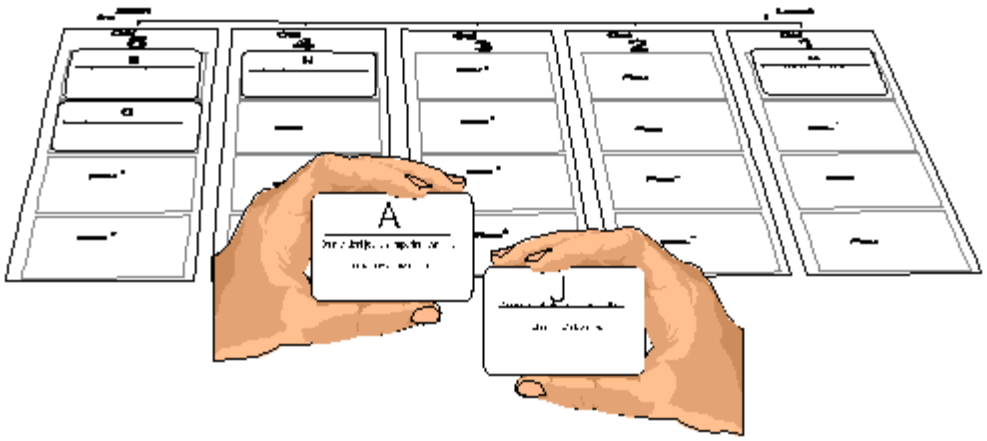
- Find the "Work Values Card Sorting Page".
- Notice the five columns printed under the importance scale that is at the top of the page. Each column has a number that refers to a point on the importance scale.
- Rank the cards by putting each one in the column that best matches how important it would be for you to have a job like the one described on each card.
- Put ***exactly*** 4 cards in each column.
- When all the cards are ranked, go to Step 3.

For example:

If card A describes something that is more important to you than what is on the other cards, put card A in Column 5. On the other hand, if what is written on card A is least important to you compared to the other cards, put card A in Column 1. If the importance of the card is in between "most important" and "least important" put the card into the column that best matches how you feel.

Do the same thing for all of the cards, but remember, put four cards in every column.

When you are done, the four most important statements will be in Column 5, the four next-most important will be in Column 4, and so on. The four least important statements should be in Column 1.



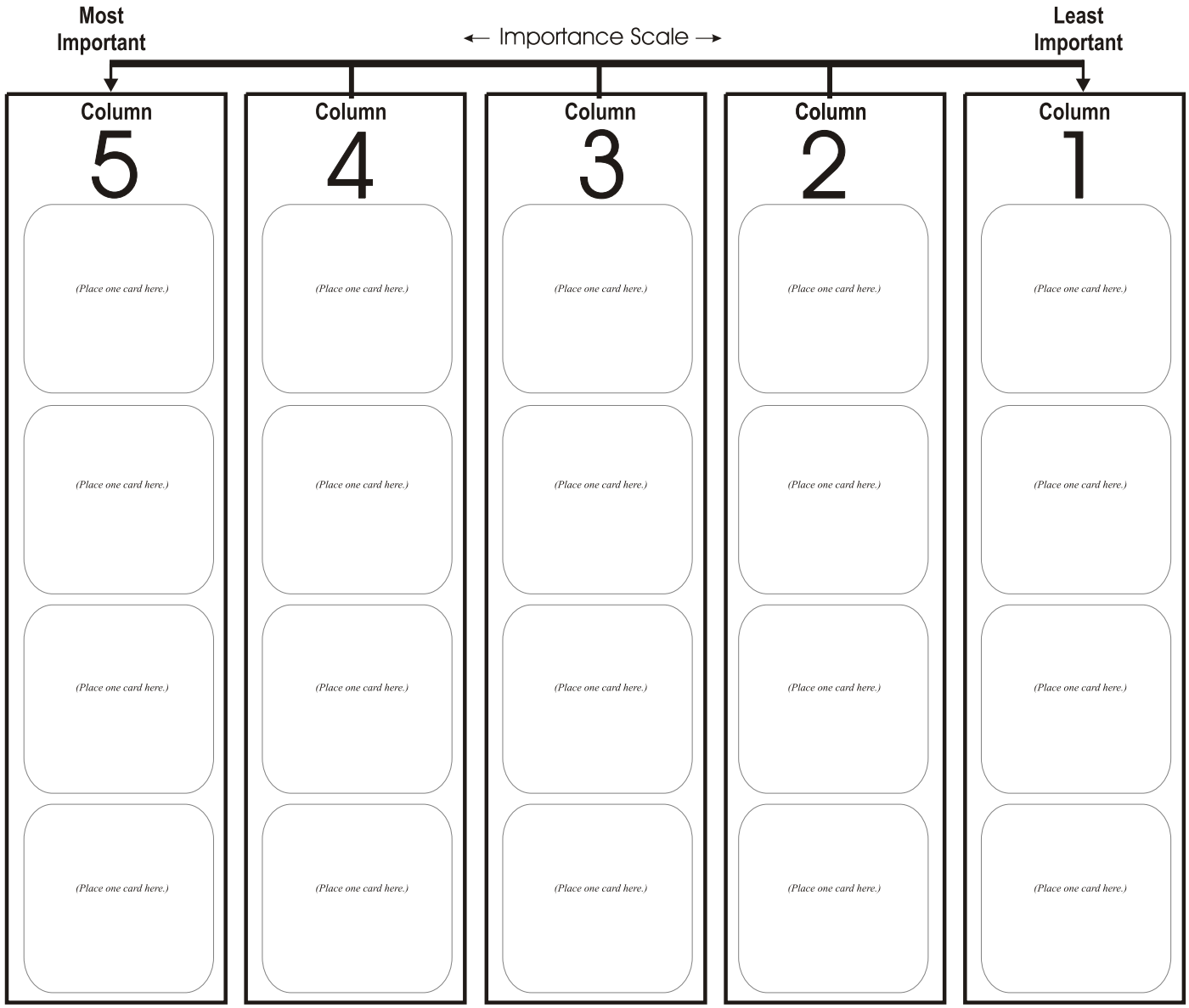
When all cards are ranked,
Go To Step 3

Work Value Card Sorting Sheet

Step 2. Sort the Cards

- Notice the five columns printed under the Importance Scale at the top of the sheet. Each column has a number from 5 (Most Important) to 1 (Least Important).
- Put each card in the column that best matches how important it is for you to have a job like the one described on the card.
- Put exactly 4 cards in each column.
- When you are done, the four most important statements should be in Column 5, the four next most important should be in Column 4, and so on. The four least important statements should be in Column 1. You may have to shuffle the cards around until you have exactly 4 cards in each column.

After you have sorted all of the cards, go to **Step 3** and figure out your scores.



Step 3. Figure Out Your Scores

- Your score for each card is the number of the column you put it in. For example, the cards in Column 5 each get a score of 5; the cards in Column 4 each get a score of 4, and so on.
- For example, find the column containing card A on the Sorting Page. Record the number of that column in the space provided for card A in the Work Value 1 chart below. Follow the same procedure to record the score for each card.
- **Add up the numbers in each chart and write the total in the space provided.**
- **Multiply each total score by the number shown at the bottom of each chart. Work Value 2 does not need to be multiplied.**

Work Value 1	
Card	**Score**
A	
F	+
Add Scores for TOTAL ☛	
Multiply TOTAL by 3 ☛	x 3
Work Value 1 Score ➪	=

Work Value 2	
Card	**Score**
C	
G	+
J	+
N	+
R	+
S	+
ADD scores for TOTAL	=
Work Value 2 Score ➪	

Work Value 3	
Card	**Score**
D	
E	+
L	+
Add Scores for TOTAL ☛	
Multiply TOTAL by 2 ☛	x 2
Work Value 3 Score ➪	=

Work Value 4	
Card	**Score**
H	
K	+
O	+
Add Scores for TOTAL ☛	
Multiply TOTAL by 2 ☛	x 2
Work Value 4 Score ➪	=

Work Value 5	
Card	**Score**
B	
P	+
Q	+
Add Scores for TOTAL ☛	
Multiply TOTAL by 2 ☛	x 2
Work Value 5 Score ➪	=

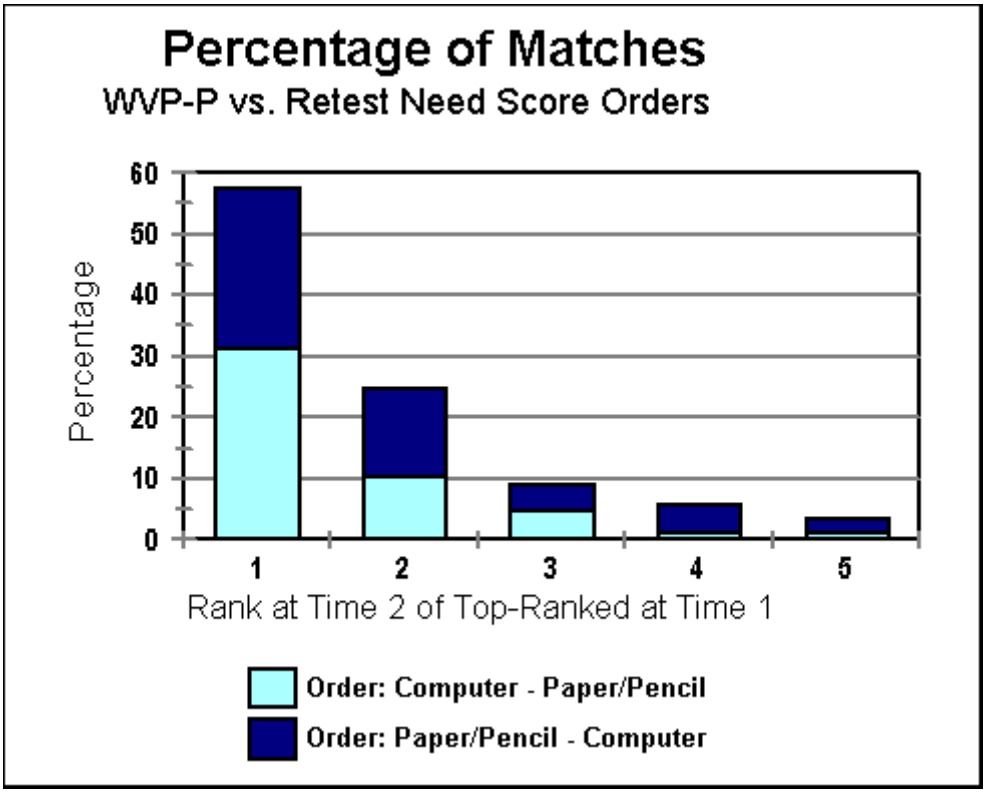
Work Value 6	
Card	**Score**
I	
M	+
T	+
Add Scores for TOTAL ☛	
Multiply TOTAL by 2 ☛	x 2
Work Value 6 Score ➪	=

Appendix B

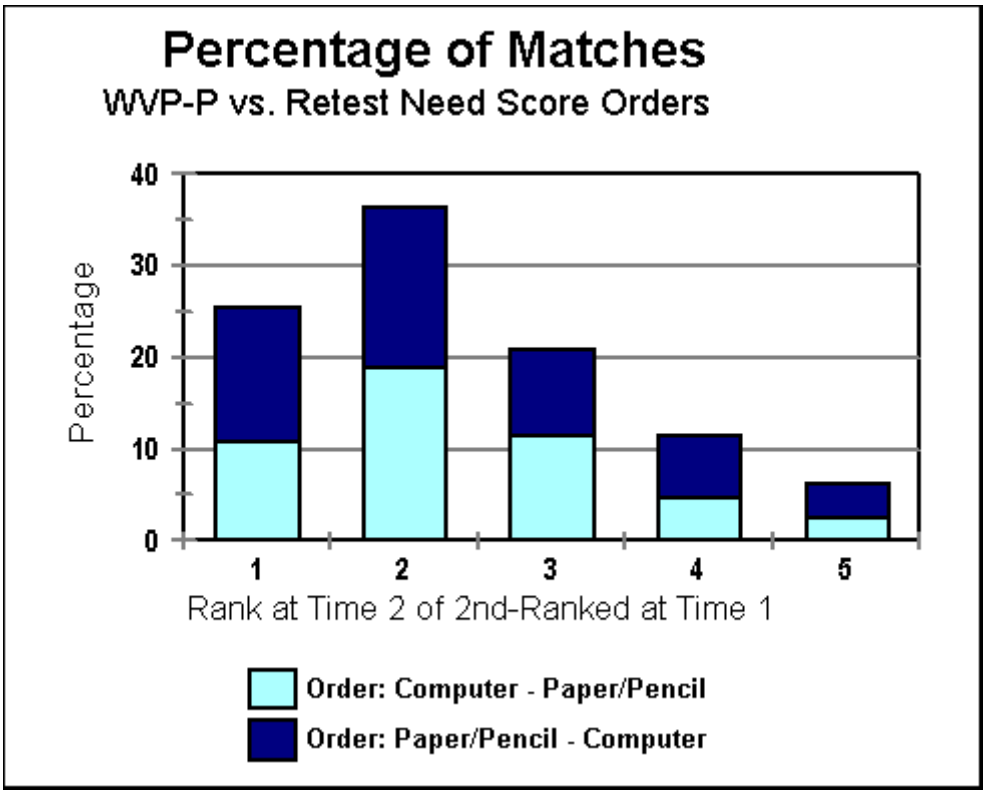
Histograms Showing Test-Retest Results (WVP-P = WIL-P&P; WVP-C = WIP-C)

Histograms Showing Test-Retest Results for Need Statements on the WVP-P

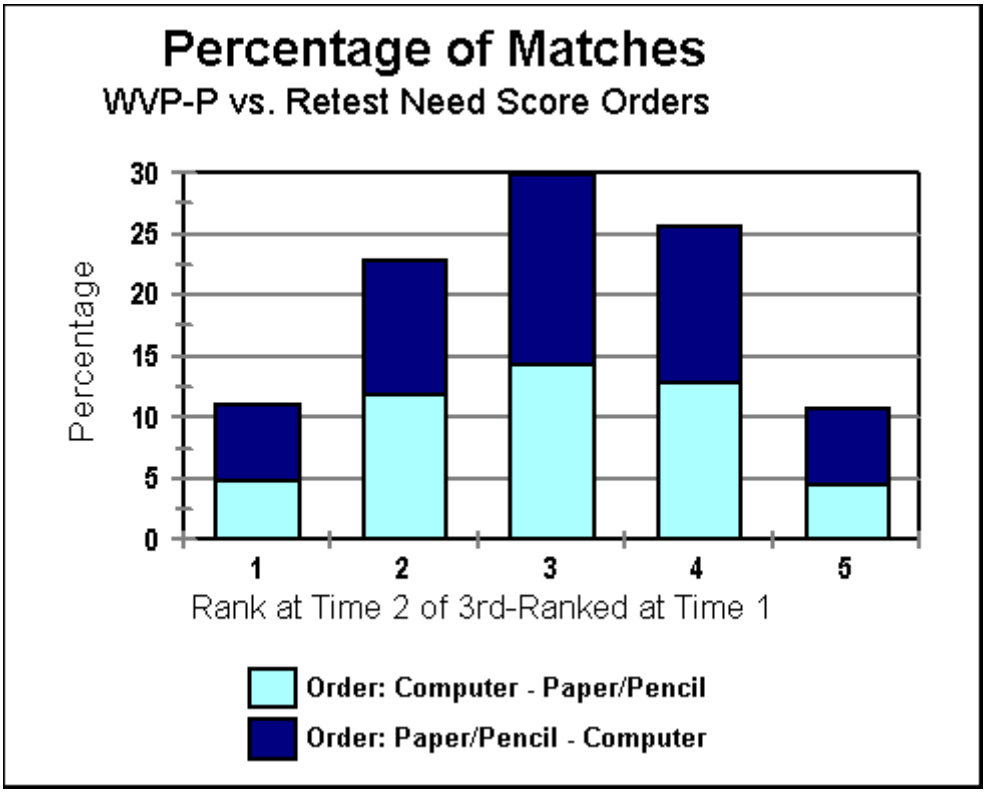
The following set of histograms indicates whether the cards sorted by the respondents were placed in the same pile the first time the respondent took the WVP-P and the second time the respondent took the WVP-P. Cards were sorted into five piles based on the importance to the respondent of the need statements shown on the cards. Thus, these piles represent a rank ordering of the need statements into five ranks of four need statements in each rank or pile. The first histogram examines whether or not the need statements placed in pile 1 at Time 1 were also placed in pile 1 at Time 2. The second histogram examines the same question for need statements in pile 2 at the two times. The third, fourth, and fifth histograms examine the same question for the third, fourth, and fifth piles. The frequency tables that follow the histograms show the numerical values on which the histograms were based. The histograms are stacked to indicate the order in which the examinees had taken the WVP-P and the WVP-C. The histograms demonstrate that the version of the WVP an examinee completed first had little effect on the similarity of ranking or piling of need statements across time (i.e., there is little order effect).



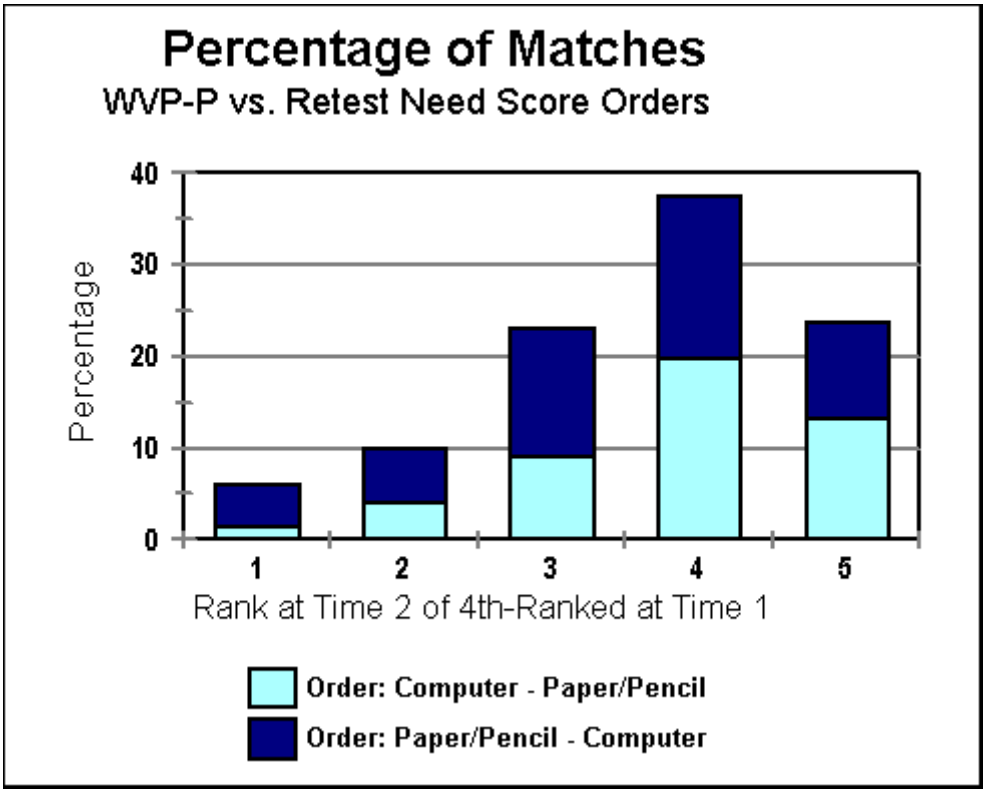
Histogram 1: Rank at Time 2 of Top-Ranked WVP-P Need at Time 2



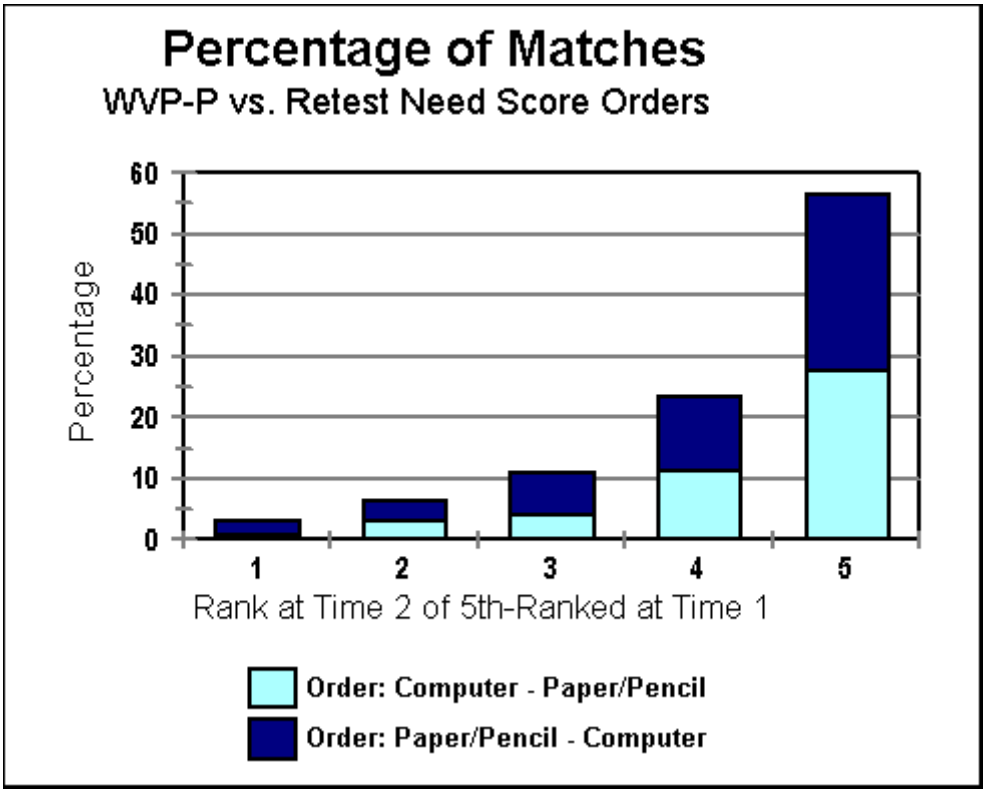
Histogram 2: Rank at Time 2 of 2nd-Ranked WVP-P Need at Time 2



Histogram 3: Rank at Time 2 of 3rd-Ranked WVP-P Need at Time 2



Histogram 4: Rank at Time 2 of 4th-Ranked WVP-P Need at Time 2



Histogram 5: Rank at Time 2 of 5th-Ranked WVP-P Need at Time 2

Percentage of Matches of WVP-P Test vs. Retest Need Score Orders

Frequency Table for Histogram 1

Test Order: WVP-C then WVP-P			Pile at Time 2 of Pile 1 at Time 1 (n = 373)	
P12RNK1	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
1	1141	63.0	1141	63.9
2	381	21.3	1522	85.2
3	170	9.5	1692	94.7
4	52	2.9	1744	97.6
5	43	2.4	1787	100.0
Test Order: WVP-P then WVP-C			**Pile at Time 2 of Pile 1 at Time 1 (n = 501)**	
1	954	51.3	954	51.3
2	514	27.6	1468	79.0
3	160	8.6	1628	87.6
4	149	8.0	1777	95.6
5	82	4.4	1859	100.0

Frequency Table for Histogram 2

Test Order: WVP-C then WVP-P			Pile at Time 2 of Pile 2 at Time 1 (n =486)	
P12RNK2	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
1	370	22.1	370	22.1
2	648	38.7	1018	60.8
3	403	24.1	1421	84.9
4	169	10.1	1590	95.0
5	84	5.0	1674	100.0
Test Order: WVP-P then WVP-C			**Pile at Time 2 of Pile 2 at Time 1 (n = 589)**	
1	503	28.4	503	28.4
2	602	34.0	1105	62.4
3	315	17.8	1420	80.2
4	222	12.5	1642	92.7
5	129	7.3	1771	100.0

Frequency Table for Histogram 3

Test Order: WVP-C then WVP-P			Pile at Time 2 of Pile 3 at Time 1 (n =384)	
P12RNK3	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
1	181	10.2	181	10.2
2	429	24.3	610	34.5
3	520	29.4	1130	64.0
4	469	26.6	1599	90.5
5	167	9.5	1766	100.0
Test Order: WVP-P then WVP-C			**Pile at Time 2 of Pile 3 at Time 1 (n =501)**	
1	220	11.8	220	11.8
2	395	21.2	615	33.1
3	564	30.3	1179	63.4
4	457	24.6	1636	88.0
5	223	12.0	1859	100.0

Frequency Table for Histogram 4

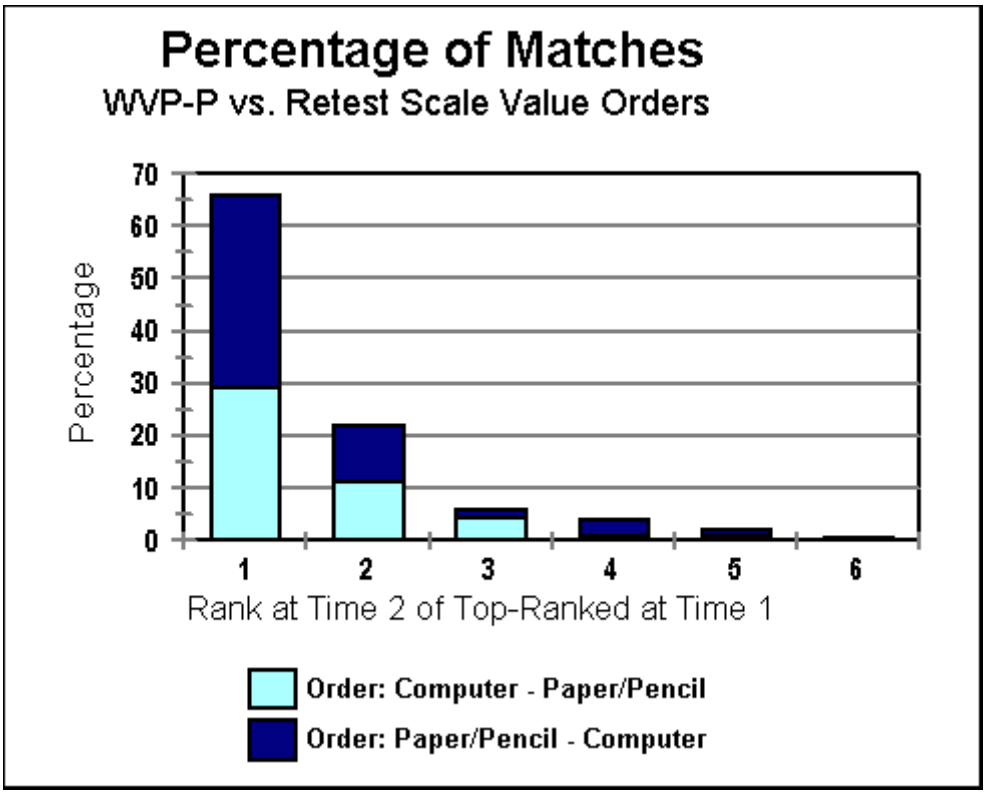
Test Order: WVP-C then WVP-P			Pile at Time 2 of Pile 4 at Time 1 (n =318)	
P12RNK4	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
1	59	3.2	59	3.2
2	164	8.9	223	12.1
3	355	19.3	578	31.4
4	756	41.0	1334	72.4
5	508	27.6	1842	100.0
Test Order: WVP-P then WVP-C			**Pile at Time 2 of Pile 4 at Time 1 (n =378)**	
1	171	8.6	171	8.6
2	213	10.7	384	19.4
3	527	26.6	911	46.0
4	673	34.0	1584	79.9
5	398	20.1	1982	100.0

Frequency Table for Histogram 5

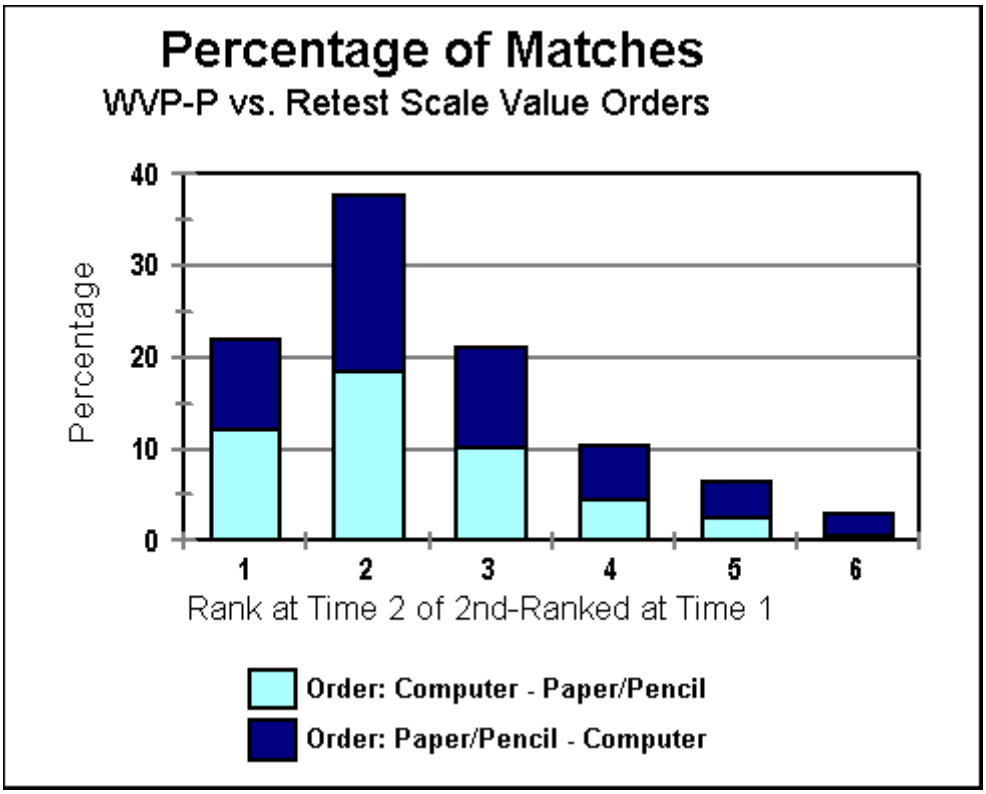
Test Order: WVP-C then WVP-P			Pile at Time 2 of Pile 5 at Time 1 (n =202)	
P12RNK5	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
1	41	2.1	41	2.1
2	128	6.5	169	8.6
3	176	9.0	345	17.6
4	459	23.4	804	41.1
5	1154	58.9	1958	100.0
Test Order: WVP-P then WVP-C			**Pile at Time 2 of Pile 5 at Time 1 (n =162)**	
1	88	4.0	88	4.0
2	127	5.8	215	9.8
3	275	12.5	490	22.3
4	509	23.2	999	45.5
5	1199	54.5	2198	100.0

Histograms Showing Test-Retest Results for the Six Values on the WVP-P

Each of the six values on the WVP-P were based on the combination of the individual need statements which were related to that value based on the theoretical development of the MIQ. The following set of histograms shows whether values were given the same rank the first time the respondent took the WVP-P and the second time the respondent took the WVP-P. The first histogram examines whether or not the value which a respondent rated at the top at Time 1 was also top ranked by the respondent at Time 2. The second histogram examines the same question for values ranked second at the two times. The frequency tables that follow the histograms show the numerical values on which the histograms were based. The histograms are stacked to indicate the order in which the examinees had taken the WVP-P and the WVP-C. The histograms demonstrate that the version of the WVP an examinee completed first had little effect on the similarity of ranking or piling of values across time (i.e., there is little order effect).



Histogram 1: Rank at Time 2 of Top-Ranked WVP-P Value at Time 2



Histogram 2: Rank at Time 2 of 2nd-Ranked WVP-P Value at Time 2

Percentage of Matches of WVP-P Test vs. Retest Scale Value Orders

Frequency Table for Histogram 1

Test Order: WVP-P then WVP-C			Rank at Time 2 of Top-Ranked at Time 2	
P12RANK1	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
1	65	60.2	65	60.2
2	22	20.4	87	80.6
3	8	7.4	95	88.0
4	4	3.7	99	91.7
5	7	6.5	106	98.1
6	2	1.9	108	100.0
Test Order: WVP-C then WVP-P			**Rank at Time 2 of Top-Ranked at Time 2**	
1	74	62.7	74	62.7
2	20	16.9	94	79.7
3	7	5.9	101	85.6
4	7	5.9	108	91.5
5	10	8.5	118	100.0

Frequency Table for Histogram 2

Test Order: WVP-P then WVP-C			Rank at Time 2 of Top-Ranked at Time 2	
P12RANK2	**Frequency**	**Percent**	**Cumulative Frequency**	**Cumulative Percent**
1	16	14.8	16	14.8
2	36	33.3	52	48.1
3	27	25.0	79	73.1
4	14	13.0	93	86.1
5	12	11.1	105	97.2
6	3	2.8	108	100.0
Test Order: WVP-C then WVP-P			**Rank at Time 2 of Top-Ranked at Time 2**	
1	20	16.9	20	16.9
2	42	35.6	62	52.5
3	23	19.5	85	72.0
4	17	14.4	102	86.4
5	11	9.3	113	95.8
6	5	4.2	118	100.0