# PREVUE ASSESSMENT™

# TECHNICAL MANUAL

Published by:

ICES ASSESSMENT SYSTEMS INC.

Authored by
David Bartram, BA, DPhil

Newland Park Associates Ltd.
Chartered Occupational Psychologists

Fourth Edition: June 2000 (NA)

# FOREWORD

The Prevue Assessment utilizes the ICES Plus test battery developed at the University of Hull, England by View Assessments International Inc. and Dr. David Bartram of Newland Park Associates Ltd. ("NPAL").

This technical manual describes the development of the ICES Plus test battery. The development and data analysis have been carried out by NPAL on behalf of View Assessments International Inc. All the data was collected by View Assessments International Inc. (or their agents) and transferred to NPAL in computer-readable form.

Technical Manual Author: Dr. Dave Bartram

NPAL Test Development Team:
Dr. Dave Bartram,
Dr. Patricia Lindley,
Dr. Iain J. Coyne

The first edition of the Technical Manual was published in 1992. A second edition was published in 1994 following completion of the third phase of studies undertaken in the development of the ICES Plus battery. Continued research, development and validation studies prompted a third edition in 1998.

The fourth edition reports on a major new initiative to explore the relationship between personality traits and risk-taking, adaptability to change and a person's focus on work. This research is the basis for the new "Working Characteristics Report" that is available from the Prevue Assessment System. This fourth edition also incorporates further validation studies undertaken since 1998.

# Table of Contents

# PART I: CONSTRUCTION OF THE ICES PLUS SCALES

## 1. GENERAL INTRODUCTION

### 1.1 Contents of the ICES Plus Battery

The ICES Plus battery was designed to provide a reliable but rapid means of assessing Interest, Ability and Personality. It includes:

- **A Personality Assessment instrument (ICES)** which covers four major personality dimensions (Independence, Conscientiousness, Extraversion and Stability) each of which is represented by two "minor" scales;
- **An inventory of interests (ICES Plus Interest Inventory)** which assesses occupational interest in relation to working with People, Data and Things (PDT);
- **An Ability scale (ICES Plus Ability)** represented by a set of three Ability tests-designed to assess Numerical, Verbal and Spatial ability (NVS).

### 1.2 Outline of the Development Program

The development program was designed to achieve three main aims.

1. To create a battery of tests (ICES Plus) containing a balanced mix of Interest, Ability and Personality scales (ICES Plus scales).
2. To obtain a broad sample of data from different people working in different job areas in the United States, Canada, the United Kingdom and the Far East.
3. To provide norms, reliability data and initial validation data on the scales.

Development of ICES Plus began in 1991 and continued through to the end of 1993. It took place in three main Phases, the first two of which can each be divided into two stages.

Phase One concentrated on the development of the ICES and PDT scales and used a large sample, mainly from one North American organization. This sample was divided into two parts. The first (Stage One) was used for initial scale development work and the second (Stage Two) acted as a "hold-out" sample for cross-validation.

Phase Two, Stage One focused on broadening the range of organizations and job types covered. Phase Two, Stage Two continued this with the additional aims of providing re-test and construct validity data for the ICES scales.

The main focus for Phase Three was the development and initial validation of the Ability scales. Work on the Numerical Reasoning test (Working with Numbers) was carried on throughout Phases One and Two, with the final format and item content being fixed for Phase Three. The Spatial and Verbal scales were piloted and developed during Phase Three. Further data on ICES and PDT were also obtained during this Phase.

# 2. CONSTRUCTION OF THE TRIAL ICES PLUS BATTERY

## 2.1 The ICES Personality Inventory

### 2.1.1 Rationale

There have been many debates over the factor structure of personality. However, there is an increasing consensus on the identity and generality of a relatively small number of domains of personality (e.g. McRae & Costa, 1987; Digman, 1990) and on their utility as a framework for studies of validity (Barrick & Mount, 1991; Tett, Jackson & Rothstein, 1991; Robertson, 1993; Robertson & Kinder, 1993). Often referred to as the "Big Five" factors, these have been labeled as Extraversion; Agreeableness/Independence; Conscientiousness; Anxiety; and Openness to Experience.

Barrick & Mount (1991) report a meta-analysis of 117 criterion-related validity studies. They classified Personality data from these studies in terms of the "Big Five" factors. They then examined the prediction of three different job performance criteria (job proficiency, training proficiency and personnel data - such as tenure) for five occupational groups (from professional to skilled/semi-skilled). The main findings were that Conscientiousness was a valid predictor of all criteria across all occupational groups; Extraversion was a valid predictor for managers and sales personnel across all criteria; other scales were valid predictors of some criteria for certain occupational groups.

Tett et al (1991) analyzed 97 samples representing data on 13,521 people. They found average corrected validities for the "Big Five" factors ranging from 0.16 (for Extraversion) to 0.33 (for Agreeableness). What is more, the average validity for studies that had used confirmatory research approaches was 0.29, with the mean rising to 0.38 for studies that had used a job analysis as the basis for selecting relevant personality variables. Robertson and Kinder (1993) report criterion-related validity coefficients for personality variables of up to 0.33 - without corrections for range restriction or unreliability. The highest validities are obtained for criteria such as creativity, analysis and judgment. As well, they conclude that personality measures add to the level of prediction obtained with ability measures on their own.

The approach adopted for the development of the ICES scales has been to accept the view that there are four or five major dimensions. In addition, it is also accepted that there is a need for scales which are conceptually distinct but which may be quite highly correlated. Height and weight, for example, are highly correlated (and would be difficult to "identify" as factorially distinct) and yet they describe very different qualities of people. In a similar manner, it is important to have Personality scales that are well defined *a priori* in content terms.

In designing the inventory, it was decided to focus on the development of good measures of four of these "Big Five" personality factors (the Extraversion, Independence, Conscientiousness, and Emotional Stability factors).

To have a short (i.e. about 100 items) but reliable inventory which covers the main domains of personality involves making certain trade-offs. It was decided that the priority should be the development of measures of the four main factors described above which were reliable, had minimal overlap with each other and were sufficiently broad to give good coverage of each domain. High reliability can always be obtained by generating items that focus on a very narrow aspect of the relevant content domain. Such scales measure what Cattell has called "bloated specifics": they look good on the surface in psychometric terms, but they may have little power or generality as personality descriptions.

If one is to tap the full range of a broad personality domain - and avoid the production of "bloated specifics" - it is inevitable that the items generated will have a relatively low mean inter-item correlation (i.e. between 0.10 and 0.20). As a result, broad scales need to have a relatively large number of items if they are to have acceptable internal consistencies.

For the present purposes, it was decided that internal consistencies should be at least 0.70 (and no more than 0.85) with each of the four scales having between 20 and 25 items. Further analyses of these scales would be carried out by examining item content and by item clustering to produce two sub-scales for each of the four main factors:

| Major Scales | Minor Scales |
|---|---|
| Independence: | |
| I1 | Competitive, tough-minded |
| I2 | Assertive, forthright |
| Conscientiousness: | |
| C1 | Conventional, traditional, concern with moral values |
| C2 | Organized, attention to detail, neatness |
| Extraversion: | |
| E1 | Group-oriented, sociable |
| E2 | Outgoing, group dependent |
| Stability: | |
| S1 | Poised, unruffled, not easy to annoy or upset |
| S2 | Relaxed, not anxious |

To summarize, the intention was to design an instrument that would provide measures at two "levels":

- Four major scales at the higher-order factor level providing a description of the location of a person within the main dimensions of personality factor space as defined by four of the "Big Five" personality factors;
- Eight minor scales to provide information for a richer descriptive interpretation of personality.

### 2.1.2 Production of items

Development of the eight personality sub-scales started from the definition of a larger number (eleven) of more specific putative "content" based scales (P1 to P11):

| | | | |
|---|---|---|---|
| P1 | Relaxed | P7 | Competitive |
| P2 | Conventional | P8 | Hard-headed |
| P3 | Proactive | P9 | Trusting |
| P4 | Group-dependent | P10 | Optimistic |
| P5 | Self-controlled | P11 | Outgoing |
| P6 | Assertive | | |

Descriptions of each of these were generated as lists of four or five adjectives describing hypothetical "low" and "high" scores. These were used as the basis for item generation and for an item-sorting task that was used to make initial identification of poor or ambiguous items. A large trial inventory was then constructed with the "best" items identified by the sorting task.

Four item writers generated an initial set of 270 items. These were written working from the P1 to P11 content definitions with the brief that:

(a) between 15 and 20 usable items would be needed for each of the categories;
(b) item content should give good coverage of the relevant domain;
(c) care should be taken to avoid items, which were invasive of a person's privacy.

The normal practice of keeping items as short and unambiguous as possible was followed.

The 270 items were then content-analyzed using a sorting task. Item sorting was carried out by four experienced psychologists. The items were printed on individual cards and presented to each judge for sorting into one of the 11 content categories - P1 to P11 - a category for items, which were judged to be Motivational Distortion (MD) scale items, and a final "undecided" category. Each of the P1 to P11 categories was defined by a card containing the descriptive adjectives for the low and the high ends of the dimension.

The judges sorted the items independently of each other and of the item writer. Judges were free to put as many or as few cards as they wished in each category. The item writer's assignment of each item was used together with that of four other judges. The majority of the items chosen for inclusion in the trial version of the inventory had three or

four judges agreeing with each other and the item writer. In some instances items have been included where only two judges agreed with the item writer (so long as the other two disagreed with each other).

A total of 179 of the 270 items met the criteria for inclusion in the trial inventory scales. The numbers below show that this resulted in certain categories dominating. A further iteration of the item generation process was carried out with 30 additional items written for these scales being sorted (with a set of filler items) with respect to all eleven categories.

This produced sufficient new items to bring all scales up to a minimum of 15 items each. Where there were more than 15 items, the items were examined for overlap. Items were discarded from overlapping pairs. The final trial inventory contained 196 items (Table 2.1 shows the breakdown of items for each scale). The hypothesized relationships between the proposed ICES scales and the eleven content categories are illustrated in Table 2.2. These hypotheses are based on commonality of scale construct definition and item content.

***Table 2.1: Construction of the initial ICES inventory.***

| Scale | ICES $_1$ | ICES $_2$ | Total | Final selection |
|---|---|---|---|---|
| P1 | 17 | | 17 | 15 |
| P2 | 16 | | 16 | 15 |
| P3 | 16 | | 16 | 15 |
| P4 | 20 | | 20 | 15 |
| P5 | 18 | | 18 | 15 |
| P6 | 19 | | 19 | 15 |
| P7 | 9 | 6 | 15 | 15 |
| P8 | 12 | 3 | 15 | 15 |
| P9 | 9 | 6 | 15 | 15 |
| P10 | 17 | | 17 | 15 |
| P11 | 19 | | 19 | 15 |
| MD | 7 | | 7 | 7 |
| Totals | 179 | 12 | 191 | 172 |
| Selected From | 270 | 30 | 300 | 300 |

***$ICES_1$ items are those remaining after the content analysis of the first set of 270 items.***
***$ICES_2$ items are those remaining after the content analysis of the second set of 30 new items.***

***Table 2.2. Hypothesized relationships between the content definitions and the proposed "Big Five" factor based ICES scales.***

| ICES | Content-based categories | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Scales | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P11 |
| Independence | | | | | | ** | ** | * | | | |
| I1 | | | | | | | ** | | | | |
| I2 | | | | | | ** | | | | | |
| Conscientiousness | | ** | ** | | | | | | | | |
| C1 | | ** | | | | | | | | | |
| C2 | | | ** | | | | | | | | |
| Extraversion | | | | ** | | | | | | * | ** |
| E1 | | | | ** | | | | | | | |
| E2 | | | | | | | | | | * | ** |
| Stability | ** | | | | ** | | | | * | | |
| S1 | | | | | ** | | | | | | |
| S2 | ** | | | | | | | | * | | |
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P11 |

**** Clear strong relationship hypothesized on the basis of content.**
*** Possible or weak relationship hypothesized on the basis of content.**

## 2.2 The ICES Plus Interest Inventory

### 2.2.1 Rationale

The US Department of Labor's Dictionary of Occupational Titles (DOT) is probably the most extensive and systemic categorization of occupations available to the human resource specialist. It uses an occupational coding system that incorporates three digits relating to the worker functions ratings of the tasks performed in the occupation. These three areas of function are referred to as People, Data and Things respectively. For each one, different numbers are used to indicate different types of activity. For example, computing is a Data task with a code of 4, driving-operating is a Things task with a code of 3. The codes are assigned such that lower numbers are used for more complex tasks, tasks that require more responsibility, or which are more complicated.

Interest inventories provide a useful indicator of whether a person's preferences are such as to be likely to suit them to particular areas of work. Someone who had no interest in working with data, whatever their ability would be unlikely to be satisfied working in an occupation with a low DOT Data code. Similarly someone who preferred working with things rather than people would be ill advised to enter occupations where negotiating with others was a vital part of the work.

Thus, in helping people make choices about possible areas of work, it is useful to have a means of assessing their interests in terms of work activities relating to the three areas: People, Data and Things. This, together with information about the skills and abilities, will provide basic information on which positive guidance can be based.

The new ICES Plus Interest Inventory was intended to provide a simple measure of a person's preferences for activities in the three general areas (People, Data and Things). This should be useful both in vocational guidance and in helping in placement decisions. It must be stressed that this inventory is designed to provide structured guidance - it is not a "test". It is intended to provide help in making occupational choices. It should not be used as the basis of simple occupational-matching decisions.

### 2.2.2 Construction of the inventory

A total of 51 items (17 for each of the three scales) were written. Each item described an activity or area of activities that focused on working with people, working with data or working with things. Each item was to be rated on a one-to-five scale with "1" meaning "*I would dislike it a lot"* to "5" meaning, "*I would like to do it a lot"*

A small pilot study was carried out (with 28 UK subjects), which indicated that the complete scales all had good internal consistency (0.80, 0.83 and 0.81 for the full 17-item scales; and 0.85, 0.88 and 0.81 for reduced 13-item ones). The intention was that the final version should have about a dozen items per scale with internal consistencies of around 0.80, with inter-scale correlations below 0.50.

## 2.3 The ICES Plus Ability Scales

### 2.3.1 Design of the new ICES Plus Ability Scales

As the first step in this process, during Phase One a new Numerical Reasoning test was developed (Working with Numbers: WWN). This was constructed with three main item-types:

- Series (e.g. 1,2,4,8,....);
- Manipulation (e.g. Multiply the middle figure by the last one: 4 6 7 5 2);
- Analogies (e.g. 4 is to 8 as 3 is to ?).

Thirty-five items were produced and pilot tested on 26 subjects. Following this, 32 items were chosen for the trial version. As the test is somewhat speeded, two versions were produced. The items were divided into two sets of 16 (Set A and Set B). One version was composed of Set A followed by Set B, while the other had Set B followed by Set A. This enabled estimates of reliability to be obtained using internal consistency measures based on the first 16 of each set and provided better estimates of facility levels for items which might otherwise have had low response rates due to their always coming at the end of the test.

Development of the full set of three Ability tests (Numerical, Verbal and Spatial), however, did not begin until Phase Three. Data from the earlier phases was used to define a final version of the Numerical test, which consisted of 24 items, 22 of which are counted towards the scale score.

Initial pilot work for the Verbal and Spatial tests was carried out in the UK. Subsequently, revised versions of the tests were trialed in the USA using a sample of 197 people. The results of these pilots were used as the basis for constructing the scales for the main Phase Three standardization study. The final versions of the two scales were as follows:

> The Verbal scale (Working with Words: WWW) contained 14 anagrams; 17 “Hidden Words” and 17 “Letter Sequences”.
>
> The Spatial scale (Working with Shapes: WWS) contained one sub-test of 14 items which involved simple spatial matrices, non-verbal reasoning and spatial series completion and a second sub-test of 8 items which involved following spatial manipulation instructions.

The ICES Plus Ability measure is derived from the three specific Ability scores by weighting them in relation to the numbers of items in each test:

> WWW + 2 x WWS + 2 x WWN.

# PART II: DEVELOPMENT OF THE ICES PLUS SCALES

## 3. THE PHASE ONE DEVELOPMENT SAMPLE

The development sample consisted of 1,518 people. Information about age, gender, and ethnic origin was available for 1,511 people. Of these, the majority (1,416) came from a major US wholesaler with the rest (95) coming from a Canadian financial organization. 62.5% (945) were male, and 37.5% (566) female.

The sample was collected in two stages. The first 816 were used for all the scale development work. The second set of 695 people formed the "hold-out" sample that was used to cross-validate (i.e. to check for shrinkage of internal consistency reliability) the scale construction.

***Table 3.1: Stage 1 sample age distribution.***

| AGE | | | |
|---|---|---|---|
| **Age in years** | **Frequency** | **Percent** | **Cumulative Percent** |
| Under 35 | 202 | 24.8 | 24.8 |
| 35-39 | 179 | 21.8 | 46.6 |
| 40-44 | 177 | 21.7 | 68.3 |
| 45-49 | 131 | 16.1 | 84.4 |
| 50-54 | 71 | 8.7 | 93.1 |
| 55 and over | 56 | 6.9 | 100.0 |
| TOTAL | 816 | 100.0 | |

***Table 3.2: Stage 1 Sample. Breakdown by gender and ethnic origin.***

| | ETHNIC ORIGIN | | | | |
|---|---|---|---|---|---|
| **GENDER** | **White** | **Black** | **Hispanic** | **Other** | **Total** |
| Male | 396 | 86 | 25 | 7 | 514<br>63.0% |
| Female | 271 | 17 | 7 | 7 | 302<br>37.0% |
| Total | 667<br>81.7% | 103<br>12.6% | 32<br>3.9% | 14<br>1.7% | 816<br>100.0% |

The Stage One and Stage Two Development Samples were combined to look at effects of age, gender and ethnic origin group and to develop provisional ICES Plus norms.

Tables 3.1 and 3.2 show the breakdown of the Stage 1 sample in terms of age, ethnic origin and first language. They show the distribution of age within the sample and the breakdown by gender and ethnic origin. Of the sample, 21 of the Hispanic group said Spanish was their natural language while all others said it was English.

The breakdown by age, gender and ethnic origin group for the whole sample of 1,511 was very similar to that of the Stage 1 sample. Table 3.3 shows the age distribution and breakdowns by gender and ethnic origin for the whole sample.

***Table 3.3: Whole Sample: Breakdown of ethnic origin by gender and by age.***

| ETHNIC ORIGIN | | | | | |
|---|---|---|---|---|---|
| **AGE** | **White** | **Black** | **Hispanic** | **Other** | **Total** |
| Under 35 | 340 | 57 | 25 | 11 | 433<br>28.7% |
| 35-39 | 248 | 68 | 19 | 4 | 339<br>22.4% |
| 40-44 | 233 | 46 | 17 | 7 | 303<br>20.1% |
| 45-49 | 177 | 35 | 11 | 4 | 227<br>15.0% |
| 50-54 | 106 | 12 | 3 | 2 | 123<br>8.1% |
| 55 and over | 78 | 4 | 4 | | 86<br>5.7% |
| Total | 1182<br>78.2% | 222<br>14.7% | 79<br>5.2% | 28<br>1.9% | 1,511<br>100.0% |
| **ETHNIC ORIGIN** | | | | | |
| **GENDER** | **White** | **Black** | **Hispanic** | **Other** | **Total** |
| Male | 698 | 169 | 64 | 14 | 945<br>62.5% |
| Female | 484 | 53 | 15 | 14 | 566<br>37.5% |
| Total | 1182<br>78.2% | 222<br>14.7% | 79<br>5.2% | 28<br>1.9% | 1511<br>100.0% |

All the Black and Hispanic subjects were from the American firm. Of those in the "Other" category, 20 were Asian, with 17 of these being American and 3 of these being Canadian. Of the 1,182 Whites, 1,094 were from the US wholesale distributor and 88 from the Canadian financial organization. In all the following analyses of ethnic origin effects, those in the "Other" category have been excluded, as there are too few for separate analysis.

# 4. ANALYSIS OF THE ICES PERSONALITY SCALE DATA

Analysis of data from the Stage 1 sample followed a sequence of steps. First the original content-based item groupings (P1 to P11) were examined. Then attention was focused on the development from these of the four ICES major higher order factor scales and the related development of their pairs of sub-scales.

Once this stage had been completed, the Stage 2 sample of data was examined to check for shrinkage in the internal consistencies and to assess the robustness of the item assignments.

## 4.1 Phase One, Stage One Data Analysis

### 4.1.1 Item analyses and examination of the content-based ICES scales

1. The means and standard deviations ("SD's") of each item were checked. It was noted that a few items had low SDs and hence poor discrimination.
2. The first analysis included all items - regardless of means and SD's. Internal consistencies, item inter-correlations and scale inter-correlations were computed for the eleven 15-item scales (P1 to P11). Inspection of the results of these analyses led to the removal of 58 items from the putative scales. These were excluded on the grounds of low or negative item-total correlation or small SD.
3. Reliability analyses were carried out again on the 11 scales, but with the reduced item sets. The results from Steps 2 and 3 are summarized in Table 4.1. It can be seen that despite the reduction in number of items per scale the internal consistencies were generally the same or better for the shorter scales. In particular, scale P9 shows a clear improvement.

***Table 4.1: Internal consistencies of the 11 scales before and after removal of "poor" items.***

| | | Reduced item sets | |
|---|---|---|---|
| | **Initial 15 items- Alpha**<br><br>**Initial 15 items - Alpha** | **Alpha** | **Items** |
| P1 | 0.65 | 0.66 | 9 |
| P2 | 0.56 | 0.53 | 8 |
| P3 | 0.70 | 0.65 | 11 |
| P4 | 0.69 | 0.71 | 11 |
| P5 | 0.68 | 0.62 | 9 |
| P6 | 0.65 | 0.65 | 12 |
| P7 | 0.70 | 0.74 | 11 |
| P8 | 0.42 | 0.48 | 8 |
| P9 | 0.23 | 0.57 | 8 |
| P10 | 0.52 | 0.53 | 8 |
| P11 | 0.74 | 0.77 | 12 |

4. Scale scores were produced for these eleven scales. Examination of the correlations between them and of the correlations between each item and all eleven-scale scores indicated four clear clusters: P4+P10+P11; P1+P5+P9; P2+P3 and P6+P7+P8. This was confirmed by principal components analysis of the correlation matrix. Examination of the eigen-value plot shows a clear break between the first four components (which accounted for 65.5% of the total variance) and the "scree" from component five onwards (see Table 4.2 for rotated loadings).

***Table 4.2: Varimax rotated component loadings (accounting for 65.5% of the total variance).***

| Scale | I | II | III | IV | Communality |
|---|---|---|---|---|---|
| P4 | **.82** | .13 | -.15 | .09 | .72 |
| P10 | **.62** | .25 | .25 | -.20 | .55 |
| P11 | **.83** | .08 | .15 | -.02 | .72 |
| P1 | .15 | **.78** | .17 | .08 | .66 |
| P5 | .20 | **.81** | .15 | .10 | .73 |
| P9 | -.04 | **-.68** | .31 | -.01 | .56 |
| P6 | .45 | .03 | **.59** | -.02 | .56 |
| P7 | .25 | -.16 | **.75** | .18 | .68 |
| P8 | -.18 | .19 | **.75** | .07 | .64 |
| P2 | .07 | .01 | .01 | **.85** | .72 |
| P3 | -.14 | .16 | .21 | **.76** | .66 |

Examination of the adjective descriptors used to construct the items, confirmed that these four scales corresponded to four of the "Big Five" personality factors: Extraversion (P4, P10, P11), Emotional Stability (P1, P5, P9), Independence (P6, P7, P8) and Conscientiousness (P2 and P3).

5. Scale scores on P1 to P11 were correlated with all those items that had been excluded at Step 2 on the basis of poor fit to their intended scale - rather than low SD. The results of this indicated a number of items showing high item-scale correlations with scales other than those to which they had been assigned on the basis of content by the judges. In a few cases, where these "switches" were between one of the "Big Four" factors, item response keys also needed to be reversed.

#### 4.1.2 Developing the ICES major and minor scales

1. As earlier analyses had suggested, the new data strongly supported the presence of the four major ICES factors, with minimal inter-scale overlap. This part of the analysis focused on developing these four scales from the available set of new items and then developing a pair of sub-scales for each. The rationale was that the four main scales would provide measures of the overall structure of personality. They would require good reliability as recent meta-analyses have shown they have good potential as predictors of occupational performance, training outcome and other personnel data (e.g. see Barrick and Mount, 1991; Tett et al, 1991). The sub-scales would be developed for use in providing a richer description of personality. As such they were to reflect facets of the main factors which, while overlapping in terms of variance, embody important conceptual distinctions.

   Within the original development plan (see above) it was intended to develop the following sub-scale for each of the four main factors:

   Independence:
   - I1 Competitive, tough-minded
   - I2 Assertive, forthright

   Conscientiousness:
   - C1 Conventional, traditional, concern with moral values
   - C2 Organized, attention to detail, neatness

   Extraversion:
   - E1 Group-oriented, sociable
   - E2 Outgoing, group-dependent

   Stability:
   - S1 Poised, unruffled, not easy to annoy or upset
   - S2 Relaxed, not anxious

2. From the previous analyses, four sets each of twenty items were selected for the four scales. Reliability analyses confirmed that these had the required properties. The internal consistencies and the goodness of fit of these item sets to the original content analysis can be seen in Table 4.3.

***Table 4.3: Fit of 20-item scales to original content groups.***

| | | Number of items from each content group | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **ICES Scales** | **alpha** | **P1** | **P2** | **P3** | **P4** | **P5** | **P6** | **P7** | **P8** | **P9** | **P10** | **P11** |
| Independence | 0.77 | | | | | | 8 | 8 | 3 | | | [1] |
| Conscientiousness | 0.71 | | 6 | 11 | | | [1] | | | | [2] | |
| Extraversion | 0.82 | | | | 10 | | [1] | | | | 2 | 7 |
| Stability | 0.78 | 9 | | | | 9 | | | | 2 | | |

Only four items (those in square brackets) were included in scales other than those hypothesized on the basis of item content analysis (see Table 2.2).

3. Through inspection of the item content, principal components analyses and item cluster analyses for each set of 20 items, a number of item parcels were defined. Each parcel had from three to seven items in it. Correlations between these item parcel scores indicated how these might be grouped into pairs of

sub-scales (10 items each) for each scale. The results of reliability analyses on these sub-scales are shown in Table 4.4

***Table 4.4: Item content and reliability of the four 20-item scales and eight 10-item sub-scales.***

| | | Number of items from each content group | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **ICES Scales** | **alpha** | **P1** | **P2** | **P3** | **P4** | **P5** | **P6** | **P7** | **P8** | **P9** | **P10** | **P11** |
| Independence | 0.77 | | | | | | 8 | 8 | 3 | | | 1 |
| I1 [P7] | 0.74 | | | | | | 1 | 8 | | | | 1 |
| I2 [P6] | 0.57 | | | | | | 7 | | 3 | | | |
| Conscientiousness | 0.71 | | 6 | 11 | | | [1] | | | | [2] | |
| C1 [P2] | 0.61 | | 6 | 4 | | | | | | | | |
| C2 [P3] | 0.53 | | | 7 | | | 1 | | | | 2 | |
| Extraversion | 0.82 | | | | 10 | | [1] | | | | 2 | 7 |
| E1 [P4] | 0.72 | | | | 10 | | | | | | | |
| E2 [P11] | 0.72 | | | | | | 1 | | | | 2 | 7 |
| Stability | 0.78 | 9 | | | | 9 | | | | 2 | | |
| S1 [P5] | 0.68 | 3 | | | | 7 | | | | 0 | | |
| S2 [P1] | 0.63 | 6 | | | | 2 | | | | 2 | | |
| Total items used | | 9 | 6 | 11 | 10 | 9 | 10 | 8 | 3 | 2 | 4 | 8 |
| Total items available | | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |

4. While the above solution met most of the criteria set for the design of the inventory, some of the sub-scales still had rather low alphas. For short Personality scales, it was felt that minimum alphas should be 0.60. In addition, the correlations between each pair of sub-scales were relatively high. With corrections for attenuation, these indicated that some sub-scales shared 70 or 80 percent of their true score variance.

   Correlations of the new scales, with items excluded during earlier stages of the analysis, indicated that the scale length could be readily increased to 24 items per scale - with sub-scales of 12 items each - and that this could both improve the reliabilities of some of the sub-scales and reduce their overlap.

   Thus four items were added to each of the 20-item scales and allocated to sub-scales on the basis of content. Subsequent cluster analyses resulted in the interchange of a small number of these items between sub-scales. Table 4.5 shows the final improved alphas for the sub-scales and the distribution of items across the original eleven sets of 15-items (P1 to P11). It can be seen that the hypothesized relationships between the content-based groups and ICES scales (see Table 2.2) were strongly supported by the data.

***Table 4.5: Item content and reliability of the four 24-item scales and eight 12-item sub-: Stage 1 sample analysis.***

| | | Number of items from each content group | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **ICES Scales** | **alpha** | **P1** | **P2** | **P3** | **P4** | **P5** | **P6** | **P7** | **P8** | **P9** | **P10** | **P11** |
| Independence | 0.79 | | | | | | 10 | 9 | 3 | | 1 | 1 |
| I1 [P7] | 0.72 | | | | | | 2 | 7 | 2 | | 1 | |
| I2 [P6] | 0.71 | | | | | | 8 | 2 | 1 | | | 1 |
| Conscientiousness | 0.75 | | 8 | 13 | | | [1] | | | | [2] | |
| C1 [P2] | 0.62 | | 8 | 2 | | | | | | | 2 | |
| C2 [P3] | 0.68 | | | 11 | | | 1 | | | | | |
| Extraversion | 0.85 | [1] | | | 10 | | [1] | | | | 2 | 10 |
| E1 [P4] | 0.73 | 1 | | | 10 | | | | | | 1 | |
| E2 [P11] | 0.79 | | | | | | 1 | | | | 2 | 9 |
| Stability | 0.80 | 9 | | | | 11 | | | | 4 | | |
| S1 [P5] | 0.71 | 3 | | | | 8 | | | | 1 | | |
| S2 [P1] | 0.64 | 6 | | | | 3 | | | | 3 | | |
| Total items used | | 10 | 8 | 13 | 10 | 11 | 12 | 9 | 3 | 4 | 5 | 11 |
| Total items available | | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |

## 4.2 Phase One, Stage Two Data Analysis: Scale Cross-Validation

Having completed this stage of the scale development, the second tranche of data was analyzed to see how well these new scales would perform with a new sample of people. Any lack of robustness would show up by "shrinkage" in the alpha coefficients for the new sample.

The alpha coefficients for the scales developed with the Stage 1 sample for the Stage 1 sample itself, the Stage 2 sample and the combined samples are shown in the first three columns of Table 4.6. In most cases, there is very little shrinkage - with alphas dropping by about 0.02 from Stage 1 to Stage 2. The least stable of the scales were the Independence scales and the C1 sub-scale of Conscientiousness. Examination of the data indicated that for the latter, one item that had been "marginal" in the Stage 1 sample was a low negative item-whole correlation in the Stage 2 sample. This was discarded in favor of one of the unused items from the P2 content scale. The discarded item was from the P10 content scale.

This change, therefore, slightly improved the fit of the final scales to those predicted from the content analysis (compare Table 2.2 and Table 4.7) and also resulted in all the final alpha values for the combined sample being above 0.60 for the sub-scales and above 0.70 for the four main scales.

***Table 4.6: Reliabilities for the Stage 1, Stage 2 and total Phase One Development Sample.***

| | Stage 1 n=815 | Stage 2 N=696 | Combined n=1511 Initial | Final | | |
|---|---|---|---|---|---|---|
| **ICES Scales** | **Alpha** | **Alpha** | **Alpha** | **Alpha** | **MEAN** | **SD** |
| Independence | 0.79 | 0.74 | 0.77 | 0.77 | 49.46 | 8.37 |
| I1 [P7] | 0.72 | 0.67 | 0.70 | 0.70 | 23.61 | 4.87 |
| I2 [P6] | 0.71 | 0.65 | 0.68 | 0.68 | 25.84 | 5.16 |
| Conscientiousness | 0.75 | 0.73 | 0.74 | 0.74 | 53.68 | 7.85 |
| C1 [P2] | 0.62 | 0.60 | 0.58 | 0.60 | 26.53 | 4.30 |
| C2 [P3] | 0.68 | 0.66 | 0.67 | 0.67 | 27.15 | 5.00 |
| Extraversion | 0.85 | 0.83 | 0.84 | 0.84 | 53.29 | 9.59 |
| E1 [P4] | 0.73 | 0.70 | 0.72 | 0.72 | 26.23 | 5.18 |
| E2 [P11] | 0.79 | 0.77 | 0.79 | 0.79 | 27.06 | 5.60 |
| Stability | 0.80 | 0.80 | 0.80 | 0.80 | 52.38 | 8.81 |
| S1 [P5] | 0.71 | 0.69 | 0.70 | 0.70 | 27.18 | 4.94 |
| S2 [P1] | 0.64 | 0.66 | 0.65 | 0.65 | 25.20 | 4.86 |

***Table 4.7: Item content and reliability of the four 24-item scales and eight 12-item sub-scales: Final combined sample analysis. The only change from Table 4.5 is to Sub-scale [P2].***

| | | Number of items from each content group | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **ICES Scales** | **Alpha** | **P1** | **P2** | **P3** | **P4** | **P5** | **P6** | **P7** | **P8** | **P9** | **P10** | **P11** |
| Independence | 0.79 | | | | | | 10 | 9 | 3 | | 1 | 1 |
| I1 [P7] | 0.72 | | | | | | 2 | 7 | 2 | | 1 | |
| I2 [P6] | 0.71 | | | | | | 8 | 2 | 1 | | | 1 |
| Conscientiousness | 0.75 | | 9 | 13 | | | [1] | | | [1] | | |
| C1 [P2] | 0.62 | | 9 | 2 | | | | | | | 2 | |
| C2 [P3] | 0.68 | | | 11 | | | 1 | | | | | |
| Extraversion | 0.85 | [1] | | 10 | | | [1] | | | | 2 | 10 |
| E1 [P4] | 0.73 | 1 | | | 10 | | | | | | | 1 |
| E2 [P11] | 0.79 | | | | | | 1 | | | | 2 | 9 |
| Stability | 0.80 | 9 | | | | 11 | | | | 4 | | |
| S1 [P5] | 0.71 | 3 | | | | 8 | | | | 1 | | |
| S2 [P1] | 0.64 | 6 | | | | 3 | | | | 3 | | |
| Total items used | | 10 | 9 | 13 | 10 | 11 | 12 | 9 | 3 | 4 | 4 | 11 |
| Total items available | | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |

## 4.3 Response Bias Measures

### 4.3.1 Motivational distortion - Social Desirability

It was recognized that a separate scale assessing motivational distortion was needed to complete the inventory and provide an indication of the extent to which respondents were making “socially desirable” rather than “honest” responses to the items. Development of this was carried out in Phase Two and is discussed in the next part of this Manual.

#### 4.3.2 Analysis of the number of “in-between” responses

People may differ systematically in their tendency to use the “in-between” option for each item. Over use of this can, of course, reduce the effective scale variance as well as being indicative of general cautiousness in terms of how one describes oneself.

The reliability of the count of number of “in-between” responses was estimated in the following way. The number of “in-between” responses selected for each of the eight sub-scales was counted. These eight scores for each person were then treated as item scores and subjected to reliability analysis. For the full Stage One and Stage Two sample, the average inter-item correlation was 0.53, with an alpha coefficient of 0.90. Thus, it seems that this is a very stable measure: people are very consistent across the scales in their tendency to use or avoid the “in-between” response options. Furthermore, this measure is quite independent of the other personality measures.

The median number of “in-between” responses was 6, with only 5% of the sample making 28 or more. The fact that this is uncorrelated with scale scores implies that the use of the “in-between” category does not, in itself, bias scores. However, one might still wish to look at this score if it is very high (e.g. over 30) as this might indicate either a general difficulty in responding to the items or a very “defensive” or “cautious” approach to answering the inventory.

#### 4.3.3 Other response bias measures

A major source of bias can arise when people omit to respond to one or more items. There are various ways in which omitted responses can be dealt with. In the present instance the most neutral action is to assume an “in-between” response. In practice, one could set a limit on the number of omitted items one would allow. The present data suggest that setting the maximum at two would result in only one in every 200 candidates being invalidated. Certainly, assigning “in-between” responses to no more than two items per candidate would not have any substantive effect on people's scale scores.

Further development work on response bias measures is discussed later in relation to Phase Two. In future, it may be useful to build more subtle keys into the scoring procedure, which produce measures that are independent of the main scale scores. Norms were not generated for these simple response bias measures. They are best used in a diagnostic fashion: where scores are high, care needs to be exercised in the interpretation of the scales and possible reasons for the bias should be explored with the candidate. Proposed cut-offs for use with ICES are discussed later in the Manual.

# 5. THE ICES PLUS INTEREST INVENTORY

As with the Personality scales, initial development was carried out on the Stage 1 sample only. Once the scale items had been identified, these were cross validated against the Stage 2 sample. Finally, norms were produced and the effects of age, gender and ethnic origin were examined for the combined sample.

## 5.1 Phase One, Stage One Data Analysis

In all the following Stage 1 data analyses, people who omitted to answer one or more items are excluded from analyses. Hence the "n" values given will vary.

Principal component analysis was used to identify the best items for each scale. This was carried out with extraction of three components and then Varimax rotation. The plot of eigenvalues showed a clear change of slope between three and four components, with components four onwards-representing "scree". Initial choice of items was based on the criteria that each item should have a loading of 0.30 or better on the scale to which it would be assigned and less than 0.30 on both of the other scales. This criterion produced 19 items for People, 11 for Data and 8 for Things. It was clear that the People scale was over-inclusive (as items designed to indicate Data or Things preferences tended to load on People while the reverse was not the case). Items that had two loading above 0.30 were examined and adopted if one was substantially higher than the other (e.g. 0.30 and 0.60) and the higher one was on the intended scale. This second step enabled the number of items in Data and Things to be increased to 12 each. The number of items in People was reduced to 12 to make the scale equal in length.

The results of reliability analyses of the three sets of items are shown in Table 5.1. (People with missing item responses are excluded from the relevant analyses.)

***Table 5.1: Reliability of the three Interest scales.***

| Stage 1 sample: | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Inter-item correlations | | | Scale | | | |
| | Mean | Min | Max | Mean | SD | Alpha | N |
| PEOPLE | 0.31 | 0.10 | 0.61 | 43.32 | 8.40 | 0.84 | 784 |
| DATA | 0.26 | 0.06 | 0.59 | 29.72 | 7.94 | 0.80 | 797 |
| THINGS | 0.33 | 0.06 | 0.68 | 32.80 | 9.57 | 0.86 | 793 |

Correlations between the scales are reasonably low - well below 0.50. People and Data correlated 0.38, People and Things 0.19 and Data and Things 0.33 (n=767 in each case).

## 5.2 Phase One, Stage Two data analysis: scale cross-validation

Reliability analyses of the holdout sample data using the scales developed with the Stage 1 sample are shown in Table 5.2.

***Table 5.2. Stage 2 (holdout) sample.***

| Stage 1 sample: | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Inter-item correlations | | | Scale | | | |
| | Mean | Min | Max | Mean | SD | Alpha | N |
| PEOPLE | 0.31 | 0.10 | 0.61 | 43.55 | 8.50 | 0.84 | 688 |
| DATA | 0.26 | 0.06 | 0.63 | 30.73 | 8.25 | 0.81 | 684 |
| THINGS | 0.33 | 0.06 | 0.68 | 33.05 | 9.85 | 0.85 | 685 |

These are very similar to the results obtained with the scale development sample and imply that these scales are very robust. The People scale showed no change in alpha, while Data actually increased by 0.01. Only the Things scale showed a small degree of shrinkage to 0.85.

# 6. DEVELOPMENT OF THE ICES PLUS NUMERICAL ABILITY TEST

## 6.1 Phase One, Stage One Data Analysis

Development work on the Ability tests in Phase One was limited to the Numerical Reasoning test (Working with Numbers: WWN). Development of the Verbal and Spatial scales is described later in relation to Phase Three. Data was obtained from all those in the Development sample for the new test (WWN). The procedure adopted for each test was as follows.

- Items with very high or very low facility values (greater than 0.90 or less than 0.10) and any others with very restricted variance were removed.
- The remaining items were subject to principal component analysis. Items with loadings of more than 0.30 on the first un-rotated component were retained for further analysis.

The results of this analysis are shown in Table 6.1. Completion rates on the last four or five items were fairly low. This did not affect the computation of alpha values, as facility level estimates for each item were available from the first 16 items in the test (nb: the test was administered in two formats such that in one items 1-16 were administered first and in the other items 17-32 were administered first).

***Table 6.1. Stage 1-sample item analysis results and scale inter-correlations.***

| | Initial n of items | Final n of items | Inter-item correlations Mean | Min | Max | Scale Mean | SD | Alpha | N |
|---|---|---|---|---|---|---|---|---|---|
| WWN | 32 | 24 | .26 | 0.01 | 0.75 | 16.56 | 5.71 | 0.89 | 815 |

## 6.2 Phase One, Stage Two Data Analysis: Scale Cross-Validation

Results of the reliability analyses of the holdout sample data using the scales developed with the Stage 1 sample are shown in Table 6.2.

***Table 6.2: Stage 2 (holdout) sample.***

| | Inter-item correlations Mean | Min | Max | Scale Mean | SD | Alpha | N |
|---|---|---|---|---|---|---|---|
| WWN | 0.23 | 0.02 | 0.67 | 16.42 | 5.48 | 0.88 | 703 |

These are very similar to the results obtained with the scale development sample and imply that these scales are robust. The WWN scale alpha only decreased by 0.01. For the total Phase One sample, alpha is 0.89. For the Stage 2 data collection, all 32 items were retained in the test, with the 8 unscored ones acting as fillers and practice.

# 7. PHASE TWO: BROADENING THE SAMPLE

Following the initial development phase, further samples of people were obtained to increase the variety of jobs covered and to provide a more representative balance in the overall sample in terms of gender and ethnic origin. The total number of people tested during Phase Two was 1,840.

The Phase Two data collection can be divided into two stages.

### Phase Two, Stage One

For Stage One, a number of job-related groups totaling 1,236 people were obtained with the intention of broadening the range of jobs and organizations sampled. Organizations in Canada, the United States, Great Britain and the Far East were sampled.

### Phase Two, Stage Two

Stage Two had a number of specific objectives that needed to be met in order to complete the development work on the ICES Personality scales.

1. 30 of the ICES scale items needed to be revised. Revisions of these would need to be checked for equivalence with the original items.
2. A “Social Desirability” scale needed to be constructed.
3. In addition, it was planned to obtain:
    - a re-test sample to examine re-test reliability for the scales; and
    - a construct validation sample to examine correlations with the Cattell 16 Personality Factor Questionnaire (16PF).

604 people were tested during Stage Two, using just the ICES scales. Special booklets were constructed which contained the full ICES scale item set (96 items) plus the revised versions of each of the 30 items which had been changed and a set of 14 new items for the Social Desirability scale. Inclusion of both the original and the revised items enabled a direct check to be made of the equivalence of the new items.

147 people from the 604 completed a re-test on average one-week later. A further 151 completed Form A of the Cattell 16PF in addition to the ICES.

## 7.1 Equivalence of the Modified ICES Items

Table 7.1. gives the means, SDs and alpha consistencies for each scale for scales containing the old items and scales containing the revised items. The new items have generally improved the internal consistencies without substantially affecting the means or SDs. Correlations between the two versions of each scale are all very high (from r=0.92 to r=0.99). Table 7.2 shows the correlations between the scales containing the revised items. The pattern of correlations is very similar to that obtained with the old items.
It can be concluded that the item revisions - which were designed to make the test content more acceptable to users and reduce some instances of item ambiguity - were successful. The scales were improved in terms of reliability without any changes in scale score distributions. This is important as a high degree of equivalence across modifications is needed if information from the combined data set (all Stages) is to be used for normative purposes. Any major change in scale structure would have necessitated discarding much of that previous data. Based on these results, the old items were discarded.

***Table 7.1. Phase Two, Stage Two ICES sample (n=604): comparisons between scales with previous and amended items.***

| Scale | Scale with old items | | | Scale with new items | | | Correlation |
|---|---|---|---|---|---|---|---|
| | Mean | SD | alpha | Mean | SD | alpha | Old with new |
| I1 | 21.90 | 4.89 | 0.70 | 21.88 | 4.95 | 0.71 | 0.95 |
| I2 | 25.55 | 5.22 | 0.68 | 25.32 | 5.40 | 0.71 | 0.92 |
| C1 | 22.75 | 5.13 | 0.70 | 22.38 | 5.22 | 0.72 | 0.93 |
| C2 | 24.98 | 5.35 | 0.68 | 24.81 | 5.50 | 0.70 | 0.96 |
| E1 | 24.76 | 5.56 | 0.73 | 24.53 | 5.62 | 0.74 | 0.98 |
| E2 | 24.75 | 5.97 | 0.79 | 24.30 | 6.19 | 0.80 | 0.98 |
| S1 | 25.19 | 5.22 | 0.69 | 25.21 | 5.33 | 0.70 | 0.97 |
| S2 | 24.77 | 4.68 | 0.61 | 25.18 | 5.02 | 0.65 | 0.93 |
| | | | | | | | |
| INDEP | 47.45 | 8.49 | 0.77 | 47.19 | 8.75 | 0.79 | 0.95 |
| CONSC | 49.51 | 10.36 | 0.77 | 48.83 | 10.71 | 0.78 | 0.99 |
| EXTRAV | 47.72 | 8.71 | 0.85 | 47.19 | 9.00 | 0.86 | 0.96 |
| STABLE | 49.96 | 8.95 | 0.78 | 50.39 | 9.43 | 0.81 | 0.97 |
| SocDes | | | | 22.57 | 5.94 | 0.78 | |

***Table 7.2. Scale inter-correlations for the ICES scales containing the revised items (n=604).***

| Scale | I1 | | I2 | | C1 | | C2 | | E1 | | E2 | | S1 | | S2 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I1 | 1.00 | | .43 | ** | -.04 | | .03 | | .03 | | .21 | ** | -.09 | | -.02 | |
| I2 | .42 | ** | 1.00 | | -.20 | ** | -.05 | | .27 | ** | .45 | ** | .19 | ** | .20 | ** |
| C1 | -.04 | | -.20 | ** | 1.00 | | .41 | ** | -.20 | ** | -.20 | ** | .08 | | .06 | |
| C2 | .03 | | -.05 | | .41 | ** | 1.00 | | -.09 | | -.18 | ** | .18 | ** | .16 | ** |
| E1 | .03 | | .27 | ** | -.20 | ** | -.09 | | 1.00 | | .64 | ** | .16 | ** | .17 | ** |
| E2 | .27 | ** | .45 | ** | -.20 | ** | -.18 | ** | .64 | ** | 1.00 | | .13 | ** | .15 | ** |
| S1 | -.09 | | .17 | ** | .08 | | .18 | ** | .16 | ** | .13 | ** | 1.00 | | .66 | ** |
| S2 | -.02 | | .20 | ** | .06 | | .16 | ** | .17 | ** | .15 | ** | .66 | ** | 1.00 | |
| | | | | | | | | | | | | | | | | |
| SocDes | -.05 | | -.09 | | .46 | ** | .30 | ** | -.03 | | -.11 | * | .20 | ** | .24 | ** |
| **Scale** | **IND** | | **CONSC** | | **EXT** | | **STAB** | | | | | | | | | |
| INDEP | 1.00 | | -.09 | * | .33 | ** | .10 | * | | | | | | | | |
| CONSC | -.09 | * | 1.00 | | -.22 | ** | .15 | ** | | | | | | | | |
| EXTRAV | .33 | * | -.22 | ** | 1.00 | | .18 | ** | | | | | | | | |
| STABLE | .10 | * | .15 | ** | .18 | ** | 1.00 | | | | | | | | | |
| SocDes | -.08 | | .45 | ** | -.08 | | .24 | ** | | | | | | | | |
| * p<.05, ** p<.01 | | | | | | | | | | | | | | | | |

## 7.2 The New Social Desirability Scale

Table 7.1 also shows results for the new Social Desirability (SocDes) scale. This consisted of 14 items and has good internal consistency (alpha=0.78). From Table 7.2, it can be seen that people who score high on SocDes tend to be high on Conscientiousness (especially C1, which indicates a concern for traditional high moral values) and S2. The pattern of correlations indicates one of the problems of interpreting SocDes-type scales. People who really are very "good" (never tell lies, are always nice to people, always do the "right" thing) will get a high SocDes score - so will people who are "faking good". For this reason social desirability scales can only be used as indicators of faking good behavior.

High scores should be taken as a warning sign and not as proving that the person is lying. A high scorer might be faking good or might really be a very stable conscientious person giving a valid self-report. More subtle types of scale are needed to distinguish the two. As being "Stable" and "Conscientious" are socially desirable characteristics, so those who really are stable and conscientious will score high on a Social Desirability scale. Similarly, those who are "faking good" tend to do so by presenting themselves as more stable and conscientious than they are.

Apart from the expected relationships with Conscientiousness and Stability, the SocDes scale is relatively independent of both Extraversion and Independence. Based on the results of the analyses carried out on this sample, it was decided to adopt the 14-item SocDes scale for inclusion in the ICES inventory. Thus the number of items in the final version of the Inventory is 110 (12 for each of the eight minor scales and 14 for the SocDes scale).

## 7.3 ICES Re-test Reliability

The results for the re-test sample (n=147) are presented in Table 7.3. In most cases, re-test correlations were higher than internal consistencies (alpha). For the minor Scales, S1 has a lower re-test correlation (0.60) than its alpha (0.70), as does E2 (re-test r=0.74, alpha=0.80). This is reflected in the results for the major scales, where the alpha values for both Extraversion and Stability are higher than the re-test correlations. The relatively low re-test correlation for S1, given its reasonable internal consistency could indicate that S1 is measuring more variable aspects of emotional stability (i.e. state-related anxiety) than S2.

Examination of Table 7.3 shows that the means and SDs for each scale are not affected by re-testing. This means that the same norms can be used for interpreting first and subsequent administration of the ICES scales.

***Table 7.3. Test-re-test data for the final ICES scales. (n=147 from the Phase Two, Stage Two sample).***

| SCALE | TEST Mean | SD | RE-TEST Mean | SD | test-re-test correlation | |
|---|---|---|---|---|---|---|
| I1 | 22.28 | 5.02 | 22.01 | 4.87 | 0.81 | |
| I2 | 24.68 | 5.17 | 24.93 | 5.35 | 0.80 | |
| C1 | 23.25 | 5.01 | 23.63 | 5.36 | 0.79 | |
| C2 | 25.67 | 5.85 | 25.92 | 5.77 | 0.86 | |
| E1 | 22.88 | 5.26 | 22.82 | 5.43 | 0.74 | |
| E2 | 23.18 | 5.53 | 23.12 | 6.04 | 0.74 | |
| S1 | 24.99 | 4.82 | 25.48 | 4.89 | 0.60 | |
| S2 | 25.20 | 4.67 | 25.64 | 5.12 | 0.74 | |
| IND | 46.96 | 8.70 | 46.95 | 8.81 | 0.83 | Independence |
| CONSC | 48.92 | 8.91 | 49.54 | 9.22 | 0.84 | Conscientiousness |
| EXT | 46.07 | 9.65 | 45.93 | 10.35 | 0.76 | Extraversion |
| STAB | 50.19 | 8.61 | 51.12 | 8.76 | 0.69 | Stability |
| SocDes | 24.10 | 6.23 | 24.34 | 6.56 | 0.82 | Social Desirability |

## 7.4 Equivalence of the French translation of the ICES scales

A Canadian Government agency provided a sample that included 64 people who completed a French version of ICES and 116 who completed the English version. Individuals were given a free choice of which version they wanted to complete. Comparison of the two groups indicates that scale scores are not affected by the version used. The only scale showing signs of a difference between groups is the Social Desirability scale (SocDes).

***Table 7.4. Means and SDs for the English (n=116) and French (n=64) language versions of the ICES scales.***

| | English | | French | | | |
|---|---|---|---|---|---|---|
| Scale | Mean | SD | Mean | SD | F Ratio | p |
| I1 | 22.68 | 4.55 | 22.70 | 4.58 | <1 | ns |
| I2 | 25.07 | 5.60 | 24.59 | 5.30 | <1 | ns |
| C1 | 24.65 | 5.14 | 25.66 | 4.24 | 1.8 | ns |
| C2 | 26.42 | 5.32 | 27.03 | 4.44 | <1 | ns |
| E1 | 22.71 | 5.25 | 21.66 | 4.86 | 1.7 | ns |
| E2 | 22.42 | 6.01 | 22.11 | 5.09 | <1 | ns |
| S1 | 25.74 | 5.48 | 24.56 | 4.32 | 2.2 | ns |
| S2 | 25.07 | 5.17 | 26.36 | 4.88 | 2.7 | ns |
| INDEP | 47.76 | 8.38 | 47.30 | 8.22 | <1 | ns |
| CONSC | 51.07 | 8.58 | 52.69 | 6.72 | 1.7 | ns |
| EXTRAV | 45.13 | 10.06 | 43.77 | 8.94 | <1 | ns |
| STABLE | 50.81 | 9.81 | 50.92 | 8.38 | <1 | ns |
| SocDes | 54.45 | 6.35 | 28.45 | 5.51 | 18.08 | <.001 |

# 8. DESCRIPTIVE STATISTICS FOR THE COMBINED PHASE ONE AND PHASE TWO SAMPLE

Together, the 1518 people testing during Phase One and those tested during Phase Two provide a combined sample data set of 3358 cases. Relevant job code information was missing for seven of the Phase One sample, so the effective number in the combined data set was 3351. Data on the ICES Personality scales were available for the whole sample - major and minor scales. While there had been variations made to some of the items during the development process, analysis had indicated that these would not have an effect on scale score averages or on scale inter-correlations.

Data on the Social Desirability (SocDes) scale were only available for the Stage Two Phase Two sample. Data on the Interest scales (PDT: People, Data and Things) and on the Numerical Reasoning test (Working with Numbers: WWN) were available for all but the Stage Two Phase Two sample.

## 8.1 Characteristics of the Phase One and Phase Two Combined Data Set

The total number of people tested was 3351, of whom 58.1% were male, 41.9% female; 75.6% were White, 8.2% Black, 6.1% Oriental, 6% Asian and 2.4% Hispanic. English was the native language of 79.2% of the sample; 6% of the sample had French as their first language and completed a French version of the ICES Plus battery. All the others completed the English version.

The samples were drawn from Canada, the United States, the United Kingdom and Singapore. The type of work carried out by each person was classified using the DOT (*Dictionary of Occupational Titles*) system. A breakdown of the sample by type of job is also presented in Appendix A**.**

## 8.2 Scale Raw Score Inter-Correlations

### 8.2.1 Personality

Table 8.1 shows the correlations between the ICES scales. For the major scales, all the correlations are low, the largest (0.35) being between “Extraversion” and “Independence”. Some interesting findings emerge from the correlations between the minor scales. First, the correlations between pairs of minor scales are substantively lower than their reliabilities (the between-sub-scale correlations being 0.40 for “Independence”, 0.41 for “Conscientiousness” and 0.60 for both “Extraversion” and “Stability”).

Second, there are clear differences in the patterns of correlations between sub-scales across the four main scales. Both Stability minor scales correlate with I2 (“Assertiveness”) but not with I1 (“Tough-mindedness”). The same is true for E1 (“Group-dependence”). However E2 (“Sociability”), correlates with both I1 and I2. Such differences support the value of breaking descriptions down to the sub-scale level, as they indicate that the sub-scales provide useful information additional to that provided by the four major scales.

***Table 8.1: Correlations (decimal points omitted) between the four main scales and between the eight sub-scales (n=3284).***

| | **INDEP** | | **CONSC** | | **EXTRAV** | | | |
|---|---|---|---|---|---|---|---|---|
| INDEPENDENCE | -- | | | | | | | |
| CONSCIENTIOUS | -05 | | -- | | | | | |
| EXTRAVERSION | 35 | | -08 | | -- | | | |
| STABILITY | 13 | | 14 | | 26 | | | |
| | I1 | I2 | C1 | C2 | E1 | E2 | S1 | S2 |
| INDEP | 82 | 85 | -08 | -02 | 21 | 41 | 08 | 15 |
| I1 | -- | 40 | 00 | 01 | 06 | 21 | -10 | -01 |
| I2 | 40 | -- | -12 | -04 | 28 | 46 | 22 | 25 |
| CONSC | 01 | -09 | 82 | 86 | -03 | -11 | 15 | 11 |
| C1 | 00 | -12 | -- | 41 | -03 | -08 | 11 | 08 |
| C2 | 01 | -04 | 41 | -- | -02 | -10 | 14 | 10 |
| EXTRAV | 15 | 42 | -06 | -07 | 88 | 91 | 24 | 22 |
| E1 | 06 | 28 | -03 | -02 | -- | 60 | 22 | 20 |
| E2 | 21 | 46 | -08 | -10 | 60 | -- | 21 | 18 |
| STABLE | -06 | 26 | 11 | 13 | 24 | 22 | 90 | 89 |
| S1 | -10 | 22 | 11 | 14 | 22 | 21 | -- | 60 |
| S2 | -01 | 25 | 08 | 10 | 20 | 18 | 60 | -- |

### 8.2.2 Interests

Table 8.2 shows the correlations between the three Interest scales (PDT). These are all fairly low. The low inter-scale correlations, together with the good internal consistencies supports the view that these three scales are measuring aspects of interest which are relatively independent of each other.

***Table 8.2: Correlations between the ICES Plus PDT Interest scales (n=2559).***

| | DATA | THINGS |
|---|---|---|
| PEOPLE | 0.38 | 0.21 |
| DATA | | 0.33 |

## 8.3 Gender-related differences

Table 8.3 gives means and SDs for all scales for males and females.

It is always difficult to interpret gender differences - especially as the present sample, though large, was structured to represent the working population rather than the population as a whole. It contains a variety of work-related samples, which may contain biases in terms of gender and other factors. However, the differences noted in relation to Personality and Interests are typical of those found on other inventories (e.g. Cattell's 16PF, Holland's Vocational Preference Inventory, etc.).
The gender differences noted above must be considered in the light of possible confounding with both age and ethnic origin group. These sources of variance are examined individually below and then possible interaction effects are explored.

***Table 8.3: Means and SDs for the males and the females. The eta values give the correlation between Gender and each scale, with the Common Variance (eta squared as a percentage) indicating the amount of scale score variance which could be predicted from knowing a candidate's gender. The F-ratio indicates whether the difference between the sexes is statistically greater than zero (p<.05) or not (ns).***

| | Males | | Females | | | Common | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | eta | Var | F-ratio | p |
| N | 1936 | | 1394 | | | | | |
| INDEP | 50.43 | 8.10 | 45.21 | 8.17 | 0.30 | 09.1% | 332.51 | <.001 |
| I1 | 24.29 | 4.68 | 20.59 | 4.33 | 0.37 | 14.0% | 539.80 | <.001 |
| I2 | 26.15 | 5.16 | 24.62 | 5.33 | 0.14 | 02.0% | 69.09 | <.001 |
| CONSC | 52.98 | 8.24 | 49.92 | 8.08 | 0.18 | 03.3% | 112.47 | <.001 |
| C1 | 25.88 | 4.65 | 24.56 | 4.66 | 0.14 | 01.9% | 64.93 | <.001 |
| C2 | 27.11 | 5.10 | 25.39 | 5.13 | 0.16 | 02.7% | 90.72 | <.001 |
| EXTRA | 51.15 | 9.99 | 52.11 | 9.94 | 0.05 | 00.2% | 7.57 | <.01 |
| E1 | 25.45 | 5.40 | 26.00 | 5.09 | 0.05 | 00.3% | 8.97 | <.01 |
| E2 | 25.70 | 5.79 | 26.10 | 6.04 | 0.03 | 00.1% | 3.69 | ns |
| STAB | 51.88 | 8.87 | 49.77 | 9.02 | 0.12 | 01.3% | 45.32 | <.001 |
| S1 | 26.35 | 5.13 | 25.62 | 5.19 | 0.07 | 00.5% | 16.44 | <.001 |
| S2 | 25.53 | .81 | 24.15 | 4.88 | 0.14 | 02.0% | 66.28 | <.001 |
| N | 318 | | 285 | | | | | |
| SocDes | 22.90 | 6.29 | 22.28 | 5.37 | 0.05 | 00.3% | 1.67 | ns |
| N | 1600 | | 1086 | | | | | |
| PEOPLE | 42.99 | 8.09 | 42.58 | 8.49 | 0.02 | 00.0% | 1.64 | ns |
| DATA | 30.26 | 7.98 | 30.67 | 7.87 | 0.02 | 00.0% | 1.67 | ns |
| THINGS | 35.05 | 8.99 | 28.90 | 9.40 | 0.31 | 09.8% | 291.63 | <.001 |
| N | 1628 | | 1117 | | | | | |
| WWN | 16.11 | 5.93 | 13.30 | 6.34 | 0.22 | 04.9% | 140.22 | <.001 |

## 8.4 Age differences

Analysis of age-related differences was carried out after coding people into one of six age bands: under 35; 35 to 39; 40 to 44; 45 to 49; 50 to 54; 55 and above. Effects of Age were examined using analysis of variance. The results of these analyses are summarized in Table 8.4. ETA coefficients and R coefficients are both shown as the former are sensitive to both linear and non-linear relationships between age and scale score while the latter indicate the degree of linear relationship.

The age differences noted here are cross-sectional and are best interpreted as cohort-effects, or between-group differences. That is, they relate to differences between different groups of people who happen to vary in age. It does not follow that the same variations in scale score would be found across time within the same individuals (i.e. longitudinally). That is, they do not necessarily imply or reflect changes that occur within individuals as they get older.

***Table 8.4: For each scale the table shows the overall relationship between the six age bands (see text for explanation) and scale scores, expressed as eta, and the linear relationship, expressed as R. (NB. Negative values of R indicate that scale score decrease as age increases.) The final column (r) gives the simple linear correlation between age (in years) and scale score.***

| Scale | eta | Common Var% | Overall F ratio | p | R | Linear Var% | Linear F ratio | p | [r] |
|---|---|---|---|---|---|---|---|---|---|
| INDEP | 0.05 | 00.3% | 1.64 | ns | -.03 | 00.1% | 2.68 | ns | -.02 |
| I1 | 0.02 | 00.0% | <1.00 | ns | -.00 | 00.0% | <1 | ns | .01 |
| I2 | 0.07 | 00.5% | 2.99 | <.05 | -.04 | 00.2% | 6.47 | <.05 | -.04 |
| CONSC | 0.14 | 02.1% | 14.09 | <.001 | 0.14 | 01.8% | 62.29 | <.001 | .16 |
| C1 | 0.17 | 02.9% | 19.60 | <.001 | 0.16 | 02.7% | 91.56 | <.001 | .18 |
| C2 | 0.08 | 00.6% | 4.29 | <.001 | 0.07 | 00.5% | 16.00 | <.001 | .08 |
| EXTRA | 0.15 | 02.4% | 16.07 | <.001 | -.14 | 02.0% | 68.30 | <.001 | -.14 |
| E1 | 0.15 | 02.4% | 15.97 | <.001 | -.13 | 01.7% | 58.77 | <.001 | -.14 |
| E2 | 0.13 | 01.6% | 11.08 | <.001 | -.12 | 01.5% | 50.73 | <.001 | -.12 |
| STAB | 0.07 | 00.5% | 3.41 | <.01 | .05 | 00.3% | 9.87 | <.01 | .06 |
| S1 | 0.07 | 00.5% | 3.52 | <.01 | .06 | 00.3% | 10.21 | <.01 | .06 |
| S2 | 0.06 | 00.3% | 2.29 | <.05 | .04 | 00.2% | 6.01 | <.05 | .05 |
| | | | | | | | | | n=3267 |
| SocDes | 0.28 | 08.1% | 10.51 | <.001 | .27 | 07.4% | 47.91 | <.001 | .28 |
| | | | | | | | | | n=603 |
| PEOPLE | 0.15 | 02.1% | 11.65 | <.001 | -.14 | 02.0% | 53.21 | <.001 | -.15 |
| DATA | 0.14 | 01.9% | 10.42 | <.001 | -.14 | 01.9% | 50.77 | <.001 | -.14 |
| THINGS | 0.09 | 00.8% | 4.14 | <.001 | .03 | 00.1% | 1.99 | ns | .01 |
| | | | | | | | | | n=2597 |
| WWN | 013 | 01.6% | 9.06 | <.001 | -.09 | 00.7% | 18.08 | <.001 | -.07 |
| | | | | | | | | | n=2731 |
| Degrees of freedom for F ratios: Overall = 5, (n-6); Linear = 1, (n-6) | | | | | | | | | |

# 9. PHASE THREE: DEVELOPMENT OF THE ICES PLUS ABILITY SCALES

## 9.1 Scale construction

The ICES Plus Ability scales were designed to cover the three main content areas of general ability: Numerical, Verbal, and Spatial (NVS).

### 9.1.1 Numerical - Working with Numbers (WWN)

Development of this scale began during Phase One of the program. However, the rationale is reviewed here for completeness. The test contains the following item-types:

- Arithmetic operations - e.g. “Multiply the third figure by the first: 5, 10, 15, 20”.
- Number series - e.g. “What number comes next: 4, 6, 10, 16, ...”
- Number analogies - e.g. “13 is to 26 as 7 is to ....?”

An open response format rather than multiple choice was used for all three types.

### 9.1.2 Verbal - Working with Words (WWW)

Care was taken in designing this test to minimize the impact of specific vocabulary knowledge and to focus mainly on measures of word fluency. This was done to reduce as much as possible any impact of cultural differences. Three item-types were used:

- Hidden Words (HIDWD) - 5-letter words were embedded in letter strings that also contained other longer or shorter words. e.g. “KREDGREENICEX” (the answer is GREEN).
- Anagrams (ANAG) - e.g. “GTHFI” (the answer is FIGHT). All the anagrams were five or six letters.
- Letter sequences (LSEQ) where one of a set of four three-letter sequences is the odd-one-out: e.g. “CBA EFG ZYX RQP”, where EFG is the odd-one-out because the others are all in reverse alphabetical order.

An open response format was used for the hidden words and anagrams with multiple choices for the letters sequences.

### 9.1.3 Spatial - Working with Shapes (WWS)

The emphasis in this test is on the manipulation of figural material and reasoning with shapes. Four quite distinct item-types were designed for use in this test:

- Pattern Sequences (PS). These are series completion items, where one of five response alternatives has to be selected to fill the gap in a given sequence.
- Pattern Grids (PG). These are matrices of either four, six or nine cells. In each case one of the cells is left blank and the correct missing cell has to be selected from a set of five alternatives.
- Features in Common (FIC). A pair of shapes are presented which have certain features in common. Examination of them and the five response alternatives provides sufficient information to deduce the rule defining the relevant common feature.
- Shape Manipulations (SM). A sequence of operations has to be performed on a given shape (e.g. “rotate 90 degrees clockwise”, “make smaller” etc.) and the resultant shape selected from a set of five alternatives.

## 9.2 Pilot studies

After initial development trials in the UK with small samples of subjects, a subset of the original items was identified and a trial test booklet produced for data collection in North America. The trial booklet contained:

**28 test items for WWN.**
**65 test items for WWW, composed of:**
25 anagrams;
20 letter sequences;
20 hidden words.
**32 test items for WWS, composed of:**
8 pattern sequences;
8 pattern grids;
8 features in common;
8 shape manipulations.

To facilitate item analysis and scale construction, each of the item-types for WWW and WWS were presented in separately timed blocks. Instructions for the test taker and the administrator and example items were also provided.

A sample of 197 people was tested, of whom 100 were male and 97 female. The average age was 40.7 years (SD = 11.5) with a range from 16 to 67 years old.

### 9.2.1 Analyses of the complete item sets for each scale

Means and SDs for each of the scales and for the individual WWW and WWS item-types, are shown in Table 9.1. All items are scored "1" for correct and "0" for wrong or omitted. While mean inter-item correlations were acceptable for the WWW item-types, they were low for three of the four WWS types (PS, PG and FIC). Correlations between the item-types are shown in Table 9.2. In all but one of the cases, the correlations between item-types are 0.40 or higher. PS and PG, each of which have low internal consistencies, only correlated 0.35.

***Table 9.1. Means and SD for the US pilot sample (n=197), with mean inter-item correlations and alphas for the WWW and WWS item-types.***

| Scale<br>WWN | Mean<br>17.12 | SD<br>7.06 | Alpha<br>0.92 | $\bar{r}$<br>0.29 | Items<br>28 |
|---|---|---|---|---|---|
| WWW | 27.15 | 11.36 | | | 65 |
| ANAG | 6.01 | 4.54 | 0.84 | 0.18 | 25 |
| LSEQ | 9.39 | 4.15 | 0.86 | 0.24 | 20 |
| HIDWD | 11.75 | 5.17 | 0.89 | 0.28 | 20 |
| WWS | 14.34 | 5.46 | | | 32 |
| PS | 2.50 | 1.62 | 0.50 | 0.10 | 8 |
| PG | 3.63 | 1.41 | 0.47 | 0.11 | 8 |
| FIC | 4.47 | 1.67 | 0.53 | 0.13 | 8 |
| SM | 3.73 | 2.38 | 0.76 | 0.28 | 8 |

***Table 9.2. Correlations between the item-types for the Verbal and Spatial scales (n=197).***

| Working with Words: | | | | Working with Shapes: | | | | |
|---|---|---|---|---|---|---|---|---|
| | ANAG | LSEQ | HIDWD | | PS | PG | FIC | MS |
| ANAG | 1.00 | 0.57 | 0.49 | PS | 1.00 | 0.35 | 0.40 | 0.54 |
| LSEQ | | 1.00 | 0.47 | PG | | 1.00 | 0.41 | 0.46 |
| HIDWD | | | 1.00 | FIC | | | 1.00 | 0.49 |
| | | | | MS | | | | 1.00 |

Simple addition of the complete item sets to produce overall scale scores for WWS and WWW provided measures which correlate about 0.50 with each other and with WWN (see Table 9.3). Principal components analysis of the three scales showed that 67% of the variance (eigenvalue 2.01) could be accounted for by one general factor. Commonalties for each of the three scales are shown in Table 9.3. These indicate that between 65% and 70% of the variance in each scale can be accounted for by a common factor (i.e. general ability).

***Table 9.3. Correlations between the total scale scores for each of the three scales (n=197), with their commonalties (see text for explanation).***

| Correlations: | WWN | WWW | WWS | Communality |
|---|---|---|---|---|
| WWN | 1.00 | .53 | .51 | .70 |
| WWW | .53 | 1.00 | .48 | .66 |
| WWS | .51 | .48 | 1.00 | .65 |

The item data was analyzed in order to reduce the number of items per scale while maintaining or increasing the scale reliabilities. Principal components analyses were used to identify those items which loaded (0.30 or greater) on the first un-rotated factor for all the items in WWS. 22 of the 32 items met this criterion. For this subset, the mean inter-item correlation was 0.17 and alpha was 0.82.

Principal components analyses were also used to identify items loading on the first unrotated component for each of the Verbal and Spatial item-types and for the combined sets of item-types for each of the two scales (WWW and

WWS). From these analyses, a total of 48 items were retained for the Verbal scale (14 anagrams, 17 letter sequences and 17 hidden words) and 22 for the Spatial scale (4 PS items, 4 PG items, 6 FIC items and all 8 of the SM items).

### 9.2.2 Final scales

Reliability analyses of the final scales were carried out. Descriptive statistics are reported in Table 9.4. For both the Verbal and Numerical scales, alpha reliabilities of 0.92 were obtained. For the Spatial scale, the alpha is lower (0.82). This is, at least in part, a reflection of the greater diversity of item-types in the spatial test. For all three scales, the scale means are close to the mid-point of the scale and scores are well-distributed from the minimum to maximum possible values, with SDs equal to about one quarter of the total scale range.

Correlations between WWN, WWW and WWS are not affected by the removal of items from the scales. They remain at around r=0.50 (compare Tables 9.3 and 9.5). Effects of age are quite small, with WWN scores increasing slightly with age and WWW and WWS scores decreasing slightly. However, in all cases, these effects are statistically non-significant.

***Table 9.4. Means, SDs and reliabilities of the items selected for each of the three scales.***

| Scale | Mean | SD | Minimum | Maximum | Alpha | $\bar{r}$ | Items |
|---|---|---|---|---|---|---|---|
| WWN | 13.21 | 6.21 | 0.00 | 22.00 | 0.92 | 0.34 | 22 |
| WWW | 19.21 | 9.63 | 1.00 | 44.00 | 0.92 | 0.20 | 48 |
| WWS | 10.32 | 4.46 | 0.00 | 22.00 | 0.82 | 0.17 | 22 |

***Table 9.5. Correlations between the three scales and correlations with gender and age.***

| | Gender | Age | WWN | WWW | WWS |
|---|---|---|---|---|---|
| WWN | .21 | .11 | 1.00 | 0.52 | 0.51 |
| WWW | -.06 | -.12 | 0.52 | 1.00 | 0.48 |
| WWS | .16 | -.14 | 0.51 | 0.48 | 1.00 |

The reduced item sets (22 for WWN, 48 for WWW and 22 for WWS) were used to construct the standardization test booklets.

- The three item-types for WWW were intermixed in one 8-minute test.
- The time limit for the WWN test was set at 3 minutes, and the test included 24 items - of which the first two were not scored.
- The WWS items were divided into two subtests. The first contained the PS, PG and FIC items (time limit 7 minutes) while the second contained just the SM items (time limit 2 minutes).

The Standardization Test booklet also contained the ICES Plus Interest inventory and the ICES inventory.

## 9.3 The Standardization Sample

A standardization sample of 516 people was drawn from 17 different organizations in North America (see Appendix E)**.** Seventeen different job functions were sampled across the 17 organizations - with one job function being sampled in each organization. The numbers sampled from each organization ranged from a minimum of 8 to maximum of 114. The average age of the sample was 37.3 years (SD=10.4); 51.4% were female and 48.6% male. The majority (96.9%) had English as their first language, with 10 people having Spanish as their first language and 6 reporting some other first language. 80.6% of the sample reported their ethnic origin as White, 9.3% as Black and 6.8% as Hispanic.

Item analyses of all the scales (Abilities, Interests and Personality) produced the alpha reliability estimates shown in Table 9.6. For the Ability scales, these are very similar to those found for the final versions of the scales in the pilot study (see Table 9.4 above). The ICES Plus Ability scale is a weighted sum of the three specific Ability scales. To allow for the fact that there are approximately twice as many items in the Verbal scale, WWW, the ICES Plus Ability score is computed as: 2xWWN + WWW + 2xWWS.

***Table 9.6. Means SDs and alpha reliabilities for all the Ability, Interests and Personality scales (n=516).***

| Scale | Mean | SD | Alpha | items |
|---|---|---|---|---|
| I1 | 22.83 | 4.43 | 0.63 | 12 |
| I2 | 25.33 | 5.12 | 0.67 | 12 |
| C1 | 26.31 | 4.19 | 0.60 | 12 |
| C2 | 25.43 | 5.00 | 0.63 | 12 |
| E1 | 23.34 | 5.30 | 0.73 | 12 |
| E2 | 23.78 | 6.09 | 0.80 | 12 |
| S1 | 24.50 | 5.02 | 0.69 | 12 |
| S2 | 24.35 | 4.84 | 0.61 | 12 |
| INDEP | 48.16 | 8.20 | 0.76 | 24 |
| CONSC | 51.74 | 7.59 | 0.71 | 24 |
| EXTRAV | 47.12 | 10.22 | 0.85 | 24 |
| STABLE | 48.85 | 8.83 | 0.78 | 24 |
| SocDes | 23.47 | 5.40 | 0.73 | 14 |
| PEOPLE | 39.29 | 8.63 | 0.82 | 12 |
| DATA | 30.30 | 7.87 | 0.76 | 12 |
| THINGS | 33.23 | 10.34 | 0.87 | 12 |
| WWN | 10.85 | 5.59 | 0.90 | 22 |
| WWW | 24.03 | 9.90 | 0.94 | 48 |
| WWS | 12.14 | 4.62 | 0.82 | 22 |
| GENERAL | 70.13 | 25.52 | 0.95 | 92 |

### 9.3.1 Comparisons with the Phase One and Phase Two Samples

The Interests and Personality scales are largely the same as those used during the second stage of Phase Two. Some minor modifications to some of the personality items were made between these phases and a small study carried out to check their equivalence.

For Interests (PDT), the Phase Three sample is similar to the others on Data and People, but has a somewhat lower average on People. When the differences in variance are taken into account, the scales have comparable reliabilities to those found in the earlier phases.

The Personality scales (ICES) show some minor variations between samples. However, the most important factor is the variations in SDs - as this has a direct effect on the reliability estimates. Comparisons between the present data and those for the other Phases suggest that the reliability of the scales has been maintained across successive samples. (Details of these comparisons are presented in the next chapter).

***Table 9.7: Means and SDs for the males (n=232) and the females (n=226). The eta values give the correlation between gender and each scale, with the Common Variance (eta squared as a percentage) indicating the amount of scale score variance which could be predicted from knowing a candidate's gender. The F-ratio indicates whether the difference between the sexes is statistically greater than zero (p<.05) or not (ns).***

| | MALES | | FEMALES | | Common | | | |
|---|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **eta** | **Var** | **F-ratio** | **p** |
| INDEP | 50.97 | 8.08 | 45.22 | 7.67 | 0.34 | 11.8% | 60.95 | <.001 |
| I1 | 24.91 | 4.37 | 20.72 | 3.60 | 0.46 | 21.5% | 124.85 | <.001 |
| I2 | 26.06 | 5.05 | 24.50 | 5.27 | 0.15 | 02.3% | 10.47 | <.01 |
| CONSC | 52.48 | 7.33 | 51.06 | 7.81 | 0.09 | 00.9% | 4.03 | <.05 |
| C1 | 26.35 | 4.10 | 26.31 | 4.15 | 0.00 | 00.0% | <1 | ns |
| C2 | 26.13 | 4.78 | 24.75 | 5.22 | 0.14 | 01.9% | 8.73 | <.01 |
| EXTRA | 47.89 | 10.58 | 46.32 | 10.22 | 0.08 | 00.6% | 2.59 | ns |
| E1 | 23.55 | 5.53 | 23.05 | 5.30 | 0.05 | 00.2% | <1 | ns |
| E2 | 24.34 | 6.18 | 23.27 | 6.11 | 0.09 | 00.8% | 3.47 | ns |
| STABLE | 50.21 | 8.98 | 47.34 | 8.80 | 0.16 | 02.6% | 11.94 | <.001 |
| S1 | 24.99 | 5.06 | 23.87 | 5.06 | 0.11 | 01.0% | 5.65 | <.05 |
| S2 | 25.21 | 4.91 | 23.47 | 4.76 | 0.18 | 03.2% | 14.87 | <.001 |
| SocDes | 22.84 | 5.28 | 23.70 | 5.38 | 0.08 | 00.7% | 3.00 | ns |
| PEOPLE | 40.02 | 7.74 | 38.66 | 9.47 | 0.08 | 00.6% | 2.85 | ns |
| DATA | 29.40 | 8.01 | 30.89 | 7.69 | 0.09 | 00.9% | 4.11 | ns |
| THINGS | 36.92 | 9.68 | 29.43 | 9.50 | 0.36 | 13.3% | 69.82 | <.001 |
| WWN | 12.34 | 5.55 | 10.02 | 5.25 | 0.21 | 04.4% | 21.06 | <.001 |
| WWW | 23.34 | 9.54 | 25.50 | 9.84 | 0.11 | 01.2% | 5.66 | <.05 |
| WWS | 12.84 | 4.50 | 11.62 | 4.58 | 0.13 | 01.8% | 8.29 | <.01 |
| GENERAL | 73.83 | 24.43 | 68.80 | 25.53 | 0.10 | 01.0% | 4.61 | ns |

## 9.4 Age differences

Analysis of age-related differences was carried out after coding people into one of ten age bands: from "under 20", through "20 to 24", "25 to 29", and so on in 5-year intervals to "55 to 59" and finally "59 and above". Effects of age were examined using analysis of variance. The results of these analyses are summarized in Table 9.8. Eta coefficients and R coefficients are both shown as the former are sensitive to both linear and non-linear relationships between age and scale score while the latter indicate the degree of linear relationship.

As noted in the previous discussion of age effects, any effects of age are cohort-effects; that is, they relate to different groups of people who happen to vary in age. It does not follow that the same variations in scale score would be found across time within the same individuals (i.e. longitudinally).

### 9.4.1 Personality

While some of the scales show statistically significant linear relationships with age, in most cases the magnitudes of these effects are quite small - generally less than 2% of the variance.

As with the Phase One and Two data, the effect of age in relation to Conscientiousness is confined to the minor scale C1. This reflects an expected increase in concern for traditional values across the age groups. Examination of Table 9.8 shows that this relationship is a strongly linear one, with about half of the age-related variance being accounted for by the linear relationship.

An effect of age was again found for the SocDes scale. As suggested before, this may reflect the fact that there are age-related differences in the social norms associated with the behaviors used to provide the content for the items in this scale.

### 9.4.2 Interests

None of the scales showed clear linear relationships with age.

### 9.4.3 Ability

Performance on the Spatial Ability scale (WWS) decreases significantly with increases in age. The other two scales are relatively unaffected. The effect of age on WWS is responsible for the small effect of age on the ICES Plus Ability scale.

***Table 9.8: For each scale the table shows the overall relationship between the six age bands (see text for explanation) and scale scores, expressed as eta, and the linear relationship, expressed as R. (NB. Negative values of R indicate that scale scores decrease as age increases.) The final column (r) gives the simple linear correlation between age (in years) and scale score. N=458.***

| Scale | Eta | Common Vary% | Overall F ratio | P | R | Linear Vary% | Linear F ratio | p | [r] |
|---|---|---|---|---|---|---|---|---|---|
| INDEP | 0.16 | 02.5% | 1.26 | ns | -.08 | 00.7% | 3.12 | ns | -.09 |
| I1 | 0.17 | 02.8% | 1.44 | ns | -.08 | 00.7% | 3.23 | ns | -.09 |
| I2 | 0.14 | 01.9% | <1 | ns | -.06 | 00.4% | 1.65 | ns | -.06 |
| CONSC | 0.18 | 03.3% | 1.70 | ns | 0.11 | 01.1% | 5.17 | <.05 | .09 |
| C1 | 0.22 | 04.7% | 2.45 | <.01 | 0.15 | 02.1% | 10.01 | <.01 | .12 |
| C2 | 0.13 | 01.6% | <1 | ns | 0.04 | 00.2% | <1 | ns | .04 |
| EXTRAV | 0.18 | 03.4% | 1.74 | ns | -.13 | 01.6% | 7.54 | <.01 | -.13 |
| E1 | 0.17 | 02.9% | 1.48 | ns | -.09 | 00.8% | 3.83 | ns | -.09 |
| E2 | 0.18 | 03.2% | 1.63 | ns | -.14 | 01.8% | 8.49 | <.01 | -.14 |
| STABLE | 0.23 | 05.2% | 2.72 | <.01 | 0.11 | 01.2% | 5.73 | <.05 | .11 |
| S1 | 0.23 | 05.4% | 2.83 | <.01 | 0.12 | 01.5% | 7.21 | <.01 | .11 |
| S2 | 0.19 | 03.4% | 1.77 | ns | 0.07 | 00.6% | 2.53 | ns | .09 |
| SocDes | 0.19 | 03.7% | 1.93 | <.05 | 0.17 | 0.23% | 13.74 | <.001 | .11 |
| PEOPLE | 0.16 | 02.6% | 1.30 | ns | -.04 | 00.1% | <1 | ns | -.04 |
| DATA | 0.21 | 04.2% | 2.20 | <.05 | -.05 | 00.2% | 1.23 | ns | -.07 |
| THINGS | 0.18 | 03.2% | 1.64 | ns | -.00 | 00.0% | <1 | ns | -.02 |
| WWN | 0.12 | 01.4% | <1 | ns | 0.02 | 00.1% | <1 | ns | .05 |
| WWW | 0.12 | 01.5% | <1 | ns | -.09 | 00.8% | 3.59 | ns | -.08 |
| WWS | 0.26 | 06.6% | 3.52 | <.001 | -.24 | 05.8% | 27.85 | <.001 | -.24 |
| GENERAL | 0.14 | 02.1% | 1.04 | ns | -.11 | 01.3% | 5.81 | <.05 | -.09 |
| Degrees of freedom for F ratios: Overall = 9, (n-10); Linear = 1, (n-10) | | | | | | | | | |

# PART III: NORMS AND VALIDATION

## 10. DESCRIPTIVE STATISTICS FOR THE COMBINED SAMPLE

Together, the 1518 people tested during Phase One, those tested during Phase Two and those tested in Phase Three provide a combined sample data set of 3874 cases. Data on the ICES Personality scales were available for the whole sample - major and minor scales. While there had been variations made to some of the items during the development process, analyses indicated that these had no systematic effect on scale score averages nor on scale inter-correlations.

Data on the Social Desirability (SocDes) scale were only available for the Phase Two Stage Two and Phase Three samples. Data on the Interest scales (PDT: People, Data and Things) and on the Working with Numbers test was available for all but the Stage Two Phase Two sample, while data was only available on the other two Ability scales (Working with Words and Working with Shapes) for the final Phase Three sample.

In addition to these, a further set of job-related samples (n=871) were obtained as part of a series of validation studies. These studies were carried out prior to the development of the full set of Ability scales. Item analyses were not carried out on these samples, as only scale raw scores were collected (together with criterion data - see later sections). Thus, separate reliability analyses are not reported for these data, but they are included in the data set for normative purposes.

The sizes of the various samples are summarized in Table 10.1.

***Table 10.1: ICES Plus development and validation samples (to the end of December 1993).***

| Sample | Number |
|---|---|
| Development samples | |
| Phase One | 1511 |
| Phase Two Stage One | 1236 |
| Phase Two Stage Two | 604 |
| Phase Three | 516 |
| Additional samples (validation studies) | 871 |
| Total | 4738 |

### 10.1 Characteristics of the Combined Data Set of n=4738

Information about gender was available for most (4735) of the sample: of these, 59.8% of the sample were male, 40.2% female. Information on ethnic origin was available on 3866 (81.6%) of the total sample: of these 76.3% were White, 8.4% Black, 5.3% Oriental, 5.3% Asian, 3.0% Hispanic. Information about the person's native language was available for 3859 (81.4%) of the total sample: English was the native language for 81.7% of these; 5.2% of the sample had French as their first language and completed a French version of the ICES Plus battery. All the others completed the English version. These included 4.8% for whom Chinese was their native language; 3.3% for whom it was Malaysian and 2.2% for whom it was Spanish.

### 10.2 Scale Raw Score Means and the Derivation of the ICES Plus Norm Tables

Overall means and SDs for each ICES Plus scale are shown in Table 10.2. These can be used to produce standard score conversions for each of the scales - e.g. non-normalized z-score, sten or T-score. The norm tables developed for generating the ICES Plus standard scores, however, produce normalized sten scores with age-corrections. These may differ slightly from those produced using the means and SD given in Table 10.2 - with any differences occurring mainly at the boundaries between stens at each end of the distribution.

***Table 10.2: Means and SDs for the combined sample. (NB. Complete data is available for ICES scales on 4716 of the sample and on the SocDes scale for 1120 of those.)***

| Scale | Mean | SD | N |
|---|---|---|---|
| I1 | 23.02 | 4.87 | 4716 |
| I2 | 25.81 | 5.19 | 4716 |
| C1 | 25.33 | 4.63 | 4716 |
| C2 | 26.07 | 5.19 | 4716 |
| E1 | 25.42 | 5.25 | 4716 |
| E2 | 25.71 | 5.94 | 4716 |
| S1 | 26.15 | 5.12 | 4716 |
| S2 | 25.28 | 4.92 | 4716 |
| INDEP | 48.83 | 8.50 | 4716 |
| CONSC | 51.40 | 8.26 | 4716 |
| EXTRAV | 51.14 | 10.04 | 4716 |
| STABLE | 51.43 | 9.03 | 4716 |
| SocDes | 22.98 | 5.72 | 1120 |
| PEOPLE | 42.57 | 8.28 | 4075 |
| DATA | 29.99 | 7.96 | 4075 |
| THINGS | 33.11 | 9.74 | 4075 |
| WWN | 14.73 | 6.35 | 4134 |
| WWW | 24.03 | 9.90 | 516 |
| WWS | 12.14 | 4.62 | 516 |
| GENERAL | 70.13 | 25.52 | 516 |

Normalized sten scales are now very widely used - especially for Personality inventories. There is much to be said for adopting a measurement scale, which is as “standard” as possible - as this will facilitate test interpretation.

Age-correct normalized sten conversion tables were produced for the personality and interests scales using the combined Phase One and Phase Two sample (“General Working Population norms”: GWP norms) and for breakdowns of the sample by gender. In addition, norm tables have been constructed from the Phase Three data for all the ICES Plus scales (Interest, Ability and Personality). Unless otherwise stated, all analyses reported in Chapters 10 through 13 which report sten scores were carried out using the Phase One and Phase Two GWP composite norm table.

### 10.2.1 Number of norm tables

In general, it is only likely to be necessary to produce separate norm tables for a particular group, where that group is known to have a different distribution of scores from other groups. On this basis, it might be appropriate - given the results reported above - to have separate norm tables for males and females; for each ethnic origin group and for each combination of the two variables. This would require the production of a large number of norms tables: Both sexes, male only and female only for all the ethnic origin groups combined and for each one on its own (White, Black, Asian, Oriental and so on) separately.

There is not sufficient data in the present samples to provide all these norm tables. In any case, in practice it is best to use combined norms and then make allowances, where appropriate, for any group differences, which are not job-related. Where group differences (whatever they are related to) are job-related, then one should always-used combined norms - not separate group norm tables.

For descriptive and interpretation purposes it is sometimes useful to make comparisons with a number of different norm groups - especially where these are occupationally defined groups. The ICES Plus computer scoring is always done using the combined norm table.

### 10.2.2 Controlling for age

We have seen that there are age-related effects on scores on some of the scales. These need to be taken into account when norm tables are produced so that effects of age are not confounded with individual scale differences. The procedure followed in producing the norm tables was to first “remove” linear age effects from raw scale scores (see Appendices B and D for details) and then to construct normalized sten scale cut-off points from age-regressed raw scores.

The norm tables are contained in Appendices B and D, together with detailed instructions for the production of age-corrected raw scores.

### 10.3 Scale Raw Score Means, Reliability and Standard Errors of Measurement

The scale raw score means are broken down by sample and shown together with reliability estimates (Cronbach's alpha) and re-test reliability in Table 10.3. Raw score Standard Errors of Measurement (SEMs) are shown in Table 10.4 and sten score SEMs in Table 10.5. All SEMs apart from the re-test SEMs are based on alpha coefficients (see Table 10.3).

For the Personality scales, sten score SEMs are around one sten for the main scales and up to one and a quarter stens for the sub-scales. Ideally one would like to see these all below one sten. However, such increased accuracy could only be obtained either by increasing the number of items in each scale or by increasing the within-scale item similarity. Neither of these options is desirable.

The three Interest scales (People, Data and Things) show good reliability and have SEMs well below one sten score. The Ability scales all have good reliabilities and small SEMs. The SEM for the combined ICES Plus Ability scale is less than half a sten.

***Table 10.3. Descriptive statistics for the ICES Plus Scales for each sample from the three Phases of the ICES Plus development program, with alpha and re-test scale reliabilities.***

| | PHASE ONE | | | PHASE TWO | | | PHASE TWO | | | | PHASE THREE | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Stages One & Two | | | Stage One | | | Stage Two | | | | | | |
| | n=1518 | | | n=1236 | | | n=604 | | | (n=147) | n=516 | | |
| Scale | Mean | SD | alpha | Mean | SD | alpha | Mean | SD | alpha | re-test | Mean | SD | alpha |
| II1 | 23.60 | 4.89 | 0.70 | 22.08 | 4.75 | 0.66 | 21.88 | 4.95 | 0.71 | 0.81 | 22.83 | 4.43 | 0.63 |
| I2 | 25.84 | 5.19 | 0.68 | 25.15 | 5.37 | 0.68 | 25.32 | 5.40 | 0.71 | 0.80 | 25.33 | 5.12 | 0.67 |
| C1 | 26.53 | 4.29 | 0.60 | 25.23 | 4.34 | 0.56 | 22.38 | 5.22 | 0.72 | 0.79 | 26.31 | 4.19 | 0.60 |
| C2 | 27.18 | 5.00 | 0.67 | 26.16 | 5.14 | 0.65 | 24.81 | 5.50 | 0.70 | 0.86 | 25.43 | 5.00 | 0.63 |
| E1 | 26.25 | 5.17 | 0.72 | 25.48 | 5.21 | 0.69 | 24.53 | 5.62 | 0.74 | 0.74 | 23.34 | 5.30 | 0.73 |
| E2 | 27.06 | 5.60 | 0.79 | 25.12 | 5.91 | 0.78 | 24.30 | 6.19 | 0.80 | 0.74 | 23.78 | 6.09 | 0.80 |
| S1 | 27.21 | 4.94 | 0.70 | 24.95 | 5.14 | 0.66 | 25.21 | 5.33 | 0.70 | 0.60 | 24.50 | 5.02 | 0.69 |
| S2 | 25.21 | 4.85 | 0.65 | 24.49 | 4.88 | 0.62 | 25.18 | 5.02 | 0.65 | 0.74 | 24.35 | 4.84 | 0.61 |
| INDEP | 49.43 | 8.44 | 0.77 | 47.23 | 8.45 | 0.75 | 47.19 | 8.75 | 0.79 | 0.83 | 48.16 | 8.20 | 0.76 |
| CONSC | 53.69 | 7.84 | 0.74 | 51.39 | 7.82 | 0.71 | 48.83 | 10.71 | 0.78 | 0.84 | 51.74 | 7.59 | 0.71 |
| EXTRAV | 53.32 | 9.58 | 0.84 | 50.60 | 9.86 | 0.83 | 47.19 | 9.00 | 0.86 | 0.76 | 47.12 | 10.22 | 0.85 |
| STABLE | 52.42 | 8.82 | 0.80 | 49.44 | 8.86 | 0.77 | 50.39 | 9.43 | 0.81 | 0.69 | 48.85 | 8.83 | 0.78 |
| SocDes | not applicable | | | not applicable | | | 22.57 | 5.94 | 0.78 | 0.82 | 23.47 | 5.40 | 0.73 |
| PEOPLE | 43.43 | 8.44 | 0.85 | 42.14 | 7.96 | 0.80 | not administered | | | | 39.29 | 8.63 | 0.82 |
| DATA | 30.19 | 8.10 | 0.81 | 30.71 | 7.76 | 0.77 | not administered | | | | 30.30 | 7.87 | 0.76 |
| THINGS | 32.91 | 9.75 | 0.86 | 32.23 | 9.48 | 0.84 | not administered | | | | 33.23 | 10.34 | 0.87 |
| WWN | 16.50 | 5.59 | 0.89 | 13.12 | 6.51 | 0.90 | not administered | | | | 10.85 | 5.59 | 0.90 |
| WWW | not administered | | | not administered | | | not administered | | | | 24.03 | 9.90 | 0.94 |
| WWS | not administered | | | not administered | | | not administered | | | | 12.14 | 4.62 | 0.82 |
| GENERAL | not administered | | | not administered | | | not administered | | | | 70.13 | 25.52 | 0.95 |

***Table 10.4: Raw Score Standard Errors of Measurement (SEMs).***

| Scale | *Phase I* Stage 1+2 SEM | *Phase II* Stage 1 SEM | *Phase II* Stage 2 SEM | *Phase III* Re-test-SEM | *SEM* |
|---|---|---|---|---|---|
| I1 | 2.45 | 2.77 | 2.67 | 2.19 | 2.69 |
| I2 | 2.94 | 3.03 | 2.91 | 2.31 | 2.94 |
| C1 | 2.71 | 2.87 | 2.76 | 2.30 | 2.64 |
| C2 | 2.87 | 3.04 | 3.01 | 2.12 | 3.04 |
| E1 | 2.74 | 2.90 | 2.87 | 2.68 | 2.75 |
| E2 | 2.57 | 2.77 | 2.77 | 2.82 | 2.72 |
| S1 | 2.71 | 2.99 | 2.92 | 3.05 | 2.80 |
| S2 | 2.87 | 3.01 | 2.97 | 2.38 | 3.02 |
| INDEP | 4.05 | 4.23 | 4.01 | 3.59 | 4.02 |
| CONSC | 4.00 | 4.82 | 5.02 | 3.56 | 4.09 |
| EXTRAV | 3.83 | 4.07 | 3.67 | 4.73 | 3.96 |
| STABLE | 3.94 | 4.25 | 4.42 | 4.79 | 4.14 |
| SocDes | -- | -- | 2.79 | 2.64 | 2.81 |
| PEOPLE | 3.27 | 3.56 | -- | -- | 3.66 |
| DATA | 3.53 | 3.72 | -- | -- | 3.86 |
| THINGS | 3.65 | 3.79 | -- | -- | 3.73 |
| WWN | 1.85 | 1.95 | -- | -- | 1.77 |
| WWW | -- | -- | -- | -- | 2.42 |
| WWS | -- | -- | -- | -- | 1.96 |
| GENERAL | -- | -- | -- | -- | 5.71 |

***Table 10.5: Sten Score Standard Errors of Measurement (SEMs).***

| Scale | Phase I Stage 1+2 SEM | Phase II Stage 1 SEM | Phase II Stage 2 SEM | Phase III Re-test-SEM | SEM |
|---|---|---|---|---|---|
| I1 | 1.00 | 1.17 | 1.08 | 0.87 | 1.21 |
| I2 | 1.13 | 1.13 | 1.08 | 0.89 | 1.15 |
| C1 | 1.26 | 1.32 | 1.06 | 0.92 | 1.26 |
| C2 | 1.15 | 1.18 | 1.09 | 0.72 | 1.22 |
| E1 | 1.06 | 1.11 | 1.02 | 1.02 | 1.04 |
| E2 | 0.92 | 0.94 | 0.89 | 1.02 | 0.89 |
| S1 | 1.10 | 1.16 | 1.10 | 1.27 | 1.12 |
| S2 | 1.18 | 1.23 | 1.18 | 1.02 | 1.25 |
| INDEP | 0.96 | 1.00 | 0.92 | 0.83 | 0.98 |
| CONSC | 1.02 | 1.23 | 0.94 | 0.80 | 1.08 |
| EXTRAV | 0.80 | 0.83 | 0.82 | 0.98 | 0.77 |
| STABLE | 0.89 | 0.96 | 0.94 | 1.11 | 0.94 |
| SocDes | -- | -- | 0.94 | 0.85 | 1.04 |
| PEOPLE | 0.77 | 0.89 | -- | -- | 0.85 |
| DATA | 0.87 | 0.96 | -- | -- | 0.98 |
| THINGS | 0.75 | 0.80 | -- | -- | 0.72 |
| WWN | 0.66 | 0.60 | -- | -- | 0.63 |
| WWW | -- | -- | -- | -- | 0.49 |
| WWS | -- | -- | -- | -- | 0.85 |
| GENERAL | -- | -- | -- | -- | 0.45 |

# 11. CONTENT AND CONSTRUCT VALIDITY

## 11.1 Content Validity

The scale construction process has assured good content validity. As was seen earlier (see Tables 2.2 and 4.7), the scales which have been developed have strong relationships with the prior content-based item groups which were based on rational scale definitions rather than on exploratory data analysis techniques (such as factor or cluster analysis).

Good content validity provides a sound base for construct validity, and reduces the likelihood of over-interpreting chance effects which might arise in the scale construction process.

## 11.2 Construct Validity: Relationships Between the ICES Plus Interests, Ability and Personality Scales

The patterns of correlations between scales can be examined to provide some evidence for construct validity. Certain patterns of relationship between interests and personality can be expected and would provide support for the construct validity of the scales. For example, we would expect people who have high People scores to be “Outgoing” (E2); those with interests in Things to be “Competitive” (I1) and those with interests in Data to be “Organized” (C2).

Correlations between the scales are shown in Table 11.1. The correlations are based on 3,946 people - the number from the combined data set who produced complete and usable data on all of these scales. With this sample size, the Standard Error for each correlation is less than plus or minus 0.016 and hence correlations greater than plus or minus 0.03 are statistically significant. From Table 11.1, it can be seen that the above expectations are generally supported:

- The highest correlations between People and Personality scales are with Extraversion (E1 and E2) and with I2 (Assertive).
- For Things, the highest correlation is, as predicted, with I1 (Competitive). People who are interested in things also tend to be slightly more self-sufficient (negative E1). However, the general trend is that while those interested in people are Extravert, Interest in Things is relatively independent of Extraversion.
- The pattern for Data is also clear. The largest correlations are, as one would expect, with Conscientiousness (C1 and C2), particularly C2, which concerns attention to detail. In addition, there are correlations between Data and the Stability scales (particularly S1).

***Table 11.1: Correlations between the Interest and Personality scales (n=3946).***

| | INTERESTS | | |
|---|---|---|---|
| | **PEOPLE** | **DATA** | **THINGS** |
| PERSONALITY | | | |
| I1 | .11 | -.06 | .17 |
| I2 | .33 | -.01 | .09 |
| C1 | .01 | .15 | .01 |
| C2 | .07 | .18 | .01 |
| E1 | .30 | .07 | -.06 |
| E2 | .37 | .03 | .01 |
| S1 | .23 | .11 | .07 |
| S2 | .22 | .06 | .11 |

A clearer picture of the relationships between these scales can be obtained through Principal Components analysis. Analysis of the eight Personality sub-scales and the three Interest scales produced a five-factor solution, which accounted for 72.9% of the total variance. Scale commonalties varied from a minimum of 0.68 (I2) to a maximum of 0.82 (I1). Similar patterns of loadings were obtained for both Varimax and Oblimin rotations. The pattern of loadings after Varimax rotation is shown in Table 11.2.

***Table 11.2: Varimax rotated principal component pattern matrix loadings for the 11 scales (n=3946). Loadings less than 0.30 have been omitted.***

| | FACTOR | | | | |
|---|---|---|---|---|---|
| **Scale** | **I** | **II** | **III** | **IV** | **V** |
| I1 | | | | 0.90 | |
| I2 | 0.41 | | | 0.67 | |
| C1 | | | | | 0.83 |
| C2 | | | | | 0.83 |
| E1 | 0.85 | | | | |
| E2 | 0.82 | | | | |
| S1 | | 0.88 | | | |
| S2 | | 0.88 | | | |
| PEOPLE | 0.52 | | 0.54 | | |
| DATA | | | 0.78 | | |
| THINGS | | | 0.74 | | |

The five components are readily interpreted from the pattern of loadings shown in Table 10.2.

1. **Extraversion**, with its main loading on E1 and E2, which is also linked to an interest in People and Independence - particularly I2 – Assertive and Forthright.
2. **Emotional Stability**, with loadings of S1 and S2
3. **General Interest**. The main loadings on this are the non-People scales (Data and Things). We might take this factor as indicating an overall level of interest in work-related activities of various sorts. With this in mind, it may be worth looking at the potential of a total Interest score (i.e. People+Data+Things) as a predictor of performance and tenure.
4. **Independence**, defined by I1 and I2.
5. **Conscientiousness**, defined by C1 and C2.

This analysis of the ICES Plus scales, provides good support for the “internal” construct validity of scales. In particular, the separation in factor space of the four major Personality scales is clearly shown.

## 11.3 Relationships between Ability, Interests and Personality - the Phase Three Data

While interests and personality are often related, “pure” ability, it is argued, should be independent of both of them (e.g. Kline, 1991). We would expect, therefore, to find relatively low correlations between ability scales (e.g. WWN) on the one hand, and the Personality and Interest scales on the other.

Data are available from Phase Three of the ICES Plus development, which can be used to examine the relationship between ability on the one hand and interests and personality on the other. Correlations between the Ability scales and of the Ability scales with Personality and Interests are shown in Table 11.3. In addition, this Table shows correlations with gender. For these, a positive correlation indicates that males score higher, on average, than females; while a negative correlation indicates that females tend to score higher than males.

Despite the fact that “pure” ability is held to be independent of personality and interests, there are correlations between the ICES Plus Ability scales and some of the Interest and Personality scales. Correlations between ability and interests and personality show that those of higher general ability tend to have higher levels of interests in Things, to be more relaxed and to be less concerned about matters of detail and traditional values. They also score lower on the SocDes scale, suggesting they have less concern for presenting a “socially correct” view of themselves.

***Table 11.3. Correlations between the Ability scales and Personality and Interest scales, together with correlations with gender (n=516).***

| | WWN | WWW | WWS | GENERAL | GENDER |
|---|---|---|---|---|---|
| INDEP | .12 * | .01 | .10 | .09 | .35 ** |
| I1 | .12 * | -.01 | .11 | .09 | .47 ** |
| I2 | .09 | .03 | .06 | .07 | .16 ** |
| CONSC | -.16 ** | -.19 ** | -.28 ** | -.25 ** | .10 |
| C1 | -.12 * | -.13 * | -.26 ** | -.20 ** | .00 |
| C2 | -.14 * | -.18 ** | -.21 ** | -.21 ** | .14 * |
| EXTRAV | .11 * | .07 | .06 | .10 | .08 |
| E1 | .10 | .05 | .01 | .07 | .05 |
| E2 | .10 | .08 | .09 | .11 | .09 |
| STABLE | .17 ** | .02 | .06 | .10 | .16 ** |
| S1 | .11 * | -.02 | -.01 | .04 | .11 |
| S2 | .20 ** | .05 | .11 | .15 ** | .18 ** |
| SocDes | -.15 ** | -.17 ** | -.20 ** | -.20 ** | -.10 |
| PEOPLE | .09 | .05 | -.00 | .06 | .07 |
| DATA | .01 | .07 | .04 | .04 | -.09 |
| THINGS | .20 ** | .05 | .27 ** | .20 ** | .37 ** |
| WWN | 1.00 | .57 ** | .55 ** | .86 ** | .21 ** |
| WWW | .57 ** | 1.00 | .54 ** | .83 ** | -.10 |
| WWS | .55 ** | .54 ** | 1.00 | .82 ** | .14 * |
| GENERAL | .86 ** | .83 ** | .82 ** | 1.00 | .11 |
| * p<.05 ** p<.01 | | | | | |

The pattern of relationships between the full set of ICES Plus scales was formally examined using principal components analysis. Raw scores from the eight minor ICES scales, the three Interest scales and the three Ability scales were analyzed and a six-factor solution selected (on the basis of a screen test) which accounted for 70.1% of the variance. All selected factors had eigenvalues greater than 1.0. Varimax rotated loadings, together with scale commonalties ($h^2$) as shown in Table 11.4.

If we consider loadings of 0.40 or greater, then a clear structure emerges:

Factor I is “General Ability”;
Factor II is “Extraversion” and includes a loading on Interest in People;
Factor III is “Stability”;
Factor IV is “Conscientiousness” and “Social Desirability”;
Factor V is “Independence”;
Factor VI is “Interests”.

***Table 11.4. Varimax rotated factor loadings and commonalties (n=516).***

| | FACTOR | | | | | | Communality |
|---|---|---|---|---|---|---|---|
| **Scale** | **I** | **II** | **III** | **IV** | **V** | **VI** | **$h^2$** |
| I1 | .06 | .10 | -.19 | .02 | .88 | -.02 | .79 |
| I2 | -.01 | .43 | .19 | -.09 | .66 | .05 | .67 |
| C1 | -.07 | -.13 | .11 | .80 | .00 | .00 | .68 |
| C2 | -.10 | -.05 | -.11 | .74 | .09 | .08 | .58 |
| E1 | .05 | .86 | .06 | -.06 | .01 | .06 | .75 |
| E2 | 06 | .78 | .10 | -.16 | .30 | .02 | .74 |
| S1 | .01 | .08 | .87 | .12 | -.11 | .10 | .81 |
| S2 | .11 | .11 | .86 | .03 | .15 | .00 | .79 |
| SocDes | -.15 | -.02 | .23 | .59 | -.25 | .00 | .49 |
| PEOPLE | -.05 | .43 | .22 | -.01 | .18 | .55 | .57 |
| DATA | .06 | .10 | -.09 | .15 | -.21 | .82 | .77 |
| THINGS | .17 | -.31 | .21 | -.10 | .39 | .61 | .70 |
| WWN | .84 | .05 | .14 | -.03 | .10 | .01 | .73 |
| WWW | .84 | .11 | -.06 | -.07 | -.10 | .01 | .74 |
| WWS | .79 | -.08 | .03 | -.23 | .11 | .10 | .71 |

This analysis provides further support for the distinction drawn between the four ICES major scales. They remain well separated in this analysis at a level where both Interests (Factor VI) and Abilities (Factor I) are combined into group factors.

### 11.4 Construct Validity: Relationships Between the ICES Personality Scales and the Cattell 16 Personality Factor Questionnaire

In addition to completing the ICES Plus scales, 151 of the final Phase Two Stage Two sample also completed Form A of the Cattell 16 Personality Factor Questionnaire. Recent research (e.g. Bartram, 1992) has tended to confirm the view that while some of the first order 16PF scales may be rather unreliable, the second order factor structure is clear and robust. As the 16PF is one of the best established Personality tests currently being used in occupational (as opposed to clinical) assessment, it was decided that it would provide useful information on the construct validity of the ICES scales.

Table 11.5 shows the mean raw scores for the ICES scales and the mean 16PF-sten scores for this sample. (16PF stens are based on UK general population male+female norms.) In general, this group is somewhat high on Extraversion and Independence and low on Anxiety (i.e. high on Stability).

Table 11.5 also shows correlations between the ICES SocDes scale and the 16PF scales. This shows an expected pattern of positive correlations with G and Q3. High SocDes scores are associated with people who are not expedient (low Q1), who are low in dominance (low E) and are open and lacking in suspicion of others (low L). The ICES SocDes scale was compared with the 16PF Faking Good and Faking Bad scales (see Appendix C for details). ICES SocDes was positively correlated with 16PF Faking Good (0.25) and negatively correlated with 16PF Faking Bad (-.14). Both SocDes and Faking Good had correlations of around 0.36 with ICES Conscientiousness; however, the ICES SocDes scale was less strongly related to Stability (r=0.19 as opposed to 0.45 for the 16PF Faking Good scale).

The full matrix of correlations of all ICES scales with all 16PF scales is contained in Appendix C. The correlations between the 16PF second order factors and the ICES major scales are shown in Table 11.6. All four ICES major scales have their highest correlations with the appropriate 16PF scale (shown in bold print in Table 11.6). Note that the negative correlation between Anxiety (16PF) and Stability (ICES) is expected as high scores on the former indicate lack of emotional stability while the opposite is the case for ICES.

A factor analysis (principal components) was carried out on all the 16PF first order factors (A to Q4) and the eight ICES minor scales (I1 to S2). Five factors, accounting for 52.5% of the variance, were identified and obliquely rotated (Oblimin rotation). The resulting factor loading and factor score correlations are shown in Table 11.7. This further supports the correspondence between the two instruments. All five factors are readily identifiable from the pattern of loadings on them. Each of the first four factors contains the relevant pair of ICES minor scales together with the main 16PF first order components of each of the 16PF second order factors. The only exception is Factor V that is defined solely by the 16PF scale B (“Intelligence”).

These comparisons with the 16PF provide very clear and strong support for the construct validity of the ICES scales. The ICES major scales provide robust measures of four of the 16PF second order factors.

***Table 11.5 Means and SDs for the 16PF (sten) and ICES (raw score) scales (n= 151).***

| 16PF stens | Mean | SD | 16 PF Scale Description | Correlation with ICES SocDes scale | |
|---|---|---|---|---|---|
| 16PF-A | 5.68 | 2.54 | Warm | .03 | |
| 16PF-B | 6.11 | 1.81 | Abstract-thinking | -.04 | |
| 16PF-C | 6.34 | 2.12 | Emotionally stable | .16 | |
| 16PF-E | 6.69 | 2.12 | Dominant | -.29 | ** |
| 16PF-F | 6.70 | 2.30 | Enthusiastic | -.12 | |
| 16PF-G | 5.75 | 2.25 | Conscientious | .31 | ** |
| 16PF-H | 6.56 | 2.47 | Bold | -.04 | |
| 16PF-I | 5.13 | 1.78 | Tender-minded | .09 | |
| 16PF-L | 5.23 | 2.13 | Suspicious | -.30 | ** |
| 16PF-M | 5.54 | 2.04 | Imaginative | -.04 | |
| 16PF-N | 4.56 | 1.93 | Shrewd | .10 | |
| 16PF-O | 4.96 | 2.26 | Apprehensive | -.08 | |
| 16PF-Q | 16.05 | 2.09 | Experimenting | -.21 | * |
| 16PF-Q | 24.84 | 1.99 | Self-sufficient | -.08 | |
| 16PF-Q | 35.82 | 2.12 | Following self-image | .19 | * |
| 16PF-Q | 44.68 | 2.12 | Tense | -.15 | |
| | | | | | |
| 16PF-EXT | 6.61 | 2.55 | Extravert | -.02 | |
| 16PF-ANX | 4.60 | 2.09 | Anxious | -.19 | |
| 16PF-POISE | 5.82 | 1.61 | Tough poise | -.03 | |
| 16PF-IND | 6.67 | 2.49 | Independence | -.29 | ** |
| 16PF-CTRL | 5.82 | 2.17 | Superego/control | .31 | ** |

| | Mean | SD | Description |
|---|---|---|---|
| ICES Scales: | | | |
| I1 | 21.87 | 4.36 | Competitive, tough-minded |
| I2 | 25.21 | 5.52 | Assertive, forthright |
| C1 | 23.17 | 5.23 | Conventional, traditional, concern for moral values |
| C2 | 25.02 | 5.18 | Organized, attention to detail, neatness |
| E1 | 24.81 | 6.07 | Group-oriented, sociable |
| E2 | 24.36 | 6.56 | Outgoing, talkative |
| S1 | 25.54 | 5.43 | Poised, unruffled, unflappable |
| S2 | 24.60 | 4.91 | Relaxed, not anxious |
| INDEP | 47.07 | 8.07 | Independence |
| EXTRAV | 49.17 | 11.56 | Extraversion |
| CONSC | 48.19 | 8.90 | Conscientiousness |
| STABLE | 50.13 | 9.41 | Stability |
| SocDes | 22.46 | 5.72 | Social desirability/distortion |

* p <.05 ** p<.01

***Table 11.6 Correlations between 16PF Second Order factors and ICES major scales (n=151).***

| | ICES Major Scales | | | |
|---|---|---|---|---|
| | INDEP | CONSC | EXTRAV | STABLE |
| 16PF-EXT | .33 | -.27 | **.76** | .19 |
| 16PF-ANX | -.15 | -.15 | -.14 | **-.73** |
| 16PF-POISE | -.05 | .11 | -.22 | -.19 |
| 16PF-IND | **.65** | **-.48** | .55 | .10 |
| 16PF-CTRL | -.01 | **.64** | -.23 | .25 |

***Table 11.7 Factor analysis (principal components with Oblimin rotation) of the 16PF first order factors and the eight ICES minor scales (n=151).***

| | Factor loadings (Oblimin Pattern Matrix): | | | | | |
|---|---|---|---|---|---|---|
| | | **I** | **II** | **III** | **IV** | **V** |
| I | Extraversion | | | | | |
| | ICES-E1 | **.70** | .02 | .01 | -.10 | -.10 |
| | ICES-E2 | **.72** | .07 | .17 | -.16 | -.04 |
| | 16PF-A | **.71** | .15 | -.06 | -.07 | -.13 |
| | 16PF-F | **.79** | -.09 | .00 | -.08 | .25 |
| | 16PF-H | **.67** | .20 | .30 | .09 | .14 |
| | 16PF-Q2 | **-.70** | .04 | .02 | -.14 | -.02 |
| II | Stability | | | | | |
| | ICES-S1 | -.04 | **.85** | -.05 | .05 | -.10 |
| | ICES-S2 | .03 | **.78** | .04 | -.03 | -.15 |
| | 16PF-C | .13 | **.59** | -.07 | .04 | .24 |
| | 16PF-O | -.06 | **-.63** | -.23 | -.11 | -.22 |
| | 16PF-Q4 | -.02 | **-.82** | .03 | -.08 | -.01 |
| III | Independence | | | | | |
| | ICES-I1 | -.08 | -.10 | **.63** | .09 | -.04 |
| | ICES-I2 | .20 | .26 | **.59** | -.08 | .03 |
| | 16PF-E | .11 | .13 | **.67** | -.29 | .06 |
| | 16PF-I | .31 | .02 | **-.35** | -.09 | -.15 |
| | 16PF-L | .21 | -.36 | **.46** | -.13 | -.10 |
| | 16PF-Q1 | .19 | .04 | **.49** | -.10 | .00 |
| IV | Conscientiousness | | | | | |
| | ICES-C1 | -.03 | -.06 | -.35 | **.65** | .07 |
| | ICES-C2 | -.08 | .16 | .05 | **.58** | -.33 |
| | 16PF-G | .06 | -.02 | .13 | **.72** | -.09 |
| | 16PF-Q3 | -.06 | .27 | -.10 | **.63** | .13 |
| V | Intelligence | | | | | |
| | 16PF-B | -.09 | .02 | -.04 | -.05 | **.54** |
| | 16PF-M | .03 | .25 | -.03 | -.14 | **.30** |
| | 16PF-N | -.23 | .10 | -.23 | -.01 | **-.39** |
| | FACTOR I | | | | | |
| | FACTOR II | .18 | | | | |
| | FACTOR III | .24 | -.04 | | | |
| | FACTOR IV | -.31 | .15 | -.13 | | |
| | FACTOR V | .04 | .19 | .17 | -.14 | |

## 11.5 Construct validity: Relationships between the ICES Plus Interest scales and Holland's Vocational Preference Inventory (VPI)

Holland's six occupational types (Realistic, Investigative, Artistic, Social, Enterprising and Conventional) are widely used in interest inventories both as the basis for their design (e.g. the Self-Directed Search and Vocational Preference Inventory) and for their interpretation (e.g. the Strong-Campbell Interest Inventory). To explore relationships between the two methods of classifying interests, a study was carried out in which people were asked to complete both the ICES Plus PDT inventory and Holland's VPI.

Unlike the PDT, which describes job-related activities, the VPI contains a list of job titles. From descriptions of the Holland scales and examination of the job titles associated with each scale a number of predictions can be made about the likely relationships between the PDT scales and the Holland types:

1. Interest in People should be most strongly related to the Social scale. However, it would also be expected to show relationships with the Enterprising and Artistic scales. It should either be unrelated or negatively related to the Realistic scale.
2. Interest in Data should be related to the Holland Investigative and Conventional scales.
3. Interest in Things should be related to the Holland Realistic and Investigative scales. It should be either unrelated or negatively related to the Social scale.

A study was carried out in the UK, in which 79 people completed the two inventories. Of these 59% were female and 41% male. About half (52%) were undergraduate students. The non-students were drawn from a wide range of ages and occupations. The average age of the sample was 25.8 years (SD=9.78) with a range from 15 to 59.

Descriptive statistics and inter-scale correlations for the two instruments are shown in Table 11.8. From this it is clear that the sample is somewhat biased in terms of their PDT inventory data. First, they have below average scores on the three scales. Second, they show an elevated level of correlation between the People and Data scales.

Anovas were carried out on each of the three PDT scales with age as a covariate and gender and occupation (student versus non-student) as independent variables. The covariate was non-significant in all cases and there were no interaction effects. There were only two main effects:

- For the People scale, the non-students (mean=3.08) had a significantly lower mean score than the students (mean=4.73) - F=9.75, df 1,74, p<.01; eta=0.39.
- For the Things scale, males (mean=5.63) had a significantly higher score than the females (mean=4.53) - F=6.18, df 1,74, p<.05; eta=0.27.

For the Holland scales, there were significant effects of gender on the Realistic (eta=0.40) and Social (eta=0.30) scales, with males scoring higher than females on Realistic and females higher than males on Social. In addition, the students scored lower than the non-students on Realistic (eta=0.28).

***Table 11.8 Results of the comparison between the PDT inventory and Holland's VPI. (n=79).***

| **Scale** | **Mean** | **Std Dev** | |
|---|---|---|---|
| ICES Plus Scales | | | |
| Stens | | | |
| PEOPLE | 3.94 | 2.27 | |
| DATA | 4.32 | 1.92 | |
| THINGS | 4.97 | 2.01 | |
| Raw scores: | | | |
| PEOPLE | 35.13 | 10.89 | |
| DATA | 26.05 | 7.33 | |
| THINGS | 29.48 | 10.04 | |
| Holland Scales: | | | |
| R | 3.25 | 3.07 | |
| I | 4.00 | 3.71 | |
| A | 5.59 | 3.87 | |
| S | 5.33 | 4.52 | |
| E | 3.81 | 3.21 | |
| C | 2.66 | 3.42 | |
| | **People** | **Data** | **Things** |
| PDT Scales: | | | |
| People | 1.00 | .49 ** | -.05 |
| Data | .49 ** | 1.00 | .07 |
| Things | -.05 | .07 | 1.00 |
| | | | |
| VPI Scales: | | | |
| R | -.09 | -.02 | .52 ** |
| I | .32 * | .28 * | .36 ** |
| A | .51 ** | .32 * | -.02 |
| S | .68 ** | .43 ** | -.26 |
| E | .33 * | .32 * | -.03 |
| C | .22 | .33 * | .00 |
| Significance levels: * p<.05, ** p<.01 | | | |

While the total sample size is rather small, the ratio of subjects to variables is sufficient to carry out a principal components analysis to examine the relationships between the scales. This revealed three main clusters of scales (accounting for 72.3% of the variance) - see Table 11.9. After Varimax rotation, both People and Data were linked to Artistic and Social, with Things being linked as predicted with Realistic and Investigative. Direct Oblimin rotation produced the same pattern of loadings.

***Table 11.9 Principal components loadings for the two sets of scales.***

| Rotated Factor Matrix | | | | |
|---|---|---|---|---|
| **Scale** | **Communality** | **I** | **II** | **III** |
| Things | .77 | -.05 | .86 | -.17 |
| R | .76 | -.13 | .82 | .26 |
| I | .71 | .38 | .67 | .34 |
| | | | | |
| People | .72 | .84 | -.02 | .12 |
| Data | .43 | .60 | .08 | .27 |
| A | .67 | .81 | .08 | .02 |
| S | .79 | .87 | -.19 | .21 |
| | | | | |
| E | .77 | .32 | .03 | .82 |
| C | .89 | .09 | .16 | .92 |

The results of this study were broadly in line with the prediction made about the relationships between these two instruments. There does, however, appear to be some bias in this sample as the level of correlation between People and Data was higher than one would expect from a random sample. This is likely to be a function of the high proportion of university students in the group.

# 12. JOB-RELATED VALIDITY: PHASES ONE AND TWO

## 12.1 Relationships Between ICES Plus Scale Scores and Job Groups as Defined by Dictionary of Occupational Titles (DOT) Codes

Using the US Department of Labor Dictionary of Occupational Titles (DOT) coding system, each person's job was assigned a DOT Classification. Two aspects of these classifications were analyzed. The Occupational Code Number (the first digit of the code) and the Worker Function Ratings (the middle three digits) which refer to the complexity of the job functions in relation to working with People, Data and Things. The DOT worker function ratings range from "0" for the most complex functions to 6, 8, or 7 (for People, Data and Things respectively) for the least complex.

Only levels 0/1 (Professional, Technical and Managerial occupations) and 2 (Clerical and Sales operations) of the Occupational Code Numbers were covered by the sample. For the Worker functions, the range of levels for People was well covered: from Taking Instructions (level 8) through to Negotiating (level 1). Only the more complex levels (1 to 4) were covered for Data functions and most of the sample was rated at the least complex level for Things (level 7: Handling). However, there was a reasonable sized group (n=369) whose jobs were rated "2" (Operating-Controlling) for Things.

***Table 12.1. Numbers of people in the combined sample, broken down by DOT classifications of their jobs.***

| Description | Level | Frequency | Percent |
|---|---|---|---|
| Occupational Code | | | |
| Professional, Technical, Managerial | 1 | 1297 | 39.1 |
| Clerical and Sales | 2 | 2023 | 60.9 |
| Worker function: Data | | | |
| Co-ordinating | 1 | 1399 | 42.1 |
| Analyzing | 2 | 382 | 11.5 |
| Compiling | 3 | 1493 | 45.0 |
| Computing | 4 | 46 | 1.4 |
| Worker function: People | | | |
| Negotiating | 1 | 459 | 13.8 |
| Instructing | 2 | 19 | .6 |
| Supervising | 3 | 255 | 7.7 |
| Persuading | 5 | 1685 | 50.7 |
| Speaking-signaling | 6 | 827 | 24.9 |
| Serving | 7 | 64 | 1.9 |
| Taking instructions | 8 | 12 | .4 |
| Worker function: Things | | | |
| Precision working | 1 | 4 | .1 |
| Operating-controlling | 2 | 369 | 11.1 |
| Handling | 7 | 2948 | 88.8 |

It should be noted that people were not sampled in any systematic way in order to cover all the DOT coding levels, nor can we be sure that there are not other factors confounded with membership of these groups (e.g. gender or ethnic group). However, it is possible to see whether the present data support hypotheses, which can be made about the relationships between interests and personality on the one hand, and aspects of the functional complexity of one's job on the other.

Given that DOT levels represent variations in the complexity of various aspects of work, we would not necessarily expect to find a simple linear relationship between Personality or Interest measures on the one hand and DOT levels on the other. For example, Computing (Data level 4) may involve a greater degree of interest in handling data than Co-ordinating (Data level 1) even though the latter is more complex. Similarly, an interest in working with "things" may not be reflected in the complexity of that work. There may be very little difference in the level of interest in working with things between people handling goods in a warehouse and doing fine precision engineering work. Nevertheless, a study of the patterns of relationships between ICES Plus Interest and Personality scales and people's DOT job coding should provide useful construct validation information.

***Table 12.2. Relationships between ICES Plus raw scale scores and categorizations of people in terms of DOT Occupational Code Number and Worker Functions (overall n=3321).***

| | | DOT groupings (number of groups) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Occupational Code | | | Worker Functions | | | | | | | |
| | | | | | PEOPLE (7) | | | DATA (4) | | | THINGS (3) | |
| | | r | F | | r | F | | r | F | | r | F | |
| ICES PLUS INTERESTS | | | | | | | | | | | | | |
| PEOPLE | eta | .04 | 5.6 | * | .13 | 7.0 | *** | .07 | 4.1 | ** | .14 | 28.2 | *** |
| | R(lin) | -.04 | 5.6 | * | -.10 | 24.6 | *** | -.04 | 4.4 | * | .14 | 56.4 | *** |
| DATA | eta | .01 | <1 | ns | .09 | 3.2 | ** | .08 | 6.3 | *** | .01 | <1 | ns |
| | R(lin) | .01 | <1 | ns | .05 | 7.2 | ** | .03 | 1.7 | ns | -.01 | <1 | ns |
| THINGS | eta | .00 | <1 | ns | .09 | 3.3 | ** | .07 | 4.2 | ** | .10 | 11.6 | *** |
| | R(lin) | .00 | <1 | ns | -.07 | 13.9 | *** | -.04 | 4.6 | * | .09 | 23.2 | *** |
| ICES PERSONALITY | | | | | | | | | | | | | |
| INDEP | eta | | | | | | | | | | | | |
| | R(lin) | .01 | <1 | ns | .26 | 38.5 | *** | .13 | 17.7 | *** | .22 | 83.7 | *** |
| | | .01 | <1 | ns | -.19 | 126.6 | *** | -.06 | 11.6 | *** | .22 | 167.3 | *** |
| CONSC | eta | .13 | 54.2 | *** | .17 | 16.7 | *** | .19 | 42.8 | *** | .05 | 4.2 | * |
| | R(lin) | .13 | 54.2 | *** | -.03 | 3.7 | ns | .16 | 92.1 | *** | .05 | 8.4 | ** |
| EXTRAV | eta | .07 | 14.3 | *** | .24 | 32.3 | *** | .16 | 28.2 | *** | .14 | 32.5 | *** |
| | R(lin) | .07 | 14.3 | *** | -.01 | <1 | ns | .12 | 49.9 | *** | .14 | 63.3 | *** |
| STABLE | eta | .06 | 12.2 | *** | .18 | 17.8 | *** | .08 | 6.6 | *** | .10 | 17.6 | *** |
| | R(lin) | .06 | 12.2 | *** | -.09 | 30.3 | *** | .06 | 13.5 | *** | .10 | 34.5 | *** |
| df for Within Group error varies from 2624 to 2643; df for R = 1<br>df for eta = number of groups-1; *** p<.001; ** p<.01, * p<.05. | | | | | | | | | | | | | |

Table 12.2 summarizes the results of a number of analyses of variance. The eta values (interclass correlations) indicate the strength of the relationship between the ICES Plus scales and the categorization of people by DOT codes. Where there are more than two groups, it is possible to ask whether ICES Plus scores increase or decrease systematically in relation to the DOT categorization of functional complexity. This is expressed by R(lin), the linear correlation between scale score and DOT classification.

As the DOT complexity levels are low for complex jobs and high for simple ones, this means that R(lin) will be positive if high ICES Plus scores are associated with the more simple functions and negative if high ICES Plus scores are associated with the more complex functions.

The first point to note is that there are, in general, a large number of relationships (between .15 and .25) between ICES Plus measures and DOT classifications. Second, for the most part, these relationships are readily interpreted. For example, there are relationships which one would expect between ICES Plus PDT (People, Data and Things) scales and the DOT job levels relating to each of those areas of work. These include the following:

- For the People Interest scale, the relationship is very clear: the higher the level of interest expressed in working with people, the lower the DOT level (i.e. the more complex the nature of the interaction with people). An interest in working with people is also related to DOT Things - though here the relationship is the other way round. A high interest in people is related to jobs that have a high DOT Things level. However, it must be noted that the bulk of the jobs in the present sample were at Things, level 7.
- An Interest in Data is related to both the DOT People and DOT Data classifications.
- Interest in Things is related to all the DOT classifications, but most strongly with DOT Things. Here the relationship is one where a high expressed interest in working with things is associated with the more simple levels of jobs involving working with things. Interpretation of this should bear in mind the caveat given earlier that the sample was not selected to represent all levels of the worker functions with all other factors controlled. In the present instance, we have really two groups. A managerial/clerical group for whom working with things is not important (hence the "7" coding for the majority of the sample) and a subset of the sample for whom working with things is a part of their job. Given this dichotomization, it is not surprising that there is not a clear relationship between ICES Plus Things and the DOT codings.

There are some clear relationships between the ICES major scales and the DOT classifications.

- People in Clerical and Sales job have significantly higher Conscientiousness, Extraversion and Stability scores than those in Professional and Technical jobs.
- High Independence is strongly related to the more complex DOT People function levels and the less complex DOT Things function levels.

- High Conscientiousness is associated with DOT Data function levels, with the more simple data-related work (computing) being associated with higher Conscientiousness (i.e. attention to detail) than the more complex levels (which focus more on uses of data and strategic concerns).
- For the DOT Things function, those in the "Handling" category are more Independent, Conscientious, Extravert and Stable than those in the "Operating-Controlling" category.

## 12.2 Predictions of DOT Occupational Code and DOT Worker Functions Using The ICES Plus Scales

Discriminant function analyses were carried out using the eight ICES minor scales (I1, I2, C1, C2, E1, E2, S1 and S2) and the ICES Plus Interest scales (People, Data and Things) as predictors and the categorizations into DOT-related groupings as the criteria.

### 12.2.1 Prediction of the Occupational Code

A significant canonical correlation of r=0.25 was found between the Discriminant function score, based on the 11 ICES Plus scales, and group membership (Professional, Technical and Managerial versus Clerical and Sales). Those in the Clerical/Sales group tended to have higher scores on E2, C1, I1, E1, S1, S2 (in order of importance, all significant $p<.05$). Overall, Clerical and Sales people appear to be rather more Extravert, more Stable and more Conscientious than those in the Professional/Technical/Managerial group.

### 12.2.2 Prediction of worker function groupings: DOT People groups

DOT People function levels 3 to 6 were included in the analysis. These cover: Supervise (Level 3, n=57); Negotiate (Level 4, n=372); Persuade (Level 5, n=1544) and Speak-signal (Level 6, n=494). The analysis produced three significant independent functions for discriminating between these four groups. These functions had canonical correlations of r=0.33, r=0.20 and r=0.09 respectively.

The main scales contributing to each Discriminant function were as follows:

Function 1: Scale E2, but not scale E1. Low scores on this were associated with "supervise", "persuade" and "negotiate" functions while high scores were associated with the "speaking-signaling" group.

Function 2: Scales I1, I2 and S1. High scorers are more likely to be jobs involving persuading, supervising or negotiating than those involving just speaking-signaling.

Function 3: The People Interests scale. This is the least clear Discriminant function, mainly discriminating between those in jobs labeled as "supervise" (low score) and those in jobs labeled as "negotiate". The other two worker functions scored in between these.

### 12.2.3 Prediction of worker function groupings: DOT Data groups

Four levels were examined: Co-ordinating (Level 1, n=809); Analyzing (Level 2, n=351); Compiling (Level 3, n=1327); Computing (Level 4, n=29). As for People, three significant functions were found. However, the canonical correlations were more uniform across the three functions in this case: r=0.28, r=0.23 and r=0.15 respectively.

The main scales contributing to each Discriminant function were as follows:

Function 1: Scales C1, S1 and E2. Lower scores on this function were associated with the "Computing", "Co-ordinating" functions while higher ones were associated with the "Analyzing" and "Compiling" worker function.

Function 2: Scale C2, and the Interests Data scale, with the following having negative loadings: E1, E2, I2, S2. High scores on this Discriminant function were associated with "Computing" and low scores with "Analyzing".

Function 3: Scales I2 and I1, with the Interest scales People, Things and Data (the latter having negative loading). Thus a high score on the Interests Data scale is associated more with "Computing" or "Compiling" functions than with "Co-ordinating" or "Analyzing". Low and high scores on this Discriminant function distinguished most strongly between the "Computing" and "Co-ordinating" worker function groups.

### 12.2.4 Prediction of worker function groupings: DOT Things groups

Just two levels were examined: Operating-Controlling (Level 2, n=271) versus Handling (Level 7, n=2244). A single Discriminant function, with a canonical correlation of r=0.28 (p<.001) was obtained. High scores on this Discriminant function indicated people who had higher scores on Things and People interests scales and higher scores on the ICES scales I1, C2, E2. They also tended to have lower scores on the Data Interest scale. Such people were more likely to be in the "Handling" group than in the "Operating-Controlling" group.

These analyses of combinations of ICES Plus scales as predictors of the various DOT-defined job groupings produced validities in the 0.20-0.30 region. This demonstrates the potential of the tests to discriminate between people in terms of criteria such as the DOT classification. Such Discriminant validity is very important where tests are to be used in classification or guiding placements in alternative jobs. For the most part, the scales involved and the patterns of loadings are readily interpretable in terms of the meanings of those scales.

## 12.3 Categorization in Terms of Job Types

Using the DOT code information and the job descriptions supplied by the participating organizations, a set of generic job-type groupings was formed. These, together with their shortened "code-names", are listed in Table 12.3.

***Table 12.3a. Job type groupings.***

| Codename | N | Description |
|---|---|---|
| ACCTNT | 110 | Accountant |
| ADMASS | 141 | Administrative assistant |
| ADMCLE | 195 | Administrative clerk |
| DPENG | 24 | Data Processing, Engineer/analyst |
| DPOP | 28 | Data Processing, Operator |
| MANTRN | 105 | Management Trainee |
| MANBRD | 500 | Manager, Branch/Department |
| MANRET | 90 | Manager, Retail Store |
| MANSAL | 52 | Manager, Sales |
| MANSEN | 131 | Manager, Senior level |
| SALFIN | 341 | Sales agent, Real estate/financial services |
| SALACC | 90 | Sales rep, account manager |
| SALCOM | 85 | Sales rep, commercial accounts |
| SALROU | 1129 | Sales rep, route |
| SUPADM | 235 | Supervisor, Administration |
| SUPHOU | 181 | Supervisor, hourly workers |

In addition to the job-type groups listed in Table 12.3a, three job-type groups were defined by employer.

***TABLE 12.3b. Employer based groups.***

| Codename | N | Description |
|---|---|---|
| 1xxxxx | 281 | An investment group |
| 2xxxxx | 181 | A Canadian government corporation |
| 3xxxxx | 47 | A Canadian bank |

Two distinct approaches were adopted to the development of measures, based on the ICES_Plus scale scores, which could be used to predict membership of these groups. The first is a profile matching procedure that produces a Percent Job-fit score. The second develops Adjustment Specification Equations, using Discriminant function analysis.

### 12.3.1 Percent Job-fit score rationale

This procedure defines upper and lower sten score limits for each job group. These limits are a function of the difference between the group's mean sten and the population mean, and take account of the group's SD. In addition, a scale weight is defined which is a function of the difference between the group's SD and the population SD. The more restricted the range for the group, the higher the weight.

Job-fit is assessed by comparing a person's sten scores, scale by scale, with the defined upper and lower limits for the reference job group. Where the score falls outside the limits, a deviance score is computed which is a function of the distance of the score from the limit. These individual deviance scores are multiplied by the scale weights and summed to produce the total weighted deviance score. The ratio of this value to the maximum possible total weighted deviance score is subtracted from one and expressed as a score out of 100 (hence the name “Percent Job-fit”).

This fit measure has the property of being non-linear, as the degree of fit reduces as scores depart in either direction from the reference group's mean. However, no account is taken of scale inter-correlations. The fit measure produced is a simple function of the sum of the weighted deviance scores.

## 12.4 Prediction of Job Group Membership Based on the Percent Job-Fit Formulae

Percent Job-fit scores were computed for each person in the sample in relation to each of the groups listed in Table 12.3a and 12.3b. Average Percent Job-fit scores for those who belong to the group (In-Group) and those who do not (Out-Group) are shown in Table 12.4. The F-ratio tests the significance of the difference between the mean Percent Job-fit scores for the two groupings. The eta coefficients are a measure of the magnitude of the relationship between Job-fit scores and membership of the In Group. As such, they can be interpreted as being analogous to validity coefficients.

In all but one instance, the relationship between Job-fit score and group membership is statistically significant. The validity of the Job-fit measure varies from group to group, averaging eta=0.10 for the ICES scales on their own and eta=0.13 for all the scales together. To some extent the magnitude of this type of Job-fit measure is a function of the differences between the various job groupings (the more similar they are, the smaller the eta values will be) and the extent to which a particular group's profile differs from that of the combined sample. For that reason, the eta values should not be considered as “absolute” values. Rather they show the way in which fit to some groups is more predictable than others.

***Table 12.4. Summary of analyses of Percent Job-fit scores***

| | **Based on the 8 ICES Minor scales only** | | | | | **Based on the 8 Minor scales, the 3 Interest scales and WWN** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Group** | **Out** | **In** | **eta** | **F** | | **Out** | **In** | **eta** | **F** | |
| 1xxxxx | 90.05 | 92.44 | 0.13 | 54.8 | *** | 82.51 | 87.18 | 0.18 | 88.4 | *** |
| 2xxxxx | 75.76 | 83.54 | 0.13 | 60.1 | *** | Insufficient data | | | | |
| 3xxxxx | 86.16 | 91.29 | 0.08 | 20.3 | *** | 74.53 | 84.45 | 0.11 | 33.6 | *** |
| | | | | | | | | | | |
| ACCTNT | 86.69 | 90.83 | 0.10 | 32.9 | *** | 76.56 | 84.95 | 0.14 | 47.7 | *** |
| ADMASS | 84.55 | 89.58 | 0.12 | 51.5 | *** | 77.96 | 84.73 | 0.14 | 50.5 | *** |
| ADMCLE | 80.38 | 88.60 | 1.18 | 112.4 | *** | 66.05 | 83.02 | 0.28 | 217.9 | *** |
| DPENG | 84.42 | 90.06 | 0.06 | 11.9 | *** | 74.22 | 84.23 | 0.10 | 22.3 | *** |
| DPOP | 85.00 | 91.76 | 0.10 | 32.2 | *** | Insufficient data | | | | |
| MANRET | 88.63 | 90.74 | 0.05 | 8.4 | ** | 67.20 | 79.30 | 1.10 | 25.3 | *** |
| MANSEN | 89.53 | 90.82 | 0.05 | 6.5 | * | 81.32 | 83.84 | 0.05 | 7.6 | * |
| SALFIN | 89.79 | 91.86 | 0.12 | 46.3 | *** | 82.74 | 86.50 | 0.16 | 66.5 | *** |
| SALACC | 88.09 | 90.53 | 0.06 | 11.3 | *** | 81.17 | 85.93 | 0.09 | 18.8 | *** |
| MANTRN | 95.43 | 96.31 | 0.07 | 17.0 | *** | 64.40 | 82.97 | 0.08 | 17.9 | *** |
| MANBRD | 90.11 | 91.27 | 0.08 | 18.6 | *** | 82.92 | 84.93 | 0.08 | 18.5 | *** |
| MANSAL | 91.35 | 93.14 | 0.04 | 6.2 | * | 76.48 | 86.58 | 0.10 | 23.7 | *** |
| SALCOM | 91.80 | 92.82 | 0.03 | 3.8 | NS | 81.22 | 86.83 | 0.10 | 21.1 | *** |
| SALROU | 86.30 | 89.23 | 0.20 | 134.7 | *** | 82.91 | 85.29 | 0.16 | 64.7 | *** |
| SUPADM | 85.15 | 89.43 | 0.13 | 52.8 | *** | 77.31 | 83.93 | 0.15 | 62.2 | *** |
| SUPHOU | 75.76 | 83.54 | 0.13 | 60.1 | *** | Insufficient data | | | | |

**In** = mean job fit for those in the relevant group.
**Out** = mean job fit for the rest of the sample.
In Group sizes are as shown in Table 12.3a, 12.3b with Out Groups containing the remainder of the sample.
**eta** = relationship between percent Job-fit score and membership of job group.
**F** = F ratio. Statistical significance of relationship: * p<.05, ** p<.01, ***p<.001

## 12.5 Criterion-Related Validity: The Prediction of Supervisor Ratings

Supervisor ratings were obtained for 3142 of the people from the Phase One and Two sample. These were a relatively crude criterion measure. Supervisors were simply asked to rate each person in terms of their overall performance in the job on a 5-point scale. The scale ranged from: 1 "Very Poor" through 2 "Poor", 3 "Reasonable" and 4 "Good" to 5 "Excellent". The distribution of these ratings for the whole sample is given in Appendix A**.** Over 80% of the ratings were either "3" or "4". The distribution indicated the usual problem with unanchored supervisor ratings: both a central tendency and leniency.

Table 12.5 shows the results of multiple regression analyses using the ICES Plus scales as predictors and the supervisor ratings as the dependent variable. In most cases the group sizes are too small, relative to the number of variables involved as predictors to produce statistically significant levels of prediction. However, the actual magnitude of the effects is encouraging. When the eight ICES minor scales are used on their own as predictors, the average multiple correlation with supervisor ratings is 0.32. For ICES Personality, Interests and Numerical Reasoning together, this increases to an average of r=0.41.

***Table 12.5: Means for the ASE Discriminant scores for each job group for the whole sample and for the In and Out Groups. ASE scores are based on Personality (ICES major scales) and the three Interest scales.***

| | **ICES only** | | | | **ICES Plus Interests and WWN** | | | |
|---|---|---|---|---|---|---|---|---|
| **Job-group** | **R** | **F** | **p** | **n** | **R** | **F** | **p** | **n** |
| 1xxxxx | .20 | 1.5 | ns | 279 | .26 | 1.6 | ns | 272 |
| 3xxxxx | .46 | 1.2 | ns | 47 | .58 | 1.4 | ns | 46 |
| ACCTNT | .41 | 2.5 | <.05 | 110 | .43 | 1.19 | ns | 76 |
| ADMASS | .28 | 1.31 | ns | 136 | .35 | 1.19 | ns | 115 |
| ADMCLE | .15 | <1 | ns | 193 | .22 | ,1 | ns | 167 |
| DPENG | .74 | 2.2 | ns | 23 | .84 | 2.1 | ns | 23 |
| DPOP | .43 | <1 | ns | 27 | Not available | | | |
| MANTRN | .19 | <1 | ns | 105 | Not available | | | |
| MANBRD | .26 | 4.3 | <.001 | 481 | .28 | 2.5 | <.01 | 375 |
| MANRET | .32 | 1.2 | ns | 90 | .48 | <1 | ns | 42 |
| MANSAL | .45 | 1.3 | ns | 51 | .65 | <1 | ns | 28 |
| MANSEN | .38 | 2.4 | <.05 | 124 | .50 | 2.4 | <.01 | 98 |
| SALFIN | .16 | 1.1 | ns | 338 | .20 | 1.1 | ns | 327 |
| SALACC | .29 | <1 | ns | 90 | .45 | 1.1 | ns | 65 |
| SALCOM | .32 | 1.0 | ns | 81 | .48 | 1.0 | ns | 54 |
| SALROU | .19 | 4.4 | <.001 | 986 | .20 | 3.1 | <.001 | 940 |
| SUPADM | .23 | 1.6 | ns | 225 | .30 | 1.4 | ns | 188 |
| SUPHOU | No ratings | | | | | | | |
| POCCAN | No ratings | | | | | | | |
| df = 8 and n-9 | | | df = 12 and n-13 | | | | | |

# 13. FURTHER JOB-RELATED VALIDITY STUDIES

## 13.1 Relationships Between ICES Plus Interests, Job Choice and Choice of Degree Subject

Interests affect people's choices: their choice of subject to study; their choice of work; their leisure activities. A study was carried out to test the hypothesis that people's choice of degree course at university and their subsequent job choice would be related to their profile of interests on the ICES Plus People, Data and Things scales. Specifically, three subject areas were chosen: Engineering; Mathematics and Accounting; and Social Science. It was predicted that those in Engineering would show a relatively strong interest in Things, while those in Mathematics and Accounting should show a stronger interest in Data. The People scale has a rather different status - as all types of work may involve working with people to a greater or lesser degree. However, it was expected that the Social Science students should be lower in their interests in both Data and Things and high in People.

Fifty-six final-year University of Hull students from the three subject areas (Maths/Accounts; Engineering and Social Science) were asked to complete the ICES Plus Interests Inventory. They were also asked to rate the degree to which they were interested in working with people, data or things, about their main hobby and what job they would like to go into.

The jobs and hobbies that they described were independently rated by each of five judges in terms of the degree to which they thought they involved an interest in working with People, Data and Things (using a 1 to 5 scale). The jobs and hobbies were assigned the average of the five judges' ratings (see Figure 13.2).

While the overall group contained equal numbers of males and females, there were biases within subject areas: more males in Engineering and more females in Social Science. These reflect the current biases in student intake to those areas (which are in themselves a reflection of gender-related differences in patterns of choice).

***Table 13.1. ICES Plus Interests study: gender by subject area.***

| | Maths/Accounts | Engineering | Social Science | Row |
|---|---|---|---|---|
| Male | 9 | 15 | 4 | 28 |
| Female | 11 | 4 | 3 | 28 |
| Column | 20 | 19 | 17 | 56 |
| Total | 35.7 | 33.9 | 0.4 | 100.0 |

Average sten scores for the three groups are shown in Figure 13.1. It can be seen that the hypotheses were supported. The Engineering students showed most interest in Things, less in Data and least in People, while the Maths and Accounting students showed most in Data, less in Things and least in People. The Social Science students had a generally low level of expressed interest. However, the highest scale was, as predicted, People.

Overall, the relationships between each scale and group membership were 0.30 for People, 0.52 for Data and 0.54 for Things. It was interesting to find that while there were marked differences between Departments, these differences were not strongly gender-related. With age controlled, analyses of variance showed that gender did not have a significant effect on any of the scales. However, for Things, there was a small but significant gender by subject area interaction ($F=6.9$, $df=6$ & $49$, $p<.05$). This arose from the fact that there was a difference between the sexes for the Social Science group but not for the other two groups.

***Figure 13.1: ICES Plus Interests University of Hull Study***

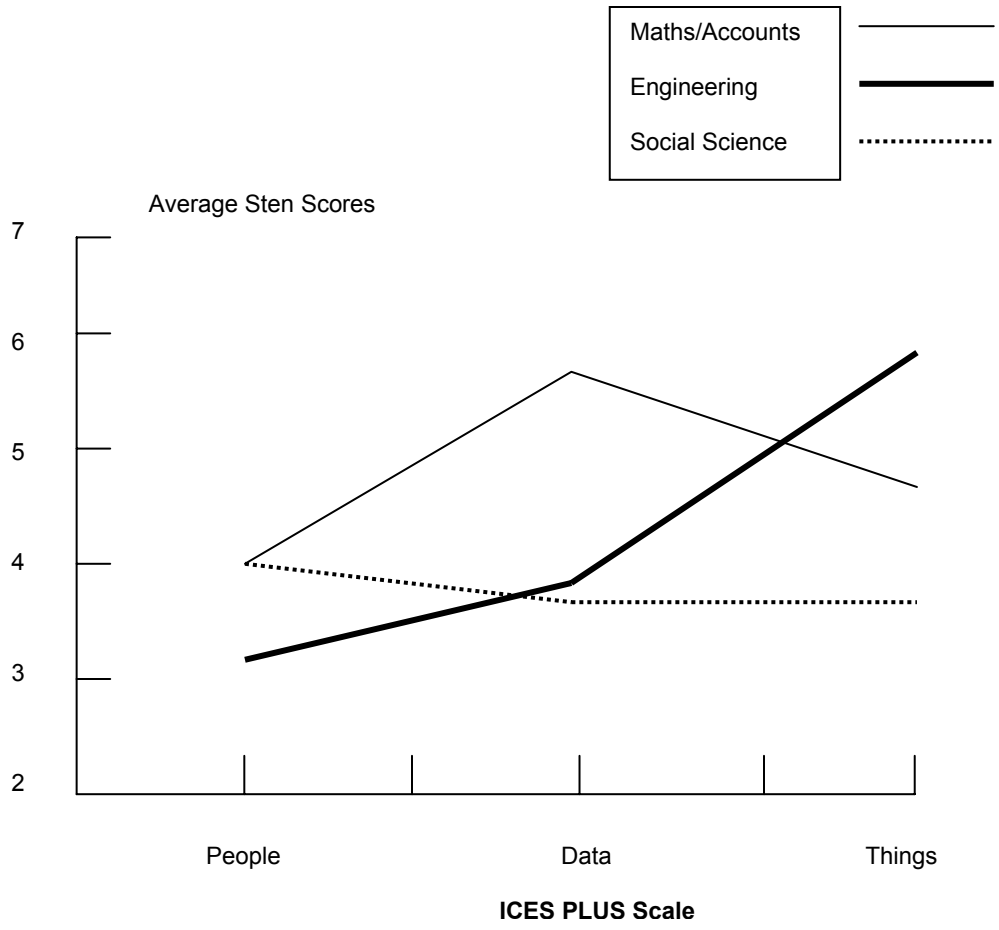


***Figure 13.2: ICES Plus Interests University of Hull Study, Preferred Hobbies and Occupations***

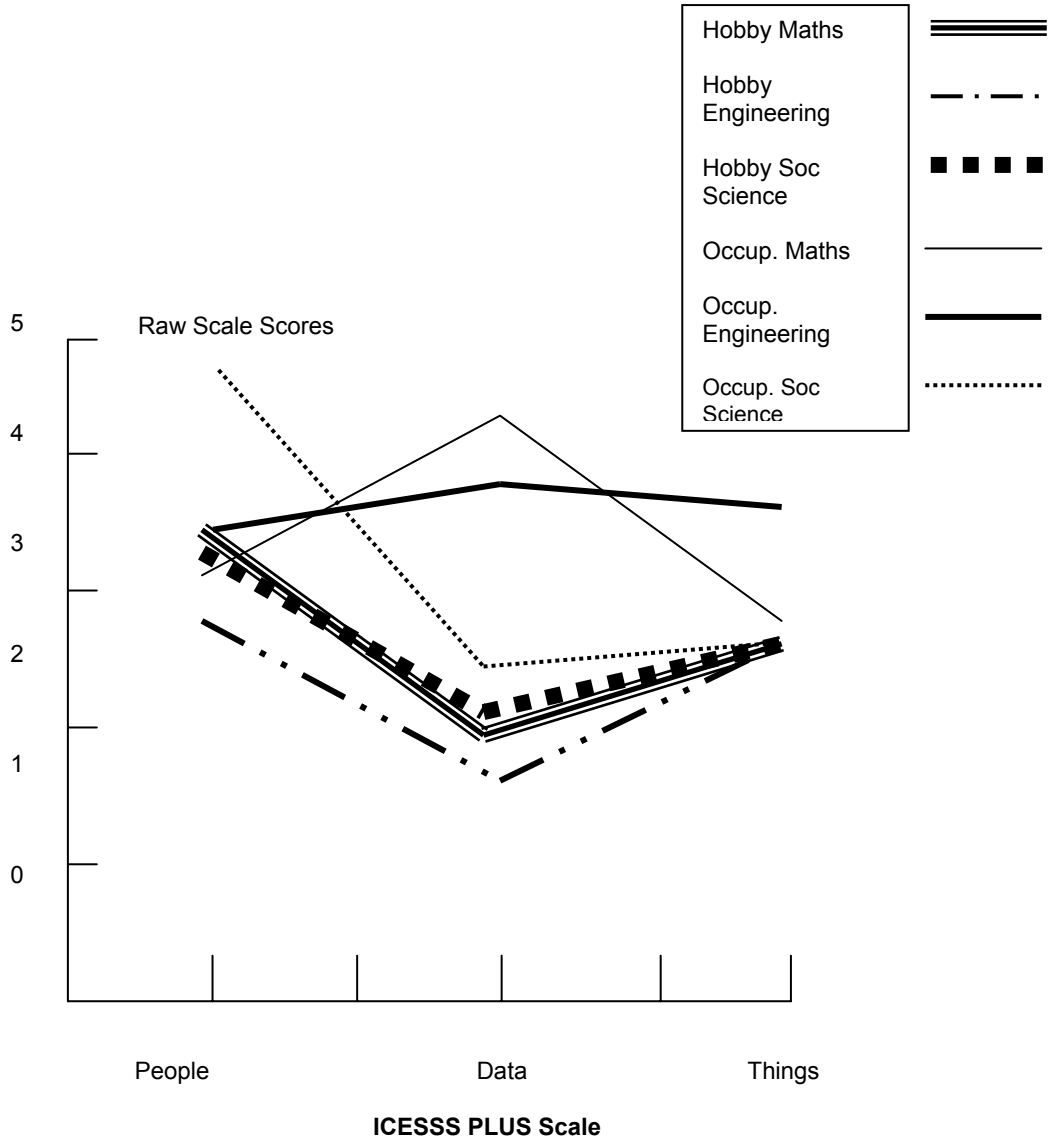



Figure 13.2 shows the average ratings assigned to the occupations and hobbies in which each group expressed interest. It must be noted that these are not normative scales and they are based on the judges' stereotypical notions about the jobs and hobbies that were mentioned by the people in the study. For occupations, this shows a very similar pattern to the PDT scale scores. The main differences are that the Engineering profile is "flatter" and the Social Science one is more differentiated.

The hobbies do not reflect any clear relationship to work-related interests. All three groups have very similar "preferred hobby" profiles, with hobbies involving other people being rated highest in all cases.

The ICES Plus scales Data and Things correlated strongly with the preferred occupations in terms of Data (r=0.37) and Things (r=0.53). The People scale was less able to discriminate (r=0.11). However, an interest in People had a strong negative correlation with occupations rated high on Data (r=-0.42).

Discriminant function analysis using the three ICES Plus scores as predictors of group membership (Engineering, Maths/Accounts or Social Science) produced two functions. Function scores had canonical correlations of 0.67 and 0.49 respectively with group membership. The main contributor to Function One was the Things scale while the main contributor to Function Two was the Data scale. The functions ordered the three groups as follows:

- Function One (from high to low): Engineering, Maths/Accounts, Social Science.
- Function Two (from high to low): Maths/Accounts, Social Science, Engineering.

Based on these function scores, predictions were made concerning which of the three groups each person should belong to. The results of these predictions are shown in Table 13.2. It can be seen that one can correctly predict the degree subject of 71.43% of the sample from their PDT scores (by chance we would expect to correctly predict only 33.3%).

***Table 13.2. Prediction of degree subject using ICES Plus PDT scales as predictors.***

| | | Predicted | | |
|---|---|---|---|---|
| **Actual Group** | **N** | **Maths/Acc** | **Engineer** | **Soc. Science** |
| Maths/Accounts | 20 | 15<br>75.0% | 3<br>15.0% | 2<br>10.0% |
| Engineering | 19 | 3<br>15.8% | 14<br>73.7% | 2<br>10.5% |
| Social Science | 17 | 3<br>17.6% | 3<br>17.6% | 11<br>64.7% |
| Percent of cases correctly classified: 71.43% | | | | |

In conclusion, this study provides very strong support for the PDT scales. They clearly differentiated in a predictable way between people working in different subject areas and showed systematic relationships to the occupations in which people expressed an interest.

## 13.2 Differences in Patterns of Interest and Personality Among Criterion Job Groups

Following completion of Phase One and Two, the ICES Plus computer-based scoring and interpretation system was modified to include provision for a structured criterion performance rating procedure. This provides the means whereby test scores and criterion data can be built up over time into a large database for future criterion-related validity analyses.

Data has so far been obtained from seven distinct groups (see Table 13.3) using this system. In each group, people have been rated on up to eight different criteria (defined by their employing organization). The rating procedure is constrained to provide an approximately normal distribution on a six-point scale. When mean criterion rating reliability is estimated, using the alpha internal consistency of the criterion measures for all seven groups, alphas of 0.90 or higher are obtained.

***Table 13.3. Post Phase Two criterion groups.***

| Job Title | Group | N | DOT code |
|---|---|---|---|
| SPECIALIST | 1 | 36 | 203.582-058 |
| TEST MARKETER | 2 | 124 | 251.257-014 |
| TRAVEL DIRECTOR | 3 | 59 | 238.167-014 |
| TELEMARKETING REP | 4 | 107 | 299.357-014 |
| PROCESS SPECIALIST | 5 | 136 | 262.357-010 |
| SALES REP | 6 | 209 | 250.257-010 |
| SENIOR/EXECUTIVE MANAGER | 7 | 200 | Not stated |

These data allow us to consider a number of questions. First, do the groups differ in terms of the patterns of Personality and Interest? Second, are differences in criterion ratings within each group related to differences in Personality and Interest?

### 13.2.1 Accounting for differences between groups in terms of Interests and Personality

Highly significant differences between groups are found on each of the four Personality major scales and the three Interest scales (see Table 13.4). Between group differences account for between 6.75% (STABLE: Eta=0.26) and 18.92% (INDEP: Eta=0.43) of scale score variance.

***Table 13.4: Differences between groups for each of the major Personality scales and the three Interest scales (mean sten scores).***

| | INDEP | CONSC | EXTRAV | STABLE | People | Data | Things | N |
|---|---|---|---|---|---|---|---|---|
| Average | 6.36 | 5.41 | 5.58 | 6.48 | 5.66 | 4.92 | 6.03 | 871 |
| Group 1 | 3.81 | 5.00 | 3.64 | 5.14 | 4.00 | 5.81 | 5.14 | 36 |
| Group 2 | 6.86 | 4.17 | 6.09 | 6.36 | 6.35 | 4.42 | 5.97 | 124 |
| Group 3 | 5.53 | 4.61 | 6.24 | 5.95 | 5.98 | 5.10 | 5.75 | 59 |
| Group 4 | 5.12 | 5.27 | 5.37 | 5.73 | 5.76 | 5.81 | 5.09 | 107 |
| Group 5 | 6.61 | 5.45 | 5.25 | 6.82 | 5.45 | 4.14 | 7.05 | 136 |
| Group 6 | 6.80 | 6.09 | 5.80 | 6.56 | 5.24 | 5.03 | 5.85 | 209 |
| Group 7 | 6.80 | 5.84 | 5.54 | 7.02 | 5.96 | 4.96 | 6.30 | 200 |
| eta | 0.43 | 0.33 | 0.27 | 0.26 | 0.28 | 0.26 | 0.30 | |
| % var | 18.92 | 10.74 | 7.21 | 6.75 | 7.83 | 6.94 | 9.00 | |
| F ratio | 33.60 | 17.33 | 11.18 | 10.43 | 12.24 | 10.75 | 14.25 | |
| df for F ratios = 6 and 864, p<.001 in all cases. | | | | | | | | |

Discriminant function analyses were used to examine further the differences between the seven groups in terms of the three Interest scales and four Personality scales. These analyses show how weighted combinations of scales scores can be used to predict group membership (see Table 13.5). When all seven scales are used, five Discriminant functions can be identified. When scores on these are used to assign people to groups, 41.22% of the sample are assigned to their correct groups. This is nearly three times as many people as would be correctly classified by chance.

Using either the three Interest scales or the four Personality scales on their own, reduces the correct classification rates to 32.15% and 30.08% respectively (in each case, these are still more than double the rates expected by chance). Two Discriminant functions were identified when just Interest scales are used, and three when Personality scales are used. These five functions together correspond very closely to the five identified when all seven scales are used together.

***Table 13.5. Prediction of group membership by Discriminant function analysis based on either interests, personality or both as predictors (n=871, groups=7). Coefficients shown are varimax rotated standardized Discriminant function coefficients. Canonical correlations (CanCor) are for unrotated functions. In all cases, the percentage of correct classifications expected by chance is 14.29%.***

| | DISCRIMINANT FUNCTIONS | | | | | |
|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **% correct classification** |
| ALL SCALES | | | | | | 41.22% |
| % Variance | 31.99 | 31.79 | 18.88 | 10.42 | 6.91 | |
| CanCor | 0.58 | 0.39 | 0.30 | 0.17 | 0.13 | |
| PEOPLE | -.07 | -.09 | -.25 | .99 * | -.13 | |
| DATA | -.74 * | -.14 | .17 | -.05 | -.08 | |
| THINGS | .92 * | -.09 | .13 | -.04 | .01 | |
| INDEP | -.03 | 1.00 * | .05 | -.02 | -.08 | |
| CONSC | .03 | .03 | .90 * | -.03 | .03 | |
| EXTRAV | .05 | -.20 | .18 | -.08 | 1.04 * | |
| STABLE | .23 | .22 | .23 | .19 | -.25 * | |
| INTERESTS: | | | | | | 32.15% |
| % Variance | 68.66 | | | 31.34 | | |
| CanCor | 0.44 | | | 0.29 | | |
| PEOPLE | -.02 | | | 1.03 | | |
| DATA | -.79 * | | | -.34 | | |
| THINGS | .90 * | | | -.17 | | |
| PERSONALITY: | | | | | | 30.08% |
| % Variance | 60.40 | 25.23 | | 14.37 | | |
| CanCor | 0.49 | 0.32 | | 0.21 | | |
| INDEP | 0.92 * | .05 | | -.01 | | |
| CONSC | -.01 | 1.01 * | | .01 | | |
| EXTRAV | -.19 | .07 | | 1.00 * | | |
| STABLE | .49 * | -.06 | | -.27 * | | |
| * p<.05 | | | | | | |

### 13.2.2 Relationships between mean criterion scores and Interests and Personality

The relationships between individual scales and job performance criteria are examined in detail in a later section. For comparison with the between group Discriminant analyses, multiple regression analyses were carried out to assess the degree to which mean criterion ratings within groups could be predicted from either Interest scales or Personality scales or both. In all cases, forced entry of all scales, rather than stepwise entry, was used. Table 13.6 shows that average levels of R=0.20 are obtained for Interests alone, R=0.25 for Personality alone and R=0.34 for both together.

***Table 13.6: Prediction of the mean criterion ratings using either interests (People, Data and Things), personality (INDEP, CONSC, EXTRAV and STABLE) or all seven scales as predictors. R=multiple correlation; F=F ratio.***

| | **INTERESTS** | | **PERSONALITY** | | **BOTH** | | |
|---|---|---|---|---|---|---|---|
| | **R** | **F** | **R** | **F** | **R** | **F** | **N** |
| Group 1 | 0.45 | 2.70 | 0.53 | 3.07 * | 0.57 | 1.88 | 36 |
| Group 2 | 0.09 | <1 | 0.24 | 1.89 | 0.27 | 1.30 | 124 |
| Group 3 | 0.19 | <1 | 0.30 | 1.37 | 0.43 | 1.62 | 59 |
| Group 4 | 0.18 | 1.21 | 0.28 | 2.13 | 0.31 | 1.53 | 107 |
| Group 5 | 0.26 | 3.15 * | 0.18 | 1.08 | 0.30 | 1.86 | 136 |
| Group 6 | 0.13 | 1.13 | 0.13 | <1 | 0.20 | 1.22 | 209 |
| Group 7 | 0.23 | 3.81 * | 0.12 | <1 | 0.28 | 2.28 * | 200 |
| N of Scales | 3 | | 4 | | 7 | | |
| * p<.05 | | | | | | | |

## 13.3 Validity Generalization Analysis of the Post Phase Two Criterion Groups and Phase Three Sample Job Group

A job performance rating was obtained for all those tested in Phase Three. As this was a single supervisor rating, its reliability and construct validity will be lower than the mean criterion ratings used in the previous analysis (see Table 13.6). Uncorrected correlations, ignoring differences in work groups, for the whole sample are given in Table 13.7. All the Ability scales have significant positive correlations with the mean rating, as does S2 (Relaxed). While we would

generally expect correlations with Ability scales to be positive for a job performance criterion, the same is not true for Personality and Interests. The direction of the relationship (positive or negative) will vary from job to job.

***Table 13.7. Phase Three: correlations between the Personality, Interest and Ability scales, and a job performance rating (n=516).***

| Scale | Related job performance | Scale | Related job performance |
|---|---|---|---|
| INDEP | .10 | PEOPLE | .07 |
| I1 | .08 | DATA | -.06 |
| I2 | .09 | THINGS | .04 |
| CONSC | -.06 | | |
| C1 | -.10 | WWN | .19 ** |
| C2 | -.01 | WWW | .18 ** |
| EXTRAV | .08 | WWS | .12 * |
| E1 | .07 | GENERAL | .20 ** |
| E2 | .08 | | |
| STABLE | .10 | | |
| S1 | .06 | | |
| S2 | .12 * | | |
| SocDes | -.05 | | |
| * p<.05<br>** p<.01 | | | |

The Phase Three sample contained 17 distinct job group sub-samples. Post hoc judgements were made for each of these on the likely direction of correlation expected for the four Personality scales and three Interest scales. Correlations were then classified as positive if they were in the expected direction and negative if in the non-expected direction. The same process was carried out for the seven groups discussed in the previous section. Thus a total of 26 distinct job group samples were available with data on the Personality scales, Interest scales and WWN, with 17 also having data on WWW and WWS. The validity generalization analysis procedures described in Hunter and Schmidt (1990) were applied to the sample of validity coefficients with corrections applied for attenuation (scale and criterion unreliability) and sampling error. The results of this analysis are summarized in Table 13.8.

In all cases, the estimated sampling error variance was equal to or greater than the actual variance in validity coefficients. Hence “worst case” (lower bound of the 90% credible interval) true correlations are equal to the mean corrected true validities shown in the Table.

***Table 13.8. Validity generalization analyses. $r_u$ is the mean uncorrected validity; $r_t$ is the mean corrected (true) validity; k is the number of correlations and n is the total sample size.***

| Scale | $r_u$ | $r_t$ | k | n |
|---|---|---|---|---|
| INDEP | .097 | .182 | 26 | 1367 |
| CONSC | .097 | .179 | 26 | 1367 |
| EXTRAV | .123 | .210 | 26 | 1367 |
| STABLE | .101 | .183 | 26 | 1367 |
| PEOPLE | .139 | .251 | 26 | 1367 |
| DATA | .119 | .214 | 26 | 1367 |
| THINGS | .135 | .222 | 26 | 1367 |
| WWN | .124 | .221 | 26 | 1367 |
| WWW | .191 | .369 | 17 | 496 |
| WWS | .156 | .275 | 17 | 496 |

These results show that when direction of prediction is taken into account, there is no additional situational specific variance associated with differences in validity coefficients for any of the scales - at least for the samples considered here. Furthermore, while the Ability scales have the highest individual levels of validity, both Personality and Interest scales show consistent patterns of prediction.

Following on from the studies described above (1994), further work has been carried out that further supports the evidence for the job-related validity of the ICES Plus scales. Some of this data is reported in Chapters 15 and 16. These further studies establish very clearly the construct validity of the ICES Plus scales and show how they can be used to predict job-related performance factors.

## 13.4 Job Description Survey Data Analysis

### 13.4.1 Introduction

As part of the development of a system which enables users to clearly specify the nature of the "benchmark" to which they are selecting people, a Job Description Survey (JDS) tool was developed. This asks the test user questions about the personal characteristics required for a job. Answers to these questions are used to guide the process of assessing the potential "Job-fit" of applicants, and can be used to provide a basis for generating hypothesis for testing criterion-related validity.

During 1997, the JDS was extensively revised and its links to the ICES Plus scales clarified. (Details of the revised JDS items are contained in Appendix K). As part of this evaluation and revision process, a study was carried out to examine expert judgements of the fit between items on the original ICES Plus Job Description Survey (JDS) and the scales on the ICES Plus inventory. The original JDS had 42 items that were designed to relate to 14 of the ICES Plus scales (three Abilities, three Interests and eight Personality scales). The JDS was designed as an easy-to-use job analysis procedure to enable managers to develop a job benchmark. These benchmark patterns allow the fit of job applicants to be assessed.

This study was undertaken, as it is important to ensure that the JDS items relate well to the ICES Plus scales.

### 13.4.2 Method

Seven subject-matter experts undertook the study. Expertise was designated on the basis of test user qualification level and psychological knowledge and experience (all the experts were experienced psychologists).

Each of the 42 JDS statements was printed on separate cards. In addition, short definitions of each of the scales from the ICES Plus Inventory were written. Scale descriptions were devised for both high and low scores and printed on a piece of card (see Table 13.9).

To control for any possibility of order effects, all of the 42 JDS cards were shuffled for each person and the scales were placed in a different order for each person (although they were allowed to move them around if they wished to).

Each expert was asked to place each of the 42 statements from the JDS underneath the scale description they thought it was most strongly linked to (either the high or low end). They were asked to do this quickly at first so that initial judgements were reached. After all the cards had been placed, the experts were allowed to examine their choices and to change them if they wished.

Finally, the experts were asked to write one or two more features of a job similar to those from the JDS that they felt would fit into the high and low ends of each scale (Table 13.10). For analysis purposes these items were adapted in style, although not content, so as to relate to features in a job.

***Table 13.9 – Descriptions ICES Plus scales***

| WORKING WITH NUMBERS | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Likely to be quicker and more accurate than over 80% of their norm group in reasoning with information derived from simple numbers. | Likely to take longer and be less accurate than other people in dealing with information derived from numbers. |

| WORKING WITH WORDS | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Unlikely to have any problems related to the use and understanding of written language, should find it easy to follow instructions, etc. | Tend to be less accurate and take longer than the majority of working adults in dealing with written information. Difficulty understanding written instructions, technical manuals, etc. |

| WORKING WITH SHAPES | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Likely to be quicker and more accurate than most people in dealing with information that involves mentally manipulating shapes and objects in space. Find it easy to work with plans and diagrams. | Likely to take longer and be less accurate than other people in dealing with information that involves mentally manipulating shapes and objects in space. Difficulty in relating figures (plans, diagrams) to actual operations. |

| INTEREST IN PEOPLE | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Interested in work that involves a lot of contact with people, including aspects such as negotiating or persuading others. | Content to work in a job with little or no contact with people. Do not necessarily avoid contact, it just does not play a major work role. |

| INTEREST IN DATA | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Interested in working with data, figures, symbols, statistics, accounts and language. Likely to enjoy working with information systems, technical documents, contracts, etc. | Tend to avoid jobs that involve dealing with figures, statistics, accounts, etc. |

| INTEREST IN THINGS | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Interested in work that deals with inanimate objects such as machinery, tools, equipment. Likely to be interested in engineering work, handling goods and warehouse work. | Tends to avoid work that involves dealing with machinery, computers, etc. |

| COMPETITIVE (I1) | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Single-minded, competitive and play to win. Strive hard to reach personal goals. Relatively little concern if other people get upset along the way. | Co-operative and non-competitive people who are team players and enjoy co-operative ventures. Unlikely to be concerned about winning or losing. |

| ASSERTIVE (I2) | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Rational, assertive and outspoken people. Often become group leaders and can be controversial, unafraid of arguments and ensure their opinions are known. | Valued for diplomacy and tact; can play a peacemaker role. Tend to be submissive, non-controversial and avoid conflict. |

| CONVENTIONAL (C1) | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Conventional, traditional, rule-bound and reliable. Best working in highly structured, clear and unambiguous environments. May find it difficult to adapt to new situations. | Innovative, flexible, with a casual attitude to guidelines, rules and regulations. Likely to seek new ways to solve problems rather than traditional ones. Work well in changing and challenging environments. |

| ORGANIZED (C2) | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Orderly, meticulous and plans ahead, thinking through all aspects before acting. Tend not to like thinking on their feet. Dependable and predictable. | Creative, spontaneous people who prefer to react to events. Focus on the overall picture rather than fine details. Do not worry about attention to detail. |

| GROUP-ORIENTED (E1) | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Strong need for others and their approval and support. Happy working in environment with a reasonable amount of contact with others. Like to be part of a group but not necessarily the leader. | Happy to work on their own and avoid noisy situations and group activities. Feel more at ease with own company and thoughts and controlling the amount of stimulation that reaches them. |

| OUTGOING (E2) | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Outgoing and talkative. Enjoy risky, action-packed lives. Tend to be impulsive and like doing exciting and stimulating things. Routine may become boring to them. | Quiet and reserved, preferring to live a quiet, orderly life. Do not like being the center of attention. |

| POISED (S1) | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Able to cope with most situations in life without getting upset or irritated. Rational approach to life which enables them to shrug off criticism and cope with adversity. | Irritable and easily upset, often losing their temper. Find it hard to cope with embarrassing situations and have difficulty coping with setbacks and personal criticism. |

| RELAXED (S2) | |
|---|---|
| **HIGH SCORE** | **LOW SCORE** |
| Very relaxed, untroubled and able to cope with life's pressure. Accept people at face value and leave job worries behind. Cope well with demanding high-pressure jobs but can be exploited. | Excitable, anxious people who are wary of others. Find high levels of pressure difficult to cope with and can be suspicious of others they do not know well. |

### 13.4.3 Interpretation of item-scale links

There were two criteria for assigning an item to a scale:

1. If more than 50% (4 or more) of the subject-matter experts judged an item to be in the scale then it was accepted as being in the scale identified.
2. If 3 or less experts placed an item in the scale then this particular item was considered not linked to the scale.

Those items determined to be unrelated to a particular scale were recommended for omission. Replacement items were obtained from experts.

Using the above rules there was agreement between the JDS scoring key and the experts on 26 out of the 42 items (62% agreement rate). Out of the 16 judgements that did not agree with the scoring key, eight items could not be placed into any scale as rule 2 was broken. The remaining eight items were found to fit better to scales other than those to which they had originally been assigned.

Overall, there was a fairly high level of agreement between the experts and the JDS scoring key on the relationships between items and scales. The main area of confusion arose around the distinction between Group-Oriented (E1), Outgoing (E2) and Interest in People. Out of the 16 disagreements between the experts and the scoring key outlined above, 8 were due to differences of opinion over items and their relationship to E1, E2 and People. The second area of confusion was between items designed to relate to Numerical Reasoning, Interests in Data and Conscientiousness.

### 13.4.4 Changes made to the JDS

Based on the analysis of the expert sorting, and the new items generated during the process, it was decided to increase the number of items for the Personality scales, and to divide the new JDS into three discrete sections: Abilities, Interest, and Personality. This will help the user to focus on the relative type of attribute. This in turn should reduce the confusion occurring between Interests in People and Extraversion, and between Numerical Reasoning, Interest in Data and Conscientiousness.

Within each of the three sections, items are randomly ordered.

**Ability Scales (four items per scale):**

1. Working with Numbers (WWN): Two (#1, #4) of the original three items are retained and a third item (#7) is moved from another scale. The original item #7 is discarded and a new item added.
2. Working with Words (WWW): All three of the original items are retained, and a fourth new item added.
3. Working with Shapes (WWS): Two items are retained (#3, #5) and two new ones added.

**Interest Scales (three items per scale):**

1. Interest in People: Item #14 is retained and two new items are added.
2. Interest in Data: Item 13 is retained and two additional items are added.
3. Interest in Things: All three of the original items are retained.

**Personality Scales (four items per scale):**

1. Competitive (I1): The original three items are retained and a new one added.
2. Assertive (I2): Two of the original three items are retained together with two new ones.
3. Conventional (C1): This is now composed of items #29, #42, #44, and #48.
4. Organized (C2): Two items are retained, one is moved to this scale (#50) and a new one is added.
5. Group-Oriented (E1): Item #53 is transferred from another scale. Two new items are added
6. Outgoing (E2): Item #40 is used together with three new items.
7. Poised (S1): Item #34 is retained and joined by three new items.
8. Relaxed (S2): The original items are retained with one new item added.

The full set of 53 items for the New JDS, some of which have minor amendments from the first version, are listed in the Appendix K.

# 14 . INTERPRETATION OF THE ICES PLUS SCALES

## 14.1 Introduction

A detailed guide to interpretation of the ICES Plus scales is contained in Appendix G. This section discusses interpretation of the ICES scales. In relation to that, it also discusses interpretation of the ICES SocDes scale and the uses of ICES Plus for assessing “integrity”. The procedure followed to develop text for the computer-based narrative report is described and rules for identifying “valid” responses on the ICES inventory are presented.

Table 14.1 contains description of all the ICES scales. Note how the minor scales represent differentiation of aspects, or facets, of the major scale. Necessarily there is considerable overlap between minor scales relating to the same major. However, there are also differences of interpretation. For example, E1 focuses on the need for other people while E2 is more concerned with being outgoing and impulsive. Often the two go together, but it is quite possible to have someone who is very outgoing and yet is quite self-sufficient (High E2 with Low E1). Similarly, C1 and C2, while correlated, reflect very different facets of personality. One could be very attentive to detail (High C2) without necessarily being very traditional in ones views (High C1).

***Table 14.1. Description of the ICES Scales.***

| | |
|---|---|
| **Independence:** | **Competitive vs. Co-operative** |
| *High scores:* | Competitive, active, physical, hard-headed, skeptical, proud, plays to win, rational, assertive, forthright and socially bold. |
| *Low scores:* | Co-operative, diplomatic, good-natured, compassionate, non-assertive and retiring, sensitive to the needs of others. |
| **Independence: I1** | |
| *High scores:* | Competitive, physical, plays to win. |
| *Low scores:* | Co-operative, sensitive to the needs of others. |
| **Independence: I2** | |
| *High scores:* | Assertive, rational, forthright and socially bold. |
| *Low scores:* | Submissive, diplomatic, passive, non-assertive and retiring. |
| **Conscientiousness:** | **Conscientious vs. Flexible** |
| *High scores:* | Conscientious, well-organized, traditional, concerned with rules and high standards, detail-conscious, forward planning, neat. |
| *Low scores:* | Flexible, easy-going, expedient, innovative and radical, responsive, spontaneous, reactive, concerned with the overall picture rather than the details. |
| **Conscientiousness: C1** | |
| *High scores:* | Conventional, rule bound, traditional and concerned with moral values, practical |
| *Low scores:* | Innovative and radical, expedient, open to new experiences |
| **Conscientiousness: C2** | |
| *High scores:* | Organized**,** detail conscious, forward planning, neat |
| *Low scores:* | Reactive, responsive, spontaneous, and concerned with the overall picture |

***Table 14.1. Description of the ICES Scales (continued).***

| | |
|---|---|
| **Extraversion:** | **Sociable vs. Self-sufficient** |
| *High scores:* | Sociable, group-oriented, energetic, seeks out people and busy places, outgoing, talkative, sensation seeking, enjoys being the center of attention. |
| *Low scores:* | Self-sufficient, prefers one's own company, quiet, reserved, mild-mannered, prefers to stay in the background. |
| **Extraversion: E1** | |
| *High scores:* | Group-oriented, sociable, seeks out people and busy places |
| *Low scores:* | Self-sufficient, those who like their own company and prefer quiet places |
| **Extraversion: E2** | |
| *High scores:* | Outgoing, talkative, impulsive, sensation-seeking and enjoys being the center of attention. |
| *Low scores:* | Reserved, quiet, mild-mannered and avoids risk. |
| **Stability:** | **Relaxed vs. Emotional** |
| *High scores:* | Relaxed, calm, unruffled, not easily worried by people or adverse events, accepting of people and able to leave their worries behind. |
| *Low scores:* | Emotional, sensitive, easily upset, irritable, anxious and suspicious of other people. |
| **Stability: S1** | |
| *High scores:* | Poised, calm, unruffled, not easily worried by people or adverse events |
| *Low scores:* | Restless, sensitive, easily upset and irritable. |
| Stability: S2 | |
| *High scores:* | Relaxed, accepting of people and able to leave their worries behind. |
| *Low scores:* | Excitable, emotional, anxious and suspicious. |

## 14.2 Using the ICES Scales as an Indicator of Integrity

Measures of integrity can be divided into two main groups: those that overtly assess "honesty" and those whose purpose is disguised from the person completing the questionnaire. Guion (1965) described these as "clear-purpose" as opposed to "disguised-purpose" instruments. More recently Sackett et al (1989) have made a similar distinction between overt and personality-based measures respectively. Overt measures of integrity include the Reid Report, the Personnel Selection Inventory and Stanton Survey. Personality-based measures include the Personnel Reaction Blank, the Personnel Decisions Inc. Employment Inventory and the Employee Productivity Index. The latter are not claimed to measure attitudes to theft and deviant acts directly. However, they were designed to predict a range of counter-productive behaviors including what is referred to as "property deviance" and "production deviance" (absenteeism, taking over-long work breaks, etc.).

While integrity testing is somewhat controversial, recent reviews of research have been generally positive, suggesting validities typically in the region of 0.20 to 0.40. In particular, Ones et al (1992) reported a meta analysis of 25 different instruments, a total sample size of over 200,000 people and over 300 validity coefficients which showed that validity generalization could be established for all integrity tests. Test validities were positive and sufficient in magnitude for the purpose of predicting counter-productive behavior across a range of settings and situations.

An important study has been reported by Woolley and Hakstian (1993). They noted that what seems to be emerging from this research is a consensus on the importance of a general construct - variously labeled Integrity, Conscientiousness, Reliability, Delinquency or Responsibility - as an indicator of counter-productive behavior in employees. Woolley and Hakstian carried out a study comparing overt measures with personality-based measures of integrity and with personality scales from three well known personality inventories: the California Personality Inventory, The Cattell 16PF and the NEO Personality Inventory. The latter (Costa & McCrae, 1985; 1989) is designed to explicitly provide measures of the "Big Five" personality factors.

Woolley and Hakstian's criterion measure was self-reported deviance (adapted from Hollinger and Clark, 1983). Research has shown that these self-report measures have true correlations of over 0.80 with external "objective" measures of deviance (Viswesvaran, 1992). Their results showed that the personality inventories were as good predictors as the personality-based integrity tests. Somewhat higher validities were obtained for the overt integrity measures, but this could be due to capitalization on method variance, as both these tests and the criterion involved self-reports of deviant behavior.

In relation to the 16PF scales, negative correlations were found for G (moral integrity, conscientiousness) and Q3 (conscientiousness and regard for social and self-reputation). For the NEO, Conscientiousness and Agreeableness were both negatively correlated with the criteria. For the CPI, negative correlations were found for all four predicted scales, Socialization, Tolerance, Self-Control and Responsibility, especially the last two.

The results of this work, together with the evidence supporting the construct validity of the ICES scales as measures of the "Big Five", provide a means of making very strong predictions about the use of ICES Plus as an indicator of counter-productive behavior. Low scores on Conscientiousness - especially C1 - and high scores on Independence - especially I1, can be expected to correlate with counter-productive indicators. Indeed, a look at Table 11.4 will show how the ICES SocDes scales has a pattern of relationships with the 16PF very close to that described by Woolley and Hakstian for an integrity measure.

Thus, very low scores on SocDes may indicate two possibilities. First, the person is "faking bad", or trying to create a bad impression. Second, they may be more prone to counter-productive behavior than others. "Faking bad", in an occupational assessment context is unusual. Thus, while high SocDes scores may be more likely to indicate faking good than "true good" low scores are more likely to indicate a lack of integrity than "faking bad".

More research is needed to develop an ICES Plus integrity scale measure that is criterion-referenced. However, it is clear that Conscientiousness will be a major component of that scale.

## 14.3 Rules for Identifying "Valid" Responses on the ICES Inventory

As part of the process of developing a computer-based interpretation system, it was necessary to formulate rules for identifying patterns of responses that might be "invalid". That is, ones where the candidate may have responded at random, or followed some systematic but non-meaningful response pattern.

Analyses of the distributions of "a", "b", "c" and omitted responses for the 96 ICES minor scale items were carried out to identify cut-off points for discarding questionnaires which had not been completed fully or which might not have been taken seriously. Full details of the distributions, based on the Phase Two Stage One sample data, are presented in Appendix F. On the basis of this information, the following rules were derived:

- Omissions**:** If a person omits an item in a scale, it is coded as a "b" (in-between) response for scoring purposes. If they omit more than 3 items on any one scale, the questionnaire is rejected and not scored. If there are any omitted responses, a message is printed giving a warning that this could affect the validity of the report - by making the person appear more "average" than they are.
- In between – "b" – responses: **-** If someone makes 36 or more "b" (in-between) responses to the ICES scale items (not the SocDes items), the questionnaire is rejected and not scored. If they are in the sten 9-10 region (between 17 and 35 b's) a warning is given that this could affect the validity of the report by making the person appear more "average" than they are.
- **"**a" and "c" responses: If there is a pattern of "a" and "c" responses that meets certain criteria, then a warning is printed saying that the distribution of responses is very unusual and the scores might be distorted by this. The accuracy would need to be checked through feedback with the candidate.

Apart from these response factors, the SocDes scale score is used to draw the user's attention to possible distortion effects. Sten scores of 8 or more on this scale are taken to indicate possible "faking good". However, as will all scales of this sort, high scores are always open to two interpretations: first that the person is "faking good"; second that they are "good".

## 14.4 Development of the Text for the Narrative Report

Development of the text units for the report involved the writing of two types of material. For each part of the report, sten-related descriptions were produced for each scale (Ability, Interests and Personality). For ICES, these were based on the descriptive adjectives in Table 14.1. No attempt was made in drafting this material to consider scale combinations - each scale was treated independently of the others.

The second type of text unit referred to scale combinations. These were written for ICES, and for Interests. For the present, no attempt was made to produce text units that represented overall battery interpretations in terms of combinations of ICES Personality, Interests and Ability.

In addition to the text units described above, descriptions were produced to accompany the SocDes scale.

The medium term strategy for the development of the narrative report involves the following stages:

1. Initial drafting of text units.
2. Refinement of test units and tailoring of the language to meet user requirements.
3. Validation of the reports.

Validating reports poses a number of complex methodological problems. Work on this has been started and the results of one study are reported below. In due course, Stage Three will involve a number of studies, including:

- Content analysis (to examine the accuracy with which text units can be sorted into their scale and score bands).
- Assessment of "Barnum" effects. These studies will assess the degree to which the reports discriminate between people in a valid fashion and are recognized as being more accurate than random or systematically distorted reports.
- Studies of the reliability and validity of decisions based on the narrative reports with users of differing levels of test expertise.

#### 14.4.1 Sten-related descriptions

For the ICES major Scales**,** ICES minor Scales, Ability and Interests, text units were produced for each scale. The relationship between each text unit and sten scores was:

| Text Unit | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Stens: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Percentages of population: | 2.5% | 4.5% | 9% | 15% | 19% | 19% | 15% | 9% | 4.5% | 2.5% |

#### 14.4.2 ICES major scale code patterns

In addition to the simple sten-related description, composite "code-pattern" descriptions were produced. These consisted of a single text unit that described a particular configuration of the four major scale scores. Eighty-one code patterns were defined as follows. Each scale was divided into Low (stens 1-4), Medium (stens 5 and 6) and High (stens 7-10). This distributes people in roughly equal proportions (31%, 38% and 31%) between the three categories. There are 81 different combinations of these three levels across the four scales. Frequencies of occurrence of each code pattern within the combined sample are presented in Appendix F.

The expected frequency of occurrence of each code pattern if they were equally distributed would be 1.23%. In fact the most common code pattern (40 = Medium on all four scales) occurs 4.4% of the time. The least common code patterns are 25 (High C and High E with Low I and Low S) and 57 (High I, High S with Low C and Low E), with each occurring only 0.2% of the time.

#### 14.4.3 ICES minor scale code patterns

For each major scale, text units were prepared describing each of the nine possible combinations of minor scales (Low-Low, Low-Medium, etc. to High-High). Again, frequencies of occurrence of each pattern are given in Appendix F. Expected frequencies of occurrence with an even distribution are 11.11% for each code pattern. In practice, frequencies vary from a minimum of 1.3% (High S1 with Low S2) to a maximum of 25.5% (Medium S1 with Medium S2).

#### 14.4.4 ICES Plus Interests code patterns

Interests code patterns for the three scales (People, Data and Things) were produced in the same way. Using the same sten score cut points, twenty-seven code patterns were defined from the Low/Medium/High combinations of the three scales. The expected frequency of occurrence of each code pattern with an even distribution is 3.7%. In practice, frequencies varied from 0.7% (Low People, High Data and Low Things) to 27.3% (Medium on all three scales).

### 14.5 Validating the ICES Reports

Computer-generated personality reports need to have two key qualities. First, what they say needs to be valid. Second, the content needs to reliably discriminate between people. Consider the following two statements:

[a] “There are times when other people are likely to upset you”.
[b] “You often get very upset by the behavior of others”.

The first is likely to be correct for most people, while the second will be true for a much smaller number. Reports which contain a large number of statements of the first type have high base rate validity - that is they are likely to be endorsed as being “true” by large numbers of people. Such reports are said to be subject to the “Barnum Effect”. This effect is responsible for the fact that people can be readily convinced about the veracity of reports based on random content, astrology or graphology. While people may judge such statements as being accurate, the statements are very poor at distinguishing in a reliable way between people. As one of the main purposes of assessment is to determine the ways in which people differ from each other, reports need to include sufficient accurate statements with a low base rate validity.

One method of assessing the accuracy of this aspect of computer-generated reports is to assess the degree to which a person is able to recognize a report based on their own scores as being “like me” while judging those based on other patterns of scores as being “not like me”. If all reports contain high base rate validity statements, they will all be judged as more or less “like me”. On the other hand, if they contain more discriminating statements, but are inaccurate, they will be judged as “not like me”. Thus the key to establishing the validity of a report is to show that a report based on actual measures of a person's personality is judged as valid while one based on other (e.g. randomly selected) measures is not.

A study was carried out by the author and a colleague (Bartram & Brennan, 1993, unpublished) to evaluate the validity of the part of the ICES Plus report that deals with the ICES scales. The output used in the study included the graph showing the ICES sten profile and the text units describing each of the relevant sten scores, assembled as a single piece of continuous text.

Thirty-five students (15 male and 20 female), of whom 20 were undergraduates and 15 were postgraduates, each completed the ICES inventory and were asked to self-rate themselves on the ICES scales. For the latter task they were given short scale descriptions and a simple explanation of the sten scale intervals. Some two weeks after this, they returned for the second part of the study in which they were given five graphs and five sets of text. These consisted of one graph and one piece of text for each of five conditions:

- Self-Rated (SR): based on their self-rated sten scores
- Observed (OB): based on their actual sten scores from the inventory
- False Average (FA): based on an “average profile” containing a mixture of sten 5 and sten 6 scores
- False Extreme (FE): based on an “extreme” profile drawn from the ICES database which had sten scores of 1, 2, 9 and 10
- False Mirror (FM): based on a profile that was the mirror-image of the FE one - that is, each sten score of 1 became 10, 2, became 9, 9 became 2 and 10 became 1

The Graphs and Text passages were labeled A through E and no indication was given to the subjects in the study of their relationship to the conditions. Subjects were asked to rate each of the ten items in terms of how much it was like them, using a seven-point scale:

1 “Not at all like me”
2 “Hardly any of it is like me”
3 “Less than half of it is like me”
4 “About half of it is like me”
5 “More than half of it is like me”
6 “Nearly all of it is like me”
7 “All of it is like me”

After this, people were given the related pairs of graphs and text and asked to rank order the five “reports” from 1 (“most like me”) to 5 (“least like me”).

It was hypothesized that people might identify their self-rated profile as being most accurate - as they would be most likely to “recognize” that as the one they had produced. The graphs and text based on the inventory scores should closely match the self-rating ones, while the three “false” conditions should be rated much lower.

A prediction was also made concerning differences between the “false” conditions. The text based on average sten scores tends to contain statements of higher base rate validity than text based on more extreme scores - this is necessarily the case as, by definition, more people have “average” scores than either above or below average ones.

Hence one would expect people to rate the FA condition higher than the FE or FM ones, at least in so far as the text is concerned.

In terms of overall rank order, then, the prediction was that SR and OB should be rated high, while FA and the two extreme false profiles should be rated low.

Distributions of ratings for each of the conditions for the 35 subjects are shown in Table 14.2, together with mean ratings for each of the conditions. Analysis of variance of the rating data showed that:

- The "True" conditions (mean=5.1) - SR and OB - were rated as more accurate than the "false" conditions (mean=3.24) - FA, FE and FM - [F=132.87, df 1 & 136, <.001].
- The difference between the two "true" conditions (SR and OB) is not significant [F=1.96], nor are the differences between the three "false" ones (FA, FE and FM) [F=1.16, ns].
- Overall, Text (mean=4.10) was rated as more accurate than the Graphs (mean=3.86) - [F=4.8, df 1 & 34, p<.05].
- The difference between the two true conditions (SR and OB) is greater for the graphs than for text [F=5.01, df=1 & 136, p<.01]
- For the "false" conditions, there is a difference between the Graph and Text difference for condition FA but not for the other two conditions (FE and FM) - [F=6.02, df=2 & 136, p<.05].

Overall, the Text versions were rated as easier to understand than the Graphs by 71.4% of the sample, while the Graphs were rated as easier to understand than the Text by only 20%.

***Table 14.2. Distributions of ratings for each of the conditions for the sten profile graphs (G) and the passages of text (T).***

| | SR | | OB | | FA | | FE | | FM | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Rating** | **G** | **T** | **G** | **T** | **G** | **T** | **G** | **T** | **G** | **T** |
| 1 | 0 | 0 | 0 | 0 | 2 | 0 | 1 | 0 | 1 | 4 |
| 2 | 1 | 0 | 0 | 0 | 10 | 8 | 13 | 15 | 9 | 8 |
| 3 | 4 | 3 | 4 | 5 | 12 | 7 | 11 | 9 | 11 | 7 |
| 4 | 9 | 3 | 2 | 3 | 10 | 10 | 5 | 6 | 4 | 6 |
| 5 | 13 | 14 | 10 | 9 | 1 | 8 | 5 | 4 | 8 | 9 |
| 6 | 8 | 15 | 18 | 15 | 0 | 2 | 0 | 1 | 2 | 0 |
| 7 | 0 | 0 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 1 |
| Mean | 4.66<br>4.91 | 5.17 | 5.29<br>5.26 | 5.23 | 2.94<br>3.31 | 3.69 | 3.00<br>3.03 | 3.06 | 3.43<br>3.39 | 3.34 |

When the combinations of text and graph were rank ordered from "most like me" to "least like me", the profile composed of the Graph and Text based on the person's actual inventory scores was ranked first by 20 of the 35 people, while that based on the self-rated scores was placed first by 10. The ranks assigned to each of the five profiles are shown in Table 14.3[a]. In terms of rank order, the FA condition tends to be placed higher than the other false conditions. If these rankings are collapsed into those ranked "like me" – i.e. ranks 1 and 2 - and the others - i.e. ranks 3, 4 or 5, we see a clear separation between the "true profiles" and the "false ones" (see Table 14.3[b]).

***Table 14.3[a]. Number of people assigning each rank order to each of the five graph+text profiles.***

| | | | RANK | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Most like -------Least like | | | | | |
| **Profile** | | | **1** | **2** | **3** | **4** | **5** | **Total** |
| SR | True | Self-Rated | 10 | 16 | 8 | 1 | 0 | 35 |
| OB | True | ICES-scores | 20 | 10 | 4 | 1 | 0 | 35 |
| FA | False | Average | 2 | 4 | 13 | 9 | 7 | 35 |
| FE | False | Extreme | 0 | 2 | 7 | 8 | 18 | 35 |
| FM | False | Mirror of FE | 3 | 3 | 3 | 16 | 10 | 35 |
| Mean | | | 35 | 35 | 35 | 35 | 35 | |

**Table 14.3[b]. Number of people assigning rank orders 1 or 2 (Like me) and 3, 4 or 5 (Not like me) to each of the five graph+text profiles.**

| Profile | | | Like me 1,2 | Not like me 3,4,5 | N |
|---|---|---|---|---|---|
| SR | True | Self-Rated | 26 | 9 | 35 |
| OB | True | ICES-scores | 30 | 5 | 35 |
| FA | False | Average | 6 | 29 | 35 |
| FE | False | Extreme | 2 | 33 | 35 |
| FM | False | Mirror of FE | 6 | 29 | 35 |
| | | | 70 | 105 | |

From this study, it can be concluded that the ICES computer-generated reports have good discriminative validity because there is a clear differentiation between the two “true” profiles and the three “false” ones. This is true even of the FA condition, where one might have expected a high endorsement rate. Contrary to what was expected, despite the fact that people had produced their SR sten profile directly, they tended to rank the profile based on the ICES scores more highly than the one they had generated by self-rating.

# 15. CONSTRUCT VALIDITY STUDIES: COMPARISONS WITH OTHER PERSONALITY INVENTORIES, 1994-1998

This chapter presents the results of a number of construct validity studies carried out since the Phase III development work was completed in January 1994. This builds on the work reported in Chapter 11 on the comparison with the 16PF, by describing studies comparing the ICES scales with the NEO-PI, the EPQ, the BPI and the Hogan Personality Inventory (HPI).

Further validation studies for the ICES Plus Ability scales are reported at the end of Chapter 16.

## 15.1 Comparison of the ICES and the NEO-PI

The NEO-PI (Costa & McRae, 1985) was developed for use in clinical, counseling and industrial/organizational settings. It is one of the few Personality tests specifically designed around the "Big Five" major factors of Personality (See Chapter 2.1). As the framework provided by research on the "Big Five" was a major factor in the design of the ICES structure, a direct comparison of the ICES and NEO-PI scales provides a key test of the construct validity of the ICES scales. Both the ICES and the NEO-PI are based on the assumption that differences in personality can be explained in terms of four or five major domains. Within each domain, however, there are additional facets of variation. The NEO-PI defines five major domains, three of which are each broken down into measures of six facets. (Note: the more recent version of the NEO-PI provides six facet scales for all five of the domain scales [i.e. 30 scales in all]. Unfortunately the PI-Revised version only became available in the UK after the present study had been carried out.) ICES, on the other hand, adopts a four-domain taxonomy. Two of the Big Five scales (Conscientiousness and Openness to New Experience) are both encompassed as facets of Conscientiousness in ICES, with "Openness to New Experience" being treated as one (C2) of two facets of major scale "C". ICES has a much simpler structure than NEO-PI, with the four major scales being broken down into just two "minor" scales each. In addition, ICES has a Social Desirability scale (SocDes).

### 15.1.1 Method

59 undergraduate students (25 male and 34 female, aged between 18 and 21) completed both ICES inventory and the NEO-PI. The version of the NEO used in this study produced six facet scales for each of the three domain scales: Neuroticism (NEO-N), Extraversion (NEO-E) and Openness (NEO-O). However, no facet scales were available for scales Agreeableness (NEO-A) or Conscientiousness (NEO-C).

### 15.1.2 Hypotheses

At the NEO domain scale level, we expect to find the following pattern of relationships between the scales.

Positive correlations between:

- NEO-E and ICES-E (Extraversion)
- NEO-E and ICES-I1 (Extraversion and Competitive)
- NEO-C and ICES-C1 (Conscientious and Conventional)

Negative correlations between:

- NEO-N and ICES-S (Neuroticism and Stability)
- NEO-O and ICES-C2 (Openness and Organized)
- NEO-A and ICES-I1 (Agreeableness and Competitive)

### 15.1.3 Results

ICES mean sten scores and NEO mean T-scores are presented in Appendix I.1, together with the scale names. Correlations between the ICES and NEO scales are shown in Appendix I.2. Despite being based on different norm groups, both instruments show the sample as being slightly below average in Extraversion, below average in Stability, and low in Conscientiousness.

Significant sex-differences were found on ICES-I (r=-.30) with males scoring higher than females. For the NEO, females scored significantly higher than males on the domain scale NEO-N (r=0.31) and on the facet scales: N1 (r=0.33), N6 (r=0.38) and O3 (r=0.31).

Correlations between the NEO domain scales and ICES major scales (Table 15.1) were in line with the hypotheses. Both Extraversion and Neuroticism/Stability are highly correlated across the two inventories. The expected relationships are also found for ICES-I and ICES-C, with NEO-A negatively correlated with ICES-I and both NEO-O

and NEO-C correlated with ICES-C. Examination of the minor-scale and facet-scale correlations suggests broad agreement between the two instruments in the measurement of common underlying constructs.

***Table 15.1: Correlations between ICES major and minor scales and the NEO "Big Five" measures.***

| Correlations | NEO-N | NEO-E | NEO-O | NEO-A | NEO-C |
|---|---|---|---|---|---|
| ICES major scales: | | | | | |
| ICES-I | -.13 | .12 | .04 | -.52 ** | -.04 |
| ICES-C | -.05 | -.31 * | -.31 * | -.01 | .56 ** |
| ICES-E | -.13 | .73 ** | .41 ** | .22 | -.21 |
| ICES-S | -.80 ** | .38 * | .01 | .23 | .33 * |
| ICES minor scales: | | | | | |
| ICES-I1 | .05 | -.07 | -.08 | -.58 ** | -.12 |
| ICES-12 | -.24 | .27 | .15 | -.27 | .10 |
| ICES-C1 | .01 | -.19 | -.23 | .13 | .38 * |
| ICES-C2 | -.08 | -.28 | -.27 | -.05 | .53 ** |
| ICES-E1 | -.05 | .54 ** | .24 | .26 | -.18 |
| ICES-E2 | -.19 | .72 ** | .47 ** | .16 | -.18 |
| ICES-S1 | -.78 ** | .28 | -.03 | .25 | .36 * |
| ICES-S2 | -.76 ** | .41 ** | .06 | .17 | .32 * |
| * p<.01, ** p<.001 (1 tail) | | | | | |

The number of subjects is too small for exploratory factor analysis. However, given the strong predictions, which can be made about the expected patterns of relationships between the two instruments, it is appropriate to examine the patterns of loadings for both four and five factor solutions to see how closely those predictions are supported.

A four factor Varimax rotated solution (principal components analysis) of the ICES minor scales and NEO domain scales (Appendix I.3:1), confirmed the expected pattern, with each pair of ICES minor scales loading on a distinct factor and with the NEO N, E, A and C scales corresponding to the relevant ICES constructs. NEO-O, however, appears to be more closely connected to the Extraversion factor than to any other. (The expectation was that the strongest link between ICES and NEO-O should be through a negative correlation with ICES-C2).

A five factor solution (Oblimin rotation - Appendix I.3:2) provided the clearest structure for the analysis of the full set of NEO and ICES scales. The reason for the difference in number of factors can be explained by the fact that in the first analysis, the majority of scales are ICES scales and hence the four-factor structure dominates. For the second analysis, however, the majority of scales are NEO scales and so the five-factor structure emerges.

Multiple regression analyses were carried out to compare the degree to which each inventory could be used to predict scores on the other. The results (see Appendix I.4) suggest that, once the differences in numbers of scales is taken into account, they have very similar levels of coverage. Each can account for about half the raw scale score variance of the other.

#### 15.1.4 Conclusions

The study shows a generally clear pattern of relationships between these two instruments showing that, despite the differences in their construction, they cover very similar areas of the major Personality domains.

NEO-N, NEO-E and NEO-A are all well defined by ICES. The clearest relationship is between NEO-N and ICES-S. NEO-E appears to be a much broader factor than ICES-E, showing correlations with both ICES-S (positive) and ICES-C (negative) as well as with ICES-E. NEO-C also appears to be broader than ICES-C, correlating with ICES-S as well as ICES-C. NEO-O is the least clearly represented by ICES. It correlates with ICES-E2 (though not with ICES-E1) and negatively with ICES-C (particularly C1).

### 15.2 Comparison of the ICES and EPQ-R Inventories

The EPQ-R (Eysenck and Eysenck, 1991) is built upon the theory that there are three rather than five major dimensions of personality. It contains 106 items and measures Extraversion, Neuroticism and Psychoticism. These three major scales were developed through a lengthy series of about 20 factor-analytic studies. While he acknowledges that these can be broken down into facets (as in the NEO-PI and ICES), Eysenck argues that Openness to New Experience is ill-defined as a Personality factor and that the variance associated with Agreeableness and Conscientiousness is captured by his Extraversion and Psychoticism scales.

The EPQ-R also contains a "Lie" (dissimulation) scale. The "Lie" scale attempts to measure a tendency to "fake good". It possesses a considerable degree of factorial unity and appears to measure some stable Personality function as well as responding to conditions of dissimulation. In design it is very similar to the ICES SocDes scale.

The EPQ-R is a development of earlier Personality questionnaires, for example the Maudsley Medical Questionnaire (Eysenck, 1952), the MPI (Eysenck, 1959), the EPI (Eysenck and Eysenck 1964) and a revision of the EPQ (Eysenck and Eysenck, 1975).

The ICES structure has some obvious similarities to that of the EPQ-R. Both have “lie” scales, both focus on a small number of major scales. However, ICES adopts a four-domain taxonomy, with its four major scales being broken down into two “minor” scales each.

### 15.2.1 Method

Sixty-eight undergraduate students (41 female and 27 male) completed both ICES and the EPQ-R. The subjects were aged between 16 and 50 (mean=24.24, SD=8.08). The version of the EPQ-R used in this study produced three Personality domain scales P, N and E. In addition, it produces a Lie scale score. Norm conversion tables are not available for the EPQ-R. For the purposes of comparison with ICES, EPQ sten scores have been computed using the general population means and SD given in the EPQ manual.

We expect to find the following pattern of relationships between the scales.

Positive correlations between:

- EPQ-E and ICES-E (Extraversion)
- EPQ-E and ICES-I2 (Extraversion and Assertive)
- EPQ-P and ICES-I1 (Psychoticism and Competitive)
- EPQ-L and ICES-SocDes (Social Desirability)

Negative correlations between:

- EPQ-N and ICES-S (Neuroticism and Stability)

It is difficult to make clear predictions about the relationship between ICES-C and the EPQ scales. However, we would expect the EPQ-L scale to show a correlation with ICES-C and would also expect an inverse relationship between EPQ-P and ICES-C.

### 15.2.2 Results

ICES and EPQ mean sten scores are presented in Appendix I.5, together with the scale names. Correlations between the ICES and EPQ scales are shown in Appendix I.6. In relation to the ICES norm group, this sample is low S and low C. Relative to the EPQ norms group, however, they are high on P. (It should be noted that P tends to have a highly skewed distribution, and that, therefore, sten scores derived from the sample mean and SD will also be skewed.)

Correlations between the EPQ scales and ICES major scales were in line with the hypotheses (see Table 15.2). Both Extraversion and Neuroticism/Stability are high correlated across the two inventories as are the two “lie” scales. The expected relationships with the EPQ were also found for ICES-I1 and I2, the former being correlated with EPQ-P and the latter with EPQ-E. ICES-C showed the relationships expected, being negatively correlated with EPQ-P and positively with EPQ-L. It was also, however, negatively correlated with EPQ-E, suggesting that people who are high on Conscientiousness will tend to be seen as Introvert on the EPQ.

The ICES-SocDes scale appears to have less overlap with other scales than does the EPQ-L scale. The latter correlates significantly with ICES I and C scales as well as with SocDes, while the ICES-SocDes scale only correlates with the EPQ-L scale.

***Table 15.2: Correlations between the ICES major and minor scales and the EPQ.***

| | EPQ Scales | | | |
|---|---|---|---|---|
| | **EPQ-P** | **EPQ-E** | **EPQ-N** | **EPQ-L** |
| ICES minor scales: | | | | |
| I1 | .31 * | .17 | .03 | -.34 * |
| I2 | .14 | .58 ** | -.23 | -.18 |
| C1 | -.47 ** | -.36 * | -.13 | .38 ** |
| C2 | -.43 ** | -.43 ** | .20 | .29 * |
| E1 | .15 | .74 ** | -.06 | -.19 |
| E2 | .27 | .79 ** | -.15 | -.20 |
| S1 | -.15 | .04 | -.70 ** | .24 |
| S2 | -.11 | .17 | -.76 ** | .09 |
| SocDes | -.28 | -.08 | -.00 | .65 ** |
| ICES major scales: | | | | |
| I | .27 | .46 ** | -.12 | -.32 * |
| C | -.53 ** | -.46 ** | .05 | .39 ** |
| E | .24 | .85 ** | -.12 | -.22 |
| S | -.15 | .12 | -.83 ** | .19 |
| * p<.01, ** p<.001 (1 tail) | | | | |

Only one scale showed a difference between the sexes: ICES-I1. The average score for males on Competitiveness was 6.33 (n=27), while for females it was 4.88 (n=41). This difference is statistically significant (F=11.13, df=1, 66, p<.01; eta=0.38).

Differences in scores related to age were noted for Extraversion (r=-.34 for EPQ-E; r=-.45 for ICES-E), Conscientiousness (r=0.29 for ICES-C) and for Psychoticism (r=-.29 for EPQ-P). The age effect for Extraversion was stronger for Group-Oriented (r=-.48 for ICES-E1) than for Outgoing (r=-.33 for ICES-E2).

Again, the number of subjects is small for robust exploratory factor analysis. However, given the predictions made about the expected patterns of relationships between the two instruments, it is appropriate to examine the patterns of loadings to see how closely the ICES structure and the predicted EPQ loadings are supported.

The simplest structure was obtained with a four factor Varimax rotated solution (principal components analysis) of the ICES minor scales and EPQ scales (Appendix I.7). This confirmed the ICES scale structure, with each pair of ICES minor scales loading on a distinct factor. EPQ-E and EPQ-N load as expected on the same factors as the ICES E and S minor scales. Factor II is a complex factor containing both ICES-C and EPQ-P (the latter with a negative loading) and the two “lie” scales. C1 and C2 show the expected negative weights on the Extraversion factor (Factor I), and I2 (Forthright) the expected positive loading on Factor I. Examination of the commonalties suggests that EPQ-P variance is less well explained by this structure than is that of the other scales.

Multiple regression analyses were carried out to compare the degree to which each inventory could be used to predict scores on the other. The results (see Appendix I.8) suggest that, when variations in numbers of scales are taken into account, the ICES scales predict a higher percentage of the variance in EPQ scale scores (around 57%) than vice versa (less than 50%).

### 15.2.3 Conclusions

The study shows a generally clear pattern of relationships between these two instruments showing that, despite the differences in their construction, they cover similar areas of the major Personality domains. There is a close relationship between the two key factors of Extraversion and Neuroticism. The “lie” scales are also clearly related, though the EPQ lie scale tends to be less independent of other scales than does the ICES SocDes scale.

Psychoticism on the EPQ is most closely reflected by a pattern of low Conscientiousness (C) and Competitive (I1) on ICES. This is consistent with the high-P person as someone who is prepared to win at all costs - regardless of the rules or the effects on other people.

The multiple regression analyses suggest that both instruments provide similar levels of coverage of the Personality domains, with a very high degree of overlap between the two. However, the results of the regression analyses suggest that the coverage provided by ICES is somewhat broader than that of the EPQ.

## 15.3 Comparison of the ICES and BPI inventories

The Business Personality Indicator (BPI: Feltham & Woods, 1995) is a new instrument developed specifically for use in industrial/organizational settings. The BPI scale content was chosen to reflect those scales which people in business have found to be the most useful. The BPI is designed to measure 11 primary scales, from which a further five secondary scales are derived. The relationship between secondary and primary is similar to ICES minor and major scales. The BPI scale names are given in Appendix I.9.

### 15.3.1 Method

Ninety people (29 male and 61 female) with an average age of 28 years (with ages ranging from 18 to 64) completed both ICES and the BPI.

Unlike ICES, the NEO-PI and EPQ-R, the BPI's 11 primary scales do not relate in any direct fashion to the “Big Five”. However, on the basis of the scale definitions, predictions can be made about the relationships we would expect to find between the BPI Primary and ICES scales. Some are direct. For example, we would expect BPI-COMPET to correlate with ICES-I1 (Competitive), BPI OUT and LIMELIGHT to correlate with I2 and E2, BPI-WORRY to correlate negatively with ICES-S. BPI scales dealing with perfectionism and time management (PERF and TIME) should correlate with ICES-C (particularly C2), while BPI Risk-Taking and Change Orientation scales can be expected to correlate negatively with ICES-C and positively with E2 (Outgoing, impulsivity and sensation seeking). To summarize, we expect to find:

Positive correlations between:

- I1 and COMPET (Competitiveness)
- I2 and E2 with OUT and LIME (Assertive and socially bold)
- C2 and PERF and TIME (Organized and forward planning)
- E1 and E2 with WARM (Group-oriented/outgoing, Warm in relationships with people)
- E2 with both RISK and CHANGE (Outgoing, impulsive and sensation seeking)

Negative correlations between:

- C1 and C2 with both RISK and CHANGE (Conventional, Organized and cautious)
- S with WORRY

### 15.3.2 Results

ICES and BPI mean sten scores are presented in Appendix I.9, together with the scale names. Correlations between the ICES minor scales and the BPI primary scales are shown in Appendix I.10. The pattern of results is very close to that expected.

Correlations between the BPI Secondary scales and ICES major scales were in line with the hypotheses (see Table 15.3). Both Extraversion and Worrying/Stability are high correlated across the two inventories. In addition to the predicted relationships, at the primary scale level, the BPI Stamina scale is also correlated with ICES Stability, and both BPI Risk-taking and BPI Limelight-seeking correlate positively with ICES-S2. Examination of the ICES minor scale and BPI primary scale correlations suggests broad agreement between the two instruments in the measurement of common underlying constructs. However, the fit between the ICES major scales and the BPI secondaries is less clear for the BPI Dynamic and Work Stamina scales.

***Table 15.3: Correlations between the ICES major scales and the BPI secondary scales.***

| | BPI Secondary Scales | | | | |
|---|---|---|---|---|---|
| **Correlations** | **DYNAMIC STAMINA** | **WORK** | **CONTROL** | **EXTRAV** | **WORRY** |
| ICES-I | 0.56 ** | 0.33 ** | -.19 | 0.11 | -.36 ** |
| ICES-C | -0.44 ** | 0.29 * | 0.69 ** | 0.37 ** | 0.27 * |
| ICES-E | 0.44 ** | -0.03 | -0.10 | 0.67 ** | -0.32 * |
| ICES-S | 0.31 * | 0.25 * | -0.15 | 0.42 ** | -0.75 ** |
| *p<.01, ** p<.001 (1 tail) | | | | | |

As for the NEO-PI, it is appropriate to examine the patterns of loadings for both four and five factor solutions to see how closely the four versus five domain models are supported. Examination of the principal components eigenvalues (by scree test) suggested that only four factors should be extracted (accounting for 53.7% of the total variance).

A four factor Varimax rotated solution (principal components analysis) of the ICES minor scales and eleven BPI scales (Appendix I.11), confirmed the expected pattern, with each pair of ICES minor scales loading on a distinct factor and with the BPI scales corresponding to the relevant ICES constructs. While examination of the eigenvalues confirmed a four rather than five factor solution, a five-factor solution was also examined to see if there was a "Big Five" fit. However, this did not provide a clear, interpretable structure.

Significant sex differences were found on ICES-I (r=-.25) and on ICES-S (r=-.37), with males scoring higher than females in both cases. However, these differences were only significant at the minor scale level for S1 and S2 - not for I1 and I2. For the BPI, females scored significantly higher than males on Worry (r=0.44), while males scored significantly higher than females on Change (r=-.25) and Stamina (r=-.39).

Multiple regression analyses were carried out to compare the degree to which each inventory could be used to predict scores on the other. The results (see Appendix I.12) suggest that, once the differences in numbers of scales are taken into account, they have very similar levels of coverage. Each can account for about half the raw scale score variance of the other.

#### 15.3.3 Conclusions

The study shows a generally clear pattern of relationships between these two instruments showing that, despite the differences in their construction, they cover very similar areas of the major Personality domains.

BPI-WORRY, BPI-EXTRAV and BPI-CONTROL are all well defined by ICES. The clearest relationship is between WORRY and ICES-S. As for the EPQ, BPI-EXTRAV appears to be a much broader factor than ICES-E, showing correlations with both ICES-S (positive) and ICES-C (negative) as well as with ICES-E. BPI-DYNAMIC has correlations with all four ICES major scales, while BPI-WORK-STAMINA has correlations with three. This suggests that these two Secondary scales are factorially complex. However, they may have an advantage in providing measures that are more closely tied to aspects of work behavior rather than "pure" aspects of personality.

The structures of the BPI and ICES are very different. However, the data collected provides a basis for defining new ICES construct scales analogous to the more business-oriented BPI scales (e.g. Work Stamina, Risk-taking, etc). While such scales are not as "pure" as the ICES minor scales, they may prove useful as ways of re-presenting information to aid interpretation.

### 15.4 Comparison of the ICES and the Hogan Personality Inventory (HPI)

The objectives of the study were to:

1. To assess the relationship between the Hogan Personality Inventory (HPI) and the ICES Personality questionnaire.
2. Examine the link between the HPI and a self-report measure of counter-productive behavior.
3. Examine the link between the ICES and a self-report measure of counter-productive behavior.

#### 15.4.1 Method

Twenty-eight production workers and 37 fire-fighters completed the Hogan Personality Inventory, the ICES and a short 8-item self-report admissions scale relating to counter-productivity in the workplace. Of the participants, 71% were at an operative job level, 18.5% at general manager level and 3% senior management level. The age range of the sample was from 20 to 54 years with a mean of 31.29 years. The majority of the sample (60) were male and only 2 were females (three non-respondents).

The Hogan Personality Inventory: (Hogan & Hogan, 1995) has 7 major scales of Adjustment, Ambition, Sociability, Likeability, Prudence, Intellectance and School Success. In addition, the HPI has 6 Occupational scales of Service Orientation, Stress Tolerance, Reliability, Clerical Potential, Sales Potential and Management Potential. Each scale comprises what are termed Homogenous Item Composites (see Appendix I.13).

The 8-item counter-productivity scale asked respondents to report whether they have or have not engaged in a number of behaviors. These behaviors were: Taken unauthorized time off work; Arrived for work late; Used telephone or mailing facilities for personal use; Disregarded company rules and regulations; Left work early without permission; Intentionally worked slowly; Taken company equipment or property without permission; Taken longer breaks than allowed.

Each "Yes" answer was scored one and each "No" zero, so high scores indicate a high level of involvement in counter-productive behaviors. The mean for the scale was 3.08 (SD=2.06) and the internal consistency (Cronbach's alpha) was 0.69.

### 15.4.2 Results

Descriptive statistics for the mean raw scores and their standard deviations for ICES and Hogan scales are present in Appendix I.14. Appendix I.14 illustrates the means and SDs for sten scores on each of the ICES scales for the production workers and fire-fighters. The data on the production workers (n=28) shows that in general this group is high on Social Desirability (SocDes) and Competitive (I1) and somewhat low on Extraversion (both E1 and E2) and Stability (both S1 and S2). For the fire-fighters (n=37), in general they are somewhat high on Independence (mostly I1), and low on Conscientiousness (especially C1). In addition, the data shows that this group is also high on Relaxed (S2).

### 15.4.3 Correlations between ICES and the HPI (see Appendix I.15).

As can be seen from Table 15.4, the highest correlations exist between ICES major scales and HPI main scales where they would be expected on the basis of common constructs (these are printed in bold type). All are highly significant except for the correlation between Independence and Ambition, which is not significant at the 5% level.

As expected, Conscientiousness correlates significantly positively with HPI-Prudence. In addition, the data shows that this scale correlates negatively with HPI-Sociability. A person who scores high on Prudence is "...reliable, thorough, dignified, cautious and responsible. They are conscientious and attentive to detail...They tend to be well liked as managers..." (HPI Manual 1995).

***Table 15.4: Correlations between ICES major scales and HPI main scales (n=65)***

| | ICES Major Scales | | | |
|---|---|---|---|---|
| | **INDEP** | **CONSC** | **EXTRAV** | **STABLE** |
| HPI-ADJ | 0.12 | -0.05 | 0.43 ** | **0.70** ** |
| HPI-AMB | **0.24** ‡ | 0.05 | 0.40 * | 0.35 * |
| HPI-SOC | 0.20 | -0.35 * | **0.74** ** | 0.07 |
| HPI-LIK | -0.19 | -0.05 | 0.37 * | 0.06 |
| HPI-PRU | -0.22 ‡ | **0.45** ** | -0.10 | 0.19 |
| HPI-INT | -0.10 | -0.06 | 0.19 | -0.03 |
| HPI-SCH | -0.03 | -0.07 | 0.04 | 0.10 |
| ‡ p<0.1 * p<0.01 ** p <0.001 | | | | |

The ICES Extraversion scale correlates significantly with the majority of the HPI scales. A person scoring high on Extraversion tends to be high on Adjustment, Ambition, Sociability and Likeability. They also tend to be lower on Prudence (which links well with the data on Conscientiousness above), although this is not statistically significant.

As well as the hypothesized strong positive relationship between Stability and HPI-Adjustment, a significant positive relationship exists between Stability and HPI-Ambitious. This is also not surprising as a person high in Ambition tends to be: "energetic...self-assured, leader-like, and eager to advance".

There are no significant correlations between Independence and the HPI scales, although correlations are in the appropriate direction: e.g. there is a positive correlation with Ambition and a negative one with Prudence. Low correlations are seen between Intellectance and School success, and the ICES scales, the highest being between Extraversion and Intellectance.

The results in Table 15.5 show that the ICES minor scales correlate significantly with the appropriate scales from the HPI (in bold type). Once again there is somewhat of a different pattern for the Independence sub-scales. Competitiveness (I1) is not strongly related to Ambition but rather to Sociability. On the other hand, Assertive (I2) is positively related to Ambition.

***Table 15.5: Correlation of ICES minor scales with HPI scales (n=65).***

| | ICES minor scales | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **HPI** | **I1** | **I2** | **C1** | **C2** | **E1** | **E2** | **S1** | **S2** | **Soc Des** |
| ADJ | 0.03 | 0.16 | -0.11 | 0.01 | 0.48 *** | 0.28 * | **0.55** *** | **0.70** *** | 0.02 |
| AMB | **0.07** | **0.28** * | 0.04 | 0.05 | 0.29 * | 0.38 ** | 0.37 ** | 0.28 * | -0.06 |
| SOC | 0.28 * | 0.06 | -0.23 ‡ | -0.35 ** | **0.58** *** | **0.68** *** | -0.15 | 0.23 ‡ | -0.28 * |
| LIK | -0.19 | -0.12 | 0.06 | -0.12 | 0.46 *** | 0.20 | 0.04 | 0.06 | 0.12 |
| PRU | -0.23 ‡ | -0.14 | **0.41** ** | **0.37** ** | 0.03 | -0.19 | 0.23 ‡ | 0.12 | 0.31 * |
| INT | -0.09 | -0.07 | -0.07 | -0.03 | 0.17 | 0.16 | -0.07 | -0.00 | -0.06 |
| SCH | -0.18 | 0.10 | -0.13 | -0.02 | -0.01 | 0.08 | 0.17 | 0.04 | -0.16 |
| ‡ p <0.1 * p<0.05 **p<0.01 *** p<0.001 | | | | | | | | | |

Social Desirability (SocDes) has a significant negative relationship with Sociability and a significant positive relationship with Prudence. This is very similar to the pattern seen for Conventional (C1) and Organized (C2).

Relationships between the HPI of Homogeneous Item Composites (HICs) and the ICES major scales are presented in Appendix I.15.2. As each main scale in the HPI comprises a number of HICs, it was decided to examine the link between these composites and major scales on the ICES. As there is a reliable and robust link between the ICES and HPI scales, it was expected that those HICs within each main scale of the HPI should correlate with the appropriate ICES scales.

Appendix I.15.2 illustrates the correlations, and as expected there are significant correlations between HICs and the appropriate ICES scale (in bold). Social Desirability is positively correlated with those HICs that make up the Prudence scale on the HPI. As Prudence has already been shown to correlate significantly with Conscientiousness, the correlation of SocDes with these HICs is unsurprising.

Although ICES does not explicitly have a Likeability (Agreeableness) scale, one can see from the correlations of the HICs comprising Likeability that these people tend to be low in Independence and high in Extraversion. Further analysis using the ICES minor scales illustrates that Likeability is related to low Competitiveness (I1) and high Group-orientation (E1).

### 15.4.4 ICES predictions of the HPI Occupational Scales

Multiple regression analyses were carried out with the score on each of the HPI Occupational scales in turn as the dependent variable and the ICES minor scales as predictor variables. Details of the analyses are in Appendix I.16.

1. ***Service Orientation (SOI)*** "To identify persons who are pleasant, courteous, co-operative, and helpful in dealing with customers, clients and co-workers" (HPI Manual, 1995). The main predictors of this were Conventional (C1) and Group-oriented (E1) (both in a positive direction). Overall R=0.59 (p<.01).
2. ***Stress Tolerance (STR)*** "To identify persons who handle pressure well and are not tense or anxious" (HPI Manual, 1995). Positive predictors included Group-oriented (E1), Poised (S1) and Relaxed (S2), with an overall R=0.70 (p<.001).
3. ***Reliability (REL)*** "To identify people who are honest, dependable, and responsive to supervision" (HPI Manual, 1995). None of the regression coefficients were significant at the 5% level, although, a positive coefficient was seen for Group-oriented (E1), which was significant at the 10% level, and a negative coefficient for Assertive (I2), again significant at the 10% level. Overall, R=0.46 (not significant).
4. ***Clerical Potential* (*CLERK*)** "To identify people who are attentive to detail, congenial, and industrious" (HPI Manual, 1995). The main contributors to the prediction (R=0.58, p<.01) were Poised (S2) and Group-oriented (E1).
5. ***Sales Potential (SALES)*** "To identify persons who are socially skilled, self-assured, assertive, and can create interest in products and services" (HPI Manual, 1995). As would be expected significant and positive contributions to a strong prediction (R=0.71, p<.001) were found for Group-oriented (E1) and Outgoing (E2).
6. ***Management Potential (MANAGER)*** "To identify persons who can supervise others in a pleasant and effective fashion" (HPI Manual, 1995). The main predictors were E1 and S1, with an overall R=0.48 (p<.05).

## 15.5 Comparison of the ICES Personality Inventory and the Hogan Development Scale (HDS)

The objectives of the study were to:

- assess the relationship between the ICES Personality questionnaire major and minor scales and 11 Hogan Development Scales (HDS); and
- to examine the underlying commonalties in the structure of the ICES scales and the HDS.

### 15.5.1 Method

Eighty-three university students, working adults and retired individuals completed the HDS and the ICES personality questionnaire. The sample was comprised of 71.1% females and 28.9% males, ranging in age from 16 to 82 with a mean age of 39.16 years and a standard deviation of 18.96 years. All of those in the sample, except one, were of white ethnic origin, the exception being of chinese origin.

The Hogan Development Survey (Hogan and Hogan, 1997) has 11 scales Volatile, Mistrustful, Cautious, Detached, Passive-Aggressive, Arrogant, Manipulative, Dramatic, Eccentric, Perfectionistic & Dependent. It was designed to look at the “dark side” of human nature. It focuses on dysfunctional behavior that manifests under stress in the workplace. The 11 HDS scales can be grouped according to Horney's (1950) three underlying interpersonal behavior patterns: “Moving away” from people (Volatile, Mistrustful, Cautious, Detached & Passive-Aggressive); “Moving against” other people (Arrogant, Manipulative, Dramatic & Eccentric); “Moving towards” other people (Perfectionistic & Dependent).

### 15.5.2 Results

ICES mean sten scores are presented in Appendix I.17. In relation to the ICES norm group, the mean score of this sample is generally lower than the norm group and approximately one sten lower on ICES-E (both E1 and E2) and on the ICES-S (particularly S1). Other scores fall within a half sten either side of the mean of the norm group.

### 15.5.3 Correlations of ICES main scales with HDS scales

Table 15.5 illustrates the Pearson correlation coefficients (1-tailed) between the raw scores of the ICES main scales and HDS scales. Appendix I.18 provides correlations for the HDS scales with ICES sub-scales.
The Volatile scale correlates significantly negatively with Extraversion and Stability. The strongest of these correlations is with Stability. Those who are Volatile tend to be “moody, irritable, prone to emotional outbursts and likely to let little things bother them” and this equates well to those scoring at the low end of the ICES-Stability scale. The relationship with the Extraversion scale, specifically with Outgoing (E2), is also in the expected direction. High scorers on Volatile may become “unpredictable and impulsive at times and through this and being disappointed with people, they may have difficulty working with others”.

The Mistrustful scale correlated significantly positively with the Independence scale, but only with the Competitive (I1) and not with the Assertive (I2) sub-scale. This reflects the “ready to fight”, competitive and challenging nature of the high scorer on Mistrustful. The significant negative correlation with Stability (S) is in the expected direction, as a Mistrustful individual tends to “…mistrust others” motives and doubt their intentions, to be alert for signs that one is being deceived…they take criticism personally…they tend to be suspicious” (Hogan Development Survey Technical Manual, 1998).

A high scorer on the Cautious scale tends to be “easily embarrassed; have a fear of being criticized; be happy letting others take the initiative; shy; be unwilling to take chances; maintain order and stick to rules; be gracious and obliging”. This is reflected in the pattern of correlations on the ICES (negative correlation with the Stability (S) and Extraversion (E) scales, a positive correlation with the Conscientiousness (C) scale and a negative correlation with the Independence (I) scale, specifically with the sub-scale Assertive (I2)).

High scorers on Detached tend to be “independent, self-sufficient and spend time alone”. There is also a significant positive relationship with Competitive (I1), indicating “single-mindedness, playing to win and a lack of concern about others”. Those high on Detached tend to be low in Extraversion (E) specifically on Outgoing (E2).

***Table 15.6: Correlations between ICES main scales and HDS scales (N=83)***

| ICES Scales<br>HDS Scales | Indep<br>(I) | Conscient<br>(C) | Extrav<br>(E) | Stability<br>(S) | Social Des<br>(SD) |
|---|---|---|---|---|---|
| Volatile | 0.02 | 0.04 | -0.30** | **-0.53*** ** | -0.04 |
| Mistrustful | **0.34** ** | -0.04 | -0.10 | -0.27** | -0.09 |
| Cautious | -0.41*** | 0.32** | -0.48*** | **-0.70*** ** | 0.23* |
| Detached | 0.17 | 0.19* | **-0.48*** ** | -0.15 | 0.24* |
| Passive Agg | -0.08 | 0.32** | -0.24* | **-0.49*** ** | 0.18 |
| Arrogant | **0.28** ** | -0.02 | 0.18 | 0.11 | 0.02 |
| Manipulative | **0.41*** ** | -0.36*** | 0.38*** | 0.26** | -0.21* |
| Dramatic | 0.33** | -0.31** | **0.51*** ** | 0.24** | -0.22* |
| Eccentric | **0.27** ** | -0.13 | 0.15 | 0.09 | 0.01 |
| Perfectionistic | -0.01 | **0.53*** ** | -0.08 | -0.34** | 0.19* |
| Dependent | **-0.53*** ** | 0.33** | -0.10 | -0.31** | 0.28** |
| * p<0.05 **p<0.01 ***p<0.001<br>**Bold** – indicates strongest correlation | | | | | |

The pattern of correlations for the Passive Aggressive scale suggests that these individuals tend to be "conscientious, organized, rule-bound and traditional", a positive correlation with Conscientiousness (C); "anxious, irritable, suspicious and emotional", a negative correlation with Stability (S); and "introverted, quiet and reserved" a negative correlation with Extraversion (E).

A significant positive relationship emerged between the Arrogant scale and Independence (I), specifically with I1 (Competitive). This may be due to the fact that high scorers on Arrogant "tend to place a lot of emphasis on their accomplishments… [and are] self –absorbed" (Hogan Development Survey Technical Manual, 1998). Hence, they tend to be competitive and single-minded.

The Manipulative scale correlates negatively with Conscientiousness (C) and positively with Extraversion (E) and specifically with Outgoing (E2). Individuals high on the Manipulative scale tend to "take risks, be impatient and impulsive, seek excitement and be carefree". Further, a positive relationship exits with Independence (I). This relates to their energetic, independent and self-assured nature. The Manipulative scale also correlates positively with Stability (S) and reflects the "no regrets" and unconcerned nature of the manipulative individual.

High scorers on the Dramatic scale tend to "enjoy being the center of attention, are entertaining and usually perform well in public". This is reflected in the strong positive relationship with Extraversion (E). They tend to "lack attention to detail", reflected in the negative correlation with Conscientiousness (C) and are "active and outspoken with their vision" reflected in the positive correlation with Independence (I), specifically with the Assertiveness (I2) sub-scale.

The large positive relationship between the Perfectionistic scale and the Conscientiousness (C) scale reflects the "conscientious, detail-conscious and orderly" nature of the high scorer on the HDS scale.

A strong negative correlation occurred between the Dependent scale and the Independence (I) scale. Those high on Dependent tend to be "eager to please others, to be pleasant and agreeable and a good team member". All these traits are found in a low scorer on Independence (I) and specifically in those on the low end of the sub-scale Assertiveness, (I2).

The Eccentric scale correlated with Independence (I) with a stronger correlation with the Assertiveness scale (I2). This captures only a small part of the HDS description of the eccentric as "the tendency to think and act in ways that are unusual, different, striking and at times odd". This HDS scale appears to measure the Openness construct of the big-5 personality traits which ICES does not measure (see section 15.5.4)

**15.5.4 Factor analysis of ICES and HDS scales**

A principle components analysis using varimax rotation was carried out on the raw scores of the 4 ICES main scales and 11 HDS scales. A five-factor solution explaining 72.2% of the variance was produced. Appendix I.19 illustrates the factor loadings of each of the scales.

The factor solution produced 5-factor structure that relates to the "Big-five" structure. The ICES scales loaded onto the expected 4 distinct factors, and with the loading of the HDS scales as well, they clearly represented scales of Independence, Stability, Extraversion and Conscientiousness. Factor 4 can be interpreted as "Openness to

Experience" as the Eccentric, Passive Aggressive and Manipulative scales of the HDS load onto this factor. These scales include traits such as acting in unusual ways, doing his/her own thing, striking, fun loving, insightful, carefree, full of energy, and working to one's own timetable. As expected no ICES scale loads onto this factor as the ICES does not sample this trait.

## 15.6 Conclusions on Construct Validity Studies

Taken together, the studies involving the 16PF, NEO-PI, EPQ-R, BPI, and HPI provide substantial backing for the interpretation given to the ICES scales (see Chapter 14). The results also provide good support for the robustness of the four - rather than three or five - factor structure. For all the data sets, four-factor solutions provided the best fit to the data. The three main EPQ scales appeared to be over inclusive, covering broader domains than either the ICES or NEO scales. The NEO five factors, on the other hand, do not seem to be independent of each other. Indeed, some of the facet scales appeared to load more highly on factors other than their own. In general, it was found easier to map both instruments onto the ICES four-factor structure than to map ICES onto either three or five factors. The BPI's eleven primary scales also showed a better fit to the four-factor ICES structure than to a five-factor one.

Comparisons between the HPI and ICES provide support for occupational related interpretations of the ICES scales (from their correlations with the occupational HPI scales). The main area of difference between the two instruments lies in the interpretation of "Reliability". The HPI-Reliability does not appear to match closely to ICES Conscientiousness.

The study that incorporated the data from HDS and ICES has provided positive findings for the construct validity of the ICES scales in relation to the Hogan Development Survey. Significant correlations between ICES and HDS scales are in the directions expected by the scale definitions. The factor analysis produced a 5-factor solution, which mapped closely onto the "Big-5", with the 4 ICES scales loading onto distinct factors and 3 HDS scales loading onto a fifth factor (Openness). This analysis further confirms the robustness of the ICES four-factor solution.

# 16. CONSTRUCT VALIDITY STUDIES: Risk, Change, Focus on Work and Social Activity

This Chapter reports on a series of validation studies, carried out between 1995 and 1998, which explore the relationships between personality and risk-taking and personality and social activity. The findings from this research have been used to inform the development of the latest versions of ICES Plus interpretative software, and to enrich the quality of the information provided from the profile of ICES scores.

## 16.1 The relationship between personality, perceived risk and risk-taking

Risk is a complex concept. In order to understand, and hence predict, risk and risk-taking behavior it is necessary to distinguish between:

- *Perceived and actuarial risk.* People may choose to act in a way which they perceive as being safe, but which in reality is not, or vice versa. People may differ in terms of how risky they perceive a given act to be: for example, observing another person driving at speeds well in excess of the legal speed limit. Under the same set of road conditions, some people are likely to judge the action as more risky than others. The driver of the car may assess the risk differently – depending on perceptions of their own skill and the prevailing conditions.
- *Likelihood.* People's choices between alternatives are affected by their judgement of the likelihood of each alternative leading to either a positive or a negative outcome for them.
- *Costs and consequences.* Making choices may involve some form of cost (effort, money, or time). The outcome of whatever option was chosen, in turn will have consequences (which may be good or bad) which may be judged as being more or less worthwhile. Some consequences may be serious (death or disability) others may be trivial (being a bit late for a meeting). Some may be immediate, others may not happen for years.

A better understanding of what causes people to make different choices, and why people perceive risks differently is important if we are to improve our prediction of peoples' actions in conditions of choice and uncertainty. The study reported here provides some initial insights into the complex relationships between personality traits (as measured by ICES) and various aspects of reported behavior in risky situations and differences in people's perceptions of and attitudes to taking risks.

### 16.1.1 Method

Eighty-five undergraduates completed the ICES Personality Inventory together with two other research inventories.

1. RISK_ACTS: a set of 13 questions asking about risky behaviors. These included questions like: "Do you smoke?", "In the past two years, have you had unsafe sex?", "In the past two years, have you ever taken an illegal substance?"
2. The Personal Risk Inventory (PRI: Clough & Hockey, 1996): a questionnaire containing 20 everyday choice dilemmas. Each dilemma consisted of a description of an everyday choice situation (e.g. "You have to take an important visitor out for a meal. Do you take them to a restaurant you know well, which is good but not very special, or take them to a new place you have heard of which is supposed to be a bit different?"). For each dilemma, people have to choose which option they would go for, how strongly they favour their chosen course of action and how risky they think it is.

#### 16.1.2 Results

Principal components analysis of the RISK_ACTS identified a subset of seven items which had loadings on a common factor. These seven items had an internal consistency of alpha=0.67. Analysis of the PRI was more complex. The data on strength of preferred option was analyzed first. This was scored such that a high score (10) indicated a strong preference for the "risky" option, and a low score (1) a strong preferences for the "safe" option. Principal components analyses suggested both the presence of a common factor (which was best described by 12 of the items, alpha=0.65) and a set of 5 sub-scales PRI-1 to PRI-5. The rated riskiness of the choice (i.e. the perceived risk level) was analyzed separately. There was higher consistency within subjects for this rating, with a 12-item scale producing an alpha of 0.73 (Rated-Risk). The scale means, SDs, alphas and intercorrelations are presented in Table 16.1.

Taking part in risky activities (RISK_ACT) only correlates positively with sub-scale 2 of the PRI. The overall correlation with the PRI is near zero and there is only a small (r=0.13, not significant) correlation between RISK_ACT and the perceived riskiness of options chosen in the PRI. While perceived riskiness of choice shows positive correlations with strength of choice, this is only significant for sub-scale 4 (r=0.26, p<.05).

**Table 16.1: Descriptive Statistics (means, SDs and alphas) and correlations between the PRI and RISK_ACT measures (n=85).**

| | | | PRI Sub-scales | | | | | Rated |
|---|---|---|---|---|---|---|---|---|
| **Scale** | **RISK_ACT** | **PRI** | **1** | **2** | **3** | **4** | **5** | **Risk** |
| Correlations | | | | | | | | |
| RISK_ACT | 1.00 | -0.01 | -0.22 | 0.32 * | -0.20 | -0.05 | 0.18 | 0.13 |
| PRI | -0.01 | 1.00 | 0.78 ** | 0.69 ** | 0.47 ** | 0.50 ** | 0.10 | 0.22 |
| PRI-1 | -0.22 | 0.78 ** | 1.00 | 0.27 * | 0.32 * | 0.26 * | 0.03 | 0.16 |
| PRI-2 | 0.32 * | 0.69 ** | 0.27 * | 1.00 | 0.06 | 0.34 ** | 0.15 | 0.17 |
| PRI-3 | -0.20 | 0.47 ** | 0.32 * | 0.06 | 1.00 | 0.38 ** | -0.07 | 0.25 |
| PRI-4 | -0.05 | 0.50 ** | 0.26 * | 0.34 ** | 0.38 ** | 1.00 | 0.11 | 0.26 * |
| PRI-5 | 0.18 | 0.10 | 0.03 | 0.15 | -0.07 | 0.11 | 1.00 | 0.18 |
| Riskiness | 0.13 | 0.22 | 0.16 | 0.17 | 0.25 | 0.26 * | 0.18 | 1.00 |
| | | | | | | | | |
| Mean | 3.64 | 87.94 | 57.48 | 46.78 | 18.92 | 31.27 | 15.80 | 31.16 |
| SD | 1.97 | 15.26 | 12.31 | 11.40 | 6.38 | 8.74 | 6.54 | 7.06 |
| Alpha | 0.67 | 0.65 | 0.65 | 0.63 | 0.60 | 0.50 | 0.45 | 0.73 |
| Items | 7 | 12 | 8 | 7 | 3 | 5 | 3 | 12 |
| *p<.01; ** p<.001 (1-tail) | | | | | | | | |

A better understanding of the risk measures can be obtained by examining their patterns of correlation with ICES. While the overall PRI measure and the individual sub-scales all have rather low reliabilities, it is of interest to examine relationships with ICES to explore the links between personality traits, reports of actual risky behavior (RISK_ACTS), expressions of behavioral intentions (PRI) and their perceived riskiness (Rated-Risk).

Table 16.2 and 16.3 show correlations between the ICES scales and each of the risk measures (Table 16.2) and the results of multiple regression analyses using the risk measures as dependent variables, and ICES scales as predictors (Table 16.3). For comparison purposes, Table 16.3 also includes the regression beta weight for predicting the BPI Risk Taking scale.

Using data from the BPI construct validation study (see Chapter 15.3) three new ICES scales were constructed. Each was derived using the correlations between ICES and the BPI scales (see Appendices I.9 and I.10 for more details) and is a weighted composite of the eight ICES minor scales and the SD scale.

- RISK. People scoring high on this scale show a willingness to take risks, enjoying the excitement for its own sake, even if it does not bring rewards. People scoring low on this scale will be cautious and place a higher value on security and stability in their work.
- CHANGE. People scoring high on this scale enjoy and value change, like new experiences and want to organize their own time and work rather than be organized by others. People scoring low on this scale prefer life to be predictable and routine.
- WORK. People who score high on this scale will value work for its own sake, and tend to define themselves in relation to their work rather than out-of-work activities. Low scorers value leisure activities more than work, and see work as a means to an end rather than an end in itself.

Sten scores for each of these composites was derived using the Phase III data set. Table 16.2 gives the mean sten scores and SDs for the present sample based on the Phase III sample (n=516) as the norm reference group.

The three new composite ICES scales show strong correlations with RISK_ACT and the PRI, especially PRI-2. Both Change and Risk show similar patterns of correlation, with Work being negatively related to PRI but not related to RISK_ACT.

People's perceptions of the riskiness of their chosen options show a very different pattern of relationships with ICES than do either RISK_ACT or PRI. People who rate their chosen option as being risky tend to be non-assertive and rather timid (low I2), but outgoing and sociable (high E2). This is combined with low S1 scores, suggesting that they see the world as essentially hostile and threatening. Together, this mixture of traits explains why some people perceive a given course of action as more risky than do other people.

Examination of the results for the PRI sub-scales, suggests that PRI-2 is closest to RISK_ACT in the pattern of relationships it has with ICES scales – it is also the only sub-scale to correlate significantly with RISK_ACT ($r=0.26$, $p<.05$). Like RISK_ACT, PRI-5 correlates with ICES I1 and C1, but it lacks the correlation with Extraversion found for RISK_ACT. PRI-3 is interesting in that it correlates positively with C2 and negatively with SD. This combination of traits suggests people who are meticulous and who like everything to be planned ahead. They will go for the planned option rather than the unplanned one. However, this is combined with an openness and willingness to choose options which might be seen as "risky" in preference to those which are unplanned. PRI-4 appears to be primarily related to SD, and could represent response bias.

***Table 16.2: Correlations between ICES scales and the two Risk inventories: RISK_ACTS and PRI (n=85)***

| | | | | | PRI Sub-scales | | | | | Rated |
|---|---|---|---|---|---|---|---|---|---|---|
| **ICES** | **Mean** | **SD** | **RISK_ACT** | **PRI** | **1** | **2** | **3** | **4** | **5** | **Risk** |
| I | 45.33 | 7.98 | .33 * | .03 | .01 | .20 | -.06 | -.03 | .22 | -0.14 |
| I1 | 21.02 | 4.40 | .32 * | -.03 | -.11 | .09 | .06 | .04 | .27 * | 0.01 |
| I2 | 24.31 | 5.21 | .23 | .08 | .10 | .23 | -.14 | -.08 | .10 | -0.22 |
| | | | | | | | | | | |
| C | 42.72 | 6.70 | -.22 | -.34 ** | -.16 | -.50 ** | .09 | -.05 | -.11 | -0.08 |
| C1 | 20.20 | 3.70 | --26 * | -.21 | .06 | -.42 ** | .18 | -.15 | -.23 | -0.03 |
| C2 | 22.52 | 4.96 | -.10 | -.31 * | -.26 * | -.35 ** | -.01 | .04 | .03 | -0.09 |
| | | | | | | | | | | |
| E | 49.85 | 9.62 | .38 ** | .19 | .09 | .27 * | .17 | .02 | .05 | 0.14 |
| E1 | 25.14 | 5.00 | .28 * | .16 | -.01 | .22 | .20 | .05 | -.13 | 0.16 |
| E2 | 24.71 | 5.77 | .39 ** | .19 | .16 | .26 * | .11 | -.01 | .19 | 0.10 |
| | | | | | | | | | | |
| S | 43.67 | 8.31 | .06 | .05 | .04 | .06 | .01 | .00 | .17 | -0.17 |
| S1 | 21.84 | 4.52 | .02 | -.01 | -.03 | -.01 | .03 | .02 | .12 | -0.22 |
| S2 | 21.84 | 4.75 | .09 | .10 | .09 | .11 | -.00 | -.01 | .19 | -0.09 |
| | | | | | | | | | | |
| SD | 20.88 | 4.86 | -.27 * | -.34 ** | -.07 | -.38 ** | -.20 | -.31 * | -.13 | -0.20 |
| | | | | | | | | | | |
| **Composite scales** | | | | | | | | | | |
| RISK | 6.21 | 2.06 | .35 ** | .26 ** | .18 | .42 ** | -.07 | -.02 | .25 | -0.03 |
| CHANGE | 7.32 | 1.92 | .35 ** | .33 ** | .15 | .51 ** | -.06 | .08 | .14 | -0.02 |
| WORK | 4.49 | 2.18 | -.09 | -.36 ** | -.22 | -.37 ** | -.11 | -.09 | .06 | -0.15 |

***Table 16.3: Multiple regression results for prediction of risk using ICES minor scales as independent variables. Values in the table are beta weights from stepwise multiple regression. For RISK_ACT and PRI scales, n=85; for BPI Risk scale, n=90.***

| | | | PRI Sub-scales | | | | | Rated | BPI Risk |
|---|---|---|---|---|---|---|---|---|---|
| **ICES** | **RISK-ACT** | **PRI** | **1** | **2** | **3** | **4** | **5** | **Risk** | **scale** |
| I1 | 0.24 | | | | | | 0.29 | | 0.17 |
| I2 | | | | | | | | -0.42 | 0.26 |
| C1 | -0.28 | | | -0.38 | | | -0.25 | | -0.23 |
| C2 | | -0.22 | -0.26 | -0.23 | 0.25 | | | | -0.29 |
| E1 | | | | | | | | | |
| E2 | 0.31 | | | 0.20 | | | | 0.41 | |
| S1 | | | | | | | | -0.25 | |
| S2 | | | | | | | | | 0.18 |
| SD | | -0.27 | | | -0.30 | -0.31 | | | |
| | | | | | | | | | |
| R= | 0.52 | 0.40 | 0.26 | 0.55 | 0.32 | 0.31 | 0.37 | 0.43 | 0.70 |
| F= | 10.12 | 7.86 | 6.19 | 11.46 | 4.53 | 8.78 | 6.33 | 6.20 | 16.07 |
| Sig(p) | <.001 | <.001 | <.05 | <.001 | <.05 | <.01 | <.01 | <.001 | <.001 |

Those who report indulging in risky behaviors (high RISK_ACT scores) tend to be extravert (high E1, E2) and competitive and show relatively little regard for others (high I1). Their extraversion is strongest on E2, where high scorers are described (see Appendix G: Outgoing E2) as enjoying "risky" action-packed and challenging lives," and as being people who "often act impulsively and like meeting new people and doing exciting and stimulating things". The low C1 weight indicates that they also see themselves as flexible and have a casual attitude to guidelines, rules and regulations. It is these three factors (need for sensation and impulsivity, lack of concern for others, and lack of concern for moral and social rules) which, in combination, result in high-risk behaviors. The multiple regression analyses shows that three of the ICES scales can predict this with a correlation of R=0.52 (p<.001).

The overall PRI score, is correlated with C2 and SD. Those who say they would choose the higher risk options tend to be creative spontaneous people, who do not like to plan things in advance, and who do not like paying attention to details (low C2). The negative correlation with SD raises again the problem of interpreting scales of this sort (see Appendix G: Social Desirability). In combination with low C2, it is likely that this indicates genuine openness and frankness rather than a lack of concern over creating a good impression on this test. It is also possible, however, that the correlation with SD indicates a tendency for people who are concerned to give a good impression, responding by picking "safe" options on the PRI. If this were the case, we would expect SD to correlate negatively with Rated Risk (i.e. the person's perception of how risky the option they chose was). While there is a negative correlation, it is small and not significant (-.20).

The BPI risk-taking scale clearly has some overlap with the constructs being measured here. It is similar to both RISK_ACT and PRI-5 in its relationship with independence and with RISK_ACT and PRI-1, 2 and 5 in its relationship with conscientiousness. However, its correlations with ICES suggest that it is a less clearly focused scale than the risk measure included in the present study. BPI-RISK seems to show undifferentiated correlations with I and C and also has a positive S2 beta weight. Only PRI-5 shows any sign of having a positive relationship with stability.

### 16.1.3 Conclusions

Overall, the pattern of relationships between ICES and the risk measures presents a fascinating insight into some of the complexities of the concept of "risk". It shows that different combinations of ICES scales predict:

a. the tendency to indulge in behaviors which provide immediate reinforcement but which run a risk of serious long-term loss or damage;
b. choices people make between more or less risky options made in various type of everyday situations (where the consequences of "risky" actions are generally not life-threatening); and
c. differences in people's perceptions of how risky are the choices they make.

It would appear that, while there is some overlap, RISK_ACT and the PRI measure more specific aspects of risk than the BPI-RISK scale. It should be borne in mind that the correlations reported here considerably underestimate the

true-score correlations between ICES and the risk measures. The latter have low reliabilities, and hence correlations with them are subject to considerable attention. Improving the reliability of the measures of risk would increase the obtained validity of ICES as a predictor.

## 16.2 Personality and performance on complex, safety-critical tasks

Taking risks at work can manifest itself in a number of ways. In tasks that require attention to detail, failure to check one's work or rushing to finish in time are both "risky" strategies. Some evidence for the relationships between speed, accuracy and personality in the performance of complex tasks is provided by a small study on the design of a safety-critical system (Westerman, Shryane, Crawshaw and Hockey, 1995).

### 16.2.1 Method

Thirteen experienced engineers were asked to complete a complex work sample task. The task was a data preparation process that required engineers to express the necessary signaling principles for sections of railway track in a form which can be processed by signaling software and hardware. The system controlling each section of track is referred to as a Solid-State Interlocking System (SSI). In order to ensure that the data processed by each SSI is error free, a rigorous process of checking and testing is carried out. The work sample represented the data from an interlocking which had been prepared with 124 of the total 1006 lines of code removed. Performance on the task was broken down into knowledge-based and rule-based components and measured in terms of the time taken and errors made on each component.

In addition to the work sample, the engineers completed the ICES Personality scales and a number of other measures. The latter included tests drawn from the General Aptitude Test Battery (US Department of Labor, Employment and Training Administration, 1982) and a computer literacy subtest of the Computer Aptitude, Literacy and Interest Profile (Poplin, Drew and Gable, 1984).

### 16.2.2 Results

While the sample size is small (only 13 people), the results are of interest in that personality measures were found to be the strongest predictors of error. Of the ability measures, spatial ability correlated ($r=-.65$) with completion for the knowledge-based component of the task. The ICES scales I1 (Competitive) correlated significantly with both completion time for the rule-based component of the task ($r=-0.53$, $p<.05$) and for total completion time ($r=-0.57$, $p<.05$). Stability (S1) was negatively related to total number of errors ($r=-.65$, $p<.05$) and to knowledge-based errors ($r=-.56$, $p<.01$). Thus, it would appear that the more competitive individuals performed more quickly and stable, unruffled ones more accurately. There is also other data to support the view that stability may be associated with reduced error rates in the performance of inspection tasks (Hsu & Chan, 1995).

Extraversion (E1, Group-oriented) was positively associated with completion time ($r=0.56$, $p<.05$) for the knowledge-based component, and the Extraversion major scale has a 0.70 correlation ($p<.01$) with knowledge-based errors. This again, is mainly due to E1 ($r=0.81$, $p<.001$). Thus, introverts both perform faster and more accurately than extraverts.

### 16.2.3 Conclusions

While one must treat results from such a small sample with caution, the findings are consistent with expectations and the previously reported studies of risk. They are also surprising in showing such large effects within a highly selected group of skilled engineers. The results show that those who are high on I1 and who are Extravert tend to work faster and make more errors. In the performance of safety-critical tasks, errors are to be avoided – even if it takes more time to complete the task.

## 16.3 The new ICES Plus "Risk", and "Change" scales: Notes on interpretation

### 16.3.1 RISK scales

Data from these studies, together with the construct validation research reported in Chapter 15, provide the basis for a more complex and subtle analysis of risk, perceptions of risk and attitudes to risk than was previously possible with ICES Plus (see Appendix J.1 for details). One of the main ways in which risk-taking impacts on everyday working life, is through the choices people make between regular secure income which may be relatively low, on the one hand, and less dependable but potentially more profitable commission earnings on the other. People who seek

positions where their compensation is commission based rather than salary oriented may not perceive themselves as "taking a risk," but will be seen by those on fixed salaries as doing so.

Based on a consideration of issues relating to commission versus salary oriented individuals, a suite of composite scales have been developed to enrich and add value to the information generated by ICES Plus interpretative software.

There is a main scale on general preference for, or avoidance of, risk. This is supported by three sub-scales that focus on facets of risk-taking more specifically related to:

1. risky behaviors (whether a person is likely to take what are considered by most people to be risky actions);
2. expressed preference for risk (what levels of risk a person says he or she likes);
3. rated perceptions of risk (how safe/dangerous people see the world as being).

**16.3.1.1 Risk Main Scale: General Preference for (avoidance of) "Risk"**

The person who scores high on this composite scale is impulsive, outgoing, and enjoys taking chances. He or she is a "who dares wins" type of person. They prefer to work on commission rather than a fixed salary, liking the excitement of commission selling, and the risks involved in new business ventures. They get a buzz from pitting themselves against the world and winning out. They are not easily put off by failure or rejection, and always seem able to find a way around obstacles (even if it means bending a few rules). They are sociable and outgoing but also very independent.

On the down side, they could be a "loose cannon" in the wrong setting, and may be difficult to manage, especially in a team. People with a high level of general preference for risk need to be careful that the pleasure they get from taking risks, and living on the edge, does not lead to recklessness.

In the mid-range of scores, is the person who prefers the security of a reasonably paid job to one that is wholly dependent on commission. However, they would be prepared to take some degree of chance on profit sharing and would gamble on a portion of earnings being commission based. While they enjoy taking the occasional risk, they are unlikely to risk anything of real importance to them. They would always prefer to think things through when there might be something important at stake. Generally, such people are able to find a good balance between caution and "playing it safe," on the one hand, and taking calculated but necessary risks on the other.

At the low end of the scale are people who prefer a safe and secure existence, even if the payoffs are not great. They would prefer a regular modest salary and work in a stable environment, to the possibility of much higher but possibly insecure earnings. They are likely to be anxious in situations where there is the need to take a degree of risk, and will avoid such situations if at all possible. They do not act on impulse, but only after careful thought and consideration. As a result they may miss out on opportunities because of their cautiousness. However, such people are a "safe pair of hands".

**16.3.1.2 Risk sub-scale 1: Indulging in risky behaviors**

High scorers are likely to indulge in everyday behaviors that are risky (e.g. parking in no parking zones, speeding). They tend to take risks for their immediate gratifications, on impulse, without thinking too much about the possible down side. This is likely to be reflected in their behavior at work. They tend to work on the basis that the "ends justify the means". So long as the goals are organizational ones, this can produce creative tension and a willingness to take decisions and act in ways that others might shy away from. When the goals are personal ones, however, behavior may be counter-productive. High scorers may not see their behavior as being particularly risky (for them it is just how they are), but others tend to judge it so.

Those scoring in the mid-range, while occasionally taking a risk when other matters are pressing, would not normally indulge in risky behaviors. They will tend to think things through and balance the risks associated with various courses of action with the goals they are intended to achieve. They will tend to avoid unnecessary risk, or behaviors that could result in accidents or damage, but not to the extent of inaction. They will take considered risks when necessary.

At the low end, people are not likely to indulge in behaviors that have a risk attached to them. They would feel uncomfortable stopping in a "No Waiting" zone even for a few minutes. They may enjoy fantasizing about doing

exciting and dangerous things, but when it comes to it they would find excuses not to act. They think ahead and tend to see all the complications involved in doing things: what could go wrong rather than what the up side would be. They will only act after careful consideration and then only if the actions are safe. They will wait until the road is clear before crossing rather than risk dodging the traffic.

**16.3.1.3 Risk sub-scale 2: Expressed preference for risk**

High scores are associated with creative spontaneous people who do not like to plan ahead. They enjoy risk seeking and like to do things on the spur of the moment. They tend to be open and frank with people, irritated by details, preferring to take life as it comes. They expect people to "take me as you find me," and may, as a result, sometimes upset people in social gatherings or in discussion.

Mid-range scorers are fairly neutral in their preference for risk. Neither seeking high-risk excitement nor being careful to avoid any activity that might have a risk associated with it.

At the low end of the scale, people are careful and considered in their interactions. They are "risk-cautious," concerned on the one hand not to make a fool of themselves and on the other to be seen to be doing and saying the "right thing". They think ahead and try to check out all the possibilities so that no avoidable, unforeseen risks are taken.

**16.3.1.4 Risk sub-scale 3: Rated perceptions of risk:**

High scorers see the world as exciting but essentially safe. A fun place to be. They see the exciting things in life as like a roller-coaster ride – exciting, but not really dangerous. They will take actions that other people might judge as risky, but which they regard as being safe. They are accepting of people and events, and generally optimistic.

Mid-range scores are associated with people who see the world as a place containing some dangers, but neither being hostile nor particularly safe. They generally adopt a realistic appraisal of the risks associated with different activities.

At the low end of the score range, people see the world as a dangerous and hostile place, full of traps and dangers. As a consequence, they may appear somewhat timid to others, and over-cautious. They are not likely to do well in a volatile, changing environment, as it will make them anxious rather than excited. They may find it hard to come forward with their concerns and worries, even though they are generally fairly sociable and outgoing. They are less likely to indulge in "risky" behavior than others are, not because they see the benefits of success as being less, but because they perceive the risk of failure as being greater.

**16.3.2 Change Main Scale: Focus on change and innovation**

While people who are change-oriented tend, of necessity, to be willing to take risks, risk-taking and change-orientation are conceptually distinct. There will be some people, for example, who may score high on the general risk scale, but average or low on measures more closely focused around change. The present ICES Plus "Change" scale composite produces sten scores which can be interpreted as follows:

High scorers enjoy change and value innovation. They seek out new experiences and like to take control over events. They will often look for new ways of doing routine tasks, rather than following the usual practice. They do not like to have others impose structure on their work or working practices. They may sometimes go for change because it is more exciting rather than because it is better or necessary.

In the mid-range, people like a degree of order and regularity in their life, but also enjoy facing new challenges and changes. They do not seek change for its own sake, or because it is exciting, but because it may provide better ways of solving problems or dealing with issues. They are able to handle change and innovation when necessary. While they feel there is value in tried and tested ways of doing things, sometimes new ways are needed.

Those scoring at the low end of the scale prefer life to be well ordered, predictable and routine. They find change difficult to manage, and will try to adapt old ways to new demands where necessary, rather than going for a radical new innovation. They do not like too much variety in their work and like to work within a well-ordered and structured environment.

#### 16.3.3 Work Main Scale: Focus On Work

High scorers value work in itself, and place a high value on being employed. They see themselves as being defined as a person, by the job they do. They place great importance on social relations in and associated with work, rather than outside. In a crisis they will stay at work and see the job done, even if it means upsetting family and friends. Problems at home have to take second place, and home arrangements will always be changed in preference to changing arrangements at work.

People in the mid-range see both work and home life as being of value. Sometimes the attempt to retain a balance between these can create tensions. However, they will tend to prioritize work or out-of-work activities on the basis of their importance – sometimes placing a higher priority on one, sometimes the other. Their social relations will tend to be a mixture of people from work and those from leisure activities.

Persons scoring at the low end see work as a means to an end. They value leisure and home activities more and place more emphasis on family relations and friends outside of work. They tend to put less energy into work, preferring to put it into leisure activities and pursuits. In a crisis they will put family and friends before work. When conflicts arise between demands of work and home, family or friends, the latter will tend to take priority.

### 16.4 Personality and membership in University clubs and societies

Employers typically ask applicants for managerial positions about their interests and, for university graduates, which clubs or societies they belonged to when they were at university. There is a belief amongst employers that this information provides useful clues to an individual's personality. If this is so, then one should find that students who belong to societies differ from those who do not, in relation to their personality profiles, and that there are further differences relating to the type of society joined and the role played by the individual in that society.

#### 16.4.1 Method

A total of 80 students at the University of Hull were asked to complete the ICES inventory and asked which, if any, societies they belonged to. Fifty belonged to one or more societies while 30 did not belong to any.

#### 16.4.2 Results

Detailed results from the study are presented in Appendix J.2. Only one of the ICES scales differentiated significantly between those who choose to join societies from those who did not: C1. Thus, those who join are more willing to follow rules and abide by the group norms than are those who do not. People who joined societies also tended to be more competitive (I1) than those who did not. However, this is partly explained by differences in the types of society joined.

On its own, extraversion is not significantly correlated with society membership. However, using stepwise multiple regression, both C1 and E2 are selected as significant predictors, each contributing approximately equally to a predictor which correlates R =0.32 with membership ($F=4.42$, df 2 and 77, $p<0.05$; see Appendix J.3).

Differences between types of society joined are difficult to assess with the present sample size. However, it was possible to compare those who joined societies for social reasons or to do community work, with those who joined for sports or other activities. There was a highly significant relationship between I1 and type of society joined (eta=0.43; $F=8.60$, df 1 and 37, $p<.01$), with those joining sports societies being more competitive (I1=25.10) than those joining social/community (I1=20.78) societies. There was also a significant relationship with type of society for SD (eta=0.37; $F=5.94$, df 1 and 37, $p<.05$). Those joining sports societies had lower SD scores (SD mean = 20.52) than those joining the social/community societies (SD mean=24.28). It was also noticeable that those who joined sports societies tended to be more tense than those who joined social/community societies (S2=25.57 as opposed to 22.50: $F=3.40$, df 1 and 37, $p=0.7$, eta = 0.29). Together, I1 and SD correlated R=0.53 with type of society joined ($F=6.92$, df 2 and 37, $p<.01$; see Appendix J.3).

In short, people who join societies tend to be more outgoing and more conscientious "rule-followers" than those who do not. Amongst those who are members of a society, those joining groups for social or community work reasons tend to be less competitive and conform more to social expectations than those who join sporting and other activity groups.

## 16.5 General conclusions

The study of the relationship between ICES and measures of risk-taking, indicate the power of ICES both to predict these reported behaviors and its ability to discriminate between acts which have potentially serious long-term consequences; choices made in everyday situations where the risks are greater but the negative outcomes are much less serious; and variations in people's perceptions of risk. The three new composite ICES scales provide a useful addition to the interpretative richness of the ICES profile, representing combinations of Minor scales which have strong relationships with aspects of risk-taking.

Overall, the results from the engineers and those from the other risk studies suggests the need for people who are low I1, low E2 and high C2 for the performance of safety critical tasks.

The Change scale and the Focus on Work scale provide further interpretative depth to the use of ICES. While Change is correlated with Risk, (as in reality innovation always carries with it an element of risk), the two scales have different shades of meaning. Interpretation of these should be carried out carefully in relation to the descriptions provided earlier in this Chapter. Focus on Work tends to show negative correlations with Risk and Change. In terms of this scale, people who are overly focused on their work tend to be high on conscientiousness and fond of the status quo. They build their world around their work and they tend to want that world to remain stable and to be free from risk and change.

The study of personality differences in relationship to society and club membership provides further evidence for the construct validity of the ICES scales and reveals some interesting insights into the reasons why some people do and others do not join clubs.

# 17. CRITERION-RELATED VALIDITY AND FAIRNESS

This chapter focuses on issues of test fairness and criterion-related validity, particularly with respect to the ICES Plus Ability scales. It explores the issues surrounding adverse impact, indirect discrimination and validity. Recent data sets from ethnic minority groups are examined to provide information on the extent to which use of the ICES Plus Ability scales could impact - either adversely or positively - on such groups. The chapter concludes with data from some smaller validity studies.

## 17.1 Job-Related Validity and Ethnic Group Differences in Ability

There is widespread confusion and misunderstanding surrounding the concepts of adverse (or disparate) impact and indirect discrimination. Adverse impact occurs when selection is made from two groups of people on the basis of a measure for which the average scores of the two groups differ. Indirect discrimination occurs only when scores on the measure being used for selection are both not job-related and create adverse impact. This section of the Manual reviews these issues and the evidence relating to the impact of Ability tests on selection and their fairness (i.e. job-relevance). Without a general understanding of these issues, it is very easy to misinterpret information about any test of ability relating to differences in average scale scores between groups.

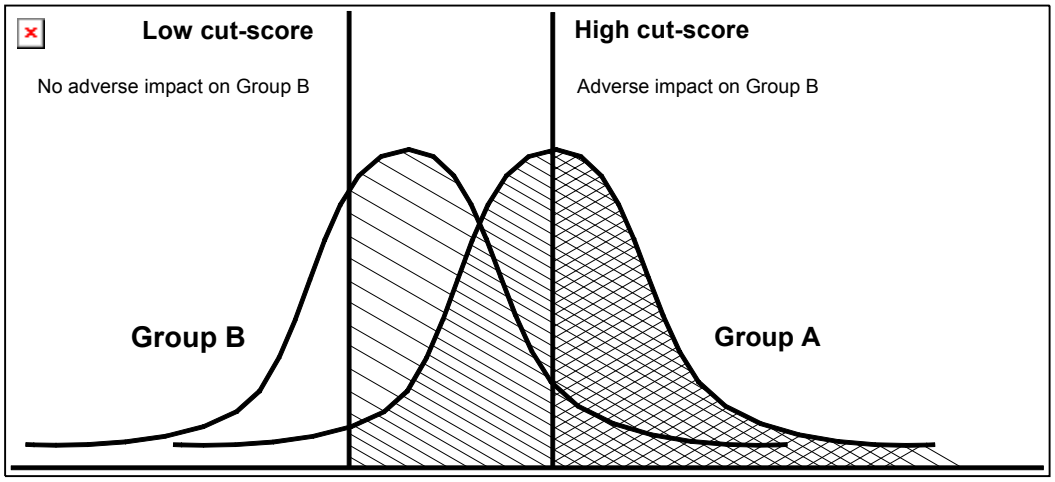
### 17.1.1 Effect sizes and adverse impact

Scores on any scale can be converted into a common or “standard” form by using the mean and standard deviation (SD) of the distribution. For example, if the distribution of scores on a test has a mean of 13 and a SD of 4, then a score of 19 on the test can be re-defined as being +1.5 SDs (i.e. it is one and a half SDs above the mean). Because the properties of normal distributions are well defined, it is common practice to convert individual scores on tests and differences between the average scores obtained by various groups into SD units.

The difference between two groups (e.g. white and black job applicants) is measured by seeing how many SDs apart are the averages for the two groups. For example, if the SD is 4 and the groups have average scores of 12.5 and 14, then the difference between them would be 1.5/4, or 0.375 SD. Differences between groups, measured in SD units, are referred to as “effect sizes”.

Adverse impact occurs when the proportion of people in one group succeeding in a selection procedure is significantly lower than that for some other group. Where there is a difference in average score between two groups, the nature of test score distributions is such that, if the test is used as a screening device with a low cut off point (i.e. a high “pass” rate), there will be less adverse impact than if a higher cut off is used. The higher the cut off is placed, the greater the difference in selection rates becomes (Bartram, 1995a). Adverse impact is a joint function of (1) the difference between two groups in average score levels and (2) the overall effective “pass” rate set for the test.

***Figure 17.1: Effect of selection ratio on adverse impact***



## 17.2 Differences on the ICES Plus Scales Between Black and White Groups Matched by Job Type

The results of validity generalization research have confirmed not only the validity of Ability tests as predictors of future job performance, but also the potential cost-benefits of using them in selection. However, significant standard deviation differences (one to one and a half SDs or more) between ethnic, gender or other distinctive groups are almost certain to result in a degree of adverse impact if the tests are used as screening devices or make a significant contribution to selection decisions.

The ICES Plus Ability scales were constructed with a view to minimizing ethnic-related between-group differences, while retaining validity as job-related measures of ability. If effect sizes could be limited to one-half SD the risk of creating adverse impact when using Ability tests in selection would be greatly reduced, if not eliminated.

### 17.2.1 Method

In order to properly measure effect sizes, it is necessary to match Black and White test takers in terms of other relevant demographic variables (such as age, gender and job type). Without this matching, differences between White and Black groups could arise due to factors other than their ethnic group membership.

A set of ICES Plus test results was collected from Black test takers working in a variety of occupations. This data was searched for sets of people who could be classified by job type into reasonable sized groups (of around 20 or more). Five such groups were identified. The Phase III data set was then searched to identify groups of White test takers in similar jobs, with the requirement that, again, the group sizes needed to be at least 20 people. It was possible to identify five job areas where there were sufficient data on Black test takers. These were:

1. Clerks and secretaries (n=80);
2. Sales representatives (n=51);
3. Managers (n=35);
4. Supervisors (n=20);
5. Analysts/programmers (n=19).

### 17.2.2 Results

Means and SDs for each group for all the ICES Plus scales are given in Appendix J.1. There were substantial differences in Ability test scores between the five job type groups. Average raw scores on ICES Plus Ability, for example ranged from 55.69 for the sales representatives up to 72.70 for the supervisor group. Indeed, the managers, supervisors and analyst/programmers all had average Ability scores similar to each other (with GEN in the region of 70) and to the Phase III norm (see Table 9.6).

Differences between the five job type groups were statistically significant for both the Numerical and Non-Verbal scales, but not for the Verbal scale. Other ICES Plus scales on which the groups showed differences were Data, Things, I2 and E2.

It was only possible to carry out analyses of White-Black group differences on the first four of the five job-related areas as there were no White analyst/programmers in the Phase III sample. The sample sizes of the White groups for the four job areas were:

1. Clerks and secretaries (n=93);
2. Sales representatives (n=36);
3. Managers (n=72);
4. Supervisors (n=25);

Means and SD for the White and Black groups are presented in Appendix J.2. Table 17.1 shows these expressed as effect sizes for all of the ICES Plus scales.

***TABLE 17.1: White-Black group differences (see Table Appendix J.2 for details) expressed as effect sizes. Positive effects indicate that the White group has a higher average than the Black group, negative effects indicate that the Black group has a higher average than the White group. Differences between the two groups for each scale were examined for statistical significance using t-tests.***

| | Clerks/Secs | | | Sales reps | | | Managers | | | Supervisors | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Scale** | **d** | **T** | | **d** | **t** | | **d** | **t** | | **d** | **T** | |
| **VERB** | 0.04 | 0.24 | | 0.04 | 0.18 | | 0.06 | 0.27 | | -0.69 | -2.42 | * |
| **NUM** | 0.45 | 3.01 | ** | 0.83 | 4.16 | *** | 0.43 | 2.11 | * | -0.53 | -1.80 | * |
| **NON-V** | 0.48 | 3.23 | *** | 0.71 | 3.46 | *** | 0.46 | 2.28 | * | -0.03 | -0.09 | |
| **GEN** | 0.40 | 2.68 | ** | 0.64 | 3.09 | ** | 0.39 | 1.89 | | -0.50 | -1.68 | |
| **PEOPLE** | -0.59 | -4.03 | *** | -0.34 | -1.59 | | -0.67 | -3.40 | *** | -0.58 | -1.99 | |
| **DAT** | -0.71 | -5.00 | *** | -0.32 | -1.49 | | -0.35 | -1.72 | | -0.90 | -3.31 | ** |
| **THINGS** | -0.37 | -2.44 | * | 0.03 | 0.14 | | 0.10 | 0.48 | | -0.07 | -0.25 | |
| **I1** | -0.77 | -5.42 | *** | 0.19 | 0.89 | | -0.16 | -0.76 | | 0.03 | 0.10 | |
| **I2** | 0.12 | 0.77 | | 0.11 | 0.49 | | -0.08 | -0.36 | | -0.32 | -1.06 | |
| **C1** | -0.12 | -0.76 | | 0.12 | 0.53 | | 0.02 | 0.11 | | 0.08 | 0.25 | |
| **C2** | -0.60 | -4.08 | *** | -0.36 | -1.69 | | -0.15 | -0.72 | | -0.76 | -2.72 | ** |
| **E1** | 0.17 | 1.14 | | 0.28 | 1.31 | | -0.08 | -0.38 | | -0.23 | -0.76 | |
| **E2** | 0.45 | 3.01 | | 0.20 | 0.93 | | -0.14 | -0.68 | | 0.07 | 0.23 | |
| **S1** | -0.16 | -1.04 | | 0.07 | 0.31 | | 0.02 | 0.11 | | -0.59 | -2.05 | * |
| **S2** | -0.07 | -0.46 | | 0.31 | 1.42 | | 0.19 | 0.93 | | -0.46 | -1.55 | |
| **SocDes** | -0.47 | -3.19 | ** | -0.41 | -1.91 | | -0.51 | -2.52 | * | -0.91 | -3.39 | ** |
| **I** | -0.35 | -2.32 | * | 0.17 | 0.78 | | -0.14 | -0.66 | | -0.20 | -0.65 | |
| **C** | -0.47 | -3.18 | ** | -0.18 | -0.82 | | -0.08 | -0.40 | | -0.46 | -1.54 | |
| **E** | 0.35 | 2.34 | * | 0.26 | 1.20 | | -0.13 | -0.62 | | -0.07 | -0.23 | |
| **S** | -0.13 | -0.83 | | 0.21 | 0.94 | | 0.12 | 0.60 | | -0.59 | -2.05 | * |
| **Df** | 171 | | | 85 | | | 105 | | | 43 | | |
| *** p<.05; ** p<.01; *** p<.001** | | | | | | | | | | | | |

The main findings can be summarized as follows:

**17.2.3 Ability**

In three of the four job areas, the White group averages are higher than the Black group's. Apart from the Sales Rep groups, the effect sizes are all less than 0.50 SDs. For the Sales reps, NUM and NON-V both show effect sizes greater than this. However, even here, the effect is well under one SD.

The Verbal scale shows no ethnic group differences for the first three groups.

The Black supervisors have average scores around 0.5 SD higher than the White supervisors, with the difference being significant for the Verbal and Numerical scales.

These results show the danger of assuming that Black applicants, as a group, will always tend to score lower than White applicants. The effects sizes for general ability in fact range from 0.64 to -0.50 SDs.

The results confirm that where there are differences between groups, for ICES Plus these effects tend to be a lot smaller than what has been established in research literature which concludes that ethnic minority groups tend to perform at lower average levels on a wide range of standardized tests of ability and attainment (Arvey and Faley, 1988; Shuey 1966; Dredger and Miller, 1960:1968). Indeed, it would appear that the Verbal scale, used on its own, would not give rise to adverse impact.

**17.2.4 Personality and interests**

Where scales do show differences between groups, the Black groups tend to score higher than the White groups: more interested in People and Data, and more Independent and Conscientious. However, it should again be noted that the pattern varies from job to job.

## 17.3 A Large-Scale 1995 Case Study

The Case Study was carried out in 1995 for a major US organization. Amongst other things, the organization wanted to know that, if used as part of their selection process, the tests would not discriminate unfairly on the basis of gender or against the members of any minority group.

### 17.3.1 Method

ICES Plus data were available for a total sample of n=1206 employees, drawn from a number of sites across the USA. The sample contained about 60% males. Of the 1206 people, 987 (81.8% of the sample) were White, 30 (2.5%) were Black, 54 (4.5%) were Hispanic, 35 (2.9%) were Oriental and one was an American Indian. The average age of the sample was 26.70 (SD=4.45, n=1105). Average age did not differ significantly as a function of either gender or race.

The samples were drawn from all parts of the US. For analysis purposes, these were grouped into the 16 regional areas used by the organization. Additional data from two regions (n=48) were obtained following completion of the analyses of the main sample of data. This sub- sample was used for cross-validation purposes, and then combined with the main sample for validity generalization analysis of differences between geographical areas.

Criterion data (manager nominations, sales performance, sales targets and other data) were collected by the organization from a number of sites across the USA. For analysis purposes, sufficient criterion data was available for 520 of the people in the total sample.

### 17.3.2 The ICES Plus Scales

Only two of the ICES Plus scales showed any statistically significant effects of ethnic group: C2 and Non-Verbal Ability. For C2, both Blacks and Hispanics had higher scores than Whites, with Orientals being between the two. The absolute size of these effects, however, was very small - accounting for less than 1% of the variance in scores on that scale. While the largest differences in Ability test scores were between the White and Black groups, these were small in terms of absolute effect size: 0.16, 0.30, 0.53 and 0.41 for VERB, NUM, NON-V and GEN respectively. The Hispanic and Oriental groups' scores were very similar and only slightly lower than those for the White group (less than 0.20 SDs in both cases).

Means and SDs for the four groups are presented in Appendix J.3. This case study confirmed that the ICES Plus Ability scales show far smaller between-group difference than the academic literature reports as typical.

While the effects of ethnic group are quite small, there are more substantial differences on some scales in relation to gender. In line with previous findings, males score higher on I1 and Interest in Things, while females score higher on Interest in People. For the present sample, males score higher than females on Numerical and Non-Verbal but there is no difference on Verbal. In addition to scoring higher on I1 (but not I2), males also score higher than females on S2, with females scoring higher than males on E1.

Within the sample, with increases in age, scores tend to be lower on Non-Verbal, I1, E1 and E2 and higher on C1 and S1. With increases in years of service in the organization, the main changes are lower Interest in People and higher C1 scores.

On almost all of the measures, this sample shows less variation in their scores than did the Phase III sample. This is to be expected, given that the present sample is a selected group all of whom have been accepted for the same type of work. This range restriction is most marked for scales I2, E1, E2, People and Ability.

Range restriction is important for a number of reasons. First it is indicative of those qualities that affected the selection of people in the first instance - the more restricted the range, the greater its influence on selection. Second, it affects other measurements - the correlation between a range-restricted measure and some other measure will always underestimate the true correlation between the two variables. This has a direct effect on two important indices: reliability and validity.

Comparisons between the alpha reliabilities and SEMs for this sample and Phase III sample showed that, while alpha values are reduced as expected with range restriction, the SEM values for the present sample are comparable to those from Phase III. In fact they tend generally to be smaller for the Personality and Interest scales, indicating a higher accuracy of measurement than estimated from the Phase III data.

### 17.3.3 1995 Case Study Conclusions

- The ICES Plus Ability scales for the 1994 Phase III sample described earlier in this manual, were fully supported by the data from this larger, later sample.
- This case study confirmed that the ICES Plus Ability scales show far smaller between-group differences than the current academic literature reports.

## 17.4 The Validity of Ability Tests as Predictors of Job Performance

Prior to the work of Schmidt and Hunter in the late 1970s (see Hunter & Schmidt, 1990), it had been thought that each selection situation was unique and that is was, therefore, necessary to validate the use of tests in every situation they were to be used. This view appeared to be supported by the fact that the validity coefficients obtained for Ability tests (and other measures) varied from situation to situation. Schmidt and Hunter revolutionized the traditional view by demonstrating that most of this variation was in fact accountable for by random factors - largely due to the generally small samples sizes used in validation studies, differences in the reliability of criterion measures and so on. They demonstrated through a series of meta-analytic studies that it was possible to estimate the true validities which underlie the varying values obtained from different samples, and that these true values were fairly stable across jobs, across employing organizations, and across tests.

The results of their work were first given approval in the US Courts in *Pegues v. Mississippi State Employment Services* (1980), where the judge stated:

> *"Empirical research has demonstrated that validity is not perceptibly changed by differences in location, differences in specific job duties or applicant populations. Plaintiff's allegations that validity is specific to a particular location, a particular set of tasks and to a specific applicant population... is not true".* [p254]

Since 1980, a wealth of meta-analysis studies by Schmidt and Hunter and others have resulted in the concept of "validity generalization" becoming the accepted view within the scientific community. This, in turn, has produced a change in the emphasis placed on test validity. Far more emphasis is now placed on establishing the construct validity of a test - i.e. that it really measures the ability or aptitude which it purports to measure - than on the conduct of small local validation studies. If a test has good construct validity, then the accumulated evidence provided by validity generalization studies tell us the level of predictive validity we can expect it to have in a known range of types of selection situation.

Pearlman, Schmidt and Hunter (1980) carried out a meta-analysis on 56 distributions of validity coefficients from 698 studies representing five families of clerical jobs and two classes of criteria: job proficiency and training success. In total, their data covered over 70,000 people. Their results showed that most of the variation in validity coefficients between studies could be accounted for by statistical artifacts; that generalization of validity to similar clerical jobs or to new settings was justified and that the tests were predictive of both classes of criteria (which were highly correlated with each other). More specifically, they provided tables giving information about the levels of validity one would expect to obtain if using particular types of test for selection in jobs covering various areas of clerical work. Grouping all clerical jobs together, their results showed:

- Tests of Verbal Ability (like the ICES Plus Verbal scale) have a mean true validity of 0.39 with job proficiency as the criterion and 0.64 with training as the criterion.
- Tests of Quantitative Ability (like the ICES Plus Numerical scale) have a mean true validity of 0.47 with job proficiency as the criterion and 0.70 with training as the criterion.
- Tests of Reasoning ability (like the ICES Plus Non-Verbal scale) have a mean true validity of 0.39 both with job proficiency as the criterion and with training as the criterion.

While the use of measures of ability is known to lead to the risk of adverse impact (especially if relatively high cut-scores are set), such measures are also known to have good levels of relevance in terms of the prediction of work performance.

What is more, the US research evidence referred to earlier shows that cognitive ability measures tend to over- rather than under-predict the performance of (at least in the US) Black minority group members. The results of the Case Study reported in this Chapter are consistent with this general finding, and also show that even variations in validity for complex measures are due mainly to non-systematic sampling effects. In short, far more reliance can be based on generalized estimates of validity than was thought to be the case in the past.

### 17.4.1 Variation in validity coefficients between areas in the Case Study data: random error or situational specificity?

With the additional 48 people from the cross-validation sample, the number of people with criterion data increased from 520 to 568. When broken down by geographical sales area, the validity coefficients varied considerably. For SALES, validities range from r=-.05 to 4=0.40 (corrected values: -.06 to 0.48) between the various areas. Validity generalization analysis (Hunter and Schmidt, 1990) can be used to assess the degree to which variations in

correlations between the Areas can be attributed to sampling error as opposed to their reflecting real differences in validity.

The validity generalization analyses were carried out (using Hunter and Schmidt's VG6 software package, 1994 version) with the reliability of the ICES Plus score estimated at 0.90 and criterion reliability was assumed to be either 0.80 or 0.50. These figures represent reasonable estimates of the upper and lower limits for the criterion measures. As discussed earlier, there is no data available from which actual estimates can be computed.

The results suggest that the true score correlation for SALES is between 0.35 and 0.41. Furthermore, the variation in correlations between regions is either the same as or less than the amount one would expect simply on the basis of sampling error. Thus it would appear that variations in validity coefficients between regions represent random variations rather than situational specificity. As a consequence, one can be confidant that ICES Plus has good generalisability as a predictor across all the regions sampled.

## 17.5 Confirmation of Validity of the ICES Plus Ability Scales

The above discussion makes clear the central importance of construct validity. Once we know what a test measures, we are able to make strong predictions about the range of situations in which it will be predictive of various outcome criteria. Without evidence of construct validity, such generalizations are dangerous. It is for these reasons that so much emphasis has been placed on establishing the relationships between the ICES Plus Personality and Interest measures and other instruments.

While the ICES Plus Ability scales are designed primarily for use in occupational assessment, evidence of their validity as measures of ability can be obtained from a variety of sources. One of the key requirements for such scales is that they should show clear relationships with accepted measures of general ability and with measures of academic attainment - which, in turn, are known to be highly correlated with general mental ability. Three sets of data provide evidence relating to the ICES Plus Ability scales' construct validity as measures of general mental ability and their potential as predictors of academic achievement.

### 17.5.1 Prediction of school achievement

The first set of data involved 89 young people (age range 15 to 23, average 17.09 years) who had completed high school in the UK, but who had not progressed on to higher education (university or college). In addition to their ICES Plus scores, data was available on their performance in the national GCSE examination in mathematics. Average scores for the group, together with the Phase III norms for comparison are shown in Table 17.2 together with correlations between maths grades and ICES Plus Ability scores. Examination of the means also confirms the view that this sample is below average in ability (relative to the Phase III norm group) and shows less variation (i.e. the smaller SDs indicate a natural restriction of range).

As one would expect, the highest correlations for performance in maths are with Numerical and Non-verbal reasoning test scores. When corrected for the effects of range restriction in the sample (using the Phase III SDs to estimate un-restricted range), these give overall validity coefficients of 0.64 for GEN as a predictor of maths performance. This is an underestimate of the true validity, as there has been no correction for range restriction in the criterion, which will be considerable in the present instance.

***Table 17.2: Relationship between ICES Plus Ability test scores and Maths grades obtained in the General Certificate of Secondary Education examinations by UK school leavers.***

| | Phase III n=516 | | School leavers n=89 | | Correlation n=89 with GCSE maths grades | | |
|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | [a] | | [b] |
| VERB | 24.03 | 9.90 | 20.97 | 7.94 | 0.14 | | 0.17 |
| NUM | 10.85 | 5.59 | 7.46 | 3.51 | 0.33 | ** | 0.49 |
| NON-V | 12.14 | 4.62 | 13.31 | 4.28 | 0.47 | ** | 0.50 |
| GEN | 70.13 | 25.52 | 62.52 | 16.08 | 0.46 | ** | 0.64 |

[a] unadjusted correlations: ** p<.001.
[b] adjusted for ICES Plus scale range restriction relative to Phase III sample.

### 17.5.2 Differences between school leavers and university students

The second set of data was obtained from a class of 40 Psychology undergraduate students. These would all have been required to obtain high grades in their final Advanced-level secondary education examinations in order to be selected for University. Furthermore, one would expect to find that the main factors discriminating them from those who have not gone on to University would be Verbal, and to a lesser degree, Numerical ability. Table 17.3 shows the results for these 40 students together with those of the 89 school leavers (from Table 17.2). Differences between these two groups are substantial, with effect sizes of one or more SDs for ICES Plus Ability, Verbal and Numerical ability, and just over half an SD for Non-verbal. These differences are all highly significant and represent validity coefficients ranging from 0.29 to 0.57A

***Table 17.3: ICES Plus Ability scale differences between school leavers and university students.***

| | University Students n=40 | | School leavers n=89 | | Differences | | Validity |
|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Effect Size | F ratio | (eta) |
| VERB | 32.53 | 8.00 | 20.97 | 7.94 | 1.21 | 58.19 | 0.56 *** |
| NUM | 11.83 | 5.46 | 7.46 | 3.51 | 0.94 | 29.70 | 0.44 *** |
| NON-V | 15.95 | 3.54 | 13.31 | 4.28 | 0.62 | 11.57 | 0.29 *** |
| | | | | | | | |
| GEN | 88.07 | 19.30 | 62.52 | 16.08 | 1.23 | 61.40 | 0.57 *** |
| *** p<.001, df 1 and 127 | | | | | | | |

### 17.5.3 Relationship between ICES Plus and Raven's Standard Progressive Matrices Study

The third set of data (Longley, 1996) explored the relationship between the ICES Plus scales and Raven's Standard Progressive Matrices (SPM). A recent survey of test use, covering 44 countries (Oakland, 1995), showed that the SPM was the second most widely used test of general ability in the world (the Weschler Intelligence Scale for Children being the most widely used). SPM is frequently used as a benchmark measure of fluid intelligence. It consists of a series of problems in the form of incomplete matrices or patterns. Alternatives for the missing section of each matrix are provided and the candidate has to select the one, which fits the pattern. While "abstract" in nature, it has been shown to provide a fairly pure measure of "g" or general ability.

Fifty-six people took part, 27 females and 29 males. Their ages ranged from 17 to 62 (mean 39) and they were drawn from a variety of occupational categories including manual unskilled, manual skilled, secretarial, administration, high-level technical and managerial. For 39 of the participants, information was also available on whether or not they were graduates. There was a higher proportion of females in the non-graduate group (9 out of 15 were female) than the graduate one (9 out of 24).

The mean SPM score for the group as a whole (see Table 17.4) was well above the average reported in the Manual for people in the age range 30-39. This indication of above-average ability is supported by the ICES Plus data. The mean raw score on ICES Plus Ability was 81.25 (SD=26.9) for this group, as against 70.15 (n=516) for the Phase III normative group. The present group was well above average on their Verbal (31.04 as opposed to 24.03 for the Phase III sample) and Non-Verbal (13.16 as opposed to 12.14 for the Phase III sample) scores, but below average on their Numerical scores (11.95 as opposed to 14.73 for the ICES Plus norm group).

***Table 17.4: Means and SDs for ICES Plus and SPM (n=56), and correlations between SPM and ICES PLUS scales.***

| Scale | Mean | SD | Correlations | | |
|---|---|---|---|---|---|
| SPM | 52.12 | 5.51 | zero order | | Estimated true |
| ICES Plus | | | | | |
| Verbal | 31.04 | 10.84 | 0.57 | ** | 0.63 |
| Numerical | 11.95 | 5.63 | 0.61 | ** | 0.69 |
| Non-Verbal | 13.16 | 4.72 | 0.53 | ** | 0.63 |
| General | 81.25 | 26.91 | 0.67 | ** | 0.74 |
| ** p<.01 | | | | | |

Correlations between the ICES Plus scales and SPM were uniformly high and significant. The results from these three sets of data provide strong support not only for the validity of GEN as a measure of general ability, but also for the differential validity of the three scales. The Verbal scale, for example, has the highest validity as a predictor of University entrance, while the other scales have the highest validities as predictors, specifically, of performance in mathematics.

The actual magnitude of the validities reported are high, well in line with the levels one would want to see for establishing the construct validity of a scale. With corrections for attenuation, the estimated correlation between true scores is 0.74 for SPM and the ICES Plus Ability measure. (This correction uses reliability estimates taken from this manual for the ICES Plus scales and a coefficient of 0.88 for the SPM).

Consistent with the results of the second study, differences between graduates and non-graduates were predicted by both tests. SPM accounts for 20.28% of the variance between groups, while ICES Plus accounts for 28.42%. These represent highly significant correlations between test scores and group membership of 0.45 and 0.53 respectively. Of the ICES Plus scales, Numerical and Non-Verbal are the best predictors of whether someone is a graduate or not (r=0.52 and r=0.59 respectively).

**17.5.4 Conclusions**

The results from these three sets of data provide strong support not only for the validity of the ICES Plus Ability measures as a measure of general ability, but for the differential validity of the three scales. The Verbal scale, for example, has the highest validity as a predictor of University entrance, while the other scales have the highest validities as predictors, specifically, of performance in mathematics.

While the sample sizes for each study are not large, the results are consistent and the actual magnitude of the validities reported are high. These are well in line with the levels one would want to see for establishing the construct validity of a scale.

## 18. REFERENCES


- Arvey, R.D. & Faley, R.H. (1988) *Fairness in Selecting Employees,* Second edition. Reading, Mass: Addison-Wesley.
- Barrick, M.R. & Mount, M.K. (1991). The Big Five personality dimensions and job performance: a meta-analysis**.** *Personnel Psychology, 44***,** 1 -26.
- Bartram, D. (1992) The Personality of UK Managers: 16PF norms for short-listed applicants. *Journal of Occupational and Organizational Psychology, 65,* 159-172.
- Bartram, D. (1995a). Predicting adverse impact in selection testing. *International Journal of Selection and Assessment,*3, 52-61
- Bartram, D. (1995b). Validation of the Micropat Battery. *International Journal of Selection and Assessment, 3*, 84-95.
- Cattell, R.B., Eber, H.W., & Tatsuoka, M.M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16PF).* Champaign, Illinois: IPAT Inc.
- Clough, P., & Hockey, G.R.J.H. (1996) *The Personal Risk Inventory.* Department of Psychology, University of Hull.
- Costa, P.T. & McCrae, R.R. (1985). *The NEO Personality Inventory Manual*. Odessa, FL.: Psychological Assessment Resources.
- Costa, P.T. & McCrae, R.R. (1989). *The NEO-PI/FFI manual supplement*. Odessa, FL.: Psychological Assessment Resources.
- Coyne, I., & Bartram, D. (1998). The relationship between the Prevue ICES and dishonest intention. *Prevue Assessments Technical Report,* 20-2-98. Vancouver, B.C.: Prevue Assessments International Inc.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41,* 417-440.
- Dredger, R.M. & Miller, K.S. (1960). Comparative psychological studies of Negroes and whites in the United States, *Psychological Bulletin, 57*, 361-402.
- Dredger, R.M. & Miller, K.S. (1968). Comparative psychological studies of Negroes and whites in the United States, *Psychological Bulletin, 70,* 1-58.
- Eysenck, H.J. (1952). *The Scientific Study of Personality,* London: Routledge and Kegan Paul.
- Eysenck, H.J. (1959). *Manual of the Maudsley Personality inventory.* London: University of London Press.
- Eysenck, H.J. & Eysenck, S.B. (1964). *Manual of the Eysenck Personality Inventory.* London: University of London Press.
- Eysenck, H.J. & Eysenck, S.B. (1975). *Manual of the Eysenck Personality Questionnaire.* London: University of London Press.
- Eysenck, H.J. & Eysenck, S.B. (1991). *The Eysenck Personality Questionnaire – Revised (EPQ-R)*. London, UK: Hodder & Stoughton.
- Feltham, R., & Woods, J. (1995). *The Business Personality Indicator (BPI) User's Guide.* Windsor: NFER-NELSON.
- Ghiselli, E.E. (1973). The validity of aptitude tests in personnel selection. *Personnel Selection, 26,* 461-477.
- Guion, R.M. (1965). *Personnel Testing* New York: McGraw-Hill.
- Hogan, R. & Hogan, J. (1995). *Hogan Personality Inventory Manual Second Edition.* Hogan Assessment Systems, Tulsa.

- Hogan, R. & Hogan, J. (1998). *Hogan Development Survey. Technical Manual, UK Edition.* Tunbridge Wells: Psychological Consultancy Limited.
- Horney, K. (1950). *Neurosis and Human Growth*. New York: Norton
- Hollinger, R and Clark, J (1983). *Theft by employees*. Lexington, MA: Lexington Books.
- Hunter, J. E. and Schmidt, F.L. (1990) *Methods of meta-analysis: correcting error and bias in research findings.* Newbury Park, Ca: Sage Publications.
- Hunter, J.E., Schmidt, F.L. & Hunter, R. (1979(. Differential validity of employment test by race: A comprehensive review and analysis. *Psychological Bulletin, 86*, 721-735.
- Kline, P. (1991) *Intelligence: the Psychometric view.* London: Routledge.
- Hsu, S-H & Chan, T. (1995). A study of inspection performance and personality of quality inspectors. *International Journal of Quality and Reliability Management. 11, 55-85.*
- Longley, K.M. (1996). *Construct validity of the Prevue General Ability Scales.* Hull University Masters thesis.
- McHenry, J.J., Hough, L.M., Toquam, J.L., Hanson, M.A. & Ashworth, S. (1990). Project A validity results: the relationship between predictor and criterion domains. *Personnel Psychology, 43,* 335-354.
- McRae, R.R. & Costa, P.T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.
- Mirels & Garrett, (1971). Cited in Cook, J.D. et al (1961). *The Experience of Work: A Compendium and Review of 249 Measures and their Use.* London: Academic Press.
- Oakland, T., (1995). 44-country survey shows international test use patterns. *Psychology International,* Winter, p7.
- Ones, D.S., Viswesvaran, C. and Schmidt, F.L. (1992). An examination of validity generalisation and moderators of integrity test validities. In F.L.Schmidt (chair), *Criterion-related validity, construct validity and legality of integrity tests.* Symposium at the Society for IO Psychology, Montreal, Canada, May 1992.
- Organ, D.W. (1988). *Organizational Citizenship Behaviour: The Good Soldier Syndrome.* Lexington, MA: Lexington Books.
- Organ, D.W. (1994). Personality and Organizational Citizenship Behaviour. Journal of Management, *20*, 465-478.
- Pearlman, K., Schmidt, F.L., & Hunter, J.E. (1980). Validity generalisation results for tests used to predict job proficiency and training success in clerical occupations, *Journal of Applied Psychology, 65*, 373-406.
- *Pegues v. Mississippi State Employment Services*, 22 FEP 392 (1980) – cited in Arvey & Faley (1988).
- Poplin, M.S., Drew, D.E. & Gable, R.S. (1984). *Computer Aptitude, Literacy and Interest Profile.* Austin, Texas: Pro-Ed.
- Reilly, R.R. & Chao, G.T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology, 35,* 1-62.
- Robertson, I. (1993). Personality assessment and personnel selection. *European Review of Applied Psychology, 43,* 187-194.
- Robertson, I. and Kinder, A (1993). Personality and job competences: the criterion-related validity of some personality variables. *Journal of Occupational and Organizational Psychology, 66,* 225-244.
- Ruch, W.W. (1972). A re-analysis of published differential validity studies. Symposium paper presented at the APA, Honolulu, September. (cited in Arvey & Faley).

- Sackett, P.R., Burris, L & Callahan, C. (1989). Integrity testing for personnel selection: A review and critique. *Personnel Psychology, 42,* 491-528.
- Schmitt, N., Gooding, R.Z., Roe, R.A. & Kirsch, N. (1984). Meta analysis of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology, 37,* 407-422.
- Schmitt, N. (1989). Fairness in employment selection. In Smith, M. and Robertson, I (Eds). *Advances in Selection and Assessment.* New York: Wiley.
- Shuey, A.M. (1966). *The Testing of Negro Intelligence.* 2nd edition. New York: Social Science press.
- Tett, R. P., Jackson, D.N. & Rothstein, M. (1991) Personality Measures as Predictors of Job Performance: a Meta-Analytic Review. *Personnel Psychology, 44,* 703-742.
- Viswesvaran, C.(1992). Appropriateness of self-reported criteria for integrity test validation: evidence from beyond. In F.L.Schmidt (chair), *Criterion-related validity, construct validity and legality of integrity tests.* Symposium at the Society for IO Psychology, Montreal, Canada, May 1992.
- Waters, !. & Whetton, C. (1986). *Differential Aptitude Tests Form W British Manual.* Windsor: The Psychological Corporation and NFER-NELSON.
- Westerman, S.J., Shryane, N.M., Crawshaw, C.M. & Hockey, G.R.J. (1995). Error analysis of the Solid State Interlocking Design Process. *Human Factors in the Design of Safety Critical Systems Work Package 1.2.* Report # SCS/04, Human Factors Group, Department of Psychology, University of Hull.
- Woolley, R.M. and Hakstian, A.R. (1993). A comparative study of integrity tests: The criterion-related validity of personality-based and overt measures of integrity. *International Journal of Selection and Assessment, 1,* 27-40.


Some of the information contained in this manual has been published in:


- Bartram D. (1993). Validation of the ICES personality inventory. *European Review of Applied Psychology, 43,* 207-218.

# PREVUE ASSESSMENT ™

# TECHNICAL MANUAL

# APPENDICES

Produced for View Assessments International Inc.
By

Professor David Bartram

Newland Park Associates Ltd.
Chartered Occupational Psychologists

Fourth Edition: April 2000

***Appendix A.1: Characteristics of the combined PHASE ONE and PHASE TWO sample***

| GENDER<br>Value Label | Code | Frequency | Percent | Valid<br>Percent |
|---|---|---|---|---|
| Male | 1 | 1946 | 58.1 | 58.1 |
| Female | 2 | 1402 | 41.8 | 41.9 |
| Missing values | 0 | 3 | .1 | |
| | Total | 3351 | 100.0 | 100.0 |
| AGE | | | | |
| Under 20 | 1 | 34 | 1.0 | 1.0 |
| 20-24 | 2 | 77 | 2.3 | 2.3 |
| 25-29 | 3 | 920 | 27.5 | 27.5 |
| 30-34 | 4 | 751 | 22.4 | 22.4 |
| 35-39 | 5 | 629 | 18.8 | 18.8 |
| 40-44 | 6 | 468 | 14.0 | 14.0 |
| 45-49 | 7 | 254 | 7.6 | 7.6 |
| 50-54 | 8 | 206 | 6.1 | 6.1 |
| 55-59 | 9 | 7 | .2 | .2 |
| 60 or over | 10 | 5 | .1 | .1 |
| | Total | 3351 | 100.0 | 100.0 |
| ETHNIC ORIGIN | | | | |
| White | 1 | 2532 | 75.6 | 75.6 |
| Black | 2 | 275 | 8.2 | 8.2 |
| Asian | 3 | 201 | 6.0 | 6.0 |
| Indian | 4 | 38 | 1.1 | 1.1 |
| Oriental | 7 | 205 | 6.1 | 6.1 |
| Hispanic | 10 | 80 | 2.4 | 2.4 |
| Other European | 9 | 3 | .1 | .1 |
| Other | 5 | 15 | .4 | .4 |
| Missing values | | 2 | .0 | |
| | Total | 3351 | 100.0 | 100.0 |
| FIRST LANGUAGE | | | | |
| English | 1 | 2655 | 79.2 | 79.4 |
| French | 2 | 202 | 6.0 | 6.0 |
| Spanish | 3 | 75 | 2.2 | 2.2 |
| Other | 4 | 74 | 2.2 | 2.2 |
| Chinese | 5 | 187 | 5.6 | 5.6 |
| Malayan | 6 | 127 | 3.8 | 3.8 |
| Indian | 7 | 24 | .7 | .7 |
| Missing values | 0 | 7 | .2 | |
| | Total | 3351 | 100.0 | 100.0 |
| **SUPERVISOR RATING** | | | | |
| Very poor | 1 | 60 | 1.8 | 1.9 |
| Poor | 2 | 243 | 7.3 | 7.7 |
| Reasonable | 3 | 1781 | 53.1 | 56.7 |
| Good | 4 | 892 | 26.6 | 28.4 |
| Excellent | 5 | 166 | 5.0 | 5.3 |
| Missing values | 0 | 209 | 6.2 | |
| | Total | 3351 | 100.0 | 100.0 |

***Appendix A.2.1: Crosstabulation of AGE by ETHNIC ORIGIN***

| | White | Black | Asian | Indian | Oriental | Oth Eur | Hispanic | Other | Row Total | Percent |
|---|---|---|---|---|---|---|---|---|---|---|
| Under 20 | 28 | 1 | 3 | | | | | 2 | 34 | 1.0 |
| 20-24 | 69 | 5 | 1 | 2 | | | | | 77 | 2.3 |
| 25-29 | 641 | 77 | 89 | 9 | 74 | 2 | 25 | 3 | 920 | 27.5 |
| 30-34 | 526 | 82 | 49 | 8 | 63 | 1 | 19 | 2 | 750 | 22.4 |
| 35-39 | 470 | 50 | 30 | 13 | 46 | | 17 | 3 | 629 | 18.8 |
| 40-44 | 377 | 40 | 16 | 3 | 18 | | 11 | 3 | 468 | 14.0 |
| 45-49 | 225 | 15 | 5 | 1 | 3 | | 4 | 1 | 254 | 7.6 |
| 50-54 | 187 | 5 | 5 | 2 | 1 | | 4 | 2 | 206 | 6.2 |
| 55-59 | 6 | | 1 | | | | | | 7 | 0.2 |
| 60+ | 3 | | 2 | | | | | | 5 | 0.1 |
| Column Total | 25.32<br>75.6 | 275<br>8.2 | 201<br>6.0 | 38<br>1.1 | 205<br>6.1 | 3<br>0.1 | 80<br>2.4 | 16<br>0.4 | 3350<br>100.0 | 100.0 |

***Appendix A.2.2: Crosstabulation of ETHNIC ORIGIN by FIRST LANGUAGE***

| | English | French | Spanish | Chinese | Malayan | Indian | Other | Row Total | Percent |
|---|---|---|---|---|---|---|---|---|---|
| White | 2292 | 202 | 4 | 1 | | | 30 | 2529 | 75.6 |
| Black | 272 | | | | | | | 272 | 8.1 |
| Asian | 54 | | 7 | 3 | 122 | | 15 | 201 | 6.0 |
| Indian | 6 | | 1 | 3 | 1 | 24 | 3 | 38 | 1.1 |
| Other | 13 | | | | 1 | | 1 | 15 | 0.4 |
| Oriental | | | | 179 | 3 | | 23 | 205 | 6.1 |
| Oth'r Eur | 1 | | | 1 | | | 1 | 3 | 0.1 |
| Hispanic | 16 | | 63 | | | | 1 | 80 | 2.4 |
| 11 | 1 | | | | | | | 1 | 0.0 |
| Column Total | 2655<br>79.4 | 202<br>6.0 | 75<br>2.2 | 187<br>5.6 | 127<br>3.8 | 24<br>0.7 | 74<br>2.2 | 3344 | 100.0 |

***Appendix A.3: Jobs represented in the sample, with frequencies of occurrence and relevant DOT codings***

Note: OCN is the ODT Occupational Code Number. DPT are the Data, People, and Things worker function complexity ratings. File code refers to the codings used to refer to each group in combined sample data file.

| OCN DPT | File Code | Job Description | Frequency of sample | Percent |
|---|---|---|---|---|
| 1—167 | 1 | Manager, Department (&11) | 5 | 0.1 |
| 1—177 | 2 | Manager, Personnel | 27 | 0.8 |
| 1—117 | 3 | Manager, Area/Branch/Region (&15) | 40 | 1.2 |
| 1—167 | 4 | Management, Trainee | 105 | 3.1 |
| 1—167 | 5 | Personnel recruiter | 8 | 0.2 |
| 1—167 | 6 | Consultant | 20 | 0.6 |
| 2—132 | 7 | Supervisor | 71 | 2.1 |
| 1—167 | 8 | Admin. Assistant/Officer | 142 | 4.2 |
| 2—357 | 9 | Sales rep, general merchandise (&36) | 4 | 0.1 |
| 1—117 | 10 | Sale promotion director | | |
| 1—167 | 11 | Manager, Department | 95 | 2.8 |
| 0—162 | 12 | Computer Programmer | 3 | 0.1 |
| 0—281 | 13 | Technical Illustrator | 2 | 0.1 |
| 1—227 | 14 | Technical Instructor | 17 | 0.5 |
| 1—117 | 15 | Manager, Area/Branch/Region | 266 | 7.9 |
| 1—167 | 16 | Manager, Retail Department | 2 | 0.1 |
| 2—357 | 17 | Administrative Assistant | 25 | 0.7 |
| 1—167 | 18 | Manager, Sales | 32 | 1.0 |
| 1—227 | 19 | Training Representative | 1 | |
| | 20 | Other | | |
| 1—167 | 21 | Manager, Administrative Serv. | 35 | 1.0 |
| 1—162 | 22 | Accountant | 68 | 2.0 |
| 1—167 | 23 | Comptroller | 9 | 0.3 |
| 2—362 | 24 | Computer operator | 25 | 0.7 |
| 1—167 | 25 | Manager, Retail Store | 92 | 2.7 |
| 1—117 | 26 | Manager, General | 35 | 1.0 |
| 1—117 | 27 | Manager, Area/Branch/Region | 59 | 1.8 |
| 1—167 | 28 | Manager, Office (Government) | | |
| | 29 | Other | 29 | 0.9 |
| 2—257 | 30 | Sales Agent, Insurance | 25 | 0.7 |
| 2—362 | 31 | Clerk-Typist | 68 | 2.0 |
| 1—117 | 32 | Vice-President | 18 | 0.5 |
| 0—167 | 33 | Systems Analyst | 18 | 0.5 |
| 0—167 | 34 | Computer Systems Engineer | 2 | 0.1 |
| 0—167 | 35 | Methods Analyst, Data Process. | 4 | 0.1 |
| 2—357 | 36 | Sales Person, Gen. Merchan. | 9 | 0.3 |
| 2—357 | 37 | Sales Rep, General Merchandise | 6 | 0.2 |
| 2—357 | 38 | Sales rep, Office Machines | 6 | 0.2 |
| 0—161 | 39 | Technician | | |
| 1—117 | 40 | Treasurer, Financial Instit. | 33 | 1.0 |
| 2—357 | 41 | Sales Agent, Real Estate | 36 | 1.1 |
| 1—267 | 42 | Manager, Financial Institution | 23 | 0.7 |
| 1—267 | 43 | Commercial Account Manager | 16 | 0.5 |

| OCN DPT | File Code | Job Description | Frequency of sample | Percent |
|---|---|---|---|---|
| 1—352 | 44 | Hospital – Admitting Clerk | 83 | 2.5 |
| 1—157 | 45 | Buyer | 7 | 0.2 |
| 1—157 | 46 | Buyer, Assistant | 8 | 0.2 |
| 2—477 | 47 | Sales Clerk | 37 | 1.1 |
| 2—357 | 48 | Sales Rep, Radio-TV-Broad | 37 | 1.1 |
| 2—257 | 49 | Sales Rep, Financial services | 240 | 7.2 |
| 2—257 | 50 | Financial Planner | 47 | 1.4 |
| 2—137 | 51 | Supervisor, Mail Carriers | 1 | |
| 2—137 | 52 | Supervisor, Mails | 184 | 5.5 |
| 1—157 | 53 | Supervisor of Sales | | |
| 2—132 | 54 | Supervisor, Accounting/clerical | | |
| 2—362 | 55 | Teller | 42 | 1.3 |
| 1—267 | 56 | Loans Officer/Manager | 10 | 0.3 |
| 2—482 | 57 | Accounting Clerk | 9 | 0.3 |
| 1—117 | 58 | Office Manager | 7 | 0.2 |
| 1—117 | 59 | Branch Manager – Financial | | |
| 9—687 | 60 | Service Rep. | 1 | |
| 1—221 | 61 | Training Manager | 1 | |
| 2—357 | 62 | Sales Rep, Tobacco products | 1152 | 34.4 |
| | | Total | 3351 | 100.0 |

***Appendix B.1: Normalized sten scores tables: Age-corrected to 30 years***

*The effects of age on each scale have been controlled for in the production of these tables. This is done by statistically "removing" age effects from each person's raw score (so that the correlation between their adjusted raw scores and age would be zero) by standardizing them all to the age of 30. The normalized sten score cut-off values are calculated using these adjusted raw scores. Thus the tables are representative of the sten scores 30-year-old people would get.*

If you are using the tables for people who are much older or younger than thirty you can correct for their age difference as follows. First you must calculate an Adjusted Raw Score for each

*Adjusted raw score = Raw score - (Age correction) x (Age - 30).*

The Age Corrections differ for each scale and are listed in the table below.

***Example: Person A, who is aged 54, obtained a raw score of 43 on Extraversion. As the Age Correction for Extraversion is -0.17, their adjusted raw score would be:***
Adjusted raw score = 43 - (-0.17) x (54-30)
= 43 + 0.17 x 24
= 47.08
= 47 rounded to the nearest whole number

Using 47 as the raw score, the Sten equivalent (from the ALL GROUPS: BOTH SEXES norm) is 5. If the score had not been adjusted for age, the person would have been assigned a more extreme Sten score of 4.

| Scale | Age Correction | Scale | Age Correction |
|---|---|---|---|
| I1 | +0.01 | PEOPLE | -0.14 |
| I2 | -0.02 | DATA | -0.13 |
| C1 | +0.10 | THINGS | +0.01 |
| C2 | +0.05 | | |
| E1 | -0.08 | | |
| E2 | -0.08 | | |
| S1 | +0.04 | | |
| S2 | +0.03 | | |
| INDEP | -0.02 | | |
| CONSC | +0.15 | | |
| EXTRAV | -0.17 | | |
| STABLE | +0.07 | SD | +0.17 |

If the Age Correction is **positive** then adjusted raw scores should be higher for people under 30 and lower for those over 30.

If the Age Correction is **negative** then adjusted raw scores should be lower for people under 30 and higher for those over 30.

***Appendix B.2: Computation of age-corrected raw scores used in production of norm tables.***

In the following equations the "x" prefix is used to denote the age-corrected raw scale score.

```
xi1=i1-0.0075*(age-30).
xi2=i2+0.0225*(age-30).
xc1=c1-0.0998*(age-30).
xc2=c2-0.0493*(age-30).
xe1=e1+0.0826*(age-30).
xe2=e2+0.0793*(age-30).
xs1=s1-0.0381*(age-30).
xs2=s2-0.0275*(age-30).
xindep=indep+0.0150*(age-30).
xconsc=consc-0.1491*(age-30).
xextrav=extrav+0.1619*(age-30).
xstable=stable-0.0657*(age-30).
xSD=SD-0.1764*(age-30).
xpeople=people+0.1421*(age-30).
xdat=dat+0.1334*(age-30).
xthings=things-0.0148*(age-30).
xWWN=wwn+0.0504*(age-30).
```

| Variable | Cases | Mean | Std Dev |
|---|---|---|---|
| XI1 | 3267 | 22.7134 | 4.8899 |
| XI2 | 3267 | 25.6068 | 5.2724 |
| XC1 | 3267 | 24.8275 | 4.6327 |
| XC2 | 3267 | 26.1468 | 5.1583 |
| XE1 | 3267 | 26.0458 | 5.2362 |
| XE2 | 3267 | 26.2425 | 5.8687 |
| XS1 | 3267 | 25.8418 | 5.1644 |
| XS2 | 3267 | 24.8185 | 4.8816 |
| XINDEP | 3267 | 48.3202 | 8.5093 |
| XCONSC | 3267 | 50.9743 | 8.2141 |
| XEXTRAV | 3267 | 52.2883 | 9.9043 |
| XSTABLE | 3267 | 50.6598 | 8.9780 |
| AGE | 3267 | 34.8757 | 8.6652 |
| XSD | 602 | 21.9975 | 5.6337 |
| AGE | 602 | 33.5266 | 9.4316 |
| XPEOPLE | 2597 | 43.6050 | 8.1504 |
| XDATA | 2597 | 31.1035 | 7.8717 |
| XTHINGS | 2597 | 32.5341 | 9.6493 |
| AGE | 2597 | 35.1001 | 8.4577 |
| XWWN | 2731 | 15.2565 | 6.2278 |
| AGE | 2731 | 35.1908 | 8.4718 |

In all cases, correlations of age-corrected scores with AGE: r=0.000

***Appendix B.3: Means, Standard Deviations, minima, maxima and sample sizes for each scale.***

Descriptive statistics for the whole sample:

| Scale | Mean | SD | Min | Max | N |
|---|---|---|---|---|---|
| I1 | 22.73 | 4.89 | 12 | 36 | 3329 |
| I2 | 25.50 | 5.29 | 12 | 36 | 3330 |
| C1 | 25.33 | 4.70 | 12 | 36 | 3324 |
| C2 | 26.39 | 5.19 | 12 | 36 | 3332 |
| E1 | 25.68 | 5.27 | 12 | 36 | 3327 |
| E2 | 25.87 | 5.90 | 12 | 36 | 3334 |
| S1 | 26.04 | 5.17 | 12 | 36 | 3336 |
| S2 | 24.95 | 4.88 | 12 | 36 | 3333 |
| INDEP | 48.23 | 8.53 | 24 | 72 | 3320 |
| CONSC | 51.71 | 5.32 | 26 | 72 | 3316 |
| EXTRAV | 51.54 | 9.99 | 24 | 72 | 3320 |
| STABLE | 51.00 | 8.99 | 24 | 72 | 3329 |
| SD | 22.61 | 5.88 | 14 | 40 | 603 |
| PEOPLE | 42.83 | 8.25 | 14 | 60 | 2675 |
| DATA | 30.43 | 7.94 | 12 | 60 | 2688 |
| THINGS | 32.56 | 9.64 | 12 | 60 | 2687 |
| WWN | 14.97 | 6.25 | 0 | 24 | 2747 |

***Appendix C.1: Correlations between ICES and 16PF scales (N=151). 16PF primary and second order factor scores are listed down the page with ICES scales being listed from left to right along the page.***

| Correlations | I1 | | I2 | | IND | | C1 | | C2 | | CONSC | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16PF-A | -.0274 | | -.2452 | ** | -.1530 | | -.2156 | * | -.0944 | | -.1816 | |
| 16PF-B | -.0006 | | .0344 | | -.0232 | | -.0373 | | -.2166 | * | -.1480 | |
| 16PF-C | -.1191 | | .2245 | | .0893 | | .0711 | | .0279 | | .0580 | |
| 16PF-E | .4068 | ** | .5420 | | .5908 | ** | -.4854 | ** | .2549 | ** | -.4335 | ** |
| 16PF-F | .0093 | | .3708 | | .2588 | ** | -.2991 | ** | -.3180 | ** | -.3608 | ** |
| 16PF-G | .0763 | | .0332 | | .0640 | | .4247 | ** | .4838 | ** | .5310 | ** |
| 16PF-H | .1520 | | .5925 | | .4877 | ** | -.2444 | * | -.1486 | | -.2300 | * |
| 16PF-I | -.1971 | * | -.0415 | | -.1349 | | .0284 | | -.0610 | | -.0189 | |
| 16PF-L | .2915 | ** | .3106 | | .3701 | ** | -.4120 | ** | -.1809 | | -.3473 | ** |
| 16PF-M | -.0617 | | .1850 | | .0933 | | -.1439 | | -.0819 | | -.1322 | |
| 16PF-N | -.1595 | | -.2338 | | -.2462 | * | .2173 | * | .2592 | ** | .2784 | ** |
| 16PF-O | -.0636 | | -.3617 | | -.2819 | ** | .0011 | | -.0233 | | -.0129 | |
| 16PF-Q1 | .3077 | ** | .4361 | | .4648 | ** | -.3873 | ** | -.0690 | | -.2677 | ** |
| 16PF-Q2 | .0075 | | -.1933 | | -.1282 | | .1364 | | .0294 | | .0972 | |
| 16PF-Q3 | -.0459 | | -.1244 | | -.1100 | | .4824 | ** | .4595 | ** | .5508 | ** |
| 16PF-Q4 | .1213 | | -.1688 | | -.0500 | | -.0706 | | -.1685 | | -.1395 | |
| PF-EXT | .0455 | | .4483 | ** | .3314 | ** | -.2790 | ** | -.1878 | | -.2731 | ** |
| PF-ANX | .0755 | | -.2852 | ** | -.1543 | | -.1226 | | -.1323 | | -.1490 | |
| PF-POISE | .1083 | | -.1551 | | -.0477 | | .1390 | | .0480 | | .1096 | |
| PF-IND | .3710 | ** | .6531 | ** | .6475 | ** | -.5245 | ** | -.2957 | ** | -.4802 | ** |
| PF-CTRL | .0308 | | -.0362 | | -.0082 | ** | .5293 | ** | .5585 | ** | .6359 | ** |
| **Correlations** | **E1** | | **E2** | | **EXT** | | **S1** | | **S2** | | **STAB** | |
| 16PF-A | .5938 | ** | .4970 | ** | .5940 | ** | .1570 | | .2078 | * | .1991 | * |
| 16PF-B | -.0691 | | -.0214 | | -.0484 | | -.0042 | | -.0099 | ** | -.0075 | |
| 16PF-C | .0978 | | .1563 | | .1401 | | .4963 | ** | .5064 | | .5507 | ** |
| 16PF-E | .2965 | ** | .4704 | ** | .4227 | ** | -.0285 | | .1042 | | .0379 | |
| 16PF-F | .5960 | ** | .6588 | ** | .6870 | ** | .0054 | | .0867 | | .0484 | |
| 16PF-G | -.1073 | | -.1223 | | -.1258 | | .1419 | | .0681 | | .1174 | |
| 16PF-H | .4433 | ** | .6938 | ** | .6267 | ** | .2003 | * | .2936 | ** | .2636 | ** |
| 16PF-I | .1271 | | .1648 | | .1603 | | .0223 | | .0435 | | .0356 | |
| 16PF-L | .2033 | * | .2748 | ** | .2628 | ** | -.3625 | ** | -.2678 | ** | -.3490 | ** |
| 16PF-M | .0224 | | .0503 | | .0403 | | .1925 | * | .1849 | | .2076 | * |
| 16PF-N | -.2373 | * | -.3042 | ** | -.2973 | ** | -.0004 | | -.0103 | | -.0056 | |
| 16PF-O | -.0989 | | -.1565 | | -.1408 | | -.5061 | ** | -.4630 | ** | -.5337 | ** |
| 16PF-Q1 | .3059 | ** | .3361 | ** | .3514 | ** | -.0034 | | .0090 | | .0027 | |
| 16PF-Q2 | -.4568 | ** | -.4969 | ** | -.5219 | ** | -.0605 | | -.1006 | | -.0874 | |
| 16PF-Q3 | -.1909 | * | -.3355 | ** | -.2907 | ** | .3809 | ** | .2332 | * | .3415 | ** |
| 16PF-Q4 | -.0545 | | -.0791 | | -.0735 | | -.7467 | ** | -.5878 | ** | -.7377 | ** |
| PF-EXT | .6433 | ** | .7348 | ** | 7550 | ** | .1327 | | .2135 | * | .1880 | |
| PF-ANX | -.0964 | | -.1519 | | -.1369 | | -.7099 | ** | -.6229 | ** | -.7347 | ** |
| PF-POISE | -.2030 | * | -.2053 | * | -.2232 | * | -.1637 | | -.1742 | | -.1854 | |
| PF-IND | .3988 | ** | .5945 | ** | .5470 | ** | .0375 | | .1465 | | .0981 | |
| PF-CTRL | -.1670 | | -.2471 | * | -.2280 | * | .2821 | ** | .1600 | | .2463 | * |
| 1-tailed Signif: *p<.01; ** p<.001 | | | | | | | | | | | | |

***Appendix C.2: Comparisons of 16PF and ICES response bias measures (n=151)***

FG and BG are the Karson and Odell 16PF Faking Good and Faking Bad scales.

| Scale | Mean | SD |
|---|---|---|
| ICES: | | |
| SD | 22.46 | 5.72 |
| I1 | 21.87 | 4.36 |
| C1 | 23.17 | 5.23 |
| E1 | 24.81 | 6.07 |
| S1 | 25.54 | 5.43 |
| I2 | 25.21 | 5.52 |
| C2 | 25.02 | 5.18 |
| E2 | 24.36 | 6.56 |
| S2 | 24.60 | 4.91 |
| INDEP | 47.07 | 8.07 |
| EXTRAV | 49.17 | 11.56 |
| CONSC | 48.19 | 8.90 |
| STABLE | 50.13 | 9.41 |
| 16PF: | | |
| FG | 6.57 | 2.59 |
| FB | 2.67 | 1.89 |

| ICES: | ICES SD | | 16PF FG | | 16PF FB | |
|---|---|---|---|---|---|---|
| SD | -- | | .25 | * | -.14 | |
| I1 | -.20 | * | .03 | | .08 | |
| C1 | .37 | ** | .30 | ** | -.11 | |
| E1 | -.03 | | -.07 | | -.16 | |
| S1 | .17 | | .43 | ** | -.37 | ** |
| I2 | -.20 | * | .10 | | -.16 | |
| C2 | .26 | ** | .32 | ** | -.12 | |
| E2 | -.13 | | -.01 | | -.20 | * |
| S2 | .18 | | .37 | ** | -.32 | ** |
| INDEP | -.24 | * | .09 | | -.06 | |
| EXTRAV | -.09 | | -.04 | | -.20 | * |
| CONSC | .37 | ** | .36 | ** | -.14 | |
| STABLE | .19 | * | .45 | ** | -.38 | ** |
| 1 – tailed Signif: *p<.01; **p<.001 | | | | | | |

***Appendix D.1: Normalized sten score tables: Age-corrected to 30 years: Phase III data***

| Scale | Age correction | Scale | Age correction |
|---|---|---|---|
| I1 | -0.04 | PEOPLE | -0.03 |
| I2 | -0.03 | DATA | -0.04 |
| C1 | +0.05 | THINGS | -0.02 |
| C2 | +0.02 | | |
| E1 | -0.05 | WWN | +0.01 |
| E2 | -0.08 | WWW | -0.08 |
| S1 | +0.06 | WWS | -0.11 |
| S2 | +0.04 | GENERAL | -0.27 |
| INDEP | -0.07 | | |
| CONSC | +0.07 | | |
| EXTRAV | -0.13 | | |
| STABLE | +0.10 | | |
| SD | +0.09 | | |

If the Age Correction is **positive** then adjusted raw scores should be higher for people under 30 and lower for those over 30.

If the Age Correction is **negative** then adjusted raw scores should be lower for people under 30 and higher for those over 30.

***Appendix D.2: Computation of age-corrected raw scores used in production of norm tables***

*In the following table the "x" prefix is used to denote the age-corrected raw scale score.*

| Variable | Cases | Mean | Std Dev |
|---|---|---|---|
| XI1 | 458 | 23.1622 | 4.5019 |
| XI2 | 458 | 25.5303 | 5.2024 |
| XC1 | 458 | 25.9000 | 4.0837 |
| XC2 | 458 | 25.3005 | 5.0404 |
| XE1 | 458 | 23.6814 | 5.3975 |
| XE2 | 458 | 24.4771 | 6.1044 |
| XS1 | 458 | 23.9542 | 5.0498 |
| XS2 | 458 | 24.0416 | 4.8902 |
| XINDEP | 458 | 48.6925 | 8.3468 |
| XCONSC | 458 | 51.2006 | 7.5560 |
| XEXTRAV | 458 | 48.1585 | 10.3346 |
| XSTABL | 458 | 47.9958 | 8.9345 |
| XSD | 458 | 22.5607 | 5.2581 |
| AGE | 458 | 37.9061 | 10.3838 |
| XPEOPLE | 458 | 39.5976 | 8.6451 |
| XDATA | 458 | 30.4314 | 7.8700 |
| XTHINGS | 458 | 33.3634 | 10.2848 |
| AGE | 458 | 37.9061 | 10.3838 |
| XWWN | 455 | 11.0934 | 5.5285 |
| XWWW | 455 | 25.1024 | 9.6849 |
| XWWS | 455 | 13.0829 | 4.4428 |
| XGENERAL | 455 | 73.4550 | 24.9242 |
| AGE | 455 | 37.9253 | 10.3922 |

***Appendix E.1:***

In all cases, correlations of age-corrected scores with AGE: 4=0.000

| GENDER | | | | |
|---|---|---|---|---|
| **Ethnic Origin** | **Count** | **Female** | **Male** | **Total** |
| White | 1 | 204 | 212 | 416<br>80.6% |
| Black | 2 | 25 | 23 | 48<br>9.3% |
| Asian | 3 | 7 | 4 | 11<br>2.1% |
| Hispanic | 4 | 26 | 9 | 35<br>6.8% |
| Amerindian | 5 | 3 | 2 | 5<br>1.0% |
| Other | 6 | | 1 | 1<br>.2% |
| Column Total | | 265<br>51.4% | 251<br>48.6% | 516<br>100.0% |

| First language | | | | | |
|---|---|---|---|---|---|
| **Ethnic Origin** | **Count** | **English 2** | **Spanish 2** | **Other 4** | **Total** |
| White | 1 | 411 | 13 | 3 | 415<br>80.6% |
| Black | 2 | 47 | | 1 | 48<br>9.3% |
| Asian | 3 | 9 | | 2 | 11<br>2.1% |
| Hispanic | 4 | 26 | 9 | | 35<br>6.8% |
| Amerindian | 5 | 5 | | | 5<br>1.0% |
| Other | 6 | 1 | | | 1<br>.2% |
| Column Total | | 499<br>96.9% | 10<br>1.9% | 6<br>1.2% | 515<br>100.0% |

Number of Missing Observations: 1

***Appendix F.1: Analysis of distributions of each response alternative for the ICES Personality scales***

*The following analyses were carried out on the data from the Phase Two Stage One sample. They are all based on the 96 items used for the 8 ICES Minor scales (12 items per scale). As the SD scale items are designed to indicate different forms of response bias they were excluded from these analyses. (N = 756). The following tables show the frequency with which each response alternative ([a], [b] or [c]) was selected. The frequency of omitted responses is also shown.*

| Score | Freq. | Cum fr | Cum% |
|---|---|---|---|
| 0 | 704 | 704 | 93.12 |
| 1 | 0 | 704 | 93.12 |
| 2 | 0 | 704 | 93.12 |
| 3 | 0 | 704 | 93.12 |
| 4 | 0 | 704 | 93.12 |
| 5 | 0 | 704 | 93.12 |
| 6 | 1 | 705 | 93.25 |
| 7 | 3 | 708 | 93.65 |
| 8 | 0 | 708 | 93.65 |
| 9 | 0 | 708 | 93.65 |
| 10 | 0 | 708 | 93.65 |
| 11 | 0 | 708 | 93.65 |
| 12 | 0 | 708 | 93.65 |
| 13 | 1 | 709 | 93.78 |
| 14 | 0 | 709 | 93.78 |
| 15 | 0 | 709 | 93.78 |
| 16 | 2 | 711 | 94.05 |
| 17 | 0 | 711 | 94.05 |
| 18 | 0 | 711 | 94.05 |
| 19 | 0 | 711 | 94.05 |
| 20 | 0 | 711 | 94.05 |
| 21 | 0 | 711 | 94.05 |
| 22 | 1 | 712 | 94.18 |
| 23 | 0 | 712 | 94.18 |
| 24 | 0 | 712 | 94.18 |
| 25 | 0 | 712 | 94.18 |
| 26 | 0 | 712 | 94.18 |
| 27 | 0 | 712 | 94.18 |
| 28 | 0 | 712 | 94.18 |
| 29 | 44 | 756 | 100.00 |
| mean | 1.812 | | |
| var | 47.509 | | |
| sd | 6.893 | | |

44 people failed to complete the questionnaire.

***Appendix F.2: Distribution of [a] responses***

| Score | Freq | Cum fr | Cum% | Score | Freq | Cum fr | Cum% |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0.00 | 50 | 22 | 611 | 80.82 |
| 1 | 0 | 0 | 0.00 | 51 | 19 | 630 | 83.33 |
| 2 | 0 | 0 | 0.00 | 52 | 22 | 652 | 86.24 |
| 3 | 0 | 0 | 0.00 | 53 | 18 | 670 | 88.62 |
| 4 | 0 | 0 | 0.00 | 54 | 11 | 681 | 90.08 |
| 5 | 0 | 0 | 0.00 | 55 | 13 | 694 | 91.80 |
| 6 | 0 | 0 | 0.00 | 56 | 17 | 711 | 94.05 |
| 7 | 0 | 0 | 0.00 | 57 | 10 | 721 | 95.37 |
| 8 | 0 | 0 | 0.00 | 58 | 4 | 725 | 95.90 |
| 9 | 0 | 0 | 0.00 | 59 | 6 | 731 | 96.69 |
| 10 | 0 | 0 | 0.00 | 60 | 4 | 735 | 97.22 |
| 11 | 0 | 0 | 0.00 | 61 | 7 | 742 | 98.15 |
| 12 | 0 | 0 | 0.00 | 62 | 0 | 742 | 98.15 |
| 13 | 1 | 1 | 0.13 | 63 | 4 | 746 | 98.68 |
| 14 | 0 | 1 | 0.13 | 64 | 4 | 750 | 99.21 |
| 15 | 3 | 4 | 0.53 | 65 | 2 | 752 | 99.47 |
| 16 | 3 | 7 | 0.93 | 66 | 0 | 752 | 99.47 |
| 17 | 1 | 8 | 1.06 | 67 | 1 | 753 | 99.60 |
| 18 | 3 | 11 | 1.46 | 68 | 1 | 754 | 99.74 |
| 19 | 1 | 12 | 1.59 | 69 | 0 | 754 | 99.74 |
| 20 | 3 | 15 | 1.98 | 70 | 2 | 756 | 100.00 |
| 21 | 2 | 17 | 2.25 | | | | |
| 22 | 6 | 23 | 3.04 | mean | 41.726 | | |
| 23 | 6 | 29 | 3.84 | var | 98.589 | | |
| 24 | 6 | 35 | 4.63 | sd | 9.929 | | |
| 25 | 6 | 41 | 5.42 | | | | |
| 26 | 6 | 47 | 6.22 | | | | |
| 27 | 11 | 58 | 7.67 | | | | |
| 28 | 17 | 75 | 9.92 | | | | |
| 29 | 8 | 83 | 10.98 | | | | |
| 30 | 20 | 103 | 13.62 | | | | |
| 31 | 13 | 116 | 15.34 | | | | |
| 32 | 15 | 131 | 17.33 | | | | |
| 33 | 20 | 151 | 19.97 | | | | |
| 34 | 24 | 175 | 23.15 | | | | |
| 35 | 18 | 193 | 25.53 | | | | |
| 36 | 29 | 222 | 29.37 | | | | |
| 37 | 28 | 250 | 33.07 | | | | |
| 38 | 30 | 280 | 37.04 | | | | |
| 39 | 36 | 316 | 41.80 | | | | |
| 40 | 22 | 338 | 44.71 | | | | |
| 41 | 29 | 367 | 48.54 | | | | |
| 42 | 37 | 404 | 53.44 | | | | |
| 43 | 24 | 428 | 56.61 | | | | |
| 44 | 18 | 446 | 58.99 | | | | |
| 45 | 36 | 482 | 63.76 | | | | |
| 46 | 23 | 505 | 66.80 | | | | |
| 47 | 36 | 541 | 71.56 | | | | |
| 48 | 24 | 565 | 74.74 | | | | |
| 49 | 24 | 589 | 77.91 | | | | |

***Appendix F.3: Distribution of [b] responses***

| Score | Freq | Cum fr | Cum% | Sten |
|---|---|---|---|---|
| 0 | 26 | 26 | 3.44 | 2 |
| 1 | 61 | 87 | 11.51 | 3 |
| 2 | 38 | 125 | 16.53 | 4 |
| 3 | 53 | 178 | 23.54 | 4 |
| 4 | 73 | 251 | 33.20 | 5 |
| 5 | 76 | 327 | 43.25 | 5 |
| 6 | 69 | 396 | 52.38 | 6 |
| 7 | 61 | 457 | 60.45 | 6 |
| 8 | 43 | 500 | 66.14 | 6 |
| 9 | 49 | 549 | 72.62 | 7 |
| 10 | 30 | 579 | 76.59 | 7 |
| 11 | 30 | 609 | 80.56 | 7 |
| 12 | 13 | 622 | 82.28 | 7 |
| 13 | 28 | 650 | 85.98 | 8 |
| 14 | 14 | 664 | 87.83 | 8 |
| 15 | 26 | 690 | 91.27 | 8 |
| 16 | 11 | 701 | 92.72 | 8 |
| 17 | 8 | 709 | 93.78 | 9 |
| 18 | 8 | 717 | 94.84 | 9 |
| 19 | 5 | 722 | 95.50 | 9 |
| 20 | 5 | 727 | 96.16 | 9 |
| 21 | 4 | 731 | 96.69 | 9 |
| 22 | 3 | 734 | 97.09 | 9 |
| 23 | 0 | 734 | 97.09 | 9 |
| 24 | 2 | 736 | 97.35 | 9 |
| 25 | 0 | 736 | 97.35 | 9 |
| 26 | 1 | 737 | 97.49 | 9 |
| 27 | 1 | 738 | 97.62 | 9 |
| 28 | 1 | 739 | 97.75 | 10 |
| 29 | 1 | 740 | 97.88 | 10 |
| 30 | 1 | 741 | 98.02 | 10 |
| 31 | 1 | 742 | 98.15 | 10 |
| 32 | 3 | 745 | 98.54 | 10 |
| 33 | 1 | 746 | 98.68 | 10 |
| 34 | 2 | 748 | 98.94 | 10 |
| 35 | 0 | 748 | 98.94 | 10 |
| 36 | 1 | 749 | 99.07 | 10 |
| 37 | 2 | 751 | 99.34 | 10 |
| 38 | 1 | 752 | 99.47 | 10 |
| 39 | 1 | 753 | 99.60 | 10 |
| 40 | 0 | 753 | 99.60 | 10 |
| 41 | 0 | 753 | 99.60 | 10 |
| 42 | 0 | 753 | 99.60 | 10 |
| 43 | 2 | 755 | 99.87 | 10 |
| 44 | 1 | 756 | 100.00 | 10 |
| mean | 7.757 | | | |
| var | 42.629 | | | |
| sd | 6.529 | | | |

***Appendix F.4: Distribution of [c] responses***

| Score | Freq | Cum fr | Cum% | Score | Freq | Cum fr | Cum% |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0.00 | 50 | 15 | 531 | 70.24 |
| 1 | 0 | 0 | 0.00 | 51 | 35 | 566 | 74.87 |
| 2 | 0 | 0 | 0.00 | 52 | 18 | 584 | 77.25 |
| 3 | 0 | 0 | 0.00 | 53 | 27 | 611 | 80.82 |
| 4 | 0 | 0 | 0.00 | 54 | 21 | 632 | 83.60 |
| 5 | 0 | 0 | 0.00 | 55 | 15 | 647 | 85.58 |
| 6 | 0 | 0 | 0.00 | 56 | 18 | 665 | 87.96 |
| 7 | 0 | 0 | 0.00 | 57 | 7 | 672 | 88.89 |
| 8 | 0 | 0 | 0.00 | 58 | 15 | 687 | 90.87 |
| 9 | 0 | 0 | 0.00 | 59 | 15 | 702 | 92.86 |
| 10 | 0 | 0 | 0.00 | 60 | 13 | 715 | 94.58 |
| 12 | 0 | 0 | 0.00 | 61 | 11 | 726 | 96.03 |
| 13 | 0 | 0 | 0.00 | 62 | 5 | 731 | 96.69 |
| 14 | 0 | 0 | 0.00 | 63 | 5 | 736 | 97.35 |
| 15 | 0 | 0 | 0.00 | 64 | 1 | 737 | 97.49 |
| 16 | 0 | 0 | 0.00 | 65 | 4 | 741 | 98.02 |
| 17 | 0 | 0 | 0.00 | 66 | 5 | 746 | 98.68 |
| 18 | 2 | 2 | 0.26 | 67 | 1 | 747 | 98.81 |
| 19 | 2 | 4 | 0.53 | 68 | 1 | 748 | 98.94 |
| 20 | 0 | 4 | 0.53 | 69 | 2 | 750 | 99.21 |
| 21 | 6 | 10 | 1.32 | 70 | 2 | 752 | 99.47 |
| 22 | 2 | 12 | 1.59 | 71 | 0 | 752 | 99.47 |
| 23 | 5 | 17 | 2.25 | 72 | 2 | 754 | 99.74 |
| 24 | 3 | 20 | 2.65 | 73 | 0 | 754 | 99.74 |
| 25 | 3 | 23 | 3.04 | 74 | 1 | 755 | 99.87 |
| 26 | 6 | 29 | 3.84 | 75 | 0 | 755 | 99.87 |
| 27 | 5 | 34 | 4.50 | 76 | 1 | 756 | 100.00 |
| 28 | 12 | 46 | 6.08 | | | | |
| 29 | 14 | 60 | 7.94 | mean | 44.705 | | |
| 30 | 10 | 70 | 9.26 | var | 103.437 | | |
| 31 | 11 | 81 | 10.71 | sd | 10.170 | | |
| 32 | 15 | 96 | 12.70 | | | | |
| 33 | 14 | 110 | 14.55 | | | | |
| 34 | 18 | 128 | 16.93 | | | | |
| 35 | 17 | 145 | 19.18 | | | | |
| 36 | 14 | 159 | 21.03 | | | | |
| 37 | 23 | 182 | 24.07 | | | | |
| 38 | 14 | 196 | 25.93 | | | | |
| 39 | 24 | 220 | 29.10 | | | | |
| 40 | 25 | 245 | 32.41 | | | | |
| 41 | 30 | 275 | 36.38 | | | | |
| 42 | 29 | 304 | 40.21 | | | | |
| 43 | 32 | 336 | 44.44 | | | | |
| 44 | 29 | 365 | 48.28 | | | | |
| 45 | 23 | 388 | 51.32 | | | | |
| 46 | 32 | 420 | 55.56 | | | | |
| 47 | 51 | 471 | 62.30 | | | | |
| 48 | 24 | 495 | 65.48 | | | | |
| 49 | 21 | 516 | 68.25 | | | | |

***Appendix F.5:*** ICES Major Scale Patterns: Code Patterns are referenced as follows. Low (Sten 1-4) is coded 0, Medium (Sten 5-6) is coded 1 and High (Sten 7-10) is coded 2. For scale I, the code is multiplied by 27; for scale C by 9; for scale E by 3 and for Scale S by 1. The sum of the products, plus one, gives the code pattern reference number (from 1 to 81). Thus High I, Medium C, High E and Low S would be 27*2+9*1+3*2+0+1 = Code Pattern 70.

Expected frequency with even distribution = 1.23% per code.

| Code Pattern | Frequency | Cum Percent | Percent |
|---|---|---|---|
| 1.00 | 47 | 1.4 | 1.4 |
| 2.00 | 38 | 1.1 | 2.5 |
| 3.00 | 10 | .3 | 2.8 |
| 4.00 | 24 | .7 | 3.6 |
| 5.00 | 35 | 1.0 | 4.6 |
| 6.00 | 13 | .4 | 5.0 |
| 7.00 | 16 | .5 | 5.5 |
| 8.00 | 20 | .6 | 6.1 |
| 9.00 | 16 | .5 | 6.5 |
| 10.00 | 66 | 2.0 | 8.5 |
| 11.00 | 96 | 2.9 | 11.4 |
| 12.00 | 30 | .9 | 12.3 |
| 13.00 | 60 | 1.8 | 14.1 |
| 14.00 | 70 | 2.1 | 16.1 |
| 15.00 | 29 | .9 | 17.0 |
| 16.00 | 10 | .3 | 17.3 |
| 17.00 | 25 | .7 | 18.1 |
| 18.00 | 22 | .7 | 18.7 |
| 19.00 | 53 | 1.6 | 20.3 |
| 20.00 | 52 | 1.6 | 21.8 |
| 21.00 | 32 | 1.0 | 22.8 |
| 22.00 | 35 | 1.0 | 23.8 |
| 23.00 | 46 | 1.4 | 25.2 |
| 24.00 | 37 | 1.1 | 26.3 |
| 25.00 | 8 | .2 | 26.6 |
| 26.00 | 17 | .5 | 27.1 |
| 27.00 | 14 | .4 | 27.5 |
| 28.00 | 34 | 1.0 | 28.5 |
| 29.00 | 36 | 1.1 | 29.6 |
| 30.00 | 13 | .4 | 30.0 |
| 31.00 | 43 | 1.3 | 31.2 |
| 32.00 | 81 | 2.4 | 33.7 |
| 33.00 | 39 | 1.2 | 34.8 |
| 34.00 | 42 | 1.3 | 36.1 |
| 35.00 | 53 | 1.6 | 37.7 |
| 36.00 | 54 | 1.6 | 39.3 |

***Appendix F.5 cont'd:***

| Code Pattern | Frequency | Cum Percent | Percent |
|---|---|---|---|
| 37.00 | 57 | 1.7 | 41.0 |
| 38.00 | 72 | 2.1 | 43.1 |
| 39.00 | 26 | .8 | 43.9 |
| 40.00 | 74 | 2.2 | 46.1 |
| 41.00 | 148 | 4.4 | 50.5 |
| 42.00 | 84 | 2.5 | 53.0 |
| 43.00 | 33 | 1.0 | 54.0 |
| 44.00 | 67 | 2.0 | 56.0 |
| 45.00 | 76 | 2.3 | 58.3 |
| 46.00 | 38 | 1.1 | 59.4 |
| 47.00 | 37 | 1.1 | 60.5 |
| 48.00 | 27 | .8 | 61.3 |
| 49.00 | 29 | .9 | 62.2 |
| 50.00 | 91 | 2.7 | 64.9 |
| 51.00 | 54 | 1.6 | 66.5 |
| 52.00 | 12 | .4 | 66.9 |
| 53.00 | 46 | 1.4 | 68.2 |
| 54.00 | 73 | 2.2 | 70.4 |
| 55.00 | 18 | .5 | 71.0 |
| 56.00 | 18 | .5 | 71.5 |
| 57.00 | 7 | .2 | 71.7 |
| 58.00 | 45 | 1.3 | 73.1 |
| 59.00 | 53 | 1.6 | 74.6 |
| 60.00 | 24 | .7 | 75.4 |
| 61.00 | 35 | 1.0 | 76.4 |
| 62.00 | 61 | 1.8 | 78.2 |
| 63.00 | 47 | 1.4 | 79.6 |
| 64.00 | 23 | .7 | 80.3 |
| 65.00 | 29 | .9 | 81.2 |
| 66.00 | 14 | .4 | 81.6 |
| 67.00 | 39 | 1.2 | 82.8 |
| 68.00 | 80 | 2.4 | 85.1 |
| 69.00 | 49 | 1.5 | 86.6 |
| 70.00 | 26 | .8 | 87.4 |
| 71.00 | 77 | 2.3 | 89.7 |
| 72.00 | 64 | 1.9 | 91.6 |
| 73.00 | 11 | .3 | 91.9 |
| 74.00 | 10 | .3 | 92.2 |
| 75.00 | 11 | .3 | 92.5 |
| 76.00 | 12 | .4 | 92.9 |
| 77.00 | 49 | 1.5 | 94.4 |
| 78.00 | 58 | 1.7 | 96.1 |
| 79.00 | 12 | .4 | 96.4 |
| 80.00 | 41 | 1.2 | 97.7 |
| 81.00 | 78 | 2.3 | 100.0 |
| TOTAL | 3351 | 100.0 | |

***Appendix F.6: ICES Minor Scale and ICES Plus Interest Code Patterns***

**ICES Minor Scale Combinations:**

*Expected frequency with even distribution = 11.11% per code*

| | | Pattern | Frequency | Percent | Cum Percent |
|---|---|---|---|---|---|
| I1 and I2 | | | | | |
| 00 | Low I1, Low I2 | 1.00 | 386 | 11.5 | 11.5 |
| 01 | Low I1, Med I2 | 2.00 | 371 | 11.1 | 22.6 |
| 02 | Low I1, Hi I2 | 3.00 | 114 | 3.4 | 26.0 |
| 10 | Med I1, Low I2 | 4.00 | 330 | 9.8 | 35.8 |
| 11 | Med I1, Med I2 | 5.00 | 754 | 22.5 | 58.3 |
| 12 | Med I1, Hi I2 | 6.00 | 383 | 11.4 | 69.8 |
| 20 | Hi I1, Low I2 | 7.00 | 118 | 3.5 | 73.3 |
| 21 | Hi I1, Med I2 | 8.00 | 428 | 12.8 | 86.1 |
| 22 | Hi I1, Hi I2 | 9.00 | 467 | 13.9 | 100.0 |
| C1 and C2 | | | | | |
| 00 | Low C1, Low C2 | 1.00 | 415 | 12.4 | 12.4 |
| 01 | Low C1, Med C2 | 2.00 | 334 | 10.0 | 22.4 |
| 02 | Low C1, Hi C2 | 3.00 | 146 | 4.4 | 26.7 |
| 10 | Med C1, Low C2 | 4.00 | 364 | 10.9 | 37.6 |
| 11 | Med C1, Med C2 | 5.00 | 562 | 16.8 | 54.3 |
| 12 | Med C1, Hi C2 | 6.00 | 411 | 12.3 | 66.6 |
| 20 | Hi C1, Low C2 | 7.00 | 137 | 4.1 | 70.7 |
| 21 | Hi C1, Med C2 | 8.00 | 421 | 12.6 | 83.3 |
| 22 | Hi C1, Hi C2 | 9.00 | 561 | 16.7 | 100.0 |
| E1 and E2 | | | | | |
| 00 | Low E1, Low E2 | 1.00 | 464 | 132.8 | 13.8 |
| 01 | Low E1, Med E2 | 2.00 | 296 | 8.8 | 22.7 |
| 02 | Low E1, Hi E2 | 3.00 | 51 | 1.5 | 24.2 |
| 10 | Med E1, Low E2 | 4.00 | 320 | 9.5 | 33.8 |
| 11 | Med E1, Med E2 | 5.00 | 743 | 22.2 | 55.9 |
| 12 | Med E1, Hi E2 | 6.00 | 389 | 11.6 | 67.5 |
| 20 | Hi E1, Low E2 | 7.00 | 61 | 1.8 | 69.4 |
| 21 | Hi E1, Med E2 | 8.00 | 426 | 12.8 | 82.2 |
| 22 | Hi E1, Hi E2 | 9.00 | 598 | 17.8 | 100.0 |
| S1 and S2 | | | | | |
| 00 | Low S1, Low S2 | 1.00 | 416 | 12.4 | 12.4 |
| 01 | Low S1, Med S2 | 2.00 | 354 | 10.6 | 23.0 |
| 02 | Low S1, Hi S2 | 3.00 | 50 | 1.5 | 24.5 |
| 10 | Med S1, Low S2 | 4.00 | 315 | 9.4 | 33.9 |
| 11 | Med S1, Med S2 | 5.00 | 855 | 25.5 | 59.4 |
| 12 | Med S1, Hi S2 | 6.00 | 333 | 9.9 | 69.3 |
| 20 | Hi S1, Low S2 | 7.00 | 44 | 1.3 | 70.6 |
| 21 | Hi S1, Med S2 | 8.00 | 441 | 13.2 | 83.8 |
| 22 | Hi S1, Hi S2 | 9.00 | 543 | 16.2 | 100.0 |
| | | TOTAL | 3351 | 100.0 | |

**INTERESTS Scale combinations**

Expected frequency with even distribution = 3.7% per code

| Pattern | | Frequency | Percent | Cum Percent |
|---|---|---|---|---|
| P D T | | | | |
| L L L | 1.00 | 158 | 4.7 | 4.7 |
| L L **M** | 2.00 | 97 | 2.9 | 7.6 |
| L L H | 3.00 | 47 | 1.4 | 9.0 |
| L **M** L | 4.00 | 95 | 2.8 | 11.8 |
| L **M** **M** | 5.00 | 139 | 4.1 | 16.0 |
| L **M** H | 6.00 | 70 | 2.1 | 18.1 |
| L H L | 7.00 | 22 | .7 | 18.7 |
| L H **M** | 8.00 | 41 | 1.2 | 20.0 |
| L H H | 9.00 | 40 | 1.2 | 21.2 |
| **M** L L | 10.00 | 103 | 3.1 | 24.2 |
| **M** L **M** | 11.00 | 130 | 3.9 | 28.1 |
| **M** L H | 12.00 | 48 | 1.4 | 29.5 |
| **M** **M** L | 13.00 | 146 | 4.4 | 33.9 |
| **M** **M** **M** | 14.00 | 916 | 27.3 | 61.2 |
| **M** **M** H | 15.00 | 179 | 5.3 | 66.6 |
| **M** H L | 16.00 | 46 | 1.4 | 67.9 |
| **M** H **M** | 17.00 | 133 | 4.0 | 71.9 |
| **M** H H | 18.00 | 116 | 3.5 | 75.4 |
| H L L | 19.00 | 46 | 1.4 | 76.8 |
| H L **M** | 20.00 | 41 | 1.2 | 78.0 |
| H L H | 21.00 | 29 | .9 | 78.8 |
| H **M** L | 22.00 | 69 | 2.1 | 80.9 |
| H **M** **M** | 23.00 | 157 | 4.7 | 85.6 |
| H **M** H | 24.00 | 91 | 2.7 | 88.3 |
| H H L | 25.00 | 53 | 1.6 | 89.9 |
| H H **M** | 26.00 | 150 | 4.5 | 94.4 |
| H H H | 27.00 | 189 | 5.6 | 100.0 |
| | TOTAL | 3351 | 100.0 | |

***APPENDIX G: ICES PLUS INTERPRETIVE INFORMATION***

*The following provide general descriptions of the characteristics of people scoring above average, around the average and below average on each scale. The descriptions relate to those in the lower 16% (stens 1-2-3), the middle 68% (stens 4-5-6-7) and the top 16% (stens 8-9-10) of the adult working population. Each description attempts to capture the typical and distinctive characteristics of people in each group. However, there will be considerable variation within each of these groups. For example, those with very low or high scores (e.g., stens of 1 or 10) will tend to show more extreme or pronounced patterns of the characteristics described for their respective groups (i.e. the 1-2-3 or 8-9-10 groups). Similarly there will be variation between those with stens of 4 and those with stens of 7. The former tending to show a balance of characteristics more like that of the low scoring group, and the latter more like those of the high scoring group. The descriptions produced by the ICES Plus software makes finer distinctions between the score levels than those given here.*

*The following descriptions are provided for guidance only. Users are reminded of the need to base selection criteria on an appropriate job analysis and to base selection decisions on a systematic appraisal of a range of relevant sources of information about the applicants – not solely on the results of the ICES Plus.*

## ABILITIES

**Working With Numbers**

High scores (stens 8-9-10) in Numerical Reasoning, show a high capacity for numerical reasoning when compared with other adults in the general working population. They are quicker and more accurate than over 80% of such people when reasoning with information which is derived from simple numbers.

In general, the middle sten range (4-5-6-7) indicates a respondent who is in the average range for numerical reasoning. This indicates that they can reason with the speed and accuracy typical of the members of the adult working population when dealing with information derived from simple numbers. Scores of 4 are in the low-average range (better than the lowest 16% of the population) while those of 7 are in the high-average range (better than the lowest 69% of the population).

The low scorers in Numerical Reasoning (stens 1-2-3) show a below average capacity for numerical reasoning when compared with others from the adult working population for speed and accuracy. They are likely to take longer and be less accurate than other people in dealing with information, which is derived from simple numbers.

**Working with Words**

*Working with words involves a number of facets of mental ability. Most important is the ability to use language as a vehicle for reasoning and problem solving. Verbal ability is more focused on reasoning with language than on communication. In particular, the focus is on the use of written language. In many occupations, the ability to work with written language is a fundamental requirement – particularly for "white collar" jobs, clerical, administrative, technical and managerial. However, it should be noted that there are many jobs where fluency or oral communication is far more important than verbal reasoning. People who score high on Working with Words are not necessarily good communicators. Oral communication and expression skills are more directly assessed by other procedures.*

Those with high sten scores (stens 8-9-10) are quicker and more accurate than the majority of people in the working population when reasoning with information which involves language. They are unlikely to have any problems related to the use and understanding of written language and should find it easy to follow written instructions, etc.

Those with average scores (stens 4-5-6-7) have average capacity for working with written language when compared to others in the general working population. Their speed and accuracy in using verbal material show they will be as able as most adult workers to deal with material, which involves written language.

Those with low scores (stens 1-2-3) show a below average capacity for working with written information when compared with others under conditions where there is a degree of time pressure. They are likely to take longer and be less accurate than the majority of people in the adult working population in dealing with information which involves using words and written material. They may find it difficult to understand written instructions, technical manuals, etc. without help. They will tend to be at a disadvantage working in any job where facility with written language is important.

**Working with Shapes**

Working with shapes involves a number of facets of mental ability. Most important is the ability to imagine how something will look when it is moved around or when its component parts are rearranged. Spatial visualization skills come into a range of working tasks: interpreting blueprints and diagrams, understanding graphs and charts, working out how to arrange objects on a shelf or set up a display stand and so on.

Those with high scores (stens 8-9-10) are quicker and more accurate than most people when reasoning with information which involves thinking about and mentally manipulating shapes and objects in space. They will feel at ease working with plans and diagrams and be able to relate working drawings and schematics to actual objects and products.

Those with average scores (stens 4-5-6-7) are as able as most adult workers to deal with information which involves thinking about and mentally manipulating shapes and objects in space.

Those with low scores (stens 1-2-3) are likely to take longer and be less accurate than the majority of people in the adult working population when dealing with information which involves thinking about and mentally manipulating shapes. They may find difficulty working in positions where they have to relate figural instruction material (plans, diagrams, schematics, etc.) to actual operations or objects.

## GENERAL ABILITY

The ICES Plus measure of General Ability is based on the combination of three specific abilities: Working with Words, Numbers and Shapes. General Ability has been found to be one of the best single predictors of success in an occupation and of performance on a wide range of training courses. High levels of General Ability are far less important in unskilled manual and semi-skilled jobs than they are in those which are skilled or white collar (i.e. clerical, administrative, technical and managerial jobs).

Those with high scores (stens 8-9-10) are quicker and more accurate in their reasoning skills than most of those in the working population. People like this are generally quick to learn and can absorb new information easily. They are likely to be very efficient and able to deal well with change in their working requirements and under conditions of high mental workload. They will find it relatively easy to absorb new information and instructions from written information, documents, plans, etc.

Those with average scores (stens 4-5-6-7) are as able as most other adult workers to deal with information which involves thinking about and mentally manipulating words, numbers and shapes. People like this can learn and can absorb new information without too much difficulty. They are likely to be efficient working in an environment which makes reasonable demands on them. However, under high levels of mental load, they may find it difficult to cope and may need assistance of specific support training.

Low scores on General Ability (stens 1-2-3) indicate that people tend to take longer than others to learn new procedures and have more difficulty in understanding new information. In occupations where these are important, they will be at a disadvantage in comparison with others. They will work best in occupations where the working environment is well structured and there are clear procedures and routines to follow. They may find frequent changes in working practices difficult to cope with.

## INTEREST SCALES

**People**

A high scorer (stens 8-9-10) is likely to be very interested in work, which involves a lot of contact with other people. This may be at a quite complex level (involving persuading others or negotiating with them) rather than just making contact. People with scores in this range are unlikely to feel satisfied in jobs in which interaction with other people does not play a major role.

The moderate scorer (stens 4-5-6-7) show an average level of interest in work which involves dealing with people. They are likely to prefer jobs, which involve a reasonable degree of contact with others, and would not be happy

working on their own. However, they are unlikely to want interaction with other people to be the major function of their work or for it to require difficult and demanding interpersonal skills.

The lower scorer (stens 1-2-3) will be quite content to work in a job in which there is little or no contact with other people. While they would not necessarily avoid contact with other people, they would not want interpersonal relations to form a key function of their responsibilities.

**Data**

A high score on Data (stens 8-9-10) indicates a high level of interest in working with data. Such people are often interested in data and information for its own sake and enjoy working with figures, symbols, statistics, accounts and language. They are likely to enjoy working with information systems, technical documents, contracts and so on. They would be unlikely to enjoy a job, which did not provide some opportunity for this type of work.

A medium score (stens 4-5-6-7) on Data would indicate an individual who has an average level of interest in working with data and information. A person like this may relate this interest in Data to its application in working with people and/or machinery and equipment. Such people are happy to handle figures and statistics, and are not "put off" by numbers. However, they would not necessarily feel the need for work with data to form the major part of their job.

The low score on Data (stens 1-2-3) indicates a person who has a below average level of interest in working with data. Individuals like this usually avoid jobs where they have to spend a lot of time dealing with figures, statistics or accounts or where such work forms an important – if infrequent – part of work.

**Things**

A high score (stens 8-9-10) on Things indicates a person with a high level of interest in work which deals with inanimate objects such as machinery, tools and equipment. Such people are likely to be interested in engineering work – both "hands-on" and in design and development. This interest may also express itself at lower qualification levels in interests in handling goods and managing equipment or stores – e.g. warehouse work.

The moderate score (4-5-6-7) on Things indicates an average level of interest in work which deals with inanimate objects such as machinery, tools and equipment. People with scores in this range are likely to be comfortable working with machinery and tools, but would not see that aspect of their work as being central (for example, word-processor operators).

The low (stens 1-2-3) scorer has a below average level of interest in work which deals with inanimate objects such as machinery, tools and equipment. Individuals like this will tend to avoid work, which involves dealing with machinery, computers, and so on.

**General comments on the interest scales**

Interpretation of the PDT profile should take account of all three scales. In general one should ask:

- Is one scale clearly higher than the others? If so, this provides a good indicator of where a person's motivation is.
- Is the high one in the 8-9-10 range or in the mid-range? If the latter then this does not indicate a clear preference and it is best to look at the lower scores (those in the 1-2-3 range) to see what sort of work the person does not like.
- Is the profile flat? If so then the person has no clear preference in terms of PDT.
- What is the level of the profile – is it all at the 1-2-3 level or 4-5-6-7 level or are all scales high? This can be taken as a general indication of overall level of interest – general motivation.

## PERSONALITY SCALES – ICES

**Independent (I)**

A high score (stens 8-9-10) on the Independence scale indicates a respondent who is very independent, single-minded and determined to win. They are likely to be assertive, forthright and confident. A person with this score will tend to be very skeptical and hard-headed, and may find other people's lack of drive irritating. They are good at

getting things done, but can be very insensitive to the needs of those around them. They do not make good "team players", but can be effective – though autocratic – leaders in the right circumstances.

A middle of the range (stens 4-5-6-7) score would indicate a balance between a desire to compete and win with a wish to collaborate with others. Individuals like this are good at getting things done while respecting the needs of those around them. They are capable of getting their own way, although typically they are considerate and co-operative people.

Individuals with low (stens 1-2-3) scores on the Independence scale are generally likeable, diplomatic and good-natured. They are considerate and co-operative people who are capable of pulling others together. They accomplish this by encouraging and persuading others, rather than forcefully asserting their own views. An individual with a low score in Independence may skirt important issues to avoid conflict. At the extreme they are very co-operative, non-competitive, compassionate, careful of relationships and sensitive to the feelings of others. These individuals are likeable, diplomatic and good-natured. While they provide good support in a team, they may lack the assertiveness and confidence needed to pull people together and provide leadership.

The **INDEPENDENT** scale is divided into two Minor scales: **Competitive** and **Assertive**.

**Competitive (I1)**

People who score high (stens 8-9-10) on this scale tend to be extremely single-minded and competitive people who play to win and are bad losers. They tend to strive hard to reach their goals, putting their own success first. In playing to win, they tend to show relatively little concern about whether other people get upset or hurt along the way. In the extreme, other people are used as the means to help the person achieve their own ends.

A middle (stens 4-5-6-7) score on this scale indicates a person with a balanced mix of competitiveness and the desire to foster team spirit and work with others. Such individuals will compromise between their own need for achievement and their need to maintain co-operative relationships with others.

Those with a low (stens 1-2-3) score on this scale will be co-operative and non-competitive people who obtain their satisfaction from contributing to collaborative efforts. They are team players and enjoy co-operative ventures and are unlikely to be concerned about winning or
losing. Such individuals concern themselves with maintaining personal relationships, foregoing their own success to help others, and they derive a great deal of satisfaction from the success of their team.

**Assertive (I2)**

The high Assertiveness score (stens 8-9-10) indicates a rational, assertive and outspoken person. They know their own mind and are not afraid to say so. Individuals like this often become group leaders and are often controversial, unafraid of argument or open debate and will make sure their opinions are known. They will stand up for their position, even if it is unpopular or likely to create conflict.

In the middle range (stens 4-5-6-7) individuals may be fairly assertive and outspoken in some situations and with some people. They are more likely to show their assertiveness in non-threatening situations, with people they know. They tend not to promote themselves as group leaders, but with some encouragement can assume most roles. They see themselves more as peacemakers than decision-makers and may appear somewhat reserved at times – being reluctant to speak out on issues.

The low scorers (stens 1-2-3) on this scale are valued for their diplomacy and tact, and can play a useful role as peacemakers and diffusers of aggression or conflict. Occasionally, they may stand up for what they see as rightly their own, but for the most part they will be a rather submissive and non-controversial person, trying to avoid conflict rather than confront it.

**Conscientious (C)**

A high score (stens 8-9-10) on the Conscientious scale (the C in ICES) indicates a respondent who is extremely conscientious, neat and tidy, and detail-conscious. This individual is careful to abide by the rules and is most comfortable working within clear guidelines to a set of well-defined values. They tend to hold to traditional moral values and not be radical or innovative. People of this type are very dependable, and often meticulous in their attention to detail. Preferring to be well prepared and planful, they are likely to be good adaptors, rather than innovators.

Those in the middle range (stens 4-5-6-7) are reasonably tidy and detail-conscious in their work habits and are generally dependable, well prepared and planful. They are comfortable with following rules and established procedures within a traditional setting. However, they are also able to work outside clear guidelines, being able to balance the need to do things well in the quickest possible way without "breaking the rules". This leads to solutions that may be innovative without implementing radical changes. Individuals like this are occasionally careless and disorganized, and they may need to be reminded of the framework in which they are operating.

The low scorer (stens 1-2-3) is often a spontaneous and innovative individual, who works well in changing situations. They are flexible and responsive to circumstances as they arise, and will produce creative and unorthodox solutions. You can expect some measure of chaos in their work habits as a consequence of the creativity and flexibility this individual brings to the job. While spontaneous, innovative and flexible, they will have little regard for the traditional ways of doing things. In fact, they will thrive in a creative, challenging situation, but may be unsuccessful in a highly structured and predictable environment. Individuals like this can be careless and not very well organized. If not channeled appropriately, their lack of conscientiousness can result in counter-productive behavior.

The **CONSCIENTIOUS** scale is divided into two Minor scales: **Conventional** and **Organized**.

**Conventional (C1)**

As a follower of the rules, a person with a high score (stens 8-9-10) on this scale will conduct themselves in a very conventional, meticulous and reliable manner. They will prefer to do things in a traditional fashion and will operate to a high moral code. Matters of principal and doing things "the right way" are seen by such people as being of prime importance. As such, they can find it difficult to adapt to new situations or new ways of working. They are at their best working in a highly structured, clear and unambiguous environment.

The mid range (stens 4-5-6-7) includes individuals who are reasonably conventional in their approach and their attitudes and values, and who have a balanced approach to change and innovation. These people can be flexible when necessary and can cope with change. Overall, though, they are likely to prefer the "status quo" to change for change's sake.

Those scoring in the low range (stens 1-2-3) regard themselves as innovative and flexible, with a rather casual attitude towards guidelines, rules and regulations. They are likely to seek new ways to solve problems rather than follow traditional methods, and are likely to enjoy change for its own sake. They operate best in fast moving and unpredictable work environments. Seeing new ways of doing things, these individuals often reach solutions by cutting corners and overlooking rules. Excelling in an ever changing and challenging environment, they will feel stifled in a highly structured and rule-bound work situation. The possible downside of this innovative approach is the risk of boredom or counter-productive behavior in over-structured work situations.

**Organized (C2)**

A person who scores high (stens 8-9-10) on this is orderly and meticulous and works well in a controlled and rational environment. They have a place for everything with everything in its place. They plan ahead and think through all the possibilities before acting. They do not like having to think on their feet or engage in unstructured verbal debate. Individuals like this are often intolerant of and irritated by others who do not share these qualities. They are dependable and predictable, and find it hard to cope in situations for which they have not had a chance to prepare.

In the middle range (stens 4-5-6-7) individuals are reasonably well organized and able to work in a controlled manner. However, they do show spontaneity and are able to respond well to unpredictable events. They are reasonably neat and tidy in their working habits without being overly fastidious. While they probably do plan ahead, they do not feel particularly uncomfortable if they have to change their plans at the last minute.

Low scorers (stens 1-2-3) on this scale regard themselves as creative, spontaneous people who prefer to react to situations as they arise rather than to plan things in advance. They like to focus on the overall picture rather than deal with the fine details and do not like to worry about the details of how things will get done. Individuals like this feel that planning and structure restrict their creative and innovative abilities. They see attention to detail as being something for other people to worry about. This can manifest itself in a disorganized work place and a failure to meet deadlines and turn up for appointments.

**Extravert (E)**

Those who score high (stens 8-9-10) on the Extravert scale (the E in ICES) are sociable and talkative individuals who often seek excitement. These people are happiest when they are the center of attention, seeking out people for fun, entertainment, company and stimulation. Others may see them as high-spirited, popular "fun" people who often act on impulse.

Individuals who score in the middle range (stens 4-5-6-7) show moderate levels of extraversion. They are generally enthusiastic and lively, contributing to social interaction without drawing undue attention to themselves. They enjoy being with others and also enjoy their own company: they have a balance between the need for companionship and the need to have time for oneself.

The low scorer (stens 1-2-3) is introverted and prefers to avoid large social gatherings and group activities. Such people are most comfortable in a quiet environment where the surroundings are familiar. They are quite content to be alone, where they can reflect on their own thoughts and ideas. They much prefer the company of a few close friends to large gatherings of acquaintances.

The **EXTRAVERT** scale is divided into two Minor scales: **Group-oriented** and **Outgoing**.

**Group-oriented (E1)**

High scorers (stens 8-9-10) on this scale have a strong need for other people. They like to be with other people and need their approval and support. They are happiest working in situations where there is a reasonable amount of contact with others and want to be seen as part of the team. They are likely to be very upset by social disapproval. Because of their need for other people, they may appear to be very sociable and seek out environments where they can meet lots of people. While they may prefer to be with other people rather than on their own, they are not necessarily particularly outgoing. They like to be part of the group, but not necessarily the leader or the most outspoken member.

Those in the average range (stens 4-5-6-7) enjoy the company of others and may seek others out, but do not need to be with other people all the time. They like to have some time to reflect and enjoy their own company. These needs are fairly evenly balanced. In general, such people will be happiest working in situations where there is a moderate amount of contact with other people and will cope well with any need for people to collaborate and work together.

Individuals who score in the low (stens 1-2-3) range of this scale are happy to work on their own and in quiet places, and tend to avoid noisy situations and group activities. They will often avoid social gatherings, group activities and busy environments. Individuals like this are generally to be found away from the social scene and feel most at ease in their own company where they can reflect on their own thoughts, and control the amount of stimulation that reaches them. They are well adapted to work situations where they might have to spend prolonged periods of time without direct contact with other people. While they can work with others, they do not feel any great need to.

**Outgoing (E2)**

A high scorer on this scale (stens 8-9-10) will be outgoing and talkative. They want to be the center of attention. Such people enjoy "risky", action-packed, and challenging lives. They often act impulsively, and like meeting new people and doing exciting and stimulating things. Routine work may become boring for them, and they often find stimulating change in their work by moving jobs more often than others. They tend to like people for the stimulation they provide rather than need people for the support they can give (compared with E1).

Those in the average range, (stens 4-5-6-7) like to lead a moderately exciting life and may act on impulse at times. They are fairly talkative and outgoing. While they find routine tasks tolerable, they would prefer some variety in their work. These people like to choose the situations in which they will take center stage as they are comfortable in the company of others, but they do not seek constant attention from others.

The low scorers (stens 1-2-3) describe themselves as people who are quiet and reserved, feeling that life is mentally stimulating enough without seeking extra exciting activities. Such people are not as readily bored by repetitive work and, while they may act impulsively at times, they prefer to live a quiet, orderly life. They do not like being the center of attention, and may therefore keep in the background at social gatherings – or avoid them altogether.

**Stable (S)**

A score in the high range (stens 8-9-10) of this scale would indicate a stable and untroubled person, who is able to accept people at face value. For the most part, they have a relaxed approach to life, taking problems, people and

circumstances in their stride. They may occasionally become anxious or suspicious, but that is the exception for them. When under normal levels of pressure or stress at work, they will remain relaxed and secure. The person high on Stability can accept criticism without feeling threatened by it and is untroubled by setbacks. As people, they are very secure in themselves and emotionally "hardy", being able to remain calm and relaxed even when under considerable stress.

In most situations, people who score in the middle range (stens 4-5-6-7) on this scale are able to accept and deal with situations in a calm and stable manner. There will be some circumstances where they become rather apprehensive and emotional, and may at times be wary about other people, particularly about their motives. In general, such people are reasonably secure in themselves, remaining fairly relaxed under moderate levels of stress.

A low score (stens 1-2-3) indicates someone who can be rather anxious. They tend to be suspicious of new people and wary of new situations. Sensitive and emotional, they appear to experience feelings of guilt and sadness more readily and openly than others. When faced with adversity, setbacks and other stressful situations, these people can become anxious and irritable and may find it difficult to cope effectively.

The **STABLE** scale is divided into two Minor scales: **Poised** and **Relaxed**.

**Poised (S1)**

People with a high score (stens 8-9-10) on this scale readily shrug off criticism. They are able to cope with most situations in life without getting upset or irritated. They have a rational approach to life and accept that few things in life proceed without challenge or setback. They can cope with adversity without "losing their cool".
People with a medium score (stens 4-5-6-7) have an average balance between calm objectivity in the face of difficult situations and a tendency to be upset and take things personally at times. In some circumstances, they have difficulty being objective and rational about situations in which they are personally involved.

Those with low scores (stens 1-2-3) can be irritable and are easily upset, often losing their temper. However, their irritation and upset is usually short-lived. Individuals with such an outlook often view the world as basically hostile and threatening, and may feel that people who do not see it this way are unreasonable or naïve. They find it hard to cope with embarrassing situations, and have difficulty coping with setbacks and personal criticism.

**Relaxed (S2)**

A high score (stens 8-9-10) indicates a person who is very relaxed, untroubled and well prepared to cope with life's pressures. They will accept people at face value, without suspecting them of ulterior motives. People like this can leave job-related troubles and worries behind them when they go home, and usually sleep well. They are not unduly bothered when things go wrong. However, their calm acceptance of life and their trust in the people around them may put them at risk of being exploited by others in some situations. They can cope well with demanding high-pressure jobs and where there is a need to work with others in an open and trusting manner.

A medium score (stens 4-5-6-7) indicates a person who remains calm and relaxed in response to most situations. For the most part, individuals like this are able to manage their problems without undue anxiety. Such individuals will not always assume the best of other people, and will feel the need to check their motives at times. They tend to worry and become somewhat anxious at times, particularly when things do not go well. However, both their level of suspicion of others and their stress under pressure are likely to be moderate and not cause any difficulties.

A person with a low (stens 1-2-3) score on this scale is likely to be a rather excitable and anxious person who is rather wary and cautious of others. Individuals like this find it difficult to cope with high levels of pressure without becoming tense and anxious. They tend to be very suspicious of others whom they do not know well and may also feel that colleagues are not to be trusted. If taken to extreme, this can cause problems in interpersonal situations. Individuals like this are best to avoid work situations in which there are likely to be prolonged periods of high pressure, or where they are expected to work with others in a very open and trusting manner.

**Social Desirability (SocDes)**

A high score (stens 8-9-10) on this scale may indicate a person who is not being totally frank in their assessment. They may be presenting what they feel to be a socially acceptable view of themselves rather than an honest picture of how they really are. Scores on other scales, particularly Conscientiousness and Stability, can be significantly influenced by this tendency. These individuals will be very certain of what is expected of them and what is proper in social situations. However, a high SocDes score can also be obtained by someone who is being honest, but who is a

genuinely “good” person. Thus a high score (particular 9 or 10) should be regarded as an indicator of a possible distorted profile – and not taken as proof of it.

A medium score (stens 4-5-6-7) on this scale indicates a person who has presented a reasonably frank picture of themselves on the other scales.

A low score (stens 1-2-3) on Social Desirability can have two interpretations. Either the person has presented a negative impression of themselves or have presented a frank picture of themselves. In either case, the meaning of low scores would need to be explored with the person.

***Appendix H:***

# What is so Important About Reliability?

The need to consider the Standard Error of Measurement

Professor Dave Bartram
Department of Psychology
University of Hull
HULL
HU67RX

The concept of reliability lies at the heart of test theory and is of great practical importance for the test user. However, the reason for its importance is often misunderstood by test users. This misunderstanding is reinforced by rules-of-thumb that are widely used as guidance on what is "acceptable" and what is not. Typically, one finds it asserted that reliability must be at least 0.70 if a test is to be any use at all. Preferably it should be above 0.80 and ideally above 0.90.

The point of the present article is not to say that these guidelines are wrong, but to reinforce the point that they are only guidelines. There are circumstances where a reliability of less than 0.70 would be quite acceptable and others where one of over 0.90 would not.

### The relationship between reliability and measurement error

The reason we are interested in reliability, in practical terms, is that it has implications for the range of error we make in measuring some psychological characteristic (be it numerical reasoning, extraversion or whatever). The lower the reliability, the larger the error is likely to be. The size of error is represented by the Standard Error of Measurement (SEM).

This is best understood as follows: Suppose you were asked to measure the length of a metal rod, which was, actually 1000 cm long (i.e. that was its true length). You make a series of measures and plot a distribution of them. If you have made a large number of measures, you would get a nice bell-shaped normal distribution curve, with a mean of 1000 cm and a standard deviation, which is the standard error of measurement. This SEM might be a few cms if the measurement process was rather crude or a few microns if it was very sophisticated. Technically, then, the SEM is the Standard Deviation (SD) of the distribution of actual scores we would obtain if we repeatedly tested the same person on the same scale (assuming this could be done without any carry-over effects and without the person changing during the process).

The size of the SEM is a function of the measurement process – of how intrinsically accurate or inaccurate it is. It is not dependent on the people you measure; the SEM does not change from sample to sample if the measurement process is kept the same. On the other hand, estimates of reliability are sample-dependent. They are affected by the degree of range restriction present in the sample of data on which they are based. Recall that the formal definition of reliability is that N is the ratio of the variance of true scores. If the sample is restricted in range (small observed variance), the reliability will appear to be low. Thus we could quite easily obtain a reliability estimate of say 0.5 if the sample we based it on was restricted in range of scores relative to the population for which the test was intended. However, the SEM is not affected by such restriction.

Recall that the formula for SEM is:

**SEM = SD x sqr (1- reliability)**

As the SD decreases and the reliability decreases, so the SEM will remain the same. The SEM is useful because it provides us with direct information about how accurate our measures are, expressed in the same units as the scores.

### When are reliabilities of less than 0.70 acceptable?

So the first point to note is that the guidelines on reliability can only be used making judgements about the reliability of a test which is based on an unrestricted sample. Often, test constructors are able to provide the SD for the

relevant population (from the general population test norms), but their reliability estimates have had to be obtained on some subset of those people. If the SD of the subset is smaller than that of the population, then the reliability will be an underestimate. The SEM, however, will not be affected.

A population reliability of 0.70 implies that the SEM is just over half of the population SD. While that represents a very useful gain in accuracy, there is really no reason why a reliability of 0.68 should be regarded as "no good" while one of 0.70 is "OK". The first represents an error of 0.56 of an SD, while the second is an error of 0.55 of an SD.

It would be better to rephrase the "rule of thumb" to say: The SEM should not be substantially greater than one half of the population SD. The following table shows the relationship between population reliability and the SEM.

| Population Reliability | SEM as a Proportion of pop SD |
|---|---|
| 0.30 | 0.84 |
| 0.40 | 0.77 |
| 0.05 | 0.71 |
| 0.60 | 0.63 |
| 0.65 | 0.59 |
| 0.70 | 0.55 |
| 0.75 | 0.50 |
| 0.80 | 0.45 |
| 0.85 | 0.39 |
| 0.90 | 0.32 |
| 0.95 | 0.22 |
| 0.99 | 0.10 |

**Can reliability be too high?**

The simple answer is "yes". When a test is constructed, there are a number of conflicting demands about which decisions have to be made. How broad is the characteristic, which is being measured? For example, Extraversion or General Ability are very broad characteristics, while Persuasiveness and Verbal Reasoning are much narrower ones. In general the broader the characteristic we are measuring, the greater the variety of the items we need to measure it and, hence, the lower the average inter-item correlation.

One common method of estimating reliability is based on the "internal consistency" of the set of items (e.g. Cronbach's alpha, split-half and odd-even reliability, KR-20 and so on). These methods produce a reliability estimate that is a function of two things: the average inter-item correlation and the number of items. As the trait being measured gets broader, so does the average inter-item correlation and the number of items. As the trait being measured gets broader, average inter-item correlation decreases and the number of items needed has to be increased if we are to keep reliability constant.

Given this, what should we think of a 12-item extraversion scale, which claims to have a reliability of 0.95? One's suspicion would be that the items in the scale were not covering the breadth of the trait adequately, as this implies an average inter-item correlation of 0.60. This is the sort of level one might expect for a fairly narrow ability test.

It is probably reasonable to say that average inter-item correlations should be in the region 0.20 for broad traits to 0.50 for narrow ones. To obtain a reliability of 0.90, this would mean 36 items for the broad trait scale but only 9 for the narrow one.

Typically, personality inventories measure a number of traits and are limited – for practical reasons – in terms of length. Between 12 and 24 items per scale are typical. This implies that for the minimum population reliability of 0.70, the average inter-item correlation needs to be 0.16 for a 12-item scale. This is not unreasonable. In practice, one might aim for nearer 0.20 in designing such a scale – which would give a reliability in the region of 0.75.

For the longer 24-item scales, the same level of inter-item correlation (i.e. about 0.20) gives a reliability of 0.85. For personality assessment, I would argue that these levels of internal consistency are about right – given the test length. One should not have higher reliability bought at the expense of breadth. The only legitimate way to increase the reliability is to increase the length of the scale. However, the reduction in measurement error (resulting from the gain in reliability) which this brings is usually more than offset by the cost in terms of testing time.

In certain applications high reliability is more important than others. For example, in differential ability testing small standard errors of difference (SED) are needed if practical use is to be made of differences between scale scores.

For two scales with the same SDs, the Standard Error of Difference is roughly 1.4 times as large as their SEM. This means that each scale needs a reliability of at least 0.87 if their Standard Error of Difference is to be no more than half of the population SD. A population reliability of 0.70 would be too low for scales in a differential test battery.

### Different estimates of reliability

The above arguments relate to using internal consistency measures as indicators of reliability. Such measures are inappropriate in some cases (e.g. with highly speeded tests). We can also use test/retest correlations and correlations between alternate forms of a test as estimators of reliability.

Different issues arise here. A measure of a broad but very stable trait might have a high retest reliability but a low internal consistency. On the other hand a measure of mood should have good internal consistency and low retest "reliability" because mood is not stable across time. The general rule of thumb that internal consistency is always higher than retest reliability is only true for narrow relatively stable traits, generally specific ability tests. For tests such as these, one expects to find internal consistencies in the 0.90s and retest reliabilities in the 0.80s.

Some of these points are illustrated in Table 1. This gives alpha coefficients and retest correlations, with SEMs based on each, for scales from ICES personality inventory. The inventory has four major scales (Independence, Conscientiousness, Extraversion, and Stability) composed of 24 items each. These are broad measures each of which is divided into two "minor" scales composed of 12 items (Independence is divided into I1 and I2 and so on). In addition, there is a 14-item Social Desirability scale (SocDes). The data shown are from a particular set of samples (from the UK, USA, and Canada) which were used for construct validation. Population data (and hence scale SDs) are available on a total of 3351 people. Scale construction was based on a "target" average inter-item correlation of around 0.20. This would produce 24-item scales with population reliabilities in the 0.85 region. While the minor scales are narrower, the smaller number of items entailed that reliabilities were expected to be in the 0.70 – 0.75 region.

***Table 1. Reliability information for the ICES Inventory, showing internal consistencies and retest correlations (sample size n=604) – one-two week retest***

| Scale | Mean | SD | Alpha | SEM | SEM/SD | r | Retest-SEM | Description |
|---|---|---|---|---|---|---|---|---|
| I1 | 21.88 | 4.95 | 0.71 | 2.67 | 0.54 | 0.81 | 2.19 | Tough-minded, competitive |
| I2 | 25.32 | 5.40 | 0.71 | 2.91 | 0.54 | 0.80 | 2.31 | Forthright, assertive |
| C1 | 22.38 | 5.22 | 0.72 | 2.76 | 0.53 | 0.79 | 2.30 | Traditional, concern for morality |
| C2 | 24.81 | 5.50 | 0.70 | 3.01 | 0.55 | 0.86 | 2.12 | Attention to detail |
| E1 | 24.53 | 5.62 | 0.74 | 2.87 | 0.51 | 0.74 | 2.68 | Sociable, outgoing |
| E2 | 24.30 | 6.19 | 0.80 | 2.77 | 0.45 | 0.74 | 2.82 | Group-dependent |
| S1 | 25.21 | 5.33 | 0.70 | 2.92 | 0.55 | 0.60 | 3.05 | Unruffled, unflappable |
| S2 | 25.18 | 5.02 | 0.65 | 2.97 | 0.59 | 0.74 | 2.38 | Relaxed, not anxious |
| INDEP | 47.19 | 8.75 | 0.79 | 4.01 | 0.46 | 0.83 | 3.59 | Independence |
| CONSC | 48.83 | 10.71 | 0.78 | 5.02 | 0.47 | 0.84 | 3.56 | Conscientiousness |
| EXTRAV | 47.19 | 9.00 | 0.86 | 3.67 | 0.41 | 0.76 | 4.73 | Extraversion |
| STABLE | 50.39 | 9.43 | 0.81 | 4.40 | 0.47 | 0.69 | 4.79 | Stability |
| SocDes | 22.57 | 5.94 | 0.78 | 2.79 | 0.47 | 0.82 | 2.64 | Social desirability/distortion |
| NB: Retest data were not available for all the sample. Retest SEMs are based on the SDs for the retest group. | | | | | | | | |

For the most part, Table 1 shows that these design requirements were met. The 24-item "major" scales all show good accuracy – with SEMs of less than half an SD. Note how in many cases, the retest correlations are higher than the internal consistencies. This illustrates the point made above, that retest correlations are higher than internal consistency for stable but broad characteristics. The Stability scale (and its minors) appear to have the lowest retest reliability. This is due partly to some range restriction in this sample. However, it also reflects the item content: S2 is "trait anxiety" (relatively broad, but stable over time), while S1 may be more affected by "state anxiety" (more narrow and less stable).

### Shrinkage of reliability estimates

When looking at reliability estimates – or SEMs based on them – you need to know whether these are from the sample which was used to construct the test scales or from a sample which was assessed after the scales had been constructed. The internal consistency obtained on the sample which was used to construct the scales in the first instance will always be higher than that of subsequent samples. Like validity coefficients, reliabilities will "shrink" in

size from the sample used for scale development to future samples. The amount of shrinkage is an indicator of the degree to which the item selection process has capitalized on random factors in the development sample.

The ICES scales, for example, were developed using an initial sample of 1,518 people. Final selection of the 12 items for each scale were made from the large pool of trial items using data from just half these people (the Scale Development Group). Reliabilities for the scales based on the selected items were then measured using the data from the remaining half of the sample (the Hold-out Group). Reliabilities for the Hold-out Group were only 0.01 smaller than those for the Development Group. This shows that the scales are robust.

**Conclusion**

It is easy to construct a test with high reliability. Just ask the same question lots of times. However, that is of little practical value, because what we then have is a very reliable answer to a very specific question. Tests are designed to measure relatively broad characteristics (even so-called “specific” abilities are not really all that specific). As such, any test needs a variety of item types and item content, which is designed to sample across the whole of the relevant domain. This entails the need for rather more complex guidelines that those often advocated.

While it is reasonable to say that 0.70 is the lower bound of the range of population reliabilities which are acceptable, sample reliabilities may well be less than this if the samples are restricted in range. On the other hand, very high levels of internal consistency should only be found for very specific traits or for scales containing very large numbers of items. It is better to focus attention on the average inter-item correlation (which gives a direct indication of “item variety”) and the SEM, which provides a direct indication of measurement scale error.

The importance of reliability depends on what purpose a measure is to serve. Consider again the process of measuring the metal bar. If this is being carried out to check machining tolerances for some precision engineering function then an accuracy of plus or minus a few centimeters will not do. However, if it is part of a process of sorting bars into 950, 1000, 1050 cm sizes, a relatively crude measure would do.

A SEM of half the population SD will suffice for many of the uses to which psychological test results are put. Attempts to get higher reliabilities with short tests are likely to result in narrow and, ultimately, less valid measures. However, test users need to be aware of the various factors, which affect reliability and how it relates to the SEM (and the SED). It is also wise to

remember that the levels of accuracy we work to in psychological testing are not comparable to the sort of tolerances we can achieve in physical measurement.

***Appendix I.1: Means and SDs for the ICES and NEO Scales (N=59)***

| Scale | Mean | SD | Scale Name |
|---|---|---|---|
| ICES Major Scales: | | | |
| ICES-I | 5.15 | 1.70 | Independence |
| ICES-C | 3.75 | 2.06 | Conscientiousness |
| ICES-E | 4.75 | 1.81 | Extraversion |
| ICES-S | 4.24 | 2.15 | Stability |
| ICES Minor Scales: | | | |
| ICES-I1 | 5.41 | 1.71 | Competitive |
| ICES-I2 | 5.34 | 1.55 | Assertive |
| ICES-C1 | 3.78 | 1.77 | Conventional |
| ICES-C2 | 4.22 | 2.17 | Organized |
| ICES-E1 | 5.17 | 1.81 | Group-oriented |
| ICES-E2 | 4.80 | 1.86 | Outgoing |
| ICES-S1 | 4.03 | 1.88 | Poised |
| ICES-S2 | 4.81 | 2.33 | Relaxed |
| | | | |
| ICES-SD | 5.03 | 1.96 | Social desirability |
| | | | |
| NEO Domain Scales: | | | |
| NEO-N | 59.03 | 13.03 | Neuroticism |
| NEO-E | 53.85 | 12.19 | Extraversion |
| NEO-O | 58.71 | 10.49 | Openness to new experience |
| NEO-A | 39.59 | 13.70 | Agreeableness |
| NEO-C | 38.32 | 13.96 | Conscientiousness |
| NEO Facet Scales: | | | |
| NEO-N1 | 54.25 | 11.14 | Anxiety |
| NEO-N2 | 56.66 | 13.27 | Angry-Hostility |
| NEO-N3 | 57.73 | 12.41 | Depression |
| NEO-N4 | 56.58 | 14.46 | Self-consciousness |
| NEO-N5 | 59.80 | 11.08 | Impulsiveness |
| NEO-N6 | 56.59 | 13.77 | Vulnerability |
| | | | |
| NEO-E1 | 45.53 | 12.28 | Warmth |
| NEO-E2 | 53.58 | 11.65 | Gregariousness |
| NEO-E3 | 47.78 | 11.88 | Assertiveness |
| NEO-E4 | 50.92 | 9.91 | Activity |
| NEO-E5 | 62.76 | 11.48 | Excitement-seeking |
| NEO-E6 | 53.46 | 11.39 | Positive emotions |
| | | | |
| NEO-O1 | 58.36 | 10.52 | Fantasy |
| NEO-O2 | 50.98 | 11.46 | Aesthetics |
| NEO-O3 | 55.27 | 9.71 | Feelings |
| NEO-O4 | 58.02 | 10.12 | Actions |
| NEO-O5 | 54.88 | 9.00 | Ideas |
| NEO-O6 | 57.92 | 9.65 | Values |

***Appendix I.2: Correlations between the ICES minor scales and the NEO domain and facet scales (N=59), and correlations between each of the scales and sex (male coded 0, female coded 1).***

| | **ICES Minor Scales** | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **I** | | | | **C** | | | | **E** | | | | **S** | | | | | | | |
| | **I1** | | **I2** | | **C1** | | **C2** | | **E1** | | **E2** | | **S1** | | **S2** | | **SD** | | **Sex** | |
| | -30 | * | | | .08 | | .25 | | -22 | | | | | | | | | | | |
| Sex | -.21 | | .30 | | -.04 | | -.00 | | .17 | | .29 | | -.18 | | -.23 | | -.16 | | | |
| NEO-N | .05 | | -.24 | | .01 | | -.08 | | -.05 | | -.19 | | -.78 | ** | -.76 | ** | -.09 | | -.31 | * |
| N1 | -.08 | | -.26 | | .15 | | .14 | | -.11 | | -.30 | | -.68 | ** | -.29 | ** | .06 | | .33 | * |
| N2 | .33 | * | -.18 | | -.07 | | -.01 | | -.08 | | -.01 | | -.56 | ** | -.44 | ** | -.23 | | -.12 | |
| N3 | .01 | | -.29 | | .05 | | -.11 | | -.20 | | -.32 | * | -.65 | ** | -.70 | ** | -.04 | | -.24 | |
| N4 | -.04 | | -.35 | * | .05 | | -.03 | | -.15 | | -.34 | * | -.65 | ** | -.66 | ** | -.08 | | .23 | |
| N5 | -.00 | | .02 | | -.28 | | -.34 | * | .24 | | .31 | * | -.45 | ** | -.35 | * | -.14 | | .15 | |
| N6 | -.04 | | -.37 | * | .03 | | -.16 | | .03 | | -.07 | | -.62 | ** | -.69 | ** | -.11 | | .38 | * |
| | | | | | | | | | | | | | | | | | | | | |
| NEO-E | -.07 | | .27 | | -.19 | | -.28 | | .54 | ** | .72 | ** | .28 | | .41 | ** | .19 | | .08 | |
| E1 | -.30 | | .10 | | .04 | | -.11 | | .41 | ** | .56 | ** | .19 | | .29 | | .35 | * | .19 | |
| E2 | -.11 | | .10 | | -.16 | | -.31 | * | .63 | ** | .62 | ** | .14 | | .18 | | .06 | | .19 | |
| E3 | .33 | * | .49 | ** | -.18 | | .02 | | .11 | | .47 | ** | .34 | * | .46 | ** | .07 | | .08 | |
| E4 | .03 | | .11 | | -.02 | | -.18 | | .26 | | .29 | | .21 | | .25 | | .24 | | -.02 | |
| E5 | -.11 | | -.00 | | -.27 | | -.38 | * | .43 | ** | .44 | ** | .01 | | .10 | | .15 | | -.02 | |
| E6 | -.22 | | .26 | | -.20 | | -.25 | | .41 | ** | .60 | ** | .26 | | .38 | * | .09 | | .08 | |
| | | | | | | | | | | | | | | | | | | | | |
| NEO-O | -.08 | | .15 | | -.23 | | -.27 | | .24 | | .47 | ** | -.03 | | .06 | | .05 | | .10 | |
| O1 | .03 | | .10 | | -.27 | | -.37 | * | .20 | | .39 | * | -.32 | * | -.17 | | -.27 | | .03 | |
| O2 | .04 | | -.05 | | -.04 | | -.16 | | .06 | | .36 | * | -08 | | -.05 | | -.02 | | .12 | |
| O3 | .03 | | .11 | | -.03 | | -.03 | | .21 | | .24 | | -.14 | | -.06 | | .07 | | .31 | * |
| O4 | -.17 | | .20 | | -.18 | | -.27 | | .16 | | .17 | | .25 | | .20 | | -.04 | | -.01 | |
| O5 | -.09 | | .20 | | -.07 | | -.07 | | .20 | | .40 | ** | .09 | | .13 | | .09 | | -.20 | |
| O6 | -.23 | | .12 | | -.34 | * | -.13 | | .18 | | .23 | | .21 | | .27 | | .06 | | .14 | |
| | | | | | | | | | | | | | | | | | | | | |
| NEO-A | -.58 | ** | -.27 | | .13 | | -.05 | | .26 | | .16 | | .25 | | .17 | | .39 | * | .15 | |
| NEO-C | -.12 | | .10 | | .38 | * | .53 | ** | -.18 | | -.18 | | .36 | * | .32 | * | .29 | | .00 | |

N=59; 1-tailed Signif: * p <.01 **p <.001

***Appendix I.3: Factor analysis (principal components) - four and five factor solutions.***

***Appendix I.3.1: Four-factor Varimax rotated factor loadings***

| | F1 | F2 | F3 | F4 | Communality |
|---|---|---|---|---|---|
| | "S" | "E" | "C" | "I" | |
| NEO-N | -.93 | .01 | -.03 | .09 | .88 |
| ICES-S1 | .87 | .07 | .18 | .12 | .62 |
| ICES-S2 | .89 | .19 | .04 | .02 | .84 |
| | | | | | |
| NEO-A | .09 | .32 | .10 | .77 | .71 |
| ICES-I1 | -.07 | -.11 | -.11 | -.85 | .75 |
| ICES-I2 | .26 | .31 | .16 | -.71 | .69 |
| | | | | | |
| NEO-C | .44 | -.22 | .63 | -.02 | .64 |
| ICES-C1 | .01 | -.17 | .82 | .22 | .75 |
| ICES-C2 | -.00 | -.27 | .82 | -.18 | .78 |
| | | | | | |
| NEO-E | .35 | .81 | -.10 | .03 | .80 |
| ICES-E1 | .00 | .73 | -.08 | .20 | .58 |
| ICES-E2 | .15 | .87 | -.12 | -.07 | .80 |
| | | | | | |
| ICES-SD | .17 | .18 | .61 | .52 | .71 |
| | | | | | |
| NEO-O | -.18 | .67 | -.21 | .08 | .53 |
| Total eigenvalue - 10.28<br>73.5% of variance accounted for. | | | | | |

***Appendix I.3.2: Five-factor Oblimin rotated pattern loadings.***

| | F1 | F2 | F3 | F4 | F5 | |
|---|---|---|---|---|---|---|
| | "N" | "O" | "A" | "C" | "E" | Communality |
| NEO-N6 | -.88 | -.03 | .08 | -.06 | -.13 | .78 |
| ICES-S2 | .83 | .04 | .07 | -.01 | -.08 | .73 |
| ICES-S1 | .83 | -.01 | .23 | .08 | .04 | .76 |
| NEO-N3 | -.82 | .18 | .11 | .03 | .22 | .78 |
| NEO-N4 | -.81 | .10 | .11 | -.04 | .29 | .80 |
| NEO-N1 | -.80 | .16 | .11 | .26 | .16 | .77 |
| NEO-E3 | .58 | .13 | -.53 | .12 | -.31 | .76 |
| NEO-N2 | -.51 | .17 | -.48 | .01 | .01 | .54 |
| NEO-N5 | -.48 | .44 | -.05 | -.24 | -.28 | .64 |
| | | | | | | |
| NEO-O5 | .11 | .83 | .04 | -.05 | .09 | .68 |
| NEO-O3 | -.21 | .70 | -.05 | .23 | -.12 | .58 |
| NEO-O2 | -.18 | .68 | .02 | -.03 | -.10 | .54 |
| NEO-O1 | -.29 | .63 | -.08 | -.33 | -.16 | .74 |
| | | | | | | |
| ICES-I1 | .02 | .00 | -.84 | -.04 | .11 | .73 |
| NEO-A | .08 | .25 | .76 | .00 | -.12 | .71 |
| ICES-I2 | .42 | .30 | -.59 | -.01 | .03 | .59 |
| | | | | | | |
| ICES-C1 | -.02 | .05 | .29 | .74 | .11 | .68 |
| ICES-C2 | .16 | .13 | -.06 | .67 | .36 | .65 |
| NEO-O4 | .28 | .25 | .27 | -.59 | .06 | .57 |
| ICES-SD | .05 | .05 | .56 | .57 | -.23 | .70 |
| NEO-C | .48 | -.08 | .05 | .54 | .22 | .63 |
| NEO-O6 | .36 | .31 | .25 | -.43 | .15 | .47 |
| | | | | | | |
| NEO-E2 | .07 | -.01 | -.01 | -.09 | -.81 | .70 |
| ICES-E1 | -.07 | -.02 | .09 | -.01 | -.75 | .57 |
| ICES-E2 | .17 | .27 | -.15 | -.08 | -.68 | .72 |
| NEO-E1 | .16 | .36 | .16 | .24 | -.66 | .77 |
| NEO-E5 | -.09 | -.14 | .10 | -.21 | -.65 | .48 |
| NEO-E6 | .33 | .34 | .10 | -.08 | -.51 | .66 |
| NEO-E4 | .18 | .05 | .00 | .11 | -.45 | .27 |
| Total eigenvalue = 19.00<br>65.50% of the variance accounted for. | | | | | | |
| F1 | 1.00 | | | | | |
| F2 | .01 | 1.00 | | | | |
| F3 | .04 | .02 | 1.00 | | | |
| F4 | .04 | -.14 | .03 | 1.00 | | |
| F5 | -.13 | -.26 | -.05 | .15 | 1.00 | |

***Appendix I.4: Multiple regression analyses***

***Appendix I.4.1: Prediction of NEO scores by ICES minor scales***

| Scale | R | Adjusted $R^2$ as a percentage |
|---|---|---|
| NEO-N | 0.85 | 67.83 |
| NEO-E | 0.83 | 63.31 |
| NEO-O | 0.62 | 27.76 |
| NEO-A | 0.71 | 41.01 |
| NEO-C | 0.68 | 35.59 |
| | | Mean = 47.10 |

**Appendix I.4.2: Prediction of ICES major scales scores with NEO facet scales as predictors**

| Scale | R | Adjusted $R^2$ as a percentage |
|---|---|---|
| ICES-I | 0.77 | 38.89 |
| ICES-C | 0.81 | 48.45 |
| ICES-E | 0.86 | 60.44 |
| ICES-S | 0.90 | 71.65 |
| | | Mean = 54.86 |

***Appendix I.4.3: Prediction of ICES minor scale scores with NEO domain scales and with NEO facet scales (the latter including NEO-A and NEO-C) as predictors***

| | Domain scales | | Facet scales | |
|---|---|---|---|---|
| Scale | R | Adjusted $R^2$ as a percentage | R | Adjusted $R^2$ as a percentage |
| ICES-I1 | 0.67 | 39.14 | 0.81 | 47.05 |
| ICES-I2 | 0.54 | 22.42 | 0.70 | 23.00 |
| ICES-C1 | 0.49 | 16.56 | 0.77 | 37.35 |
| ICES-C2 | 0.64 | 35.09 | 0.78 | 39.14 |
| ICES-E1 | 0.62 | 32.57 | 0.80 | 44.33 |
| ICES-E2 | 0.80 | 60.15 | 0.85 | 57.84 |
| ICES-S1 | 0.79 | 59.21 | 0.86 | 61.52 |
| ICES-S2 | 0.81 | 61.97 | 0.86 | 60.68 |
| ICES-SD | 0.55 | 23.80 | 0.83 | 53.31 |

***Appendix I.5: Means and SDs for the ICES and EPQ-R scales (N=68) expressed as stens***

| Scale | Mean | SD | Scale name |
|---|---|---|---|
| **Age** | **24.24** | **8.08** | |
| ICES major scales: ICES-I | 5.16 | 1.86 | Independence |
| ICES-C | 3.94 | 2.06 | Conscientiousness |
| ICES-E | 4.99 | 1.94 | Extraversion |
| ICES-S | 3.97 | 1.76 | Stability |
| ICES minor scales: ICES-I1 | 5.46 | 1.89 | Competitiveness |
| ICES-I2 | 5.04 | 1.81 | Assertive |
| ICES-C1 | 4.22 | 1.96 | Conventional |
| ICES-C2 | 4.07 | 1.96 | Organized |
| ICES-E1 | 5.19 | 1.85 | Group-oriented |
| ICES-E2 | 5.06 | 2.04 | Outgoing |
| ICES-S1 | 4.10 | 1.84 | Poised |
| ICES-S2 | 4.47 | 1.86 | Relaxed |
| ICES-SD | 5.32 | 2.10 | Social desirability |
| EPQ-R scales: EPQ-P | 6.96 | 2.39 | Psychoticism |
| EPQ-E | 6.03 | 2.01 | Extraversion |
| EPQ-N | 6.10 | 2.12 | Neuroticism |
| EPQ-L | 5.28 | 1.68 | Lie |

***Appendix I.6: Correlations between the ICES minor scales and the EPQ-R scales (n=68)***

| | ICES Minor Scales | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | I | | | | C | | | | E | | | | S | | | | | |
| | I1 | | I2 | | C1 | | C2 | | E1 | | E2 | | S1 | | S2 | | SD | |
| ICES Minor Scales | | | | | | | | | | | | | | | | | | |
| I1 | 1.00 | | .37 | ** | .01 | | -.04 | | .07 | | .09 | | -.09 | | -.20 | | -.19 | |
| I2 | .37 | ** | 1.00 | | -.05 | | -.24 | | .34 | * | .55 | ** | .10 | | .16 | | -.11 | |
| C1 | .01 | | -.05 | | 1.00 | | .18 | ** | -.41 | ** | -.31 | * | .13 | | .10 | | .28 | |
| C2 | -.04 | | -.24 | | .48 | ** | 1.00 | | -.27 | | -.35 | * | .01 | | -.13 | | .26 | |
| E1 | .07 | | .34 | * | -.41 | ** | -.27 | | 1.00 | | .63 | ** | -.05 | | -.00 | | .09 | |
| E2 | .09 | | .55 | ** | -.31 | * | .35 | * | .63 | ** | 1.00 | | .06 | | .18 | | .05 | |
| S1 | -.09 | | .10 | | .13 | | .01 | | -.05 | | .06 | | 1.00 | | .54 | ** | .10 | |
| S2 | -.20 | | .16 | | .10 | | -.13 | | -.00 | | .18 | | .54 | ** | 1.00 | | .14 | |
| SD | -.19 | | -.11 | | .28 | | .26 | | .09 | | .05 | | .10 | | .14 | | 1.00 | |
| ICES Major scales: | | | | | | | | | | | | | | | | | | |
| I | .82 | ** | .83 | ** | -.03 | | -.17 | | .25 | | .39 | * | .01 | | -.02 | | -.18 | |
| C | -.02 | | -.18 | | .85 | ** | .88 | ** | -.39 | ** | -.39 | ** | .08 | | -.02 | | .31 | * |
| E | .09 | | .50 | ** | -.39 | ** | -.35 | * | .89 | ** | .92 | ** | .01 | | .11 | | .07 | |
| S | -.16 | | .15 | | .13 | | -.07 | | -.03 | | .14 | | .88 | ** | .87 | ** | .14 | |
| EPQ scales: | | | | | | | | | | | | | | | | | | |
| EPQ-P | .31 | * | .14 | | -.47 | ** | -.43 | ** | .15 | | .27 | | -.15 | | -.11 | | -.28 | |
| EPQ-E | .17 | | .58 | ** | -.36 | * | -.43 | ** | .74 | ** | .79 | ** | .04 | | .17 | | -.08 | |
| EPQ-N | .03 | | -.23 | | -.13 | | .20 | | -.06 | | -.15 | | -.70 | ** | -.76 | ** | -.00 | |
| EPQ-L | -.34 | * | -.18 | | .38 | * | .29 | * | -.19 | | -0.20 | | .24 | | .09 | | .65 | ** |
| 1-tailed Signif: * p<.01 ** p<.001 | | | | | | | | | | | | | | | | | | |

| | ICES major scales | | | | | | | | EPQ Scales | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | I | | C | | E | | S | | EPW-P | | EPQ-E | | EPQ-N | | EPQ-L | | Age | |
| ICES Minor scales: | | | | | | | | | | | | | | | | | | |
| I1 | -.82 | ** | -.02 | | .09 | | -.16 | | .31 | * | .17 | | .03 | | -.34 | * | .11 | |
| I2 | .83 | ** | -.18 | | .50 | ** | .15 | | .14 | | .58 | ** | -.23 | | -.18 | | -.08 | |
| C1 | -.03 | | .85 | ** | -.39 | ** | .13 | | -.47 | ** | -.36 | * | -.13 | | .38 | ** | .25 | |
| C2 | -.17 | | .88 | ** | -.35 | * | -.07 | | -.43 | ** | -.43 | ** | .20 | | .29 | * | .25 | |
| E1 | .25 | | -.39 | ** | .89 | ** | -.03 | | .15 | | .74 | ** | -.06 | | -.19 | | -.48 | ** |
| E2 | .39 | ** | -.39 | ** | .92 | ** | .14 | | .27 | | .79 | ** | -.15 | | -.20 | | -.33 | * |
| S1 | .01 | | .08 | | .01 | | .88 | ** | -.15 | | .04 | | -.70 | ** | .24 | | .22 | |
| S2 | -.02 | | -.02 | | .11 | | .87 | ** | -.11 | | .17 | | -.76 | ** | .09 | | .09 | |
| SD | -.18 | | .31 | * | .07 | | .14 | | -.28 | | -.08 | | -.00 | | .65 | ** | -.13 | |
| ICES Major scales: | | | | | | | | | | | | | | | | | | |
| I | 1.00 | | -.12 | | .36 | * | -.01 | | .27 | | .47 | ** | -.12 | | -.32 | * | .01 | |
| C | -.12 | | 1.00 | | -.43 | ** | .03 | | -.53 | ** | -.46 | ** | .05 | | .39 | ** | .29 | * |
| E | .36 | * | -.43 | ** | 1.00 | | .07 | | .24 | | .85 | ** | -.12 | | -.22 | | -.45 | ** |
| S | -.01 | | .03 | | .07 | | 1.00 | | -.15 | | .12 | | -.83 | ** | .19 | | .18 | |
| EPQ scales: | | | | | | | | | | | | | | | | | | |
| EPQ-P | .27 | | -.53 | ** | .24 | | -.15 | | 1.00 | | .28 | | .08 | | -.40 | ** | -.29 | * |
| EPQ-E | .46 | ** | -.46 | ** | .85 | ** | .12 | | .28 | | 1.00 | | -.19 | | -.40 | ** | -.34 | ** |
| EPQ-N | -.12 | | .05 | | -.12 | | -.83 | ** | .08 | | -.19 | | 1.00 | | -.09 | | -.17 | |
| EPQ-L | -.32 | * | .39 | ** | -.22 | | .19 | | -.40 | ** | -.40 | ** | -.09 | | 1.00 | | .06 | |
| 1-tailed Signif: *p <.01; ** p<.001 | | | | | | | | | | | | | | | | | | |

***Appendix I.7: Factor analysis (principal components) - Varimax rotated factor loadings***

| | FACTORS | | | | |
|---|---|---|---|---|---|
| | I | II | III | IV | Communality |
| | "E" | "C" | "S" | "I" | |
| ICES-E2 | .87 | -.12 | .11 | .09 | .79 |
| EPQ-E | .87 | -.25 | .12 | .17 | .87 |
| ICES-E1 | .86 | -.07 | -.08 | -.06 | .75 |
| | | | | | |
| ICES-SD | .24 | .76 | .00 | -.29 | .71 |
| EPQ-L | -.10 | .74 | .13 | -.35 | .70 |
| ICES-C1 | -.38 | .67 | .16 | .34 | .73 |
| EPQ-P | .20 | -.67 | -.12 | .09 | .50 |
| ICES-C2 | -.36 | .64 | -.18 | .18 | .60 |
| | | | | | |
| EPQ-N | -.09 | .01 | -.93 | -.06 | .88 |
| ICES-S2 | .11 | .04 | .86 | -.10 | .76 |
| ICES-S1 | -.02 | .13 | .82 | -.02 | .69 |
| | | | | | |
| ICES-I1 | .07 | -.18 | -.13 | .82 | .73 |
| ICES-I2 | .57 | -.04 | .21 | .59 | .72 |
| Total eigenvalue = 9.46<br>72.80% of the variance accounted for. | | | | | |

***Appendix I.8: Multiple regression analyses***

***Appendix I.8.1: Prediction of EPQ scores by ICES minor scales and by ICES major scales***

| | Major scales (included SD) | | Minor scales | |
|---|---|---|---|---|
| Scale | R | Adjusted $R^2$ as a percentage | R | Adjusted $R^2$ as a percentage |
| EPQ-P | 0.59 | 29.05 | 0.65 | 23.65 |
| EPQ-E | 0.88 | 75.63 | 0.89 | 75.30 |
| EPQ-N | 0.85 | 69.98 | 0.87 | 72.38 |
| EPQ-L | 0.73 | 49.36 | 0.78 | 53.98 |
| | | Mean = 56.01 | | Mean = 58.58 |

***Appendix I.8.2: Prediction of ICES major scale scores with EPQ scales as predictors***

| Scale | R | Adjusted $R^2$ as a percentage |
|---|---|---|
| ICES-I | 0.50 | 20.30 |
| ICES-C | 0.63 | 35.73 |
| ICES-E | 0.86 | 72.80 |
| ICES-S | 0.84 | 69.21 |
| | | Mean = 49.51 |

***Appendix I.8.3: Prediction of ICES minor scale scores with EPQ scales as predictors***

| | Domain Scales | |
|---|---|---|
| **Scale** | **R** | **Adjusted $R^2$ as a percentage** |
| ICES-I1 | 0.40 | 10.30 |
| ICES-I2 | 0.60 | 31.79 |
| ICES-C1 | 0.56 | 27.54 |
| ICES-C2 | 0.57 | 28.04 |
| ICES-E1 | 0.76 | 54.41 |
| ICES-E2 | 0.81 | 63.02 |
| ICES-S1 | 0.73 | 49.64 |
| ICES-S2 | 0.76 | 55.63 |
| ICES-SD | 0.69 | 44.66 |
| | | Mean = 40.57 |

***Appendix I.9: Mean and SDs for the ICES and BPI scales (N=90)***

| Scale | Mean | SD | Scale name |
|---|---|---|---|
| ICES major scales: | | | |
| ICES-I | 5.18 | 1.58 | Independence |
| ICES-C | 4.19 | 2.04 | Conscientiousness |
| ICES-E | 4.86 | 1.49 | Extraversion |
| ICES-S | 4.57 | 2.13 | Stability |
| ICES minor scales: | | | |
| ICES-I1 | 5.49 | 1.69 | Competitive |
| ICES-I2 | 5.17 | 1.53 | Assertive |
| ICES-C1 | 4.21 | 1.80 | Conventional |
| ICES-C2 | 4.56 | 2.08 | Organized |
| ICES-E1 | 5.32 | 1.50 | Group-oriented |
| ICES-E2 | 4.86 | 1.67 | Outgoing |
| ICES-S1 | 4.32 | 1.71 | Poised |
| ICES-S2 | 5.20 | 2.32 | Relaxed |
| ICES-SD | 5.41 | 2.25 | Social desirability |
| BPI primary scales: | | | |
| CHANGE | 6.24 | 2.10 | Change-oriented |
| RISK | 6.36 | 2.24 | Risk taking |
| COMPETITIVE | 4.87 | 1.89 | Competitive |
| LIMELIGHT | 5.51 | 2.09 | Limelight seeking |
| WORK | 4.89 | 2.04 | Work-oriented |
| STAMINA | 5.13 | 2.19 | Stamina |
| PERF | 4.60 | 2.09 | Perfectionist |
| TIME | 4.58 | 2.39 | Time managed |
| WARM | 5.91 | 2.46 | Warm |
| OUT | 5.69 | 2.06 | Outgoing |
| WORRY | 5.51 | 2.76 | Worrying |
| BPI secondary scales: | | | |
| DYNAMIC | 6.02 | 1.73 | Dynamic |
| WORK/STAMINA | 5.06 | 1.83 | Work stamina |
| CONTR | 4.36 | 2.21 | Controlled |
| EXTRAVT | 5.97 | 2.44 | Extravert |
| WORRY | 5.51 | 2.76 | Worrying |

***Appendix I.10: Correlations between the ICES minor scales and the BPI scales (n=9990) and correlations between each of the scales and sex (male coded 1, female coded 2)***

| | ICES Minor Scales | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | I | | | | C | | | | E | | | | S | | | | | | | |
| | I1 | | I2 | | C1 | | C2 | | E1 | | E2 | | S1 | | S2 | | SD | | Sex | |
| CHGE | -.11 | | .34 | ** | -.49 | -- | -.41 | ** | .23 | | .41 | ** | .09 | | .18 | | -.36 | ** | -.25 | * |
| RISK | .24 | | .49 | ** | -.44 | ** | -.55 | ** | .12 | | .49 | ** | .13 | | .33 | ** | -.26 | * | -.23 | |
| COMPET | .42 | ** | .25 | * | .01 | | .18 | | -.06 | | .14 | | -.01 | | .10 | | .05 | | .07 | |
| LIME | .20 | | .51 | ** | -.18 | | -.28 | * | .11 | | .56 | * | .14 | | .41 | ** | -.05 | | -.21 | |
| WORK | .12 | | -.12 | | .33 | ** | .45 | ** | -.18 | | -.20 | | -.04 | | -.22 | | .36 | ** | .26 | * |
| STAM | .34 | ** | .34 | ** | -.03 | | -.05 | | -.00 | | .19 | | .41 | ** | .43 | ** | .11 | | -.40 | ** |
| PERF | -.22 | | .01 | | .48 | ** | .42 | ** | .01 | | -.08 | | -.21 | | -.24 | | .34 | ** | .03 | |
| TIME | -.11 | | -.18 | | -.40 | ** | .70 | ** | .02 | | -.27 | * | .05 | | -.09 | | .28 | * | .08 | |
| WARM | -.27 | * | .11 | | -.22 | | -.24 | | .46 | ** | .45 | ** | -.08 | | .15 | | -.29 | * | .05 | |
| OUT | .05 | | .49 | ** | -.33 | ** | -.39 | ** | .36 | ** | .78 | ** | -.04 | | .16 | | -.22 | | -.01 | |
| WORRY | -.16 | | -.40 | ** | .22 | | .26 | * | -.27 | * | -.28 | * | -.62 | ** | -.76 | ** | .19 | | .44 | ** |
| SEX | -.21 | | -.24 | | .00 | | -.01 | | .04 | | -.02 | | -.29 | * | -.43 | ** | .07 | | | |
| N=90; 1-TAILED Signif: * p<.01; ** p<.001 | | | | | | | | | | | | | | | | | | | | |

***Appendix I.11: Four factor analysis (principal components) with Varimax rotation***

| | I | II | III | IV | |
|---|---|---|---|---|---|
| | "C" | "E" | "S" | "I" | Communality |
| C1 | -.74 | -.20 | -.01 | .00 | .59 |
| C2 | .81 | -.21 | -.01 | -.13 | .72 |
| SD | .60 | -.15 | .02 | .14 | -.40 |
| TIME | .67 | -.02 | .01 | -.08 | .45 |
| PERF | .56 | .15 | -.20 | .02 | .38 |
| RISK | .56 | .20 | .16 | .43 | .57 |
| WORK | .48 | -.18 | -.16 | .15 | .31 |
| CHNGE | -.47 | .32 | .08 | .10 | .34 |
| | | | | | |
| EI | -.01 | .59 | .21 | -.14 | .41 |
| E2 | -.32 | .77 | .05 | .35 | .81 |
| OUT | -.23 | .77 | -.03 | .33 | .75 |
| WARM | -.18 | .63 | .01 | -.22 | .48 |
| | | | | | |
| S1 | .00 | -.04 | .82 | .02 | .67 |
| S2 | -.12 | .17 | .88 | .14 | .83 |
| WORRY | .25 | -.17 | -.77 | -.22 | .73 |
| | | | | | |
| I1 | -.04 | -.27 | .05 | .63 | .48 |
| I2 | -.19 | .36 | .14 | .56 | .50 |
| COMPET | .20 | -.02 | .03 | .57 | .36 |
| LIME | -.13 | .42 | .22 | .57 | .56 |
| STAM | .02 | .04 | .42 | .48 | .41 |

***Appendix I.12: Prediction of ICES major scale scores with BPI primaries and of BPI secondary scale scores with ICES primaries as predictors***

| | BPI secondaries as predictors | |
|---|---|---|
| **Scale** | **R** | **Adjusted $R^2$ as a percentage** |
| ICES-I | 0.79 | 44.96 |
| ICES-C | 0.81 | 64.43 |
| ICES-E | 0.73 | 51.99 |
| ICES-S | 0.79 | 60.36 |

| | ICES minors as predictors | |
|---|---|---|
| **Scale** | **R** | **Adjusted $R^2$ as a percentage** |
| DYNAM | 0.74 | 52.74 |
| WSTAM | 0.59 | 30.40 |
| CONTR | 0.59 | 55.71 |
| EXTRA | 0.70 | 47.13 |
| WORRY2 | 0.82 | 66.09 |

***BPI Primaries predicted using all nine ICES scales***

| BPI Scale | R |
|---|---|
| CHNGE | 0.59 |
| RISK | 0.71 |
| COMPET | 0.59 |
| LIME | 0.71 |
| WORK | 0.56 |
| STAM | 0.61 |
| PERF | 0.66 |
| TIME | 0.72 |
| WARM | 0.65 |
| OUT | 0.79 |
| WORRY | 0.82 |

***Appendix I.13: Description of the Hogan Personality Inventory scales***
*Scale descriptions are taken from the Hogan Personality Inventory Manual, Hogan and Hogan (1995).*

| **Adjustment** | **"Measures the degree to which a person appears calm and self-accepting or conversely, self-critical and overly self-reflective".** |
|---|---|
| **HIC** | **Definition** |
| Empathy | Emotional Identification with others |
| Not Anxious | Absence of anxiety |
| No Guilt | Absence of regret |
| Calmness | Lack of emotionality |
| Even Tempered | Not moody or irritable |
| No Somatic Complaints | Lack of health concerns |
| Trusting | Not paranoid or suspicious |
| Good Attachment | Good relations with one's parents |
| **Ambition** | **"Measures the degree to which a person is socially self-confident, leader-like, competitive, and energetic".** |
| **HIC** | **Definition** |
| Competitive | Being competitive, ambitious and persistent |
| Self-Confidence | Confidence in oneself |
| No Depression | Feelings of contentment |
| Leadership | Capacity for learning |
| Identity | Satisfaction with one's life tasks |
| No Social Anxiety | Social self-confidence |
| **Sociability** | **"Measures the degree to which a person seems to need and/or enjoy interactions with others".** |
| **HIC** | **Definition** |
| Likes Parties | Enjoys parties |
| Likes Crowds | Finds large crowds exciting |
| Experience Seeking | Preference for variety and challenge |
| Exhibitionistic | Exhibitionistic tendencies |
| Entertaining | Being witty and entertaining |
| **Likeability** | **"Measures the degree to which a person is seen as perceptive, tactful, and socially sensitive".** |
| **HID** | **Definition** |
| Easy to live with | Tolerant and easy going nature |
| Sensitive | Tends to be kind and considerate |
| Caring | Interpersonal sensitivity |
| Likes People | Enjoys social interaction |
| No Hostility | Lack of hostility |

*Appendix I.13: Description of the Hogan Personality Inventory scales (continued)*

| **Prudence** | **"Measures the degree to which a person is conscientious, conforming and dependable".** |
|---|---|
| **HIC** | **Definition** |
| Moralistic | Adhering strictly to conventional values |
| Mastery | Being hard working |
| Virtuous | Being perfectionistic |
| Not Autonomous | Concerns about others' opinions of oneself |
| Not Spontaneous | Preference for predictability |
| Impulse Control | Lack of impulsivity |
| Avoids Trouble | Professed probity |
| **Intellectance** | **"Measures the degree to which a person is perceived as bright, creative, and interested in intellectual matters".** |
| **HIC** | **Definition** |
| Science | Interest in science |
| Curiosity | Curiosity about the world |
| Thrill Seeking | Enjoyment of adventure and excitement |
| Intellectual Games | Enjoys intellectual games |
| Generates Ideas | Ideational fluency |
| Culture | Interest in culture |
| **School Success** | **"Measures the degree to which a person seems to enjoy academic activities and values educational achievement for its own sake".** |
| **HIC** | **Definition** |
| Good Memory | Having a good memory |
| Education | Being a good student |
| Math Ability | Being good with numbers |
| Reading | Enjoys reading |

***Appendix I.14.1: Means and SD for ICES and HPI raw score scales (N=65)***

| HPI Scales | Mean | SD | Description |
|---|---|---|---|
| ADJ | 21.85 | 7.16 | ADJUSTMENT |
| AMB | 19.97 | 5.96 | AMBITION |
| SOC | 14.98 | 4.81 | SOCIABILITY |
| LIK | 17.57 | 3.86 | LIKEABILITY |
| PRU | 14.63 | 4.15 | PRUDENCE |
| INT | 13.46 | 4.62 | INTELLECTANCE |
| SCH | 5.88 | 3.34 | SCHOOL SUCCESS |
| | | | |
| SOI | 8.75 | 2.73 | SERVICE ORIENTATION |
| STR | 16.82 | 5.54 | STRESS TOLERANCE |
| REL | 8.83 | 3.34 | RELIABILITY |
| CLERK | 15.14 | 3.91 | CLERICAL POTENTIAL |
| SALES | 46.15 | 8.83 | SALES POTENTIAL |
| MANAGER | 24.60 | 6.64 | MANAGERIAL POTENTIAL |
| **ICES scales** | **Mean** | **SD** | **Description** |
| INDEP | 49.02 | 6.33 | INDEPENDENCE |
| EXTRAV | 49.08 | 8.63 | EXTRAVERSION |
| CONSC | 48.98 | 7.94 | CONSCIENTIOUSNESS |
| STABLE | 47.71 | 9.16 | STABILITY |
| SOCDES | 22.83 | 4.71 | SOCIAL DESIRABILITY/DISTORTION |
| | | | |
| I1 | 23.92 | 3.39 | COMPETITIVE, TOUGH MINDED |
| I2 | 25.09 | 4.44 | ASSERTIVE, FORTHRIGHT |
| C1 | 24.00 | 3.89 | CONVENTIONAL, CONCERN FOR MORAL VALUES |
| C2 | 24.98 | 5.31 | ORGANIZED, ATTENTION TO DETAIL |
| E1 | 24.28 | 4.43 | GROUP-ORIENTED, SOCIABLE |
| E2 | 24.80 | 5.56 | OUTGOING, TALKATIVE |
| S1 | 23.75 | 4.49 | POISED, UNRUFFLED, UNFLAPPABLE |
| S2 | 23.95 | 5.66 | RELAXED, NOT ANXIOUS |

***Appendix I.14.2: Means and standard deviations on ICES scales (sten scores) for each group.***

| | Production workers (N=28) | | Fire fighters (N=37) | |
|---|---|---|---|---|
| **ICES Scales** | **Sten Mean** | **SD** | **Sten Mean** | **SD** |
| I | 5.50 | 1.29 | 6.08 | 1.53 |
| C | 5.46 | 2.41 | 4.86 | 1.40 |
| E | 4.54 | 1.50 | 5.38 | 1.67 |
| S | 4.11 | 1.40 | 5.65 | 2.06 |
| SocDes | 6.43 | 1.71 | 5.41 | 1.72 |
| I1 | 6.07 | 1.36 | 6.46 | 1.17 |
| I2 | 5.29 | 1.38 | 5.54 | 1.74 |
| C1 | 5.82 | 1.94 | 5.00 | 1.37 |
| C2 | 5.21 | 2.25 | 5.19 | 1.63 |
| E1 | 4.86 | 1.69 | 5.30 | 1.54 |
| E2 | 4.50 | 1.73 | 5.78 | 2.02 |
| S1 | 4.43 | 1.50 | 5.08 | 1.61 |
| S2 | 4.36 | 1.77 | 6.19 | 2.44 |

***Appendix I.15.1: Table of correlations between ICES minor scales and HPI Occupational scales***

| | ICES Minor Scales | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **I1** | | **I2** | | **C1** | | C2 | | **E1** | | **E2** | | **S1** | | **S2** | |
| HPI-SOI | -0.10 | | 0.02 | | 0.17 | | 0.03 | | 0.39 | ** | 0.16 | | 0.31 | | 0.43 | *** |
| HPI-STR | 0.07 | | 0.18 | | -0.16 | | -0.05 | | 0.47 | *** | 0.34 | ** | 0.49 | *** | 0.59 | *** |
| HPI-REL | -0.23 | ‡ | -0.14 | | 0.09 | | 0.12 | | 0.18 | | -0.10 | | 0.23 | ‡ | 0.23 | ‡ |
| HPI-CLERK | -0.08 | | 0.25 | * | -0.07 | | -0.06 | | 0.40 | ** | 0.32 | * | 0.45 | *** | 0.38 | ** |
| HPI-SALES | 0.17 | | 0.14 | | -0.19 | | 0.34 | ** | 0.55 | *** | 0.62 | *** | -0.01 | | 0.26 | * |
| HPI-MANAGER | -0.07 | | 0.12 | | 0.67 | | 0.11 | | 0.36 | ** | 0.21 | | 0.34 | ** | 0.27 | * |
| ‡ p<0.1 * p<.05 ** p<.01 *** p<.001 | | | | | | | | | | | | | | | | |

***Appendix I.15.2: Correlations between HICs and ICES major scales***

| | ICES Scales | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **I** | | **C** | | **E** | | **S** | | **SocDes** | |
| Adjustment | | | | | | | | | | |
| Empathy | 0.06 | | 0.03 | | 0.25 | * | **0.51** | *** | 0.23 | ‡ |
| Not Anxious | 0.35 | ** | -0.08 | | 0.30 | * | **0.66** | *** | -0.15 | |
| No Guilt | 0.04 | | -0.04 | | 0.27 | * | **0.49** | *** | 0.04 | |
| Calmness | 0.14 | | -0.11 | | 0.19 | | **0.42** | *** | -0.18 | |
| Even Tempered | 0.10 | | 0.01 | | 0.35 | ** | **0.54** | *** | -0.09 | |
| No Somatic Complaints | 0.01 | | -0.06 | | 0.38 | ** | **0.38** | ** | -0.21 | ‡ |
| Trusting | -0.06 | | 0.06 | | 0.07 | | **0.20** | | 0.17 | |
| Good Attachments | -0.02 | | -0.01 | | 0.24 | ‡ | **0.24** | ‡ | 0.26 | * |
| Ambition | | | | | | | | | | |
| Competitive | **0.10** | | 0.27 | * | 0.24 | ‡ | 0.22 | ‡ | -0.01 | |
| Self-Confidence | **0.13** | | 0.09 | | 0.11 | | 0..33 | ** | -0.06 | |
| No Depression | **0.08** | | -0.13 | | 0.51 | *** | 0.29 | * | -0.02 | |
| Leadership | **0.28** | * | 0.01 | | 0.14 | | 0.09 | | -0.08 | |
| Identity | **0.05** | | 0.10 | | 0.15 | | 0.24 | ‡ | 0.10 | |
| No Social Anxiety | **0.21** | | -0.03 | | 0.37 | ** | 0.31 | * | -0.11 | |
| Sociability | | | | | | | | | | |
| Likes Parties | 0.20 | | -0.21 | ‡ | **0.64** | *** | 0.11 | | -0.20 | |
| Likes Crowds | -0.03 | | -0.08 | | **0.51** | *** | 0.20 | | -0.11 | |
| Experience Seeking | 0.11 | | -0.25 | * | **0.48** | *** | 0.08 | | -0.08 | |
| Exhibitionistic | 0.22 | ‡ | -0.29 | * | **0.42** | *** | -0.10 | | -0.45 | *** |
| Entertaining | 0.15 | | -0.33 | ** | **0.45** | *** | -0.05 | | -0.09 | |
| Likeability | | | | | | | | | | |
| Easy to Live With | -0.28 | * | 0.05 | | 0.17 | | 0.07 | | 0.14 | |
| Sensitive | -0.13 | | 0.04 | | 0.31 | * | 0.01 | | 0.16 | |
| Caring | -0.23 | ‡ | -0.14 | | 0.35 | ** | -0.14 | | 0.86 | |
| Likes People | 0.08 | | -0.05 | | 0.45 | *** | 0.13 | | -0.02 | |
| No Hostility | -0.17 | | -0.13 | | 0.02 | | 0.15 | | 0.15 | |
| Prudence | | | | | | | | | | |
| Moralistic | 0.09 | | **0.32** | ** | -0.12 | | 0.27 | | 0.36 | ** |
| Mastery | -0.01 | | **0.35** | ** | -0.06 | | 0.16 | | 0.23 | ‡ |
| Virtuous | -0.06 | | **0.18** | | 0.14 | | 0.37 | ** | 0.33 | ** |
| Not Autonomous | -0.29 | * | **0.02** | | 0.10 | | -0.34 | ** | 0.06 | |
| Not Spontaneous | -0.09 | | **0.29** | * | 0.14 | | 0.08 | | -0.10 | |
| Impulse Control | -0.16 | | **0.36** | ** | -0.20 | | 0.11 | | 0.02 | |
| Avoids Trouble | -0.26 | * | **0.05** | | -0.04 | | 0.13 | | 0.23 | ** |

‡ p<0.1 * p<0.05 ** p<0.01 *** p<0.001
**BOLD** indicates HICs which are expected to correlate with appropriate ICES scales

NOTE: Only 5 of the HPI's 7 scales were analyzed here, as it is these 5 scales which are linked with the ICES scales.

***Appendix I.16: Regression analyses***
***Appendix I.16.1: SOI as the dependent variable***

| Predictor | Coefficient | Beta | T | Signif |
|---|---|---|---|---|
| I1 | -0.029 | -0.036 | -0.280 | 0.781 |
| I2 | -0.138 | -0.225 | -1.616 | 0.112 |
| C1 | 0.189 | 0.269 | 2.127 | **0.038** |
| C2 | -0.051 | -0.100 | -0.702 | 0.486 |
| E1 | 0.183 | 0.298 | 2.263 | **0.028** |
| E2 | 0.014 | 0.028 | 0.210 | 0.834 |
| S1 | 0.123 | 0.203 | 1.106 | 0.274 |
| S2 | 0.135 | 0.280 | 1.789 | 0.079 |
| (Constant) | (-1.299) | | | |
| Multiple R:<br>$R^2$:<br>Adjusted $R^2$:<br>F(8,56): | 0.59<br>0.35<br>0.25<br>3.70** | | | |
| ** p<.01 | | | | |

***Appendix I.16.2: STR as the dependent variable***

| Predictor | Coefficient | Beta | T | Signif |
|---|---|---|---|---|
| I1 | 0.198 | 0.121 | 1.059 | 0.294 |
| I2 | -0.256 | -0.205 | -1.681 | 0.098 |
| C1 | -0.089 | -0.063 | -0.562 | 0.576 |
| C2 | -0.067 | -0.065 | -0.519 | 0.606 |
| E1 | 0.352 | 0.282 | 2.439 | **0.018** |
| E2 | 0.083 | 0.083 | 0.705 | 0.484 |
| S1 | 0.506 | 0.410 | 2.544 | 0.014 |
| S2 | 0.248 | 0.254 | 1.846 | 0.070 |
| (Constant) | (-6.236) | | | |
| Multiple R:<br>$R^2$:<br>Adjusted $R^2$:<br>F(8,56): | 0.70<br>0.50<br>0.42<br>6.89*** | | | |
| *** p<.001 | | | | |

***Appendix I.16.3: REL as the dependent variable***

| Predictor | Coefficient | Beta | T | Signif |
|---|---|---|---|---|
| I1 | -0.100 | -0.102 | -0.711 | 0.480 |
| I2 | -0.199 | -0.264 | -1.727 | 0.090 |
| C1 | 0.056 | 0.066 | 0.471 | 0.639 |
| C2 | 0.014 | 0.022 | 0.139 | 0.900 |
| E1 | 0.192 | 0.255 | 1.760 | 0.084 |
| E2 | -0.097 | -0.162 | -1.097 | 0.277 |
| S1 | 0.159 | 0.214 | 1.058 | 0.295 |
| S2 | 0.077 | 0.131 | 0.758 | 0.451 |
| (Constant) | (6.656) | | | |
| Multiple R:<br>$R^2$:<br>Adjusted $R^2$:<br>F(8,56): | 0.46<br>0.21<br>0.10<br>1.85ns | | | |
| ns: not significant | | | | |

***APPENDIX I.17: Mean sten scores and SD's for the ICES Scales (N=83)***

| ICES Scales | Sten Mean | Sten SD |
|---|---|---|
| Independence (I) | 4.86 | 1.80 |
| Conscientiousness (C) | 5.00 | 1.89 |
| Extraversion (E) | 4.43 | 1.88 |
| Stability (S) | 4.64 | 1.92 |
| Social Desirability (SD) | 6.08 | 1.90 |
| Competitive (I1) | 5.28 | 1.60 |
| Assertive (I2) | 4.92 | 1.77 |
| Conventional (C1) | 5.08 | 1.95 |
| Organized (C2) | 5.01 | 1.72 |
| Group-Oriented (E1) | 4.66 | 1.91 |
| Outgoing (E2) | 4.64 | 1.92 |
| Poised (S1) | 4.54 | 1.90 |
| Relaxed (S2) | 5.11 | 1.87 |

***APPENDIX I.18: Correlational table of ICES sub-scales with HDS scales (1-tailed)***

| ICES Scales<br>HDS Scales | I1 Competitive | I2 Assertive | C1 Convent | C2 Organized | E1 Group-oriented | E2 Outgoing | S1 Poised | S2 Relaxed |
|---|---|---|---|---|---|---|---|---|
| Volatile | 0.19* | -0.11 | 0.07 | 0.00 | -0.22* | -0.32** | -0.47*** | -0.47*** |
| Mistrustful | 0.45*** | 0.15 | -0.03 | -0.04 | -0.13 | -0.06 | -0.24* | -0.24* |
| Cautious | -0.04 | -0.59*** | 0.30** | 0.25* | -0.27** | -0.56*** | -0.65*** | -0.59*** |
| Detached | 0.39*** | -0.05 | 0.20* | 0.13 | -0.36** | -0.49*** | -0.11 | -0.15 |
| Passive Agg | 0.11 | -0.21* | 0.28** | 0.26** | -0.18 | -0.25* | -0.48*** | -0.39*** |
| Arrogant | 0.30** | 0.19* | -0.03 | -0.01 | 0.09 | -0.22* | 0.13 | 0.07 |
| Manipulative | 0.30** | 0.38*** | -0.30** | -0.32** | 0.17 | 0.50*** | 0.24* | 0.23* |
| Dramatic | 0.15 | 0.38*** | -0.23* | -0.29** | 0.33** | 0.57*** | 0.16 | 0.27* |
| Eccentric | 0.20* | 0.26** | -0.03 | -0.18 | 0.03 | 0.22* | 0.13 | 0.02 |
| Perfectionistic | 0.14 | -0.13 | 0.42*** | 0.49*** | -0.04 | -0.10 | -0.24* | -0.37*** |
| Dependent | -0.32** | -0.54*** | 0.28** | 0.29** | 0.10 | -0.25* | -0.21* | -0.35** |
| *p<0.05 | | | **p<0.01 | | | | ***p<0.001 | |

***APPENDIX I .19: Factor loadings of the ICES and HDS scales***

| | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 |
|---|---|---|---|---|---|
| ICES-I | 0.79 | | | | |
| Depend | -0.72 | | | | |
| Mistrust | 0.51 | 0.49 | | | |
| Arrogant | 0.40 | | | 0.40 | 0.50 |
| Volatile | | 0.84 | | | |
| ICES-S | | -0.78 | | | |
| Cautious | -0.45 | 0.69 | | | |
| ICES-E | | | 0.82 | | |
| Detached | | | -0.74 | | |
| Dramatic | | | 0.65 | | |
| Eccentric | | | | 0.80 | |
| Pass_Agg | | | | 0.59 | |
| Manipulative | 0.46 | | 0.42 | 0.52 | |
| Perfectionist | | | | | 0.87 |
| ICES-C | | | | | 0.69 |

All loading less than 0.4 not selected for the final solution

***Appendix J.1: The new ICES "Risk," "Change," and "Focus on Work" scales. Descriptive statistics for Phase 3 sample (n=516):***

| Scale | Raw mean | SD | Sten Mean | SD |
|---|---|---|---|---|
| GPR | 214.26 | 30.47 | 5.50 | 1.99 |
| IRB | 254.87 | 28.39 | 5.51 | 1.98 |
| EPR | 180.72 | 19.83 | 5.51 | 1.99 |
| RPR | 265.17 | 43.18 | 5.48 | 1.98 |
| CHANGE | 258.48 | 24.51 | 5.50 | 1.98 |
| WORK | 143.19 | 27.80 | 5.51 | 1.97 |

***Correlations between the composite scales (n=516):***

| | GPR | IRB | EPR | RPR | CHANGE |
|---|---|---|---|---|---|
| GPR | | | | | |
| IRB | 0.70 | | | | |
| EPR | 0.55 | 0.37 | | | |
| RPR | 0.64 | 0.74 | 0.13 | | |
| CHANGE | 0.79 | 0.73 | 0.62 | 0.68 | |
| WORK | -0.47 | -0.54 | -0.76 | 0.01 | -0.48 |

***Appendix J2: ICES and membership in University Societies. Average sten scores***

| | Whole Group n=80 | | Not in a Society n=30 | | In a Society N=50 | | |
|---|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | |
| AGE | 21.01 | 3.00 | 21.03 | 2.70 | 21.00 | 3.21 | |
| Scale: | | | | | | | |
| I1 | 5.54 | 1.88 | 5.10 | 1.90 | 5.80 | 1.84 | Competitive |
| I2 | 5.17 | 1.63 | 4.90 | 1.47 | 5.34 | 1.71 | Forthright |
| C1 | 4.17 | 1.74 | 3.70 | 1.32 | 4.46 | 1.91 | Traditional |
| C2 | 4.48 | 1.77 | 4.10 | 1.24 | 4.70 | 2.00 | Detail-consc |
| E1 | 4.35 | 1.47 | 5.23 | 1.33 | 5.42 | 1.54 | Group-oriented |
| E2 | 5.40 | 1.74 | 5.03 | 1.75 | 5.62 | 1.71 | Outgoing |
| S1 | 4.63 | 1.63 | 4.37 | 1.67 | 4.78 | 1.59 | Unruffled |
| S2 | 4.99 | 1.89 | 4.70 | 1.62 | 5.16 | 2.02 | Relaxed |
| SD | 5.73 | 1.86 | 5.63 | 1.71 | 5.78 | 1.95 | Social Des |
| | | | | | | | |
| I | 5.17 | 1.85 | 4.70 | 1.84 | 5.46 | 1.01 | Independence |
| C | 4.14 | 1.71 | 3.53 | 1.25 | 4.50 | 1.85 | Consciousness |
| E | 5.30 | 1.48 | 5.03 | 1.40 | 5.46 | 1.51 | Extraversion |
| S | 4.59 | 1.75 | 4.40 | 1.73 | 4.70 | 1.76 | Stability |

***Differences between those in and not in societies (raw scores)***

| | Not in a Society n=30 | | In a Society n=30 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Scale** | **Mean** | **SD** | **Mean** | **SD** | **F Ratio** | **p** | | **eta** | |
| I1 | 20.97 | 4.55 | 22.84 | 4.83 | 2.94 | 0.09 | | 0.19 | Competitive |
| I2 | 23.53 | 4.34 | 24.78 | 4.55 | 1.46 | 0.23 | | 0.14 | Forthright |
| C1 | 20.10 | 3.38 | 22.06 | 4.69 | 3.99 | 0.05 | * | 0.22 | Traditional |
| C2 | 21.87 | 3.79 | 23.80 | 5.35 | 3.00 | 0.09 | | 0.19 | Detail-conscious |
| E1 | 24.87 | 3.89 | 25.32 | 4.08 | 0.24 | 0.63 | | 0.06 | Group-oriented |
| E2 | 24.43 | 5.02 | 26.02 | 4.92 | 1.92 | 0.17 | | 0.16 | Outgoing |
| S1 | 22.17 | 4.59 | 23.42 | 4.25 | 1.54 | 0.22 | | 0.14 | Unruffled |
| S2 | 22.50 | 4.23 | 23.44 | 5.36 | 0.67 | 0.42 | | 0.09 | Relaxed |
| SD | 22.40 | 4.60 | 22.72 | 4.96 | 0.08 | 0.77 | | 0.03 | Social desirability |
| | | | | | | | | | |
| I | 44.50 | 7.91 | 47.62 | 7.92 | 2.92 | 0.09 | | 0.19 | Independence |
| C | 41.97 | 5.28 | 45.86 | 8.21 | 5.39 | 0.02 | * | 0.25 | Conscientiousness |
| E | 49.30 | 7.26 | 51.34 | 7.91 | 1.32 | 0.25 | | 0.13 | Extraversion |
| S | 44.67 | 7.54 | 46.86 | 8.70 | 1.31 | 0.26 | | 0.13 | Stability |
| Df and 78, *p<.05. | | | | | | | | | |

***Appendix J3: Prediction of Society membership and type of society joined.***

***Prediction of SOCIETY membership (1=in a society, 0=not in a society).***
***Multiple R=.32; F = 4.42, df 2,36, p<.05***

| Variables in the Equation | | | | | |
|---|---|---|---|---|---|
| **Scale** | **B** | **SE B** | **Beta** | **T** | **Sig T** |
| C1 | .033147 | .012743 | .294651 | 2.601 | .0111 |
| E2 | .023904 | .011075 | .244495 | 2.158 | .0340 |
| (constant) | -.689628 | .449751 | | -1.533 | .1293 |

Prediction of TYPE of society joined (1=social/community, 2=sports/activity).

Multiple r= .53; F – 6.92, df 2,36, p<.01

| Variables in the Equation | | | | | |
|---|---|---|---|---|---|
| **Scale** | **B** | **SE B** | **Beta** | **T** | **Sig T** |
| I1 | .038177 | .014491 | .379472 | 2.635 | .0123 |
| SD | -.030042 | .014270 | -.303227 | -2.105 | .0423 |
| (constant) | 1.325098 | .506388 | | 2.617 | .0129 |

***Appendix K.1: Job-related difference in scores for job-related groups of Black test takers (data collected during 1995-96).***

| | Clerk/sec n=80 | | Sales rep n=51 | | Manager n=35 | | Supervisor n=20 | | Analyst/prog n=19 | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Scale** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| VERB | 26.70 | 8.85 | 22.59 | 9.86 | 24.80 | 8.48 | 25.60 | 10.25 | 25.37 | 9.19 |
| NUM | 8.40 | 3.85 | 7.67 | 4.12 | 10.86 | 5.02 | 11.25 | 6.41 | 10.37 | 6.01 |
| Non-V | 10.11 | 4.28 | 8.88 | 4.14 | 11.03 | 4.60 | 12.30 | 5.02 | 12.68 | 4.49 |
| GEN | 63.72 | 19.56 | 55.69 | 21.50 | 68.57 | 24.25 | 72.70 | 27.70 | 71.47 | 25.31 |
| | | | | | | | | | | |
| PEOPLE | 41.66 | 7.31 | 42.65 | 7.20 | 45.34 | 7.03 | 42.95 | 7.67 | 44.63 | 44.98 |
| DATA | 37.50 | 7.03 | 33.80 | 6.78 | 32.51 | 6.31 | 33.90 | 6.07 | 32.79 | 6.11 |
| THINGS | 34.07 | 9.18 | 33.47 | 7.33 | 35.97 | 9.10 | 63.45 | 8.86 | 40.47 | 6.19 |
| | | | | | | | | | | |
| I1 | 23.85 | 3.97 | 23.47 | 3.70 | 24.74 | 3.03 | 23.25 | 3.34 | 24.89 | 3.56 |
| I2 | 22.94 | 4.44 | 25.31 | 4.23 | 26.74 | 3.66 | 26.70 | 4.37 | 25.21 | 4.34 |
| C1 | 26.81 | 4.25 | 26.27 | 4.31 | 26.00 | 3.68 | 27.25 | 3.27 | 25.26 | 3.89 |
| C2 | 27.27 | 4.41 | 27.27 | 4.90 | 27.34 | 5.76 | 29.15 | 3.51 | 28.16 | 4.36 |
| E1 | 22.40 | 4.68 | 23.12 | 4.73 | 23.09 | 4.64 | 23.05 | 4.15 | 23.37 | 3.20 |
| E2 | 20.64 | 5.25 | 23.49 | 5.35 | 23.85 | 5.08 | 23.50 | 5.63 | 24.37 | 4.62 |
| S1 | 24.81 | 4.53 | 24.82 | 4.56 | 24.71 | 4.83 | 26.85 | 3.88 | 27.47 | 4.22 |
| S2 | 23.48 | 4.72 | 23.80 | 5.05 | 24.83 | 4.68 | 25.70 | 4.82 | 26.11 | 4.25 |
| SD1 | 26.72 | 4.47 | 26.47 | 4.35 | 25.37 | 3.88 | 27.00 | 3.43 | 25.95 | 3.70 |
| I | 46.79 | 7.43 | 48.78 | 6.78 | 51.49 | 4.74 | 49.95 | 6.44 | 50.11 | 6.67 |
| C | 54.09 | 6.53 | 53.55 | 7.98 | 53.34 | 7.87 | 56.40 | 5.29 | 53.42 | 6.47 |
| E | 43.04 | 8.88 | 46.61 | 9.25 | 46.94 | 8.55 | 46.55 | 8.76 | 47.74 | 6.38 |
| S | 48.29 | 8.05 | 48.63 | 8.85 | 49.54 | 7.95 | 52.55 | 8.19 | 53.58 | 7.92 |

| Scale | eta | %var | F | p | |
|---|---|---|---|---|---|
| VERB | 0.17 | 3.05 | 1.57 | 0.182 | |
| NUM | 0.28 | 8.01 | 4.35 | 0.002 | ** |
| NON-V | 0.27 | 7.41 | 4.00 | 0.004 | ** |
| GEN | 0.25 | 6.43 | 3.44 | 0.009 | ** |
| | | | | | |
| PEOPLE | 0.19 | 3.75 | 1.95 | 0.104 | |
| DATA | 0.30 | 9.21 | 5.07 | 0.001 | ** |
| THINGS | 0.23 | 5.48 | 2.90 | 0.023 | * |
| | | | | | |
| I1 | 0.15 | 2.23 | 1.14 | 0.338 | |
| I2 | 0.35 | 12.13 | 6.90 | <.001 | *** |
| C1 | 0.13 | 1.75 | 0.89 | 0.470 | |
| C2 | 0.12 | 1.54 | 0.78 | 0.537 | |
| E1 | 0.08 | 0.68 | 0.34 | 0.850 | |
| E2 | 0.28 | 8.03 | 4.36 | 0.002 | ** |
| S1 | 0.21 | 4.24 | 2.21 | 0.069 | |
| S2 | 0.19 | 3.76 | 1.95 | 0.103 | |
| SD1 | 0.13 | 1.59 | 0.81 | 0.521 | |
| | | | | | |
| I | 0.26 | 6.57 | 3.51 | 0.009 | ** |
| C | 0.12 | 1.46 | 0.74 | 0.564 | |
| E | 0.21 | 4.61 | 2.42 | 0.049 | * |
| S | 0.22 | 4.62 | 2.43 | 0.049 | * |

Df = 4 & 200 for all F ratios

* p<.05 ** p<.01 *** p<.001

***Appendix K.2: Means and SD for White and Black Groups matched by job type.***

| | Clerks/Secretaries | | | | Sales reps | | | | Managers | | | | Supervisors | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | White n=83 | | Black n-80 | | White n=36 | | Black n=51 | | White n=72 | | Black n-35 | | White n-25 | | Black n-20 | |
| . | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| VERB | 27.03 | 9.17 | 26.70 | 8.85 | 22.97 | 9.86 | 22.59 | 9.56 | 25.32 | 9.70 | 24.80 | 8.48 | 19.04 | 7.93 | 25.80 | 10.25 |
| NUM | 10.51 | 5.14 | 8.40 | 3.85 | 11.78 | 5.08 | 7.97 | 4.12 | 13.19 | 5.54 | 10.86 | 5.02 | 8.44 | 4.03 | 11.25 | 6.41 |
| NON-V | 12.31 | 4.81 | 10.11 | 4.28 | 12.08 | 4.39 | 8.88 | 4.14 | 13.10 | 4.32 | 11.03 | 4.60 | 12.16 | 5.21 | 12.30 | 5.02 |
| GEN | 72.67 | 2369 | 63.72 | 19.56 | 70.69 | 23.45 | 55.69 | 21.50 | 77.90 | 23.74 | 68.57 | 24.25 | 60.24 | 21.94 | 72.70 | 27.70 |
| | | | | | | | | | | | | | | | | |
| People | 36.63 | 8.85 | 41.86 | 7.31 | 39.89 | 8.96 | 42.65 | 7.20 | 40.04 | 7.81 | 45.34 | 7.03 | 37.80 | 9.31 | 42.95 | 7.67 |
| Data | 32.11 | 7.10 | 37.50 | 7.03 | 31.31 | 8.81 | 33.80 | 6.78 | 29.90 | 7.94 | 32.51 | 6.31 | 27.92 | 58.99 | 33.90 | 6.07 |
| Things | 30.63 | 9.33 | 34.08 | 9.18 | 33.75 | 11.34 | 33.47 | 7.33 | 36.92 | 9.79 | 35.97 | 9.10 | 35.88 | 11.47 | 36.45 | 8.86 |
| | | | | | | | | | | | | | | | | |
| I1 | 20.53 | 4.07 | 23.85 | 3.97 | 24.25 | 4.43 | 23.47 | 3.70 | 24.07 | 4.80 | 24.74 | 3.03 | 23.36 | 3.94 | 23.25 | 3.34 |
| I2 | 23.52 | 5.36 | 22.94 | 4.44 | 25.83 | 5.60 | 25.31 | 4.23 | 26.42 | 4.83 | 26.74 | 3.66 | 26.20 | 4.98 | 26.70 | 4.37 |
| C1 | 26.33 | 4.01 | 26.81 | 4.25 | 26.75 | 3.86 | 26.27 | 4.31 | 26.10 | 4.59 | 26.00 | 3.68 | 27.56 | 4.69 | 27.25 | 3.27 |
| C2 | 24.29 | 5.11 | 27.27 | 4.41 | 25.36 | 5.57 | 27.27 | 4.90 | 26.56 | 5.07 | 27.34 | 5.76 | 25.52 | 5.08 | 29.15 | 3.51 |
| E1 | 23.32 | 5.81 | 22.40 | 4.85 | 24.50 | 5.02 | 23.12 | 4.73 | 11055 | 5.34 | 23.09 | 4.64 | 21.92 | 5.46 | 23.05 | 4.15 |
| E2 | 23.28 | 6.16 | 20.64 | 5.25 | 24.61 | 5.86 | 23.49 | 5.35 | 23.06 | 5.99 | 23.86 | 5.08 | 23.92 | 6.29 | 23.50 | 5.83 |
| S1 | 24.05 | 4.99 | 24.81 | 4.53 | 25.14 | 4.93 | 24.82 | 4.56 | 24.82 | 4.71 | 24.71 | 4.83 | 23.84 | 5.57 | 26.85 | 3.88 |
| S2 | 23.14 | 4.87 | 23.48 | 4.72 | 25.22 | 3.80 | 23.80 | 5.05 | 25.75 | 4.87 | 24.83 | 4.88 | 23.582 | 4.80 | 25.70 | 4.82 |
| SD | 24.29 | 5.41 | 26.73 | 4.47 | 34.33 | 6.10 | 26.47 | 4.35 | 22.83 | 5.31 | 26.37 | 3.88 | 21.48 | 6.59 | 27.00 | 3.43 |
| | | | | | | | | | | | | | | | | |
| I | 44.04 | 8.05 | 46.79 | 7.43 | 50.08 | 6.77 | 48.78 | 6.78 | 50.49 | 6.27 | 51.49 | 4.74 | 48.56 | 7.68 | 49.85 | 6.44 |
| C | 50.62 | 7.83 | 54.09 | 6.53 | 52.11 | 8.14 | 53.56 | 7.96 | 52.65 | 8.49 | 53.34 | 7.87 | 53.08 | 8.38 | 56.40 | 5.29 |
| E | 46.50 | 10.83 | 43.04 | 6.88 | 49.11 | 10.01 | 46.91 | 9.25 | 45.74 | 9.87 | 46.94 | 8.55 | 45.84 | 11.03 | 46.55 | 8.76 |
| S | 47.19 | 9.06 | 48.29 | 8.05 | 50.36 | 7.82 | 46.83 | 8.85 | 50.57 | 8.46 | 49.54 | 7.95 | 48.36 | 8.96 | 52.55 | 8.19 |

***Appendix K.3: The 1995 Case Study***

| | White (1) n=987 | | Black (2) n=30 | | Hispanic (3) n=54 | | Oriental (4) n-35 | | Anova by RACE (1 to 4) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | F | P | eta | Df |
| AGE | 26.71 | 4.51 | 27.03 | 4.48 | 24.09 | 3.47 | 25.40 | 3.70 | 1.19 | 0.310 | 0.057 | 3, 1101 |
| | | | | | | | | | | | | |
| VERB | 26.25 | 8.47 | 24.87 | 9.34 | 26.30 | 7.71 | 26.29 | 8.71 | 0.26 | 0.854 | 0.027 | 3, 1102 |
| NUM | 11.72 | 5.05 | 10.23 | 4.46 | 10.85 | 5.26 | 10.37 | 4.16 | 2.02 | 0.109 | 0.074 | 3, 1102* |
| NON-V | 13.63 | 3.91 | 11.53 | 4.47 | 12.61 | 4.00 | 13.46 | 3.31 | 3.74 | 0.011 | 0.100 | 3, 1102* |
| GEN | 76.94 | 20.76 | 68.40 | 23.01 | 73.22 | 21.66 | 73.94 | 18.42 | 2.28 | 0.078 | 0.079 | 3, 1102 |
| | | | | | | | | | | | | |
| PEOPLE | 45.62 | 6.55 | 47.57 | 4.78 | 45.69 | 6.87 | 46.43 | 5.35 | 1.03 | 0.381 | 0.053 | 3, 1102 |
| DATA | 26.51 | 7.31 | 28.77 | 5.46 | 26.83 | 7.24 | 26.34 | 5.70 | 0.98 | 0.399 | 0.052 | 3, 1102 |
| THINGS | 31.70 | 8.81 | 28.17 | 6.84 | 30.26 | 9.57 | 32.43 | 9.55 | 2.06 | 0.104 | 0.075 | 3, 1102 |
| | | | | | | | | | | | | |
| I1 | 26.74 | 4.44 | 25.73 | 4.12 | 25.96 | 4.18 | 25.66 | 4.47 | 1.60 | 0.188 | 0.066 | 3, 1102 |
| I2 | 28.41 | 4.29 | 28.20 | 4.36 | 28.50 | 4.07 | 27.83 | 4.08 | 0.24 | 0.869 | 0.026 | 3, 1102 |
| C1 | 25.59 | 4.00 | 26.20 | 4.40 | 26.33 | 4.02 | 25.89 | 4.61 | 0.82 | 0.482 | 0.047 | 3, 1102 |
| C2 | 24.50 | 5.19 | 26.17 | 5.58 | 26.04 | 5.32 | 25.94 | 4.95 | 3.13 | 0.025 | 0.092 | 3, 1102* |
| E1 | 28.03 | 4.33 | 27.63 | 4.32 | 28.26 | 4.27 | 28.83 | 3.64 | 0.53 | 0.664 | 0.038 | 3, 1102 |
| E2 | 30.43 | 4.69 | 28.70 | 4.93 | 29.81 | 4.56 | 30.83 | 4.44 | 1.69 | 0.167 | 0.068 | 3, 1102 |
| S1 | 25.46 | 4.49 | 26.73 | 4.73 | 25.78 | 5.53 | 24.51 | 5.56 | 1.35 | 0.258 | 0.060 | 3, 1102 |
| S2 | 26.41 | 4.46 | 25.57 | 4.28 | 26.81 | 5.44 | 26.00 | 5.29 | 0.58 | 0.629 | 0.040 | 3, 1102 |
| SD | 22.27 | 4.81 | 21.33 | 4.69 | 23.24 | 4.82 | 22.06 | 4.43 | 1.14 | 0.332 | 0.056 | 3, 1102 |
| I | 55.15 | 7.23 | 53.93 | 7.11 | 54.46 | 6.91 | 53.49 | 7.65 | 0.97 | 0.407 | 0.051 | 3, 1102 |
| C | 50.09 | 7.78 | 52.37 | 8.13 | 52.37 | 7.17 | 51.83 | 6.96 | 2.71 | 0.044 | 0.086 | 3, 1102* |
| E | 58.46 | 7.92 | 56.33 | 8.60 | 58.07 | 7.50 | 59.66 | 7.40 | 1.03 | 0.380 | 0.053 | 3, 1102 |
| S | 51.87 | 7.90 | 52.30 | 8.13 | 52.59 | 9.92 | 50.51 | 9.72 | 0.50 | 0.682 | 0.037 | 3, 1102 |

**Appendix L: JOB DESCRIPTION SURVEY**

**JOB TITLE:** _____ **DATE:** _____

**COMPLETED BY:** _____ **COMPLETED BY:** _____

**Choose the appropriate job requirement rating as it applies to this particular job.**

**Circle either:**

| often | in moderation | rarely |
|---|---|---|

**YOUR CHOICE SHOULD ONLY REFLECT THE JOB IN QUESTION NOT THE CANDIDATE.**

**IT IS VERY IMPORTANT THAT YOUR SELECTION DOES NOT EXCEED PERFORMANCE STANDARDS.**

**Ability required to do this job**: **"The speed at which one learns new tasks".**

| | | | | |
|---|---|---|---|---|
| 1 | The use of simple arithmetic occurs... | often | in moderation | rarely |
| 2 | Good reading and writing skills are essential... | often | in moderation | rarely |
| 3 | Problem solving by mentally sorting, organizing or visualizing processes occurs... | often | in moderation | rarely |
| 4 | The processing of information derived from numbers is necessary... | often | in moderation | rarely |
| 5 | Using charts, diagrams, or contemplating logical steps in a process occurs... | often | in moderation | rarely |
| 6 | An ability to use language to reason or solve problems is Required in this job… | often | in moderation | rarely |
| 7 | A need to process numbers, files or records occurs... | often | in moderation | rarely |
| 8 | Emphasis on written communication is significant... | often | in moderation | rarely |
| 9 | Working with diagrams, plans, or models relating to objects or buildings is required... | often | in moderation | rarely |
| 10 | Speed and accuracy in working with numbers is required... | often | in moderation | rarely |
| 11 | *Good language skills are required to comprehend complex documents…* | often | in moderation | rarely |
| 12 | This job requires mental manipulation of images of shapes and objects… | often | in moderation | rarely |

**Interest: "The level of interest, desire or motivation to do certain tasks".**

| | | | | |
|---|---|---|---|---|
| 13 | The use of information systems, technical documents or data takes place... | often | in moderation | rarely |
| 14 | The workplace is frequently quiet without significant interaction with others... | often | in moderation | rarely |
| 15 | The position depends on "hands-on" interest in working with things... | often | in moderation | rarely |
| 16 | This job requires regular and frequent contact with people... | often | in moderation | rarely |
| 17 | Tasks that do not require an emphasis on data and facts occur... | often | in moderation | rarely |
| 18 | The use of machinery, keyboards or other equipment or tools happens... | often | in moderation | rarely |
| 19 | Work requires interacting with other people... | often | in moderation | rarely |
| 20 | This job requires working with numbers, data or financial information... | often | in moderation | rarely |
| 21 | **Limited use of equipment or machinery occurs…** | often | in moderation | rarely |

**Personality Traits: "Characteristics which influence behavior differently in situations".**

| | | | | |
|---|---|---|---|---|
| 22 | This job requires people to put the needs of the team before their own personal goals... | often | in moderation | rarely |
| 23 | This job involves working with others in groups or teams... | often | in moderation | rarely |
| 24 | A need for creativity and spontaneity occurs... | often | in moderation | rarely |
| 25 | There is relatively little interpersonal conflict... | often | in moderation | rarely |
| 26 | Emphasis on mediating and diplomacy occurs... | often | in moderation | rarely |
| 27 | A regular orderly job with few distractions occurs... | often | in moderation | rarely |
| 28 | This job requires dealing with challenge and criticism… | often | in moderation | rarely |
| 29 | Complying with traditional methods and guidelines is required... | often | in moderation | rarely |
| 30 | A low-pressure job with low levels of emotional demand occurs... | often | in moderation | rarely |
| 31 | This job requires people to be outgoing, talkative and easy-going... | often | in moderation | rarely |
| 32 | In this job there is minimal conflict with others... | often | in moderation | rarely |
| 33 | Little contact with others happens... | often | in moderation | rarely |
| 34 | Dealing with adversity or urgency occurs... | often | in moderation | rarely |
| 35 | This job offers a variety of stimulating and exciting activities... | often | in moderation | rarely |
| 36 | Achieving individual goals rather than team targets and goals is stressed... | often | in moderation | rarely |

| | | | | |
|---|---|---|---|---|
| 37 | Tight time frames and demanding schedules arise… | often | in moderation | rarely |
| 38 | Special emphasis on planning and prioritizing work is required... | often | in moderation | rarely |
| 39 | Teamwork and collaboration is stressed... | often | in moderation | rarely |
| 40 | The workplace is peaceful, serene and predictable... | often | in moderation | rarely |
| 41 | Emphasis on winning is important... | often | in moderation | rarely |
| 42 | Work circumstances and situations change... | often | in moderation | rarely |
| 43 | This job requires taking command of situations or people... | often | in moderation | rarely |
| 44 | A flexible application of policies or practices is allowed... | often | in moderation | rarely |
| 45 | The position could be described as calm, harmonious and non-stressful... | often | in moderation | rarely |
| 46 | This position requires working within a social environment... | often | in moderation | rarely |
| 47 | A fast-paced, high-pressure day is expected... | often | in moderation | rarely |
| 48 | A regular, consistent workload is part of this job... | often | in moderation | rarely |
| 49 | This position requires decision making, taking control... | often | in moderation | rarely |
| 50 | The position requires “thinking on your feet”... | often | in moderation | rarely |
| 51 | This job does not involve dealing with adversarial or hostile situations... | often | in moderation | rarely |
| 52 | Attention to detail is important... | often | in moderation | rarely |
| 53 | Long periods of solitary work occur... | often | in moderation | rarely |

**APPENDIX M: Professor David Bartram Publications**

***Books, technical and software manuals.***


D. Bartram, & R. Bayliss. (1984). *Test Administrator's Manual for the Microcomputerised Personnel Aptitude Tester.* Hull: Bartdale.

D. Bartram, (1985). *Chord-keyboard training system: Instructor's Manual.* Post Office Research: Swindon. 23pp.

D. Bartram, (1985). *Chord-keyboard training system*: *Trainee's Manual*. Post Office Research: Swindon. 43pp.

D. Bartram, (1985). *Chord-keyboard training system: Technical software Manual.* Post Office Research: Swindon. 20pp.

D. Bartram, (1987). *AEG-Telefunken flat pack sorter training system: Instructor's Manual.* Post Office Research: Swindon. 22pp.

D. Bartram, (1987). *AEG-Telefunken flat pack sorter training system: Trainee's Manual*. Post Office Research: Swindon. 22pp.

D. Bartram, (1987). *AEG-Telefunken flat pack sorter training system: Technical software Manual*. Post Office Research: Swindon. 20pp.

D. Bartram, (1987). *The Vocational Preference Inventory: Software documentation.* NFER-NELSON Publishing Co.: Windsor. 33pp.

D. Bartram, (1987*). Data Analysis Disk manual for the BBC microcomputer.* NFER-NELSON Publishing Co.: Windsor. 43pp.

D. Bartram, (1987). *The Lewis Counselling Inventory: Software documentation*. NFER-NELSON Publishing Co.: Windsor. 28pp.

D. Bartram, (1988). *The MICROPAT tests: software documentation for the test administration and profiler modules*. NPAL: Hull. 88pp.

D. Bartram, (1988). *The MICROPAT tests: norms supplement.* NPAL: Hull. 237pp.

D. Bartram, (1989). *Flat pack and packet sorter training system*: *Instructor's Manual*. Post Office Research: Swindon. 22pp.

D. Bartram, (1989). *Flat pack and packet sorter training system: Trainee's Manual.* Post Office Research: Swindon. 30pp.

D. Bartram, (1989). *Flat pack and packet sorter training system: Technical software Manual.* Post Office Research: Swindon. 21pp.

D. Bartram, (1989*). SCREENTEST Assessment and Data Manager User's Manual.* NFER-NELSON: Windsor. 139pp.

D. Bartram, (1989). *Sixteen Personality Factor Questionnaire: software documentation for the SCREENTEST test administration modules*. NFER-NELSON: Windsor. 14pp.

D. Bartram, (1989). *Sixteen Personality Factor Questionnaire: software documentation for the SCREENTEST profiler module.* NFER-NELSON: Windsor. 18pp.

D. Bartram, (1989). *Myers-Briggs Type Indicator: software documentation for the SCREENTEST test administration module.* NFER-NELSON: Windsor. 11pp.

D. Bartram, (1989). *Graduate and Managerial Assessment - Numerical, Verbal and Abstract: software documentation for the SCREENTEST test administration and profiler modules.* NFER-NELSON: Windsor. 32pp.

D. Bartram, P.A. Lindley and J.M. Foster (1990*) A Review of Psychometric Tests for Assessment in Vocational Training.* The Training Agency: Sheffield. c. 334pp.

D. Bartram, (1990) *SCREENTEST Test Interpretation Authoring and Narrative Report Generation Modules. Version 1.1. General system documentation.* Ministry of Defence. 39pp.

D. Bartram, P.A. Lindley and J.M. Foster. (1991) *Profile of Moods States - Bipolar form: British normative data supplement.* NFER-NELSON: Windsor.

D. Bartram, P.A. Lindley and J.M. Foster. (1991) *Culture Free Self-Esteem Inventory: British normative data supplement.* NFER-NELSON: Windsor.

D. Bartram, (1991). *MICROPAT Version 5.1: Normative Data.* Hull: Bartdale. pp 250

D. Bartram, P.A. Lindley, J. Foster. (1992) *Review of Psychometric Tests for Assessment in Vocational Training*. BPS Books: Leicester. 315pp

D. Bartram, P A Lindley, J Foster. (1992). *1992 Update to the Review of Psychometric Tests for Assessment in Vocational Training.* BPS Books: Leicester. 64pp

D. Bartram, (1992) Norms for a national sample of UK Managers and Supervisors applying for new jobs. *Data Plus, 2.7*, NFER-NELSON: Windsor. 3pp

D. Bartram, (1992) Norms for UK Managers and Supervisors applying for management positions. *Data Plus, 10.5*, NFER-NELSON: Windsor. 5pp

D. Bartram, (1992) *The British Psychological Society's Certificate of Competence in Occupational Testing (Level A): Guidance for Assessors.* Leicester: British Psychological Society pp50

D. Bartram, (1993) *PREVUE ASSESSMENT Technical Manual.* Barbados: PREVUE Assessment Inc. 97pp
D. Bartram, J Foster, P A Lindley, A Brown and S Nixon. (1993) *The Learning Climates Questionnaire: A User's Guide.* Sheffield & Hull: Employment Service and Newland Park Associates. pp 19.
D. Bartram, J Foster, P A Lindley, A Brown and S Nixon. (1993) *The Learning Climates Questionnaire: Background and Technical information.* Sheffield & Hull: Employment Service and Newland Park Associates. pp 24
D. Bartram, & P.A. Lindley (1994). *Psychological Testing: The British Psychological Society's Level A Open Learning Programme.* [Distance learning package containing 12 sections.]
L. Mitchell & D. Bartram, (1994). *The Place of Knowledge and Understanding in the Development of NVQs and SVQs.* Competence & Assessment, Briefing Series #10. Sheffield: Employment Department.
D. Bartram, (1994) *PREVUE ASSESSMENT Technical Manual: Second Edition.* Barbados: PREVUE Assessment Inc. pp
D. Bartram, (Senior Editor) with N. Anderson, D. Kellet, P.A. Lindley and I. Robertson (Consulting Editors). (1995). *Review of Personality Assessment Instruments (Level B) for use in occupational settings.* Leicester: BPS Books. 300pp. ISBN 1-85433-151-5
D. Bartram, (1996). *The involvement of Higher Education in the delivery of S/NVQ's: A stimulus paper for a series of NCVQ workshops.* London: National Council for Vocational Qualifications. 38pp.
D. Bartram, (1997). (Senior Editor, with E. Burke, R. Kandola, P.A. Lindley, L. Marshall, and P. Rasch, Consulting Editors). *Review of Tests of Ability and Aptitude (Level A) for use in occupational settings.* Leicester: BPS Books.

### *Articles in refereed journals or volumes, book chapters.*

D. Bartram, (1973). "The effects of familiarity and practice on naming pictures of objects*". Memory and Cognition, 1*, 101- 105.
D. Bartram, (1974). "The role of visual and semantic codes in object naming". *Cognitive Psychology, 6*, 325-356.
D. Bartram, (1976). "Levels of coding in picture-picture comparison tasks". *Memory and Cognition, 6*, 593-602.
D. Bartram, (1978). "Post-iconic visual storage: chunking in the reproduction of briefly displayed visual patterns". *Cognitive Psychology, 10*, 324-355.
D. Bartram, (1980). "Comprehending spatial information: the relative efficiency of different methods of presenting information about bus routes". *Journal of Applied Psychology, 65*, 103-110.
D. Bartram, (1980). "Do you really need your null hypothesis?" *Bulletin of the British Psychological Society, 33*, 318-321.
N. Sprent, D. Bartram, & C.M.Crawshaw. (1980). "Intelligibility of bus timetables". In Oborne, D.J. & Levis, J.A. (Eds.) *Human Factors in Transport Research, Vol. 1..* London: Academic Press.
D. Bartram, (1982). "Perception of semantic quality in type: differences between designers and non-designers". *Information Design Journal, 3*, 38-50.
D. Bartram, & H.C.A. Dale. (1982). "The E.P.I. as a selection test for military pilots". *Journal of Occupational Psychology, 55,* 287-296.
N. Sprent, C.M.Crawshaw, & D. Bartram, (1982). "Making timetables easier to read". *Buses,* October, 440-441.
D. Bartram, & P.A. Lindley. (1983). "Perception and Behaviour in the Flow of Events*". The New Psychologist,* January, 37-49.
D. Bartram, (1983) "EDAS: An interactive expert system for experimental design analysis". In *The User Interface: the ergonomics of interactive computing,* Loughborough: The Ergonomics Society.
N. Sprent, D. Bartram, & C.M. Crawshaw. (1983). "Structuring timetable information". *Ergonomics, 26*, 505-516.
D. Bartram, (1984). "The presentation of information about bus services". In Easterby, R. & Zwaga, H (Eds.) *Information Design: The design and evaluation of signs.* Chichester: Wiley.
D. Bartram, (1984). "Using Microsoft Basic for statistical analysis: Some cautionary notes". *Bulletin of the British Psychological Society, 37*, 297-300.
D. Bartram, & R. Bayliss. (1984). "Automated Testing: Past, Present and Future". *Journal of Occupational Psychology, 57*, 221-237.
D. Bartram, & P. Smith. (1984). "Everyday memory for everyday places". In J. Harris and P. Morris (Eds.) *Everyday memory, actions and absentmindedness.* London: Academic Press.
M. Morgan, D. Bartram, & A.D.B. Clarke. (1984). "The control of training and transfer effects in the Minnesota Spatial Relations and other tasks". *British Journal of Developmental Psychology*, *2*, 113-122.
D. Bartram, N. Banerji, D. Rothwell and P. Smith. (1985). "Task parameters affecting individual differences in pursuit and compensatory tracking performance". *Ergonomics, 28*, 1633- 1652.
D. Bartram, & H.C.A. Dale. (1985). "The prediction of success in helicopter pilot training". In M. Sorsa, (Ed.) *Report of the XVI Conference of the Western European Association for Aviation Psychology,* pp. 92-101. Helsinki, Finland.

D. Bartram, M. Poots, S. Hayes & R. Bayliss. (1985)".The effects of playing video games on performance on tests of tracking ability". In Brown, I.D., Goldsmith, R., Coombes, K. & Sinclair, M.A. (Eds.) *Ergonomics International 85.*, pp. 484-486. London: Taylor & Francis.

H.C.A. Dale, & D. Bartram, (1985). "Personality and the selection of military pilots". In M. Sorsa, (Ed.) *Report of the XVI Conference of the Western European Association for Aviation Psychology,* pp. 81-89. Helsinki, Finland.

H.C.A. Dale, & D. Bartram, (1985). "Micropat". *Proceedings of the 26th Annual Military Testing Association Conference,* Munich, November 1984, pp. 705-708.

H.C.A. Dale, & D. Bartram, (1985). "Micropat: A progress report on the validation of an automated test battery for pilot selection". In Jensen, R.S. & Adrion, J. (Eds.) *Proceedings of the 3rd Symposium in Aviation Psychology*, Columbus, Ohio, pp.533-540.

D.I. Williams, A.J.E. Crawshaw, C.M. Crawshaw & D. Bartram, (1985) "Heating: People in Control". In *Proceedings of the conference on Controlling Internal Environments*, Budapest, September 1985.

D. Bartram, & O. Feggou. (1986). "An evaluation of four different keyboard designs for foreign-destination coding desks". In D.J. Oborne (Ed.), *Contemporary Ergonomics 1986*, London: Taylor & Francis. pp335-339.

J.M. Dubery and D. Bartram, (1986) "The CLEARWAY ring as a tool in a University environment". In *Proceedings of a colloquium on Computer networks in education*, Institution of Electrical Engineers: London. 4pages.

D. Bartram, (1986). "A comparison of three different methods of presenting post-code information to coding-desk operators". In D.J. Oborne (Ed.) *Contemporary Ergonomics 1986*, pp325-329. London: Taylor & Francis.

D. Bartram, (1986) "Validation of the Micropat battery of pilot aptitude tests". In Newstead S.E., Irvine S.H. and Dann P.L. (Eds), *Human Assessment: Cognition and Motivation.* Martinus Nijhoff: Dordrecht. p.371.

D. Bartram, (1987). "Fitness for Purpose", *London Passenger Transport Research Group Journal.*

D. Bartram, (1987) The development of an automated pilot testing system for pilot selection: the MICROPAT project. *Applied Psychology: an international review*, *36*, 279-298.

D. Bartram, J.G. Beaumont, T. Cornford, P.L. Dann, and S. L. Wilson. (1987) "Summary of recommendations for the design of software for computer based assessment". *Bulletin of the British Psychological Society, 40*, 86-87.

D. Bartram, (1988) The design and evaluation of a computer-based training system for foreign destination coding desk operators. In J. Patrick and K.D. Duncan (Eds). *Training, Human Decision Making and Control.* Elsevier Science Publishers. pp.13-41.

D. Bartram, (1988) Development of a computer based training system for Post Office sorting desk operators. In E.D. Megaw (Ed).*Contemporary Ergonomics 1988*, Taylor & Francis: London. pp.401-406.

P.A. Lindley and D. Bartram, (1989). Evaluation of the psychological impact of RESTART courses on participants. *Proceedings of the Fourth meeting of the International Society for the Study of Individual Differences*: Heidelberg. p62.

D. Bartram, (1989). Computer-based assessment. In Herriot, P (Ed.) *Handbook of Assessment in Organizations*, Wiley: London.

D. Bartram, (1989). A reply to Eysenck's "Why not the 16PF". *PSL Newsletter,* November.

D. Bartram, (1990). An appraisal of the case for the adaptive assessment of knowledge and understanding in the delivery of competence-based qualifications. In H. Black and A. Wolf (Eds). *Knowledge and Competence: current issues in training and education.* HMSO, Careers and Occupational Information Centre.

D. Bartram, (1990) Chapters 1, 2 and 3 in J.R. Beech and L. Harding (Eds) *Testing People: A Guide to Psychometrics*. NFER-NELSON: Windsor. pp 1-86.

D. Bartram, and H.C.A. Dale (1991). Validation of the MICROPAT battery of pilot aptitude tests. In P.L. Dann, S.H. Irvine and J.M. Collis (Eds). *Advances in Computer-Based Human Assessment*. Dordrecht: Kluwer Academic Publishers. pp149-170.

D. Bartram, (1991). *Competence in Occupational Testing: General Information Pack.* Leicester: British Psychological Society. [Author of final draft produced for the BPS Implementation Group].

D. Bartram, (1991). Addressing the abuse of psychological tests. *Personnel Management*, April, pp34-39.

D. Bartram, (1991) Personality Assessment: its true value in management selection. *ASE News* Autumn, 1991. p2.

D. Bartram, (1992). A framework for evaluating analysis and assessment methods in competence-based approaches to human resource development. *Competence and Assessment, 18*, 15-19.

D. Bartram, (1992) The Personality of UK Managers: 16PF norms for short-listed applicants. *Journal of Occupational and Organizational Psychology, 65*, 159-172.

M. Paechter, & D. Bartram, (1993). Cross-validation in small samples. In R. Steyer, K.F. Wender & K.F. Widaman (Eds) *Psychometric Methodology: Proceedings of the 7th European Meeting of the Psychometric Society in Trier*. Stuttgart & New York: G F Verlag. pp 382-387.

D. Bartram, (1993) The validity and utility of personality assessment in occupational psychology. *European Review of Applied Psychology, 43,* 183-186.

D. Bartram, (1993) Validation of the ICES personality inventory. *European Review of Applied Psychology,43,*207-218.

D. Bartram, (1993). Verification and monitoring procedures for the British Psychological Society Certificate of Competence in Psychological Testing. *Selection and Development Review, 9,* 1-3.

D. Bartram, (1993) Development of the Level-B standards in Occupational Testing. *The Psychologist, 6*, p561.

D. Bartram, & F Kinnaird (1993) Assuring Quality. *The Psychologist, 6*, 561-562.
D. Bartram, (1993). "Emerging trends in computer-assisted assessment". In Schuler, H, Farr, J L and Smith M (eds) *Personnel Selection and Assessment: Individual and Organizational Perspectives.*, Chapter 17, pp267-288. Lawrence Erlbaum Associates: Edison, New Jersey.
D. Bartram, (1993) "Aptitude testing and selection in aviation. In Telfer R (Ed) *Aviation Training and Instruction* Chapter 3. Gower Ashgate Publishing: Aldershot, Hampshire. pp34-51.
D. Bartram, J Foster, P A Lindley, A Brown and S Nixon. (1993) The Learning Climates Questionnaire: A User's Guide. In *A Toolkit for the Learning Organization* London: Institute of Training and Development. pp 19
D. Bartram, J Foster, P A Lindley, A Brown and S Nixon. (1993) The Learning Climates Questionnaire: Background and Technical information. In *A Toolkit for the Learning Organization* London: Institute of Training and Development. pp 24
D. Bartram, (1994) What is so important about reliability? The need to consider the Standard Error of Measurement. *Selection and Development Review, 10,* 1-3.
D. Bartram, (1994). Computer Based Assessment. *International Review of Industrial and Organizational Psychology, 9*, 31-69.
D. Bartram, (1994) Desperate measures. Letter to the Editor, *The Psychologist,* August.
D. Bartram, (1994) The use of psychological tests for selection. *Recruitment International*, Number 14, p 36.
D. Bartram, & P. Baxter. (1995) Cathay Pacific Airways Pilot Selection Validation. In N Johnston, R Fuller & N McDonald (Eds) *Aviation Psychology: Training and Selection.* pp 194-199. Aldershot: Avebury Aviation, Ashgate.
D. Bartram, (1995). PART ONE: Chapters 1 through 4. In D. Bartram, Anderson, N., Kellett, D., Lindley, P.A. and Robertson, I. (1995). *Review of Personality Assessment Instruments (Level B) for use in occupational settings.* Leicester: BPS Books. pp 3-40.
J. Foster, G. Hodgkinson, P.A. Lindley & D. Bartram, (1995). The 15 Factor Questionnaire. Ibid.
L. Brindle, B. Tyler & K. Miller, P. Lindley & D. Bartram, (1995). The 16 Personality Factor Questionnaire. Ibid.
R. Drakeley, C. Fletcher, N. Anderson & D. Bartram, (1995). The Gough Adjective Checklist. Ibid.
J. James, N. Stanton, D. Kellett & D. Bartram, (1995). The Adult Personality Inventory. Ibid.
D. Adams, P. Kline, D. Kellett & D. Bartram, (1995). The California Psychological Inventory. Ibid.
R. Childs, S. Fletcher, D. Kellett & D. Bartram, (1995). Elements of Awareness - Behaviour. Ibid.
S.E. Poppleton, J.M. Smith, I. Robertson & D. Bartram, (1995). The Eysenck Personality Scales. Ibid.
R. Childs, S. Fletcher, D. Kellett & D. Bartram, (1995). The Firo B. Ibid.
B. Alimo-Metcalfe, V. Shackleton, N. Anderson & D. Bartram, (1995). The Global Gordon Personal Profile Inventory. Ibid.
R. Bayne, A. Brown, P.A. Lindley & D. Bartram, (1995). The Jung Type Indicator. Ibid.
C. Egan-Strang, N. King, N. Anderson & D. Bartram, (1995). The Kirton Adaptor-Innovator Inventory. Ibid.
N. Banerji, N. Stanton, D. Kellett & D. Bartram, (1995). The Managerial and Professional Profiler. Ibid.
M. Killcross, M. Poots, P.A. Lindley & D. Bartram, (1995). The Myers Briggs Type Indicator. Ibid.
R. Sage, R. Wood, P.A. Lindley & D. Bartram, (1995). The McQuaig Word Survey. Ibid.
D. Sharpley J. Rust, P.A. Lindley & D. Bartram, (1995). The Minnesota Multiphasic Personality Inventory. Ibid.
P. Barrett, J. James, D. Kellett & D. Bartram, (1995). The Manchester Personality Questionnaire. Ibid.
S.E. Craig, J.M. Smith, I. Robertson & D. Bartram, (1995). The NEO Personality Inventory - Revised. Ibid.
P. Kline, K. Johns, D. Kellett & D. Bartram, (1995). The Occupational Personality Profile. Ibid.
R. Drakeley, A. Hallmark, I. Robertson & D. Bartram, (1995). The Occupational Personality Questionnaire - CM5.2. Ibid.
V.J. Shackleton, C. Fletcher, I. Robertson & D. Bartram, (1995). The Occupational Personality Questionnaire - FS5.2. Ibid.
P. Jackson, L. Marshall, N. Anderson & D. Bartram, (1995). The Morrisby Objective Personality Tests. Ibid.
L. Marshall, V. Shackleton, I. Robertson & D. Bartram, (1995). The Occupational Relationships Profile. Ibid.
D. McLeod, R. Bayne, I. Robertson, & D. Bartram, (1995). The Occupational Type Profile. Ibid.
E. Ferguson, S. Newell, N. Anderson & D. Bartram, (1995). The Jackson Personality Research Form. Ibid.
L. Marshall, R. Wood, N. Anderson & D. Bartram,. (1995). The Profile Rep on Individual Style and Motivation. Ibid.
P. Barrett, J. James, D. Kellett & D. Bartram, (1995). The Rapid Personality Questionnaire. Ibid.
P. Jackson, S. Newell, N. Anderson & D. Bartram, (1995). The Gordon Survey of Interpersonal Values. Ibid.
P. Jackson, S. Newell, N. Anderson & D. Bartram, (1995). The Gordon Survey of Personal Values. Ibid.
E. Ferguson, C. Jackson, P. A. Lindley & D. Bartram, (1995). The Spielberger State Trait Anxiety Inventory. Ibid.
S. Cartwright, P. Iles, I. Robertson & D. Bartram, (1995). The Team Management Index. Ibid.
D. Bartram, (1995). The role of computer-based test interpretation (CBTI) in occupational assessment. *International Journal of Selection and Assessment, 3*, 178-185.
D. Bartram, (1995). The predictive validity of the EPI and 16PF for military flying training. *British Journal of Occupational and Organizational Psychology, 68*, 219-236.

D. Bartram, (1995). Predicting adverse impact in selection testing. *International Journal of Selection and Assessment,3,* 52-61.
D. Bartram, Validation of the Micropat Battery. *International Journal of Selection and Assessment , 3,* 84-95.
D. Bartram, Guest Editorial: Computer-based testing. *International Journal of Selection and Assessment , 3,* 73-74.
D. Bartram, Lindley, P.A., Foster, J. and Marshall, L. The selection of young people by small businesses. *British Journal of Occupational and Organizational Psychology, 68,* 339-358.
D. Bartram, (1995). The development of occupational standards in applied psychology. *The Psychologist, 8,* 33-37.
D. Bartram, (1995). The development of standards for the use of psychological tests: The competence approach. *The Psychologist, 8,* 219-223.
D. Bartram, (1995). When is a significant correlation significant and when is it not? *The Occupational Psychologist, 27,*12-115.
D. Bartram, (1995). Encouraging good practice in the use of psychological tests. *Network*, Spring, 22-24.
D. Bartram, (1995) Reducing the risk. *Training Management,* 23 March p9.
D. Bartram, (1995). The use of psychological tests for selection. *Recruitment International, 14,* p36.
D. Bartram, (1995) Quality assurance in test use: Implementing competence based standards. *Newsletter of the International Test Commission, 5,* 2-4.
D. Bartram, (1995). National Vocational Qualifications and Professional Development. *The Occupational Psychologist, 26,* 25-29..
D. Bartram, (1995). National Vocational Qualifications and Professional Development. *Clinical Psychology Forum, 81,* 38-40..
D. Bartram, (1995). National Vocational Qualifications and Professional Development. *Counselling Psychology Review, 10,* 21-24.
D. Bartram, (1995). Don't do it yourself. *Network*, Autumn, p23.
D. Bartram, (1995). International Guidelines on Core Standards for Test Use. *Newsletter of the International Test Commission,* 5, 6-8.
D. Bartram, (1995). The long long road to achieving accreditation. *Occupational Standards for Applied Psychology, 1,* 2-3 (Supplement to *The Psychologist,* ISSN 0952-8229).
D. Bartram, N. Ashley, & S.J. Wright. (1995). Test instructions: The importance of getting them right. *Selection and Development Review, 11,* 1-3. ISSN 0963-2638.
D. Bartram, (1996) The relationship between ipsatized and normative measures of personality. *British Journal of Occupational and Organizational Psychology, 69*, 25-39.
D. Bartram, (1996). Test qualifications and test use in the UK: The competence approach. *European Journal of Psychological Assessment, 12,* 62-71. ISSN 1015-5759.
D. Bartram, (1996). Occupational standards and competence-based qualifications for professional applied psychologists in the UK. *European Psychologist, 3,* 157-165. ISSN 1016 9040.
D. Bartram, (1996). Psychometric test: The good, the bad and the ugly. *Science & Public Affairs,* Spring. 46-49. ISSN 0268 490 X.
D. Bartram, (1996). National Vocational Qualifications and Professional Development. *Forensic Update, 44,* 4-7. ISSN 1356-5028.
D. Bartram, (1996). The development of standards for the use of psychological tests in occupational settings: The competence approach. *Hong Kong Institute of Human Resource Management Forum,* February 1996. 14-16.
D. Bartram, (1996). NVQ's/SVQs: Sorting out the facts from the fiction. *Occupational Standards for Applied Psychology, 2,* 2-3. (Supplement to *The Psychologist,* ISSN 0952-8229).
D. Bartram, (1996). Report on the first phase of the project; and Next steps in the Development Process. *Occupational Standards for Applied Psychology 3,* 1-4. (supplement to *The Psychologist,* ISSN 0952-8229).
D. Bartram, (1996). Entry on *Psychological Testing.* In ENCARTA [CD-ROM encyclopedia].
D. Bartram, & P. Baxter (1996). Validation of the Cathay Pacific Airway pilot selection programme. *International Journal of Aviation Psychology, 6(2),* 149-169. ISSN 1050-8414.
D. Bartram, (1997). Project on the development of international guidelines on core standards for test use. *The International Test Commission Newsletter, 6(2),* 8-9
D. Bartram, (1997). Distance assessment: Psychological assessment through the internet. *Selection and Development Review, 13 (3),* 15-19 ISSN 0963-2638.

### *Technical reports:*

D. Bartram, & H.C.A. Dale (1981). "The EPI as a selection test for AAC pilots". *Ministry of Defence Technical report, ERG/Y6536/81/8/R* , 35pp.
D. Bartram, H.C.A. Dale, R. Bayliss & D. Farnsworth (1981). "An examination of the present selection procedure for A.A.C. pilots". *ERG/Y6536/81/4R* (Confidential report to APRE, Ministry of Defence). 73pp.

H.C.A. Dale & D. Bartram, (1981) "Final report: MoD Research Agreement 2119/022 SAG(A)". (Confidential report to APRE, Ministry of Defence). 20pp.

D. Bartram, (1982). "The development of a fully automated psychometric testing system: MICROPAT". *Ministry of Defence Technical Report, ERG/Y/6536/82/7*, 16pp.

D. Bartram, H.C.A. Dale, & P. Smith. (1982). "Leconfield Trials of the Micropat system". *Ministry of Defence Technical report, ERG/Y6536/82/5* , 61pp.

D. Bartram, (1983). "A pilot investigation of criterion measures and psychometric assessment instruments". *Report for the BBC Training Centre, Wood Norton.* 18pp.

D. Bartram, N. Banerji, D. Rothwell, & P. Smith. (1983). "Individual differences in performance on tracking tasks". *Ministry of Defence, Supplementary Technical report, ERG/Y6536/83/5*, pp33.

D. Bartram, & H.C.A. Dale. (1983). "Micropat Version 3: A description of the fully automated personnel selection testing system being developed for the Army Air Corps". *Ministry of Defence Technical report, ERG/Y6536/83/7*, 14pp.

D. Bartram, H.C.A. Dale, & R. Bayliss. (1983). "Report on the concurrent validity of the Micropat test battery". *Ministry of Defence Technical report, ERG/Y6536/83/10*, 32pp.

D. Bartram, (1984). "A comparison of three different methods of presenting post-code information to coding-desk operators". *Post Office Technical Report, (A/PO.PAP/6/84)*. 18pp.

D. Bartram, & R. Bayliss. (1984). "Automated testing: past, present and future". *Ministry of Defence Supplementary Technical report, ERG/Y6536/84/1*, 31pp.

D. Bartram, & H.C.A. Dale (1984). "Validation of the MICROPAT battery of pilot aptitude tests". *Ministry of Defence Interim Technical report, ERG/Y6536/84/12*, 20pp.

D. Bartram, H.C.A. Dale, & R. Bayliss. (1984). "Analysis of personality inventory data". (Confidential report to *APRE, Ministry of Defence*) 47pp.

D. Bartram, (1985). "The Development of a general occupational assessment system". Confidential Report to *NFER-NELSON Ltd*. 22pp.

D. Bartram, & H.C.A. Dale. (1985). "The validity of Micropat and personality measures for the prediction of success in helicopter pilot training". *Ministry of Defence APRE Technical Report,* 44pp.

D. Bartram, & H.C.A. Dale. (1985). An analysis of personality test inventory data administered to AAC applicants". *APRE Technical Report,* 57pp.

D. Bartram, & O. Feggou. (1985). "A pilot investigation of the sequential keyboard". *Post Office Technical Report, ERG/Z6545/85/4a.* 8pp.

D. Bartram, & M. Lee. (1985). "A pilot investigation of the ten-key chord keyboard". *Post Office Technical Report, ERG/Z6545/85/4b.* 11pp.

D. Bartram, & O. Feggou. (1985). "An evaluation of different keyboard designs for foreign destination coding desks". *Post Office Technical Report, ERG/Z6545/85/4c.* 29pp.

D. Bartram, C.M. Crawshaw, D.I. Williams, A.J.E. Crawshaw & P.A. Lindley. (1985). "The use of heating controls". *Ergonomics Research Group Report, ERG/Y6543/85/1.* 84pp.

D. Bartram, H.C.A. Dale, K.G.G. Corkindale & D. Dennison (1985). "The validity of Micropat tests for Army helicopter pilots". *Army Personnel Research Establishment Report 85R005*.

D. Bartram, (1986). "Requirements for an integrated transport information system". In M. Roberts *The adequacy of British Rail's London Services*, Report 385/7580k/ST, Greater London Council, February 1986, pp. 93-106.

D. Bartram, (1986) "Development and evaluation of Micropat version 4.0". *Ministry of Defence, SP(N) Report TR 184.* 146 pages.

D. Bartram, (1986) "Evaluation of training data obtained using a computer-based chord-keyboard training system". *Technical Report ERG/Z6554/86/11,* Post Office Research Department: Swindon.

D. Bartram, P.A. Lindley and C. M. Crawshaw (1986). "A comparison of four different conditions for using the Rapid A101 foot stapler". *Report ERG/MSC/86/8 prepared for the Manpower Services Commission* pp 12.

D. Bartram, J.G. Beaumont, T. Cornford, P.L. Dann, and S. L. Wilson. (1986) "Recommendations for the design of software for computer based assessment". Report prepared for the *Scientific Affairs Board of the British Psychological Society*.

D. Bartram, P.A. Lindley and J.M. Foster (1987) Evaluation of the impact of RESTART courses for participants: Behavioural, cognitive and affective changes occurring over the period of the RESTART course. *Report to the Manpower Services Commission.* NPAL: Hull.

D. Bartram, (1987) Analysis of application form data for the Burton Group Plc Graduate Corporate Training Scheme. I: Analysis of the self-report section. *Report to the Recruitment and Assessment Manager, Burton Group Plc.* 43pp.

D. Bartram, (1987) Analysis of application form data for the Burton Group Plc Graduate Corporate Training Scheme. II: Use of the application form data to predict the selection decision. *Report to the Recruitment and Assessment Manager, Burton Group Plc.* 71pp.

D. Bartram, and M. Choi. (1988) Evaluation of three computer-based tests of navigational ability. *Ministry of Defence SP(N) Report TR209*. 78pp

D. Bartram, (1988) Validation of MICROPAT Version 4. I: The prediction of RN Observer and Pilot grading outcome. *Ministry of Defence SP(N) Report TR210.* 88pp

D. Bartram, P.A. Lindley and J.M. Foster. (1988). Evaluation of the impact of RESTART courses for participants: Psychometric analyses of the NPAL RESTART Assessment Battery. *Technical report to the Employment Service: Sheffield.* pp 195.

P.A. Lindley, D. Bartram, J.M. Foster and D.J. Lee. (1988). Evaluation of the impact of RESTART courses for participants: A report prepared for the Claimant Services Branch. *Final executive report, Employment Service: Sheffield.* pp 40.

D. Bartram, (1989) "An appraisal of the case for the adaptive assessment of knowledge and understanding in the delivery of competence-based qualifications". *The Training Agency: Sheffield*. pp 20.

D. Bartram, and P.A. Lindley. (1989). "Coding Desk Key Sequences: Extension of National Short Codes". *Post Office Research: Swindon*. 5pp.

D. Bartram, (1989). Training and qualification in the use of psychological tests in Occupational Psychology. *Report prepared for the BPS Steering Committee on Test Standards.* pp10.

D. Bartram, (1989). Checklist of competencies in occupational testing: Level A. *Document prepared for the BPS Steering Committee on Test Standards.* pp. 17.

P.A. Lindley, D. Bartram, and J.M. Foster. (1989). Final report on the project to produce a guide to the use of psychological testing for Training Agents in ET and Managing Agents in YTS. *Report to the Training Agency*, NPAL: Hull.

C. Tuton, P.A. Lindley and D. Bartram, (1989). An analysis of factors underlying the productivity of two Training Agents. *Report to the Training Agency*. NPAL: Hull. pp23.

D. Bartram, and L. Marshall (1989). Validation of Micropat Version 4: II. The prediction of RN Observer and RN Pilot flying training outcomes. *Ministry of Defence SP(N) Report TR 233.*

D. Bartram, and J.M. Foster (1989). A comparison of 2-digit and 3-digit key sequences for flat-pack and packet sorting. *Post Office Research: Swindon.* pp18.

D. Bartram, (1989). The Cathay Pacific Airways cadet pilot training scheme candidate selection procedure. *Report to Cathay Pacific Airways: Hong Kong.* 14pp.

D. Bartram, and J.M. Foster (1990). An analysis of training data from nineteen SWDO flat-pack sorter operators. *Post Office Research: Swindon*. pp18.

M. Paechter and D. Bartram, (1990). Validity and cross-validation in small samples. *Ministry of Defence SP(N) Report TR 240* pp53.

D. Bartram, and T. Faite (1990). The equivalence of the IBM-PC and Sirton versions of the MICROPAT tracking tasks. *Ministry of Defence SP(N) Report TR 241* pp15.

D. Bartram, and L. Marshall (1990). Validation of Micropat Version 4: III. The prediction of Operational Flying Training Outcome for RN Observers and RN Pilots. *Ministry of Defence SP(N) Report TR 248.*

D. Bartram, (1990). Cathay Pacific Airways First Officer and Cadet Pilot Selection procedures. *Report to Cathay Pacific Airways: Hong Kong.* 26pp.

P.A. Lindley and D. Bartram, (1990). Review of recruitment and selection procedures at Aycliffe Centre for Children. Phase One: Job analysis and recommendations for improvements to the selection procedure. *NPAL report to Aycliffe Centre for Children.* 21pp.

D. Bartram, and J.M. Foster (1990). An analysis of training data from 28 Coventry Packet Sorter operators. *Post Office Research: Swindon*. pp28.

D. Bartram, and L. Marshall (1990). Risk, Plane and Landing. The RNAS Culdrose Evaluation of Micropat Version 4.1. *Ministry of Defence SP(N) Report TR 251.* pp 50.

D. Bartram, and L. Marshall (1990). Risk, Plane and Landing. The RNAS Culdrose Evaluation of Micropat Version 4.1. Supplement: Reliability Analysis. *Ministry of Defence SP(N) Report TR 251 (supplement).* pp 18.

D. Bartram, and L. Marshall (1990). Navor, Navcalc and Manikin. Psychometric evaluation of Micropat Versions 4.1. and 5.0 *Ministry of Defence SP(N) Report TR 259.* pp 30.

D. Bartram, and P.A. Lindley (1990). A framework for the professional development of psychologists within the Employment Department. *Report prepared for the Employment Department Head of Profession (psychologist class).* pp21.

D. Bartram, P.A. Lindley and C. Tuton. (1990) Assessment models underlying the delivery of Initial Assessment. *Report to the Employment Department* NPAL: Hull. pp 28.

C.Tuton, D. Bartram, and P.A. Lindley (1991). An Evaluation of the use of APTICOM in Initial Assessment for ET. *Report to the Employment Department* NPAL: Hull. pp 9 [excluding appendices].

C.Tuton, P.A. Lindley and D. Bartram, (1991). An Evaluation of Programmes concerned with increasing Client Confidence and Motivation in Initial Assessment for ET. *Report to the Employment Department* NPAL: Hull. pp 17 [excluding appendices].

C.Tuton, P.A. Lindley and D. Bartram, (1991). Developing Standards for Initial Assessment. *Report to the Employment Department* NPAL: Hull. pp 12.

C.Tuton, D. Bartram, and P.A. Lindley (1991). The development of a Fitness for Purpose framework for assessing the quality of Initial Assessment provision. . *Report to the Employment Department* NPAL: Hull. pp 5.

D. Bartram, (1991). The relative merits of chord keyboards and numeric keypads for the OCR/Video foreign mail coding desk. *Post Office Research, Swindon: Technical Report.* pp 14.

D. Bartram, and L. Marshall (1991). The Officer of the Watch Course trainees: psychometric evaluation of the Micropat test battery. *Ministry of Defence (Navy) SP(N) Report TR 268.* pp 45.

D. Bartram, (1991). *Competence in Occupational Testing: General Information Pack.* Leicester: British Psychological Society. pp 26.

D. Bartram, (1991) *SCREENTEST 16PF and GMA: UK Management norms and data supplement*. Report to MSL international. pp 28 [excluding appendices].

D. Bartram, and L. Marshall (1991). The effects of departures from normality in the distribution of raw scores on the reliability and validity of Micropat test scores. *Ministry of Defence SP(N) Report TR 275.* pp 35. (c)

D. Bartram, and L. Marshall (1991). Risk, Plane and Landing. The RNAS Culdrose Evaluation of Micropat Version 4.1. II: The prediction of Observer Operational Flying Training outcome. *Ministry of Defence SP(N) Report TR 278.* pp 34. (c)

D. Bartram, and L. Marshall (1991). The criterion-related validity of the RAF Executive (Phase One) and the Micropat tests for Royal Navy samples. *Ministry of Defence SP(N) Report TR 279.* pp 14. (c)

D. Bartram, and L. Marshall (1991). The predictive validity of NAVOR and NAVCALC: Preliminary findings. *Ministry of Defence SP(N) Report TR 281.* pp 12. (c)

D. Bartram, (1991) The Development and Validation of MICROPAT for Royal Navy Aircrew Selection: 1985-1991. *Ministry of Defence SP(N) Report R 163.* pp 18. (c)

D. Bartram, (1991) Evaluation criteria for the technical rigor and practicality of Sufficiency of Evidence and UKU derivation procedures within the assessment of competence-based standards. *Employment Department KAG Paper #34*. 18pp.

D. Bartram, (1991) Knowledge Elicitation Programmes: A Programme Evaluator's Guide. *Employment Department KAG Paper #44*. 29pp.

D. Bartram, and J Foster (1991). Notes on the domain of issues and priorities relating to the use of psychological tests. *Report to the Employment Department's Standing Committee on Psychological Testing*. pp 11.

D. Bartram, (1992). Range: some thoughts on the distinction between Necessity and Sufficiency of Evidence. *Employment Department Knowledge Advisory Group Paper no. 55*. 6pp

D. Bartram, (1992). Sufficiency of Evidence Workshop outcomes: Retailing, Level 3 NVQ. *Employment Department*. 12pp

D. Bartram, (1992). Sufficiency of Evidence Workshop outcomes: Retailing, Level 4 NVQ. *Employment Department*. 10pp

D. Bartram, (1992). Sufficiency of Evidence Workshop outcomes: Institute of Maintenance and Building Management, Level 3 NVQ. *Employment Department*. 16 pp

D. Bartram, (1992). Sufficiency of Evidence Workshop outcomes: Management of Residential Estate Agency, Level 4 NVQ. *Employment Department*. 19pp

D. Bartram, (1992). Defining Sufficiency of Evidence for the Assessment of NVQs: Final report of the Sufficiency of Evidence Project. *Employment Department*. pp 58

D. Bartram, (1992). Defining Sufficiency of Evidence for the Assessment of NVQs: Workshop materials and procedures. *Employment Department.* pp 47.

D. Bartram, and H Black (1992) The Assessment of Knowledge in Standards. *Employment Department Knowledge Advisory Group Paper #62.* pp20.

D. Bartram, C. M. Crawshaw, T. Plant, & D. F. Sewell, (1992) The effectiveness of computer-based training *Report to British Aerospace (Military Aircraft), Training Systems Department*. pp 73.

D. Bartram, P A Lindley and J Foster (1992). The selection of young people by medium-sized and large organizations. *Report to the Employment Department Careers Service Branch.* (40 pages excluding appendices).

D. Bartram, and P A Lindley. (1992) An enhanced Assessment and Guidance System for Adults in the Humberside TEC Area. *University of Hull DIRECT report to Humberside Training and Enterprise Council.* (62 pages excluding Annexes)

D. Bartram, and P A Lindley (1992). Proposals for a Quality Assurance system for the provision of assessment and guidance services in the Gloucester TEC area. *Report to Gloucester Training and Enterprise Council.* 20pp

D. Bartram, and L Marshall. (1992) *Cathay Pacific Airways Cadet Pilot Selection: The validity of the MICROPAT tests*. Report to Cathay Pacific Airways pp 16.

D. Bartram, and L. Mitchell (1992) The Design Development and Delivery of National Vocational Qualifications and Scottish Vocational qualifications: The place of knowledge and understanding. *Joint Employment Department, NCVQ and SCOTVEC report.* pp 103.

D. Bartram, and C. Tuton (1992). The development of Assessment and Guidance: Value for money exemplars. *Employment Department, IASU2* 12pp [excluding appendices]

C. Tuton, D. Bartram, and P.A. Lindley (1992). Expanded Benchmarks for the Provision of Assessment and Guidance. *Employment Department, IASU2*.

C. Tuton, P.A. Lindley and D. Bartram, (1992). The Development of Summary and Expanded Benchmarks for the provision of Assessment and Guidance. *Information and Advice Services Unit, Employment Department Group.* pp24.

C. Tuton, P.A. Lindley and D. Bartram, (1992). Benchmarks for Assessment and Guidance. *Report to TEC Media*. pp13.

C.Tuton, D. Bartram, and P.A. Lindley (1992). Summary Benchmarks for the Provision of Assessment and Guidance. *Employment Department, IASU2.* 25pp

C. Tuton, D. Bartram, and P.A. Lindley (1992). The Development of Summary and Expanded Benchmarks for the provision of Assessment and Guidance within the Gateways to Learning Initiative. *Employment Department, IASU2*. 146 pp.

D. Bartram, & L. Mitchell. (1993) *The Design Development and Delivery of National Vocational Qualifications and Scottish Vocational qualifications - The place of Knowledge and Understanding.* Moorfoot, Sheffield: The Employment Department. pp 183.

D. Bartram, P.A. Lindley, L. Marshall & J. Foster. (1993) *The recruitment and selection of young people by small businesses*. Moorfoot, Sheffield: Careers Service Branch Report. pp 38 excluding annexes.

P.A. Lindley & D. Bartram, (1993) *Recruitment and selection of young people: Implications of the research findings for guidance practitioners*. Moorfoot, Sheffield: Careers Service Branch Report. pp 10.

D. Bartram, P.A. Lindley, J. Foster and L. Marshall (1993). *Evaluating the new Employment Service Recruitment and Retention System.* Interim Report to Employment Service Personnel Branch: Sheffield.

D. Bartram, P.A. Lindley, J Foster & L. Marshall. (1993) *Evaluating the new Employment Service recruitment and retention system - Final executive report.* Sheffield: Personnel Branch Report. pp 59.

D. Bartram, P.A. Lindley, J Foster & L. Marshall. (1993) *Evaluating the new Employment Service recruitment and retention system - Technical report.* Sheffield: Personnel Branch Report. pp 40 excluding annexes.

D. Bartram, (1994). *Cathay Pacific Airways Pilot Selection Validation Executive Report*. Hong Kong: Cathay Pacific Airways. 11pp.

D. Bartram, & P. Baxter (1994). *Validation of the Cathay Pacific Airways Pilot Selection Programme.* Hong Kong: Cathay Pacific Airways.35pp.

D. Bartram, (1994). *National Vocational Qualifications in Applied Psychology: A Briefing Note.* Leicester: Consultative Working Group for Applied Psychology

D. Bartram, & P.A. Lindley. (1994) *Abilities underlying GNVQ competencies: An analysis of BTEC GNVQ Standards specifications.* Report to the Further Education Unit.

D. Bartram, (1994). *National Vocational Qualifications in Applied Psychology: A Briefing Note.* Leicester: Consultative Working Group for Applied Psychology. 17 pages.

D. Bartram, (1995). *External Audit Report on the Civil Service Administrative Fast Stream Selection Procedures*. London: Cabinet Office, OPSS. 59pp.[Released 3.30 14/12/95 in response to arranged parliamentary question].

D. Bartram, L. Marshall & I.J Coyne. (1995) *Guidelines for the Design of Human Computer Interactions (HCIs) for Stand-alone Instruction Systems.* Farnborough, Hants: Defence Research Agency. 108pp.

D. Bartram, (1995). *Report on the use of psychometric testing in the restructuring of the Treasury and Benefits division of Brent Council in September, 1992*. Brent: Brent Council Core Legal Service. 30pp

D. Bartram, (1995). *Report on the use of psychometric testing in the restructuring of the Treasury and Benefits division of Brent Council in September, 1992: Technical Annex*. Brent: Brent Council Core Legal Service. 27pp.

D. Bartram, (1995). *Relationships between PREVUE and sales performance: The ADP study.* Waco, Texas: Profiles International Inc. 31pp.

P. Smith., M. Coffey., C. Fernandes, & D. Bartram, (1996). *Review of the Investment Management Certificate Examination: A Report for the Institute of Investment Management and Research*. Slough: National Foundation for Educational Research.

C. Johnson, A. Wolf & D. Bartram, (1995). *External Assessment for NVQs*. London: National Council for Vocational Qualifications. 68pp

D. Bartram, (1997). An evaluation of studies on the efficacy of candidate feedback following the Home Office Extended Interview. A report prepared for the Home Office Assessment consultancy Unit. 10pp.

### *Published software:*

*NB All the MICROPAT software is Crown Copyright and is published under a Ministry of Defence license.*

D. Bartram, (1981) *MICROPAT Test management and administration system: Version 3.0.* Hull: Bartdale.

D. Bartram, & R. Ward. (1983). *Computer simulations for different designs of foreign-destinations letter-sorting desks.* Post Office Research: Swindon.
D. Bartram, (1984) *MICROPAT Test management and administration system: Version 4.0.* Hull: Bartdale.
D. Bartram, (1984). *SCHEDULE: a test of scheduling ability.* MICROPAT (4.0), Hull: Bartdale.
D. Bartram, (1984). *RISK: a test assessing attitude to risk.* MICROPAT (4.0), Hull: Bartdale.
D. Bartram, (1984). *LANDING: a multi-dimensional time-sharing and motor-coordination test.* MICROPAT (4.0), Hull: Bartdale.
D. Bartram, (1984). *DUALTASK: a test assessing resource allocation under high levels of task demand.* MICROPAT (4.0), Hull: Bartdale.
D. Bartram, (1984). *TRACKING: a package of adaptive and non- adaptive pursuit and compensatory tracking tasks.* MICROPAT (4.0), Hull: Bartdale.
D. Bartram, (1986). *NAVOR and NAVCALC: a suite of tests of spatial orientation and navigation ability.* MICROPAT (4.1), Hull: Bartdale.
D. Bartram, & R.J. Phillips (1987). *The Vocational Preference Inventory for the BBC Microcomputer.* NFER-NELSON Publishing Co.: Windsor.
D. Bartram, & R.J. Phillips (1987). *The Lewis Counselling Inventory for the BBC Microcomputer.* NFER-NELSON Publishing Co.: Windsor.
D. Bartram, & R.J. Phillips (1987). *Data Analysis System for the BBC microcomputer.* NFER-NELSON Publishing Co.: Windsor.
D. Bartram, (1987). *Computer-based training system for the AEG-Telefunken flat-pack sorter.* Post Office Research : Swindon.
D. Bartram, (1988). *SCREENTEST Module 0010/1: MICROPAT (5.0) Schedule.* Bartdale: Hull.
D. Bartram, (1988). *SCREENTEST Module 0010/2: MICROPAT (5.0) Risk.* Bartdale: Hull.
D. Bartram, (1988). *SCREENTEST Module 0010/3: MICROPAT (5.0) Landing.* Bartdale: Hull.
D. Bartram, (1988). *SCREENTEST Module 0010/4: MICROPAT (5.0) Dualtask.* Bartdale: Hull.
D. Bartram, (1988). *SCREENTEST Module 0010/5: MICROPAT (5.0) Plane.* Bartdale: Hull.
D. Bartram, (1988). *SCREENTEST Module 0010/6: MICROPAT (5.0) Adtrack.* Bartdale: Hull.
D. Bartram, (1988). *SCREENTEST Module 0010/7: MICROPAT (5.0) Comptrak.* Bartdale: Hull.
D. Bartram, (1988). *SCREENTEST Module 0010/8: MICROPAT (5.0) Navigate.* Bartdale: Hull.
D. Bartram, (1988). *SCREENTEST Module 0010/9: MICROPAT (5.0) Profiler.* Bartdale: Hull.
D. Bartram, and R. Phillips (1989). *SCREENTEST Assessment and Data Manager.* NFER-NELSON: Windsor.
D. Bartram, (1989). *SCREENTEST Module 1203/1: Sixteen Personality Factor Questionnaire forms A and B.* NFER-NELSON: Windsor.
D. Bartram, (1989). *SCREENTEST Module 1203/2: Sixteen Personality Factor Questionnaire forms C and D.* NFER-NELSON: Windsor.
D. Bartram, (1989). *SCREENTEST Module 1203/3: Sixteen Personality Factor Questionnaire Profiler.* NFER-NELSON: Windsor.
D. Bartram, (1989). *SCREENTEST Module 4700/1: Graduate and Managerial Assessment - Numerical forms A and B.* NFER-NELSON: Windsor.
D. Bartram, (1989). *SCREENTEST Module 4700/2: Graduate and Managerial Assessment - Verbal forms A and B.* NFER-NELSON: Windsor.
D. Bartram, (1989). *SCREENTEST Module 4700/3: Graduate and Managerial Assessment - Abstract forms A and B.* NFER-NELSON: Windsor.
D. Bartram, (1989). *SCREENTEST Module 4700/4: Graduate and Managerial Assessment Profiler.* NFER-NELSON: Windsor.
D. Bartram, (1989). *Flat-pack and packet sorter training system (Version 3.0).* Post Office Research : Swindon.
D. Bartram, and E. Lau. (1991) .*SCREENTEST: Test interpretation authoring and narrative report generation modules. Version 1.1.* NFER-NELSON: Windsor.

### *IBM Versions of Micropat as SCREENTEST modules:*

D. Bartram, (1991). *SCHEDULE: a test of scheduling ability.* MICROPAT (5.1), Hull: Bartdale.
D. Bartram, (1991). *RISK: a test assessing attitude to risk.* MICROPAT (5.1), Hull: Bartdale.
D. Bartram, (1991). *LANDING: a multi-dimensional time-sharing and motor-coordination test.* MICROPAT (5.1), Hull: Bartdale.
D. Bartram, (1991). *DUALTASK: a test assessing resource allocation under high levels of task demand.* MICROPAT (5.1), Hull: Bartdale.
D. Bartram, (1991). *ADTRACK: a suite of adaptive pursuit tracking tasks.* MICROPAT (5.1), Hull: Bartdale.
D. Bartram, (1991). *COMPTRAK: a suite of compensatory tracking tasks.* MICROPAT (5.1), Hull: Bartdale.

D. Bartram, (1991). *NAVIGATE: a suite of tests of spatial orientation and navigation ability.* MICROPAT (5.1), Hull: Bartdale.
D. Bartram, (1991). *RESULTS PROFILER AND NORMATIVE DATABASE.* MICROPAT (5.1), Hull: Bartdale.
D. Bartram, and Y.Y. Lau (1991). *SCREENTEST Module Definer* NFER-NELSON: Windsor.
D. Bartram, A. Kitching, & C. Vausler. (1995) *The MICROPAT Pilot Aptitude Tests: Version 6.0.* London: Psyctech Ltd. [Single screen Windows 3.1 adaptations of Micropat Version 5.1, designed to operate within a network environment running under Psytech's TEKS operating system.]
D. Bartram, A. Kitching, & C. Vausler. (1995) *The MICROPAT Results Profiler: Version 6.0.* London: Psyctech Ltd. [Windows 3.1. adaptation of the SCREENTEST Micropat results profiler, designed to operate within a network environment running under Psytech's TEKS operating system.]
D. Bartram, A. Kitching (1996) *NAVOR-P: Version 6.0.* London: Psyctech Ltd. [Windows 3.1 adaptation of one of the Micropat navigation tests, designed to operate within a network environment running under Psytech's TEKS operating system.]

### *Tests and assessment instruments:*

D. Bartram, (1993) *The ICES Plus Assessment battery.* Barbados: Prevue Assessment Systems Inc.
D. Bartram, J Foster, P A Lindley, A Brown and S Nixon. (1993) The Learning Climates Questionnaire. In *A Toolkit for the Learning Organization* London: Institute of Training and Development.

### *Some papers presented at conferences which have appeared in formal publications of proceedings as short papers or abstracts:*

D. Bartram, & H.C.A. Dale (1982) "Personality and the selection of military pilots in military aviation. " *20th International Congress of Applied Psychology Abstracts*, p.176.
D. Bartram, & H.C.A. Dale (1982). "Micropat: a computerized selection test battery for pilots". *20th International Congress of Applied Psychology Abstracts*, p339.
D. Bartram, (1983) "Micropat: the design of a fully automated personnel selection testing system". BPS Occupational Psychology Conference, Symposium on Automated Testing (convened by the author), Warwick .*Bulletin of the British Psychological Society, 36*, A76.
D. Bartram, N. Banerji, D. Rothwell and P. Smith. (1984) "Individual differences in performance on tracking tasks". Invited paper for the 4*th Annual Symposium on Automated Psychological Testing,* London.
D. Bartram, (1984). "The development of tests of information management for the Micropat system". *Innovations in Individual Differences.* Invitational seminar, Plymouth Polytechnic.
D. Bartram, (1985). "The automation of psychological testing procedures: towards some guidelines for management and operation". *Keynote paper to BPS Conference on the Management and Operation of Computer-Based Testing Procedures,* London.
D. Bartram, M. Poots, S. Hayes & R. Bayliss. (1985) "The effects of playing video games on test of tracking ability". Invited paper for the *VIth Annual Symposium on Automated Psychological Testing*, London.
D. Bartram, (1986). "Computer-based assessment: developing rules for modeling expert judgments of performance on complex tasks". Invited paper for the *Symposium on Innovations in the Application of Microcomputers to Psychological Assessment and Therapy,* London.
D. Bartram, (1987) "Future directions in computer-based assessment". British Psychological Society's Occupational Psychology Conference, Hull. *Bulletin of the British Psychological Society, 40,* A27.
D. Bartram, (1988). "Development of computerized tests of aircrew navigational ability". British Psychological Society's Occupational Psychology Conference, Manchester. *The British Psychological Society 1988 Abstracts,* p.20.
D. Bartram, (1989). "Computer-based assessment". Fourth European Congress on the Psychology of Work and Organizations: Cambridge. *The British Psychological Society 1989 Abstracts,* p.88.
P.A. Lindley, D. Bartram, and J.M. Foster. (1989) "RESTART courses: the impact of intervention on the long-term unemployed". A symposium of four papers presented at the BPS Occupational Psychology Conference: Bowness-on-Windermere. *The British Psychological Society 1989 Abstracts,* p.60.
D. Bartram, (1989) "Emerging trends in computer-assisted assessment". Paper presented at the Hohenheim conference on *Individual and Organizational Aspects of Selection*: Hohenheim University, Stuttgart. pp 21.
D McLeod and D. Bartram,. (1991). Training and qualification in test use. BPS Occupational Conference. Cardiff. *The British Psychological Society Abstracts*, p 44.
D. Bartram, (1991). Validation of the Micropat aircrew selection battery. *NATO Panel VIII RSG 15 Workshop on Computer-based Assessment of Military Personnel,* Brussels NATO HQ, 26-28 November.
M. Paechter and D. Bartram, (1991). Cross-validation in small samples. To be published in *Proceedings of the 7th European Meeting of the Psychometrics Conference.*

D. Bartram, (1992) Development and validation of the Micropat Test Battery *British Psychological Society Abstracts,* 39
D. Bartram, P.A. Lindley & J. Foster. (1993) The selection of young people by medium-sized and large organizations. *Proceedings of the British Psychological Society, 1*, p 10.
R. Kurz, B. Lodh, & D. Bartram, (1993). Practice and performance - the effects of systematic test preparation. *Proceedings of the British Psychological Society, 1,* p 13.
D. Bartram, & Baxter, P. (1995) Cathay Pacific Airways Pilot Selection Validation. *Proceedings of the 21st WEAAP Conference* (6 pages).
D. Bartram, (1995). Personality factors in pilot selection: Validation of the Cathay Pacific Airways selection procedures. In R. Jensen & L.A. Rakovan (Eds.). *Proceedings of the Eighth International Symposium on Aviation Psychology, Columbus, Ohio,* Volume 2, 1330-1335.
D. Bartram, (1995). New developments in computer-based testing: The future for pilot selection procedures. In R. Jensen & L.A. Rakovan (Eds.). *Proceedings of the Eighth International Symposium on Aviation Psychology, Columbus, Ohio,* Volume 2, 1324-1329.
D. Bartram, (1995). Factors to consider in the evaluation of computer-based test interpretation (CBTI) systems. *Proceedings of the BPS Occupational Psychology Conference,* 157-162. Leicester: BPS.
S.M. Weaver, S. Lui, K. Taylor, I. Poulter, S. Hems, J. Robinson, S.R. Killick, and D. Bartram, (1995). Attitudes and motives of semen donors and non-donors. *Journal of Psychosomatic Obstetrics and Gynaecology, 16*, 154. [Abstract of paper presented at 11th International Congress of Psychosomatic Obstetrics and Gynaecology, May 21-24, Basel.
D. Bartram, Dixon, A., & Kalogera, S. (1996). The Learning Climate Questionnaire: Shared perceptions within work groups. *Proceedings of the BPS Occupational Psychology Conference,* 241-246. Leicester: BPS.
D. Bartram, (1996) Issues related to computer-based test interpretation. *Proceedings of the First Test User's Conference*, 2-7. Leicester: British Psychological Society.
D. Bartram, (1996). International Standards for Test Use. *Proceedings of a one-day conference on ethics and Good practice in Assessment and Psychological Testing.* Cheltenham: J. Cooper Associates.
D. Bartram, (1997). The Selection Validity Index (SVI): Measuring predictive validity without correlations. *Proceedings of the BPS Occupational Psychology Conference,* 65-70. Leicester: BPS. ISBN 1 85433 2074.
D. Bartram, P. Clough, & J. Williams, (1997). The relationship between personality, perceived risk and risk taking. *Proceedings of the BPS Occupational Psychology Conference,* 59-64. Leicester: BPS. ISBN 1 85433 207 4.
D. Bartram, (1997). Distance assessment: Psychological assessment through the internet. *Proceedings of the BPS Occupational Psychology Conference,* 197-202. Leicester: BPS ISBN 1 85433 207 4.

### *Other recent conference presentation (1994-):*

D. Bartram, Lindley, P.A., Marshall, L, & Foster, J. (1994). The recruitment and selection of young people by small businesses. *BPS Occupational Psychology Conference,* Birmingham, January, 1994.
D. Bartram, (1994). Development of Level B standards and qualifications: an update. *BPS Occupational Psychology Conference,* Birmingham, January, 1994.
R. Kurz, M. Del MarTorrijos Sanchez, & D. Bartram, (1994). Computer-based assessment: equivalence or superiority? *BPS Occupational Psychology Conference, Birmingham,* January, 1994.
D. Bartram, & P. Baxter. (1994). Cathay Pacific Airways Pilot Selection Validation. *21st WEAAP Conference,* Dublin March, 1994.
D. Bartram, (1994). The role of CBTI in occupational assessment. Invited keynote paper to *Advances in Selection and Assessment*, Nottingham, July.
D. Bartram, P.A. Lindley, L. Marshall, & J. Foster. (1994). The recruitment and selection of young people by small businesses. *23rd International Congress of Applied Psychology,* Madrid, July.
D. Bartram, (1994). The development of standards for the use of psychological tests: the competence approach. *23rd International Congress of Applied Psychology*, Madrid, July.
D. Bartram, (1994) Acquiring competence and assuring quality in the use of psychological tests. Invited paper to *IPD Conference*, Harrogate, October 94.
D. Bartram, D (1995). Implementing the Level B standards: the final steps. *BPS Occupational Psychology Conference,* Warwick, January 95.
D. Bartram, (1994). Assessing performance on complex continuous performance tasks. *Second International Conference on Social Science Information Technology,* Amsterdam, December 94.
Bartram, D (1995). Factors to consider in the evaluation of computer-based test interpretation (CBTI) systems. *BPS Occupational Psychology Conference,* Warwick, January 95.
Bartram D. (1995) Personality factors in pilot selection: Validation of the Cathay Pacific Selection Procedures. *Eighth International Symposium on Aviation Psychology,* Columbus, Ohio, April 95.

Bartram D. (1995) New developments in computer-based testing: The future for pilot selection procedures. *Eighth International Symposium on Aviation Psychology,* Columbus, Ohio, April 95.
Bartram D. and Lindley P.A. Development of the Learning Climate Questionnaire. *Paper presented at the IVth European Congress of Psychology, Athens, July 1995*
Lindley P.A. and. Bartram, D. Selecting care staff for children's homes *Paper presented at the IVth European Congress of Psychology, Athens, July 1995*
Kurz, R., Lodh, B., & D. Bartram, Test Orientation Practice. *Paper presented at the IVth European Congress of Psychology, Athens, July 1995*
D. Bartram, Dixon, A., & Kalogera, S. (1996). The Learning Climate Questionnaire: Shared perceptions within work groups. *BPS Occupational Psychology Conference,* Bournemouth, January 96.
D. Bartram, (1996) *Knowledge and Understanding: Specification and Assessment issues for higher levels vocational qualifications.* Developing S/NCVQ, July 5th.
D. Bartram, (1996) *Assuring quality in the use of psychological tests - the user-competence approach.* XXVI International Congress of Psychology, Montreal.
D. Bartram, (1996) *The International Test Commission project on international core test standards for test use.* XXVI International Congress of Psychology, Montreal.
D. Bartram, (1996) *Expert systems and their use in occupational assessment.* Manchester: UMIST, 3 June.
D. Bartram, (1996) *Higher Level NVQs and Continuing Vocational Education.* Paper presented at Brighton University, June 6th.
D. Bartram, (1997) *Variations in patterns of testing and test use in Europe.* Fifth European congress of Psychology, Dublin, July.
D. Bartram, (1997) *The contribution of new technology to testing and assessment.* ITC Symposium, Fifth European Congress of Psychology, Dublin, July.

### *Invited Keynote Addresses:*

D. Bartram, (1996) psychometric Testing: Its relevance to Business. *International Assessment Conference,* Newcastle, June 12th.
D. Bartram, (1996) International standards for test Use. *Conference on Ethics and good Practice in Assessment and Psychological Testing.* Cheltenham, July 12th.
D. Bartram, (1996) The Involvement of Higher education in the delivery of S/NVQs. *NCVQ Workshop on NVQs and Higher Education,* Coventry, November 13th.
D. Bartram, (1996) The Involvement of Higher Education in the delivery of S/NVQs. *NCVQ Workshop on NVQs and Higher Education,* Nottingham, November 27th.
D. Bartram, (1996) The Involvement of Higher education in the delivery of S/NVQs. *NCVQ Workshop on NVQs and Higher Education,* London, December 10th.
D. Bartram, (1996) The Involvement of Higher Education in the delivery of S/NVQs. *NCVQ* Workshop *on NVQs and Higer Education,* Teessede, December 4th.
D. Bartram, (1996) The Involvement of Higher Education in the delivery of S/NVQs. *NCVQ Workshop on NVQs and Higher Education,* Surrey, December 12th.
D. Bartram, (1996) The Involvement of Higher education in the delivery of S/NVQs. *SCOTVEC Workshop on SVQs and Higher education,* Glasgow December 9th.

## *POSTGRADUATE SUPERVISION*

### *Doctoral theses*

N. Sprent (1982). "The intelligibility of bus timetables". *[Jointly supervised with C.M.Crawshaw]*
P.A. Lindley (1983). "Analytic and constructive processes in the comprehension of text".
P. Smith (1985). "The cognitive representation of the large-scale environment".
O. Feggou (1988). "Theoretical and practical considerations in the design of special-purpose keyboards and operator training principles".
I. Panagopoulos (1992). "Self-report measures of risk-taking".

### *Research Projects for the MSc in Industrial/Occupational Psychology*

R. J. Bayliss (1980). "A comparison between the flying grading trial and the elementary handling test and prediction of success in flying training".
D. Farnsworth (1980). "Assessment of candidates in RAF flying selection training".
I. N. Flynn (1981). "A preliminary study in the development of an adaptive tracking task as a selection test for helicopter pilots".

N. Karaminas (1981). "Measurement of individual differences in the performance of a highly complex problem solving task".
N. Khan (1981). "Developing instructions for a complex cognitive task".
N. Banerji (1982) "A study to investigate intra-individual differences in performance on tracking tasks used in pilot selection".
M. Power (1982). "An investigation of faking behaviour on the Eysenck Personality Inventory in an automated format".
M. Warma (1982) "A study of the nature of spatial ability and its importance in the prediction of flying training success for helicopter pilots".
O. Feggou (1983). "A pilot examination of criteria and psychometric assessment data for trainee technical operators in the BBC".
S. F. Roberts (1983). "A report on the development and trial testing of an auditory selective attention test".
K. M. Walsh (1983). "An analysis of Army Air Corps fortnightly Flying Wing assessments".
M. Poots (1984). "A study to investigate the effects of playing video games on performance of tests on tracking ability". *[With Distinction]*.
M. P. L. Lee (1984). "The practicability of a 10-key chord keyboard".
B. Gresswell (1984). "An investigation into the usability of a commercial computer system - a pilot study".
L. Marsden (1985). "Reliability and development of the Woodwork Assessment Test at Hull Employment Rehabilitation Centre".
K.C. Ong (1985). An exploratory study of personality correlates of performance on tasks requiring sustained attention: a personnel selection perspective. *[With Distinction. Awarded the Ergonomics Society's Ulf Arberg prize for the best postgraduate research project in ergonomics]*.
E. Gleave (1985). "An investigation into the use of a Local Area Network in the Psychology Department of the University of Hull".
M. Green (1985). "The effects of test sophistication on the validity of general aptitude test scores during personnel selection".
J. Hasenauer (1985). "Adaptive training in chord keyboard operation: a comparative study".
*I. Panagopoulos (1986). "Computerized automated testing for pilot selection: Micropat, a case study of software reliability, aspects of construct validity and attitudes towards automated versus paper and pencil testing methods".*
B. Murphy (1986). "A study to investigate the use of signal detection theory and measures of risk-taking behaviour in the selection of helicopter pilots".
P. Jacobs (1987). "The evaluation of a computer-based training programme for the AEG-Telefunken flat-pack sorter".
M. Pineda (1987). "A survey of graduate recruitment"
S-S. Teoh (1987). "Sense of direction in navigation for the selection of helicopter pilots using automated testing".
M. Choi (1987). "Trials of three new tests of navigational ability" *[With Distinction]*.
D. Alexandratos (1988). "Item analysis of the Annett Hand Preference Questionnaire and its relationships with a spatial orientation and a numerical facility test".
J. Wilkinson (1988). "Preselection to Debenhams" Retail Management Graduate Training Scheme".
J. Dunkley (1988). "A study to validate the Topshop application form for the Retail Management Training Scheme".
C. Hill (1988). "Development of NPAL's Work Preference Questionnaire".
C. Cox (1988). "An investigation of recruitment and selection procedures at Mowlem Management Limited"
L. Marshall (1988). "The impact of RESTART on unemployed people".
D. Brough (1988). "A study to evaluate the psychometric equivalence of two modes of administering the GMA Verbal and Numerical scales".
R.K. Morgan and D. Curran (1989) "The development of a tailor-made package for the selection of computer personnel for the Sun Alliance Insurance Group".
T. Faite (1989) "Examining the equivalence of the newly developed IBM-PC version of the Micropat tracking task with the older Sirton versions".
C. White (1989) "Helping Training Agents to help themselves: a "Which" approach to psychological tests".
M. Paechter (1989) "Validity and cross-validation in small samples". *[With Distinction]*
O. Alkhader (1990) "The prediction of Grade Point Average for Kuwait University Commerce College students".
C. Clapham (1990). "Candidates" attitudes towards computer-based selection tests".
D. Goldman (1990). "A company profit-share scheme: How do employees perceive, understand and respond to it?"
R. Kurtz (1990). "Test-item theory, facet form concept and the construction of parallel items".
R. McCormick (1990). "A cross-cultural comparison of a schema for measuring managerial strengths and weaknesses".
V. Senior (1990). "Are judgements rating scale dependent?"
E. Webb (1990). "The cross-cultural validation of an American career development questionnaire".
C. Lovelock (1991). "The construction of a model of assessors" ratings of assessment centre candidates using Q methodology".

C. Tuton (1991). "The development of a “Fitness for Purpose” approach for evaluating the provision of Initial Assessment for Employment Training”.
D. Fleck (1991). "The reliability of two versions of the NAVCALC test”.
M. del M. Torrijos-Sanchez (1991). "Evaluation of the automated form of the MGIB tests”.
D. Peden (1991). "Validity of SHL's Technical Test Battery as used by Hull City Council for Apprentice Selection”.
M.M.Armstrong (1992). "Ethnic minorities and practice effects on aptitude test performance: a Zimbabwean sample”.
F. Hoghughi, F. (1992) "The impact of vicarious realistic job previews on self-selection”. *[With Distinction]*
B. Lodh, (1992) "The effects of practice on psychological test performance”.
N.J. Rowland (1992) "The use of shadowing as a realistic job preview and a selection method”.
P.T. Weldon (1992) "Prediction of Army driver training outcome using MICROPAT and the ABC tests: reliability and validity”.
O. Amoh (1993) "The re-design of a behaviourally anchored rating scale for the shadowing scheme used at Aycliffe”.
A.J. Brennan (1993). "The Barnum effect in Personality Assessment".
J. Mackaill (1993). "Evaluation of a competence based application form sifting procedure”.
J.G. Clayton (1994). "The relationship of personality factors and job performance in residential child care.
S. Cumiskey (1995) The impact of organizational changes on stress levels within an ambulance service.
T.C. Delany (1995). Development of the revised format of Prevue Assessment.
A.E. Dixon. (1995). Perceptions of Learning Climate within a NHS Trust.
L.J.M. Healy, (1995). The selection of managers in East Yorkshire Community Healthcare Trust. *[Distinction]*
S. Kalogera, (1995). Application of the Learning Climate Questionnaire in a public service organization.
P. Sinclair, (1996) A survey of the importance of personality at work.
C. Longley, (1996) Construct validation of the *Prevue Assessment* General Ability Scale.
J. Casie-Swift, (1996) Development of the Social Services recruitment and selection procedure for day-centre officers.
S. Lambie, (1996) The criterion problem as it relates to the validation of the Cathay Pacific Pilot Selection Programme.
N. Ashley, (1997) An evaluation of the effectiveness of post assessment centre feedback.