"Big Five" Personality Factors Research Using Christal's Self Description Inventory

Janet M. Collis
Human Assessment Laboratory
University of Plymouth
UK

Jerry M. Barucky
Metrica, Inc.
USA

19990608 125

Introduction

This paper recounts the latest efforts to evaluate the properties of the Air Force Self Description Inventory (AFSDI) when it is administered, using both computer-assisted and paper-and-pencil formats, to a variety of subject samples in the United States and the United Kingdom. In all cases the resultant factor structures and distributions of scores were consistent and indicate that the AFSDI is capable of measuring the five personality factors across a variety of subject cultures and administrative methods. In addition, some initial comparisons of subgroup composite scores based on gender differences and among officer and officer trainee groups have been made. Some significant effects have been found across the different subject types. Tests of the instrument's stability have been carried out using both short (one day) and long-term (13-26 months) intervals. Test-retest correlations after a one-day interval showed a high degree of response consistency. A larger study of longer-term stability, with intervals ranging between 13 to 26 months, also showed moderately high correlations. Finally, additional steps in this research line will be discussed. These initiatives include determining the relationships between personality scores and both withdrawal from officer training programs and job performance ratings.

The Air Force Self Description Inventory

Christal (1993) developed a computer-based personality inventory to measure the 'Big Five' using both adjectives, or traits, and behavioural statements. After rigorous testing and extensive analysis, the fmal inventory contained 99 behavioural statements and 64 trait words. Traits and statements were delivered by computer in random order (but blocked for item type). Subjects were required to move a mouse to indicate their response to each item (see Figure 1). Christal demonstrated a robust five-factor solution time and again with this inventory. He demonstrated internal consistency of the inventory with split-half correlations of between .89 and .95, and his development work included the exploration of subfactor measurement. A TTCP agreement and collaboration with UK psychologists resulted in additional data capture from alliterative subject groups and the establishment of a paper-and-pencil version of the AFSDI.
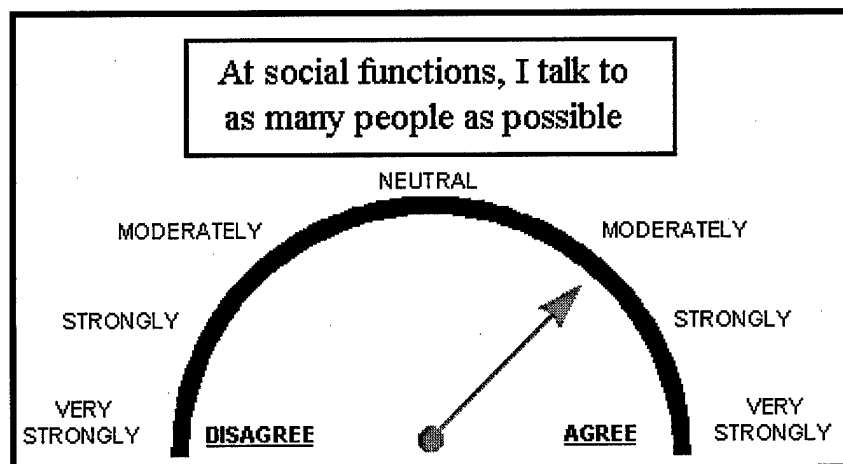

Figure 1. Christal's response scale used in the computer-delivered form of the T-SD

Despite the untimely death of Dr. Raymond Christal in April 1995, work on the AFSDI has continued both under the auspices of the U.S. Air Force's Armstrong Laboratory and the United Kingdom's Defence Research Agency.

1

AQI 99-09-1589

The first study reported here examined an AFSDI data set acquired from USAF Officer samples, since all previous US research had been undertaken using Enlisted samples, and there was a requirement to confirm that factor structure and data distributions were equivalent for different subject groups.

Officer and Enlisted comparisons

In a study of 573 Officer trainees who were given the computer-administered AFSDI, verification of the five factor solution was achieved. In addition, composite scores derived from Christal's original weightings of items loading .40 or above were correlated with factor scores produced from the new Officer data sample. These are shown in Table 1 below. This procedure has previously been employed by Christal as a form of cross-validation to identify the magnitude of concordance between two data samples. It should be noted that because of the .40 loading limitation, not all items are employed in the calculation of composite scores, but factor scores for each of the five factors are calculated using all the loadings. Since factor scores are not readily available for future test takers, composite scores would be more useful, if they can be shown to capture response differentiation to the same extent as factor scores. The table below shows that most of the variance is accounted for if composite scores are employed. Although all significant correlations are in bold type, the most meaningful comparisons are the diagonal factor/composite correlations. If all composites are forced into a multiple regression equation to predict factor scores, multiple R values are all at .98. Highly similar patterns of correlations between composite scores shown in separate data samples from Officer trainees and Enlisted subjects provided further evidence for the construct validity of the AFSDI. Bold figures denote significance at p<.05.

Table 1. AFSDI factor scores and composite scores correlations for USAF Officers .

	Neuroticism factor score	Conscientiousness factor score	Extroversion factor score	Agreeableness factor score	Openness factor score
N Composite	**.96**	**-.13**	**-.13**	**-.16**	-.03
C Composite	**-.10**	**.96**	**.09**	**.17**	.06
E Composite	**-.16**	.07	**.96**	**.17**	-.01
A Composite	**-.14**	**.17**	**.17**	**.95**	.06
O Composite	-.02	.02	.01	**.09**	**.97**

Distributions of Officer data for both composite and factor scores were normal and very similar to those obtained previously for Enlisted samples.

Paper-and-pencil comparisons

Although the computer-delivered AFSDI is a very efficient system for delivery and scoring of the inventory, there will almost always be a requirement for a paper-and-pencil version for use in contexts where computers are unavailable. Research in the UK using Christal's inventory had already begun with the construction of a paper-and-pencil version, using a 7 point scale for behavioural statements and 9 point scale for traits to represent the major divisions of the computer-delivered arch scale. Collis (1995) reported that a five-factor solution was produced using UK Officer subjects and the paper-and-pencil inventory (known in the UK as the Trait-Self-Description Inventory, or T-SD). Correlations between factor scores for the UK Officer sample and composite scores for the same sample, using factor loadings derived from Christal's USAF data were between .89 and .97, suggesting that measurement of the five factors was consistent across US/UK cultures, across delivery modes (paper-and-pencil vs. computer) and across different groups of personnel (Enlisted vs. Officer).

Comparison of the paper-and-pencil and computer-delivered AFSDI was subsequently examined with a group of 440 USAF subjects. This group was comprised of relatively junior officers in attendance at Squadron Officers School and officer candidates enrolled in Air Force Reserve Officer Training Corps (AFROTC). Once again, factor scores derived from factor analysis of this data sample were correlated with composite scores derived using Christal's earlier factor loadings. Table 2 shows the strength of the relationship between the two sets of scores. If all composites are forced into a multiple regression equation to predict factor scores, multiple R values range from .97 to .98. Inspection of correlation matrices for composite scores for the paper-and-pencil group against the computer-delivered group are also very similar.

Table 2. Correlations of AFSDI factor scores with composite scores for paper-and-pencil delivery (N=440)

	N factor score	C factor score	E factor score	A factor score	O factor score
N Composite	**.94**	**-.16**	**-.18**	**-.19**	.01
C Composite	**-.16**	**.92**	**.12**	**.24**	**.12**
E Composite	**-.23**	.06	**.94**	**.14**	-.01
A Composite	**-.13**	**.20**	**.16**	**.94**	.08
O Composite	.04	**.12**	.04	**.12**	**.97**

(Bold figures denote significance at *p<.05)*

Since scores from the computer-delivered inventory and the paper-and-pencil version have completely different scales, comparison of score distributions was made by converting both sets of scores to T-scores. The distributions of both composite T-scores and factor scores were very similar across the two groups of subjects.

Test-retest reliability

Shute and Gluck (1996) first reported 24 hour test-retest reliabilities for the AFSDI which ranged from .88 to .94. To determine the stability of the measure across longer time periods, Enlisted personnel who had first been given the AFSDI during basic training were located and given the inventory a second time during their initial duty assignments at 24 active-duty bases. The bases were selected to obtain a cross-section of Air Force specialties that closely reflected the distribution of specialties in the original sample. The data distributions were normal and very similar across the two testing sessions, and variance differences between the data samples were significant for the Agreeableness measure only. Significant differences in scores were obtained for all measures except Openness, and retest scores tended to be less positive (lower Agreeableness, Conscientiousness and higher Neuroticism).

Inspection of scatter diagrams indicated that the magnitude of a number of individual test-retest differences might reflect some careless responding. It was also noted that some individuals were showing large differences for more than one of the five measures. Temporary removal of data for reanalysis was undertaken on the basis of two criteria. One criterion was based on visual inspection of scatter diagrams and identification and removal of a small number of cases that lay well away from the main group cluster. The other was detection of substantial retest differences appearing across more than one measure, indicating subjects in whose responses little confidence could be placed. Given that for the Neuroticism measure alone, 79 subjects had a test-retest difference of 300 or more, the criteria for removal of cases was very conservative. The actual number of subjects removed was between 11 and 25, representing between 2 and 4% of the total subject sample. New correlations shown in Table 3 were found to range from .65 to .79 and all are significant at p<.001.

Table 3. Correlations of composite scores across all intervals (filtered sample with Ns shown)

Composite	r	N
Agreeableness	.65	564
Conscientiousness	.70	559
Extroversion	.69	558
Neuroticism	.79	557
Openness	.72	571

An additional analysis examined the variation across test-retest intervals and using original data, rather than data with outliers removed, subjects were grouped into three interval sets. Table 4 shows separate correlations for each retest interval. These figures show that the impressive Extroversion retest correlation in particular had been

previously masked by analysing data collapsed across all interval groups. All correlations are significant at $p<.001$.

Table 4. Composite test-retest correlations for three time intervals

Composite	13-16 months (N=84)	17-21 months (N=369)	22-25 months (N=126)
Agreeableness	.60	.64	.62
Conscientiousness	.66	.59	.55
Extroversion	.83	.71	.72
Neuroticism	.67	.67	.53
Openness	.68	.70	.63

Further breakdown of the first column of interval data (13-16 months) showed two measures (E and N) were still between .70 and .82 at a 16 month interval. Although some correlations had diminished at the longest interval (22-25 months), none had fallen below .50 and one remained above .70.

Group differences

The study with USAF Officers confirmed previous sex differences in AFSDI data first reported by Christal. Females reported themselves as significantly more agreeable, conscientious, and extroverted and significantly less neurotic and open to experience. Examination of data from different Officer groups showed the more experienced officer subjects (majors) at Air Comniand and Staff College (ACSC) subjects were clearly distinguished from the less experienced officers (captains and lieutenants) at Squadron Officer School and from the officer trainee groups at the USAF Academy and Officer Training School. ACSC subjects were significantly less agreeable, extroverted and open than the three remaining groups, who showed very similar means.

AFSDI and performance ratings

Job performance data were collected from supervisors of 71 airmen who had taken the AFSDI in basic training. Correlations between composite scores on the five AFSDI factors and performance ratings on 10 general dimensions indicated some fairly strong positive relationships between Agreeableness and ratings on all 10 dimensions. Conversely, (and as might be expected) moderate negative relationships were found between Neuroticism and five of the performance ratings that were linked more to interpersonal skills than technical ability. Moderate positive relationships were also found between Openness and four of the performance dimensions linked to interpersonal skills. Preliminary analyses of the relationships between the 22 subcomposite scores indicated much stronger relationships existing between some subcomposites than others. The results indicated that there are potential relationships between general aspects of military performance and certain personality factors. In addition, the performance rating forms employed appear to be able to capture some of the variation in subjects' active duty performance.

Table 5. Relationship of AFSDI composites and job performance dimension

Rating dimension	O	A	C	E	N
Technical Knowledge/Skill	.15	.29*	.18	.07	-.23
Initiative/Effort	.20	.39**	.11	.17	-.24*
Knowledge of and Adherence to Regulations/Orders	.26*	.35**	.07	.13	-.24*
Integrity	.27*	.36**	.20	.02	-.21
Leadership	.27*	.40**	.23	.13	-.34**
Military Appearance	.05	.27*	.14	.15	-.23
Self Development	.29*	.31**	.17	.20	-.33**
Self Control	.22	.26*	.06	-.13	-.17
Global 1: Technical Proficiency	.04	.33**	.16	.07	-.21
Global2: Interpersonal Proficiency	.35**	.47**	.20	.01	.24*

* p<.05 **p<.01

Conclusions

Since Christal's development work on the AFSDI, there have been a number of subsequent studies employing a different type of test delivery, different personnel groups and a different subject culture. Comparison of the results of all of these trials has produced more information regarding the manipulation of each variable (instrument, candidate, culture) as data become available. Table 6 shows the data available for scrutiny so far. Across all studies, a five-factor structure consistently emerges, data distributions are very similar, and correlations between factor scores and composite scores are almost always above .90.

Table 6. The current availability of AFSDI data: delivery modes and personnel tested

	Computer-delivery		Paper and pencil	
	Officers	Enlisted	Officers	Enlisted
UK	(1) Planned	(3) NO	(5) YES (Collis, 1995a)	(7) NO
USA	(2) YES (Collis, 1996)	(4) YES (Christal, 1993)	(6) YES (Collis, 1996)	(8) YES (Collis, 1995b)

The extent of this consistency across a wide range of testing sessions provides clear evidence for the reliability of the AFSDI. The next steps are to determine its validity. The preliminary analysis of relationships between performance ratings and personality in US airmen has been described. In addition to this, there is work ongoing in the UK which explores the relationship between personality factors measured by the AFSDI and likelihood of voluntary withdrawal from officer training. Christal previously extracted 22 subfactors or subcomposite scores after repeated factor analyses and suggested these could prove valuable in any studies where prediction of performance or behaviour was required.

References

Christal, R.E. (1993). R&D Summary report F33615-91-D-0010. Armstrong Laboratories, Brooks AFB.

Collis, J.M. (1995a). *The Trait-Self Description Inventory: Measurement of the 'Big Five' personality factors at the Admiralty Interview Board.* Defence Research Agency Report DRA/CHS/H53/CR95050/01.

Collis, J.M. *(1995b). The Trait-Self Description Inventory: A comparison of factor structure across delivery devices.* Defence Research Agency Report DRA/CHS/HS3/CR95067/01.

Collis, J.M. (1996). *The Air Force Self-Description Inventory (AFSDI): Comparison of select samples, test-retest reliability and comparison of delivery mechanisms.* Human Assessment Laboratory Technical Report, 1:1996: University of Plymouth, UK.

Shute, V. & Gluck, K. (1995). *How useful are personality and learning style measures as determinants of learning?* Symposium on Non-Ability Determinants of Learning; American Educational Research Association, San Francisco, CA.

Acknowledgement

Back to Table of Contents

INTERNET DOCUMENT INFORMATION FORM

A . Report Title: Big Five Personality Factors Research Using Christal's Self Description Inventory

B. DATE Report Downloaded From the Internet: 06/04/99

C. Report's Point of Contact: (Name, Organization, Address, Office Symbol, & Ph #): Navy Advancement Center
ATTN: Dr. Grover Diel (850) 452-1615
Pensacola, FL

D. Currently Applicable Classification Level: Unclassified

E. Distribution Statement A: Approved for Public Release

F. The foregoing information was compiled and provided by:
DTIC-OCA, Initials: __VM__ Preparation Date 06/04/99

The foregoing information should exactly correspond to the Title, Report Number, and the Date on the accompanying report document. If there are mismatches, or other questions, contact the above OCA Representative for resolution.