

# **Pairin Behavioral Testing**

## **Reliability and Validity**

**Pairin Executive Office 3461 Ringsby Court, Suite 335, Denver, CO 80216**

Copyright © 2014 Pairin, Inc.

**Warning:** No Part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission from the copyright holder.

**Pairin**, the **Pairin logo**, **Job Pairin**, and **Impact Pairin** are trademarks in the U.S. and/or other countries of Pairin, Inc., and its affiliate(s).

Printed in the United States of America



*For orders and inquiries:*

Pairin, Inc. 3461 Ringsby Court, Suite 335, Denver, CO 80216  
855-724-7466

# Table of Contents

<b>Acknowledgments</b> .....	4
<b>Introduction</b> .....	5
<b>Reliability</b> .....	6
Test-retest Reliability.....	6
<i>Built-in Checks</i> .....	6
<i>Reference Bias</i> .....	6
<i>Test-retest Data</i> .....	7
Inter-rater Reliability.....	9
Inter-method Reliability.....	9
Internal Consistency Reliability.....	9
<b>Validity</b> .....	10
Content Validity.....	10
<i>Normative Sample</i> .....	10
<i>Cross-cultural Validity</i> .....	10
<i>Test Structure</i> .....	11
Construct Validity.....	11
<i>Scales</i> .....	11
<i>Factors</i> .....	11
<i>Convergent and Discriminant Constructs</i> .....	12
Criterion Validity.....	15
<b>Application</b> .....	16
Behavioral Measurement Over Time.....	16
Matching Individuals to Employment.....	16
<b>EEOC Compliance</b> .....	18
<b>References</b> .....	19

# Acknowledgments

This document was prepared by the following members of the Content Development Group for Pairin Testing:

Ronald C. Young, PhD  
Chief Science Officer

Mariya Toneva, PhD Candidate  
Behavioral Science Researcher

# Introduction

Pairin's Behavioral Testing is based on Dr. Harrison Gough's Adjective Check List (ACL). The ACL is a norm-referenced standardized test consisting of 300 adjectives and adjectival phrases commonly used to describe a person's behavioral attributes. The ACL is unique in that the required number of items checked is unspecified so adjectives selected are ones that are salient for the person being evaluated. It may be administered for a multiplicity of purposes including, but not limited to, eliciting a self-evaluation from an individual.

This document provides research results and data related to the reliability and validity of the ACL. Drawn from a sample size of 9,382 individuals and normalized over fifteen years, the information in this document includes reliability and validity data based on the current normative sample, an explanation of the points of application, and a summary of findings related to EEOC compliance.

To request a copy of the ACL Technical Manual, or to offer suggestions regarding the ACL or about this or any other related publications, please contact Pairin Customer Relations at the following email address: [info@pairin.com](mailto:info@pairin.com).

# Reliability

To begin to investigate the accuracy of an instrument, one must first examine the reliability of its science. Without being reliable, a measurement can never be perfectly accurate as consecutive results may vary significantly.

The reliability of a test signifies the extent to which the assessment yields consistent results. This reliability is classically divided into four categories: test-retest, inter-rater, inter-method, and internal consistency (Trochim, 2006). To be reliable, a test must meet the standards of reliability in those categories out of the four that are relevant to the assessment.

## **Test-retest Reliability**

Test-retest reliability assesses the degree of consistency between test administrations. To evaluate this type of reliability, the measurements must be gathered from a single participant using the same methods or instruments under equivalent testing conditions (MASB, 2010).

### ***Built-in Checks***

To ensure that any inconsistencies in results between consecutive test administration is not due to false responses, the web-based system created by Pairin includes algorithms that evaluate the authenticity of each individual test score.

These algorithms detect responses that include too few (below 10%) or too many (above 90%) of the adjectives, random answers displaying no patterns (e.g. selecting adjectives similar to both "lazy" and "diligent" multiple times), and potentially too positive or too negative responses. Triggering any number of these checks decreases the authenticity of an individual score in 25% increments. In addition to utilizing these checks, the web-based system enables the removal of any scores from a test group that are less than 100% valid, 75% valid, 50% valid, 25% valid, and 0% valid to examine how these results impact the group scoring. This functionality, combined with the authenticity checks, provides unparalleled accuracy and flexibility of testing.

### ***Reference Bias***

In addition to the built-in authenticity checks, the ACL test format further reduces the possibility of inconsistent results upon retesting. While other non-cognitive skills assessments are susceptible to reference bias due to their scale question types (West et al. 2014), the ACL is much less likely to be influenced by reference bias because of its binary structure.

The reference bias alluded to by West and colleagues pertains to the influence of different reference points in self-reported measures. For instance, if a test asks two students to rate how

hard-working they are on a scale from 1 to 5, the students may imagine a stereotypical hard-working person and attempt to measure themselves against this exemplar. Even if the two students exemplify an objective hard-working score of 4, the student enrolled in a program that specifically emphasizes hard work may have higher expectations of the hard-working stereotype and thus rate herself lower than 4.

Avoiding this reference bias is especially important for any measurement that aims to accurately evaluate the growth over time of individuals involved in programs that target skill development. Because the ACL format does not include a scale, it is less susceptible to reference bias and thus can produce more accurate results upon retesting.

### ***Test-retest Data***

While the built-in authenticity checks and the low reference bias should theoretically ensure a high degree of test-retest reliability, the extent of this reliability can only be determined empirically.

To investigate the test-retest reliability of the ACL, a sample of 199 males and 45 females was tested. The test-retest data for males was gathered by administering the test twice in a six-month interval and correlating the corresponding pairs of individual scores. This data ranged from .34 for scale A-1 (high origence, low intellectence) to .77 for aggression, with a median of .65 (10 scales had retest correlations lower than .60). Test-retest correlation data for females was gathered in a one-year time interval and ranged from .45 for Femininity, A-1, and A-2, to .86 for Exhibition, with a median of .71 (nine scales had retest reliabilities below .60). The complete correlation coefficients are displayed in Table 1.

**TABLE 1**  
Reliability Coefficients for the Scales of the ACL

Scales	Alpha Coefficients		Test-Retest Coefficients		Scales	Alpha Coefficients		Test-Retest Coefficients	
	Males	Females	Males	Females		Males	Females	Males	Females
<b>No.Ckd</b>	---	---	0.64	0.85	<b>S-Cn</b>	0.65	0.71	0.72	0.76
<b>Fav</b>	0.95	0.94	0.62	0.6	<b>S-Cfd</b>	0.79	0.77	0.68	0.78
<b>Unfav</b>	0.92	0.91	0.65	0.76	<b>P-Adj</b>	0.63	0.66	0.65	0.55
<b>Com</b>	0.66	0.66	0.35	0.52	<b>Iss</b>	0.78	0.76	0.71	0.78
<b>Ach</b>	0.85	0.82	0.72	0.73	<b>Cps</b>	0.63	0.63	0.68	0.7

<b>Dom</b>	0.79	0.78	0.74	0.78	<b>Mls</b>	0.73	0.69	0.53	0.56
<b>End</b>	0.82	0.77	0.72	0.61	<b>Mas</b>	0.82	0.75	0.53	0.51
<b>Ord</b>	0.81	0.78	0.64	0.65	<b>Fem</b>	0.77	0.76	0.41	0.45
<b>Int</b>	0.79	0.77	0.62	0.59	<b>CP</b>	0.76	0.77	0.74	0.75
<b>Nur</b>	0.83	0.83	0.68	0.78	<b>NP</b>	0.81	0.77	0.71	0.73
<b>Aff</b>	0.89	0.87	0.6	0.66	<b>A</b>	0.79	0.77	0.71	0.71
<b>Het</b>	0.74	0.71	0.58	0.71	<b>FC</b>	0.72	0.76	0.7	0.82
<b>Exh</b>	0.74	0.75	0.72	0.86	<b>AC</b>	0.79	0.79	0.7	0.64
<b>Aut</b>	0.69	0.68	0.75	0.77	<b>A-1</b>	0.71	0.69	0.34	0.45
<b>Agg</b>	0.72	0.74	0.77	0.85	<b>A-2</b>	0.79	0.76	0.52	0.45
<b>Cha</b>	0.56	0.62	0.64	0.71	<b>A-3</b>	0.77	0.72	0.51	0.67
<b>Suc</b>	0.56	0.64	0.57	0.64	<b>A-4</b>	0.83	0.8	0.59	0.66
<b>Aba</b>	0.7	0.69	0.71	0.68					
<b>Def</b>	0.69	0.71	0.75	0.75					
<b>Crs</b>	0.75	0.53	0.7	0.59					

While reliability estimates are in the region commonly found for self-report inventories, several scales show variations in scores on retesting. Variations in scores from one sitting to another one may reflect changes in the individuals tested and/or an error in the assessment. According to the ACL manual, research indicates that reliability over time appears to be a meaningful psychological variable and not just a statistical property of the ACL. Respondents who have a more cheerful, outgoing, and active temperament tend to have less test-retest variance than those who are more subdued and phlegmatic.

Overall, the test-retest data in addition to the built-in authenticity checks and the low reference bias exhibited by the ACL, lead to a high degree of confidence in the test-retest reliability of the ACL.



## **Inter-rater Reliability**

Inter-rater reliability assesses the degree of agreement between the appraisals of two or more observers. Ideally, these observers would be experts in the relevant field.

To provide the necessary expert evaluations, panels of 12 to 20 staff members at the Institute of Personality Assessment and Research (IPAR) in Berkeley rated 345 men and 242 women by using the Q-sort descriptions contained in Block's California Q-Set (1961). All individual sortings made by each observer in a panel were summed into a composite, and then these sums were re-arranged into the fixed Q-sort distribution. These composite descriptions on the Q-sort were then correlated with the participants' ACL scores.

The inter-rater reliability coefficients were computed by assigning even-numbered observers to one panel and odd-numbered observers to another, and then correlating the Q-sort-ACL correspondences of the even and the odd-numbered panels. These reliability coefficients tended to be very large, usually ranging from .95 to .99. These large coefficients reveal ACL's high degree of inter-rater reliability.

## **Inter-method Reliability**

Inter-method reliability assesses the degree to which test scores are consistent when there is a variation in the methods or instruments used. Because the ACL does not employ different methods of measurement, investigating the inter-method reliability is irrelevant.

## **Internal Consistency Reliability**

Internal consistency reliability assesses the consistency of results across different items within a test. One of the most widely used methods for determining the internal consistency of an assessment is Cronbach's Alpha. Cronbach's Alpha uses the associations among a set of items to indicate how well the items, as a group, hold together. The alpha is expressed as a number between 0 and 1 and indicates the average associations among a set of items (e.g. Six Global Factors). If items comprising factors have a reliability coefficient of .50 or greater, we can have confidence that the construct is being measured reliably by all the individual items comprising that factor. ACL Reliability Coefficients rank from .53 to .95.

The high degrees of test-retest, inter-rater, and internal consistency reliability of the ACL reveal that this method is a reliable assessment. The next and final step in determining whether the ACL is accurate is testing its validity.

# Validity

The validity of a test is defined as the degree to which the test actually measures what it is intended to measure. When added to the demonstrated reliability of measurement, establishing the validity of the assessment authenticates the test as accurate. There are three main types of evidence of test validity - content, construct, and criterion validity (Guion, 1980).

## **Content Validity**

Content validity assesses the extent to which a test represents all relevant aspects of the concept the test is supposed to address. Ascertaining that the assessment addresses a wide range of facets requires investigating whether both the demographics of the original participants and the structure of the test allow for variability.

### *Normative Sample*

Ensuring that the test produces accurate results for an individual of any background requires including people of various demographics in the normative sample - the group of individuals used to determine scaled scores and percentile ranks. The most important demographics that must be represented in the normative samples are different sexes, age groups, and occupations.

For the ACL, the norm consisted of 4,144 females and 5,238 males. The male sample was drawn from groups of high school students (634), college students (936), graduate students (621), medical students (718), delinquents (293), psychiatric patients (50), and adults in various work settings (1,986). The female sample was drawn from high school students (410), college students (1,214), graduate students (336), medical students (990), law students (52), and adults (2,092) (Gough & Heilbrun, 1983, p.29).

While the manual did not disclose the exact ages for the individuals in the normative sample, the different participating occupations (e.g. high school students and medical students) reveal that the sample contained a wide array of age groups.

### *Cross-Cultural Validity*

In addition to examining the occupations and the age groups of the normative sample, it is necessary to investigate whether the test is valid across different cultures. The ACL was originally developed and studied with subjects from the United States. However in following years, several validity studies have been done. One of the books in The Plenum Series in Social/Clinical Psychological, *The Importance of Psychological Traits: A Cross-Cultural Study* by Williams, Satterwhite, and Saiz describes the use of the ACL in extensive cross-cultural studies. Specifically, the authors studied the importance of various psychological traits and their

favorability in 20 countries. Currently the ACL has been translated into 25 languages including Spanish, Chinese, Russian, French, Finnish, German, Italian, Japanese, Portuguese, Vietnamese, Cantonese (Simplified), Danish, Dutch, U.K. English, Finnish, Hungarian, Italian, Castilian Spanish, North American Spanish, Swedish, Thai (330 items), Mandarin (Simplified), Polish, and Argentine Spanish.

### ***Test Structure***

In addition to including a plethora of demographics, the test must be able to account for a wide range of personality types. To this end the test includes 300 adjectives that can be selected in a binary fashion to represent a vast array of personalities. The compilation of these exact 300 adjectives was the result of ten years of empirical research that culminated in the current version of the ACL, which includes adjectives such as "aggressive", "rational", and "optimistic".

### **Construct Validity**

While the varied normative sample and flexible test structure ensure ACL's content validity, it is necessary to also study the ACL's construct validity to be confident in its overall accuracy. Investigating the construct validity of a test determines whether the assessment measures the concept it was designed to address. Most commonly, construct validity is first assessed observationally by experts in a relevant field. These observations are then followed by empirical tests.

### ***Scales***

While the 300 adjectives provide detailed descriptions of personalities, the vast amount of information precludes effortless comparisons between individuals. To provide information that is easily digested and more structured but is still highly detailed, the ACL grouped together adjectives that covaried highly into 37 scales, such as "nurturance" and "personal adjustment". Some of these 37 scales were originally derived rationally by experts through grouping adjectives according to their inferred psychological meaning. The remaining scales were developed empirically by correlating the self-reported adjectives with external behaviors and grouping those adjectives with largest positive or negative correlation coefficients.

### ***Factors***

In addition to forming scales, the ACL further structures the psychological information by revealing the underlying behavioral factors through a statistical method called "factor analysis". Factor analysis organizes large sets of variables into smaller, more meaningful groups. When applied to the ACL data, the factor analysis groups the ACL items and scales that are most strongly correlated with one another. The specific factor analysis method used was principal components followed by an orthogonal factor rotation, a.k.a. Varimax rotation. This process

works to maximize the distinction between factors. The results yielded six factors each having an eigenvalue of 1.0 or greater. A common rule of thumb is that when a set of items has an alpha level of .70 or higher, it is considered acceptably reliable. The Six Global Factors contained in the ACL are the factors that were retained in the rotations and each of them has an eigenvalue of 1.0 or greater.

### ***Convergent and Discriminant Constructs***

To completely investigate ACL's degree of construct validity, it is necessary to consider both convergent and discriminant construct validity. Convergent validity describes the degree to which two measures of the same construct (e.g. Emotional Intelligence) are related. Discriminant validity describes the degree to which constructs that should not be related are in fact independent.

One way to examine both convergent and discriminant construct validity is to correlate the ACL scale scores to ensure that the scales that should be closely related psychologically correlate highly and those that should be independent do not.

Table 2 reveals all pair-wise correlations between the 37 scales. The correlations above the diagonal were derived from the male sample, while those below the diagonal were from the female participants. These correlation coefficients generally follow the psychologically hypothesized trends (e.g. highly positively correlated Endurance and Achievement (0.79 for males, 0.72 for females) and no correlation between Nurturance and Change (-0.03 for males, 0.01 for females)).

TABLE 2  
Pair-wise Correlation Coefficients between the 37 ACL Scales

Scales	Ckd	Fav	Fav	Com	Ach	Doal	End	Ord	Inc	Nur	Aff	Het	Exh	Aut	Agg	Cba	Sue	Aba	Def
No. Ckd		0.16	0.02	0.1	0.09	0.04	0.08	0.07	0.09	0.04	0.15	0.02	0.01	0	-0.01	0	0	-0.03	0.02
Fav	0.07		-0.68	0.63	0.66	0.49	0.65	0.57	0.79	0.67	0.82	0.49	0	-0.26	-0.31	0.07	-0.45	-0.25	0.22
UaFav	0.06	-0.68		-0.56	-0.44	-0.31	-0.57	-0.46	-0.68	-0.72	-0.6	-0.36	0.22	0.52	0.48	0.04	0.44	0.11	-0.39
Com	0.08	0.6	-0.53		0.38	0.22	0.43	0.36	0.58	0.54	0.5	0.29	-0.18	-0.37	-0.27	0.13	-0.16	0.03	0.36
Ach	0.01	0.6	-0.43	0.31		0.11	0.79	0.64	0.44	0.21	0.4	0.25	0.18	0.04	0.15	0.05	-0.5	-0.53	-0.16
Dom	-0.02	0.43	-0.26	0.14	0.73		0.49	0.33	0.22	0.14	0.33	0.42	0.56	0.3	0.47	0.27	-0.51	-0.77	-0.43
End	0.04	0.59	-0.49	0.34	0.72	0.38		0.87	0.55	0.32	0.41	0.09	-0.13	-0.27	-0.19	-0.29	-0.38	-0.24	0.2
Ord	0.06	0.44	-0.36	0.23	0.55	0.19	0.85		0.53	0.22	0.33	0	-0.23	-0.28	-0.22	-0.43	-0.33	-0.16	0.25
Inc	-0.02	0.5	-0.69	0.5	0.39	0.2	0.48	0.41		0.65	0.63	0.24	-0.23	-0.41	-0.46	-0.08	-0.3	0.01	0.38
Nur	-0.05	0.66	-0.11	0.52	0.19	0.08	0.32	0.16	0.64		0.76	0.48	-0.13	-0.6	-0.52	-0.03	-0.1	0.11	0.53
Aff	0.02	0.82	-0.59	0.51	0.39	0.33	0.4	0.26	0.63	0.74		0.53	0.04	-0.32	-0.37	0.09	-0.3	-0.15	0.31
Het	-0.02	0.5	-0.36	0.32	0.25	0.39	0.08	-0.1	0.27	0.46	0.52		0.4	-0.02	0.1	0.39	-0.2	-0.26	-0.1
Exh	-0.02	0.01	0.18	-0.17	0.18	0.61	-0.17	-0.31	-0.2	-0.11	0.08	0.4		0.53	0.69	0.48	-0.1	-0.55	-0.64
Aut	-0.02	-0.17	0.37	-0.31	0.15	0.49	-0.22	-0.28	-0.3	-0.52	-0.26	0.07	0.57		0.67	0.4	-0.21	-0.54	-0.83
Agg	0.01	-0.35	0.48	-0.28	0.13	0.48	-0.24	-0.28	-0.45	-0.51	-0.37	0.1	0.68	0.69		0.37	-0.07	-0.52	-0.76
Cha	0.05	0.12	-0.02	0.16	0.09	0.31	-0.3	-0.53	-0.01	0.01	0.12	0.42	0.49	0.42	0.34		-0.12	-0.29	-0.48
Sue	0.06	-0.49	0.48	-0.14	-0.51	-0.59	-0.36	-0.26	-0.37	-0.13	-0.35	-0.22	-0.12	-0.31	0.01	-0.11		0.66	0.22
Aba	0.03	-0.28	0.15	0.02	-0.49	-0.79	-0.21	-0.1	-0.09	0.09	-0.19	-0.26	-0.54	-0.61	-0.47	-0.25	0.66		0.63
Def	0.03	0.18	-0.35	0.31	-0.17	-0.52	0.1	0.29	0.3	0.48	0.27	-0.11	-0.63	-0.86	-0.76	-0.46	0.26	0.63	
Crs	0.06	-0.19	0.34	-0.19	0.21	0.15	0.01	0.1	-0.19	-0.56	-0.42	-0.23	0.06	0.46	0.4	0.02	-0.08	-0.23	-0.41
s-cu	0.04	0.14	-0.29	0.2	-0.06	-0.47	0.35	0.5	0.26	0.24	0.08	-0.32	-0.78	-0.66	-0.75	-0.6	0.07	0.46	0.73
5-Cfd	-0.04	0.64	-0.46	0.28	0.71	0.82	0.38	0.15	0.39	0.29	0.54	0.55	0.5	0.3	0.22	0.36	-0.63	-0.63	-0.31
P-Adj	-0.02	0.78	-0.72	0.48	0.5	0.39	0.5	0.38	0.67	0.65	0.74	0.45	-0.01	-0.26	-0.36	0.05	-0.45	-0.22	0.25
Iss	-0.01	0.73	-0.57	0.22	0.57	0.45	0.55	0.46	0.54	0.3	0.51	0.31	0.08	-0.01	-0.23	0	-0.62	-0.4	0
Cps	-0.08	0.44	-0.28	0.13	0.41	0.51	0.1	-0.07	0.32	0.08	0.26	0.44	0.42	0.44	0.2	0.47	-0.53	-0.51	-0.46
Mls	0.03	0.68	-0.65	0.55	0.66	0.41	0.7	0.65	0.64	0.44	0.5	0.17	-0.18	-0.22	-0.26	-0.12	-0.44	-0.27	0.23
Mas	0.01	0.21	0.02	-0.18	0.47	0.6	0.23	0.24	0.05	-0.23	0.08	0.01	0.35	0.5	0.38	-0.03	-0.46	-0.61	-0.44
Fem	0.06	0.49	-0.37	0.42	0.16	-0.03	0.18	0.12	0.41	0.61	0.51	0.47	-0.11	-0.37	-0.29	0.08	0.11	0.2	0.36
CP	0.12	-0.31	0.51	-0.26	0.23	0.32	-0.01	0.01	-0.4	-0.5	-0.42	-0.13	0.15	0.56	0.1	0.09	0.08	-0.31	-0.55
HP	-0.03	0.76	-0.71	0.45	0.44	0.31	0.63	0.5	0.69	0.76	0.75	0.33	-0.11	-0.39	-0.48	-0.16	-0.43	-0.15	0.4
A	0.02	0.67	-0.59	0.34	0.59	0.37	0.73	0.73	0.63	0.31	0.44	0.04	-0.19	-0.14	-0.35	-0.27	-0.58	-0.33	0.18
FC	-0.05	0.34	-0.22	0.05	0.33	0.63	-0.06	-0.24	0.12	0.15	0.34	0.56	0.74	0.46	0.41	0.6	-0.4	-0.6	-0.52
AC	0	-0.71	0.63	-0.3	-0.62	-0.54	-0.65	-0.59	-0.61	0.33	-0.53	-0.19	0.02	-0.01	0.24	0.11	0.7	0.5	-0.03
A-1	0.06	0.04	0.2	-0.12	-0.17	-0.01	-0.22	-0.27	-0.15	0.02	0.14	0.36	0.26	0.16	0.1	0.26	0.05	0.02	-0.19
A-2	0.04	-0.17	0.37	-0.14	-0.14	-0.04	-0.35	-0.38	-0.15	-0.35	-0.26	-0.03	0.19	0.49	0.3	0.4	0.06	-0.04	-0.46
A-3	0.08	0.54	-0.36	0.34	0.13	0.03	0.28	0.24	0.39	0.6	0.7	0.3	-0.03	-0.36	-0.4	-0.12	-0.08	0.04	0.4
A-4	0.1	0.51	-0.32	0.28	0.53	0.21	0.56	0.63	0.52	0.12	0.27	-0.07	-0.27	-0.12	-0.24	-0.28	-0.32	-0.18	0.15

TABLE 2 Continued  
 Pair-wise Correlation Coefficients between the 37 ACL Scales

Scales	Crs	Cu	Cfd	Adj	Iss	Cps	Mls	Mas	Fem	CP	NP	A	FC	At	A-1	A-2	A-3	A-4
No. Ckd	-0.05	-0.03	0.05	0.05	0.03	-0.01	0.07	0.14	0.11	0.09	0.04	0.05	0.03	-0.05	0.13	0.05	0.1	0.16
Fav	-0.49	0.16	0.64	0.78	0.73	0.36	0.74	0.34	0.44	-0.18	0.8	0.72	0.31	-0.74	0	-0.29	0.54	0.57
Oafav	0.28	-0.31	-0.45	-0.76	-0.6	-0.17	-0.69	-0.05	-0.26	0.46	-0.75	-0.66	-0.19	0.66	0.19	0.45	-0.42	-0.37
Com	-0.13	0.25	0.33	0.51	0.27	0.12	0.6	-0.03	0.32	-0.19	0.5	0.39	-0.02	-0.36	-0.2	-0.19	0.3	0.35
Acb	-0.58	-0.04	0.73	0.49	0.63	0.37	0.71	0.55	0.12	0.31	0.5	0.68	0.32	-0.66	-0.23	-0.21	0.1	0.59
Doa	-0.77	-0.42	0.84	0.35	0.52	0.42	0.5	0.65	-0.03	0.39	0.36	0.44	0.59	-0.57	-0.1	-0.18	0.02	0.29
End	-0.28	0.31	0.47	0.54	0.61	0.13	0.75	0.34	0.14	0.08	0.64	0.5	0.01	-0.71	-0.31	-0.35	0.21	0.65
Ord	-0.14	0.4	0.3	0.48	0.54	0.05	0.69	0.32	0.14	0.07	0.55	0.78	-0.13	-0.68	-0.27	-0.36	0.19	0.7
Iut	-0.15	0.32	0.36	0.66	0.52	0.21	0.66	0.11	0.4	-0.32	0.72	0.68	0.04	-0.62	-0.16	-0.21	0.43	0.58
Nur	-0.19	0.26	0.3	0.71	0.38	-0.01	0.5	-0.13	0.55	-0.55	0.78	0.38	0.15	-0.41	-0.01	-0.44	0.67	0.19
Afi	-0.43	0.11	0.89	0.72	0.52	0.21	0.52	0.18	0.49	-0.34	0.75	0.49	0.32	-0.56	0.13	-0.32	0.73	0.32
Bet	-0.55	-0.21	0.56	0.43	0.39	0.31	0.21	0.2	0.31	-0.1	0.36	0.08	0.54	-0.25	0.31	-0.18	0.34	-0.04
Exh	-0.62	-0.78	0.51	-0.12	0.09	0.37	-0.14	0.36	-0.1	0.38	-0.17	-0.18	0.72	0.01	0.17	0.11	-0.07	-0.25
Aut	-0.29	-0.63	0.15	-0.45	-0.13	0.38	-0.34	0.42	-0.34	0.58	-0.49	-0.26	0.35	0.15	0.16	0.5	-0.4	-0.21
Agg	-0.39	-0.73	0.26	-0.42	-0.17	0.22	-0.22	0.36	-0.3	0.68	-0.44	-0.31	0.41	0.21	0.02	0.25	-0.43	-0.21
Cba	-0.39	-0.56	0.31	-0.07	0	0.44	-0.1	0.03	-0.01	0.12	-0.16	-0.24	0.55	0.12	0.2	0.32	-0.11	-0.29
Sue	0.52	0.07	-0.56	-0.39	-0.56	-0.46	-0.41	-0.52	0.17	-0.02	-0.39	-0.56	-0.38	0.68	0.09	0.13	-0.07	-0.31
Aba	0.69	0.46	-0.62	-0.12	-0.4	-0.45	-0.26	-0.68	0.22	-0.4	-0.13	-0.3	-0.58	0.48	0.03	0.04	0.01	-0.14
Def	0.42	0.11	-0.21	0.36	0.01	-0.47	0.26	-0.42	0.35	-0.5	0.4	0.2	-0.49	-0.05	-0.16	-0.40	0.4	0.19
Crs		0.49	-0.82	-0.38	-0.49	-0.49	-0.33	-0.55	-0.03	-0.17	-0.32	-0.29	-0.15	0.47	-0.11	0.15	-0.22	-0.1
s-co	-0.17		-0.36	0.73	0.08	-0.39	0.27	-0.3	0.17	-0.44	0.31	0.34	-0.64	-0.2	-0.18	-0.31	0.18	0.27
s-cfd	0.02	-0.37		0.51	0.63	0.52	0.51	0.5	0.09	0.11	0.47	0.47	0.6	-0.5	-0.01	-0.19	0.2	0.32
P-Adj	-0.33	0.13	0.57		0.56	0.2	0.66	0.15	0.4	-0.36	0.76	0.65	0.2	-0.68	-0.14	-0.47	0.62	0.42
Iss	-0.02	0.09	0.6	0.58		0.49	0.58	0.43	0.08	-0.12	0.63	0.72	0.38	-0.72	0.05	-0.26	0.25	0.46
Cps	0.13	-0.45	0.63	0.32	0.54		0.15	0.37	-0.05	0.13	0.08	0.25	0.56	-0.29	0.09	0.29	0.01	0.16
Mls	-0.03	0.32	0.44	0.6	0.54	0.19		0.3	0.28	-0.07	0.71	0.77	0.09	-0.7	-0.31	-0.39	0.27	0.64
Mas	0.34	-0.29	0.48	0.12	0.39	0.32	0.21		-0.12	0.46	0.14	0.36	-0.33	-0.46	0.11	0	-0.02	0.35
Fem	-0.21	0.18	0.16	0.41	0.09	0	0.28	-0.2		-0.17	0.35	0.1	0.04	-0.07	0.09	-0.06	0.42	0.12
CP	0.6	-0.43	0.05	-0.41	-0.2	0.05	-0.16	0.4	-0.21		-0.41	-0.14	0.12	0.12	-0.06	0.32	-0.44	0.1
NP	-0.43	0.31	0.46	0.73	0.58	0.15	0.62	0.06	0.35	-0.51		0.71	0.14	-0.72	-0.12	-0.56	0.6	0.45
A	0.03	0.38	0.41	0.59	0.73	0.26	0.74	0.33	0.07	-0.2	0.66		0.09	-0.81	-0.26	-0.4	0.27	0.69
FC	0	-0.66	0.68	0.29	0.37	0.65	0.08	0.3	0.05	0.08	0.13	0.07		-0.25	0.16	0.02	0.1	-0.08
AC	0.07	-0.17	-0.59	-0.66	-0.73	-0.34	-0.69	-0.43	-0.06	0.22	-0.68	-0.8	-0.26		0.19	0.42	-0.33	-0.61
A-1	-0.06	-0.25	0.07	-0.06	0.01	0.14	-0.36	0.09	0.17	0.01	-0.09	-0.28	0.19	0.2		0.12	0.09	-0.25
A-2	0.32	-0.38	-0.06	-0.33	-0.16	0.28	-0.33	0.04	-0.07	0.32	-0.49	-0.34	0.15	0.3	0.16		-0.35	-0.14
A-3	-0.47	0.21	0.21	0.58	0.21	-0.03	0.79	-0.04	0.39	-0.46	0.5	0.28	0.01	-0.32	0.08	-0.78		0.15
A-4	0.25	0.3	0.28	0.4	0.45	0.16	0.6	0.3	0.08	0.01	0.35	0.65	-0.06	-0.6	-0.27	-0.12	0.16	

The high degrees of convergent and discriminant validity confirm that the ACL measures the construct it was designed to measure.

## **Criterion Validity**

Criterion validity assesses the degree to which one variable or set of variables predicts an outcome based on information from other variables. This type of validity is commonly measured by correlating the scores from the assessment in question to scores from already validated tests.

In the ACL manual, criterion validity is established by correlating the 37 scales with the California Psychological Inventory (CPI), the Minnesota Multiphasic Personality Inventory (MMPI), the Terman Concept Mastery Test, and a General Vocabulary Test. The manual further provides data showing that the social desirability response bias is "relatively inconsequential as a determiner of scores on the scales (Gough & Heilbrun, 1983, p.33).

In addition, the ACL has been correlated with measures of the Five-Factor Model (FFM), providing a useful interpretive reference point for understanding the construct validity of the ACL scales. ACL marker scales for these five major personality dimensions have also been developed and evidence good criterion validity (John, 1990; Piedmont, McCrae, & Costa, 1991; Williams, Patterson, & Fogle, 1995). Using these scales can help expand the interpretive and predictive power of the ACL.

After verifying the content, construct, and criterion validity of the ACL, it is clear that the ACL is an accurate measurement. As an authenticated test, the ACL can benefit many disciplines.

# Application

The two most impactful applications of the ACL are in education and the work force.

## Behavioral Measurement Over Time

Evidence suggests that while 89% of teachers believe most students are adequately prepared for postsecondary education or the work force, 40% of students and 67% of employers do not agree that graduates have sufficient skills (Gallup 2013). This large discrepancy demonstrates the need for a new way to evaluate student preparedness.

While teachers and employers may disagree, the consensus among policymakers, parents, students, and the public that American high schools do not adequately prepare all students for success in the 21<sup>st</sup> century is only growing. The importance of these non-cognitive skills has been further emphasized by recent evidence that while cognitive skills require extensive efforts to improve, non-cognitive skills remain malleable (Haney & Durlack 1998, O'Mara et al. 2006, Compas et al. 2001). Most importantly, improving non-cognitive skills has further been shown to increase cognitive outcomes and learning gains (Bowles & Gintis 2002, Farkas 2003, Heckman et al. 2006, Jencks 1979, Lleras 2008).

The ACL can be used to assess the development of these non-cognitive skills in educational settings. While other measurements with similar goals exist, Pairin's is unique in its quantitative and accurate approach to measuring a standardized set of non-cognitive skills. The information provided by the ACL scores can not only serve students, but also high school or college counseling centers or personnel offices.

## Matching Individuals to Employment

Understanding an applicant's behavioral tendencies and attributes enables employers to determine the fit between the individual and the position quickly and accurately. Once an employee is hired, the ACL information can also be used to design a behavioral development plan and evaluate the employee's behavioral growth accordingly.

In addition to assessing basic personality types, the ACL provides measures of coachable behavior that can be utilized to hone specific skills. The abilities necessary to succeed in a certain position can be determined by investigating the performance of the best employees (Hogan, Hogan, and Roberts, 1996). As personality traits, such as conscientiousness (Barrick and Mount, 1991), and scores on multidimensional inventories (e.g. Tett, Jackson, and Rothstein, 1991; McHenry, Hough, Toquam, Hanson, and Ashworth, 1990; Ones, Viswesvaran, and



Schmidt, 1993; McDaniel and Frei, 1994), have been shown to reliably predict a supervisor's ratings of job performance, the ACL can be used to evaluate necessary skills. Through the ACL, an employer can identify the behavioral predispositions that correlate with outstanding job performance and then determine which candidates would be most likely to exhibit similar performance.

Pairing the behavioral tendencies and attributes of an applicant or employee can be compared to the behavioral patterns of employees currently succeeding in a similar job provides many advantages. One such benefit is laser-focused development guidance for raising the percentage of the total workforce performing at the highest level of productivity.

# EEOC Compliance

The domains of occupational performance and corresponding behavioral predispositions are multifaceted. Consequently, when exploring correlations between personality and job performance it is essential to use a well-developed, multidimensional measure of normal behavior.

One of the primary concerns that arises with the use of behavioral testing in hiring is violation of federal anti-discrimination legislation. According to Title VII of the Civil Rights Act of 1964 (Title VII), the Americans with Disabilities Act of 1990 (ADA), and the Age Discrimination in Employment Act of 1967 (ADEA), employers may not use employment tests or selection procedures that prove to discriminate on the basis of disability, race, color, national origin, sex, religion, or age.

As personality and behavior testing becomes more prevalent in pre-employment selection, ensuring that no groups of people suffer from an unfair disadvantage is pivotal. Consequently, any test used in hiring must be rigorously scrutinized to determine any possible bias.

While selecting against a certain personality trait may be seen as an act of prejudice by some, section 902 of the 1992 EEOC issuance 915.002 states that "common personality traits are not impairments" and rejecting potential employees based on this information with considerable reason is not discriminatory. For instance, after "a psychological profile for a police officer determined that the applicant showed poor judgment, irresponsible behavior, and poor impulse control", the Supreme Court concluded that these personality traits do not qualify as a disability or impairment and consequently do not violate EEOC guidelines.

In addition, Hogan and Hogan found that there is no evidence that well-constructed personality inventories discriminate against any ethnic or national group (1995). Another researcher further found that measures of normal personality do not discriminate against people with disabilities (Hayes, 1996). As a reliable and valid assessment, the ACL belongs to this category of accurate non-discriminatory tests.

Furthermore, research conducted specifically on the differences in personalities of the five most predominant racial groups in the United States as measured by Gough's Adjective Checklist discovered no significant group differences in scores (Foldes, Dueher, & Ones 2008). In addition, Moore and Handal who focused on the personal adjustment scale of the ACL further concluded that there were no significant race or sex differences in ACL scores (1980).

Overall, the lack of discriminatory biases in the ACL qualifies it as EEOC compliant.

# References

- Barrick, M. R., & Mount, M. K. (1991). The Big-Five personality dimensions in job performance: A meta-analysis. *Personnel Psychology*, 44, 1 – 26.
- Block, J. *The Q-sort method in personality assessment and psychiatric research*. Springfield, Illinois: Charles C. Thomas, 1961. (Reprint. Palo Alto, California: Consulting Psychologists Press, 1978.)
- Bowles, S., & Gintis, H. (2002). Schooling in capitalist America revisited. *Sociology of Education*, 75(1), 1-18.
- Cartwright, R.D., Lloyd, S., & Wicklund, J. Identifying early dropouts from psychotherapy. *Psychotherapy: Theory, Research, and Practice*, 1980, 17, 263-267.
- Compas, B. E., Connor-Smith, J. K., Saltzman, H., Thomsen, A. H., & Wadsworth, M. E. (2001). Coping with stress during childhood and adolescence: Problems, progress, and potential in theory and research. *Psychological Bulletin*, 127(1), 87–127.
- Costantini, E., & Craik, K.H. Personality and politicians: California party leaders, 1960-1976. *Journal of Personality and Social Psychology*, 1980, 38, 641-661.
- Farkas, G. (2003). Cognitive skills and noncognitive traits and behaviors in stratification processes. *Annual Review of Sociology*, 29, 541-562.
- Foldes, H. J., Duehr, E. E., & Ones, D. S. (2008). Group Differences in Personality: Meta-Analyses Comparing Five US Racial Groups. *Personnel Psychology*, 61(3), 579-616.
- Formy Duval, D.L., Williams, J.E., Patterson, D.J., & Fogle, E.E. A “Big Five” scoring system for the item pool of the Adjective Check List. *Journal of Personality Assessment*, 1994, 65, 59- 76.
- Gallup. (2013). What America Needs to Know About Higher Education Redesign.
- Gough, H. G. (1975). *Manual for the California Psychological Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Gough, H.G., Fioravanti, M. & Lazzari, R. Some implications of self versus ideal-self congruence on the revised Adjective Check List, *Journal of Personality and Social Psychology*, 1983. 44, 1214-1220.
- Guion, R. M. (1980). "On trinitarian doctrines of validity". *Professional Psychology* 11: 385–398

- Hall, W. B., & MacKinnon, D. W. (1969). Personality inventory correlates of creativity among architects. *Journal of Applied Psychology*, 54, 332 – 326.
- Haney, P., & Durlak, J. A. (1998). Changing self-esteem in children and adolescents: A meta-analytic review. *Journal of Clinical Child Psychology*, 27, 423–433.
- Hayes, T. L. (1996). Personality correlates of performance: Does disability make a difference? *Human Performance*, 9, 121 – 140.
- Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411-482.
- Heilbrun, A.B., Jr. Validation of a need scaling technique for The Adjective Check List. *Journal of Consulting Psychology*, 1959,23,347-351
- Hofstede, G. *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage. (1980)
- Hogan, R., Hogan, J., & Roberts, B. W. (1995). Hogan Personality Inventory manual (2<sup>nd</sup> ed.). Tulsa, OK; Hogan Assessment Systems.
- Hogan, R., Hogan, J., & Roberts, B. W. (1996) Personality Measurement and Employment Decisions. *American Psychologist*, Vol. 51, No. 5, 469-477.
- Hollander, M.A., & Parker, H.J. Occupational stereotypes and self-descriptions: Their relationship to vocational choice. *Journal of Vocational Behavior*, 1972, 2, 57-65.
- Jencks, C. (1979). Who gets ahead? The determinants of economic success in America. New York: Basic Books.
- John, O.P. The search for basic dimensions of personality. In P. McReynolds, J.C. Rosen, & C.J. Chelune (Eds.), *Advances in psychological assessment*, 1990. Vol 7, 1-37. New York: Plenum.
- Kilmann, P.R., Personality characteristics of female narcotic addicts. *Psychological Reports*, 1974, 35, 485-486.
- Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social Science Research*, 37(3), 888-902.
- Marketing Accountability Standards Board (MASB). 2010. Common Language: Marketing Activities and Metrics Project.

- McDaniel, M. A., & Frei, R. L. (1994) Validity of customer service measures in personnel selection: A review of criterion and construct evidence. Manuscript submitted for publication.
- 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.
- Moore, C. D., & Handal, P. J. (1980). Adolescents' MMPI performance, cynicism, estrangement, and personal adjustment as a function of race and sex. *Journal of Clinical Psychology*, 36(4), 932-936.
- O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist*, 41(3), 181–206.
- Ones, D. S., Viswesvarian, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validation: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679 – 703.
- Simonton, D.K. Presidential personality: Biographical use of the Gough Adjective Check List. *Journal of Personality and Social Psychology*, 1990, 51, 149-160.
- 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.
- Trent, R.D., Fernandez-Marina, R. & Maldonado-Sierra, E.D. The cross-cultural application of the adjectival check list adjustment index: A preliminary report. *Journal of Social Psychology*, 1960, 51, 265-276
- Trochim, W. (2006). Types of Reliability. The Research Methods Knowledge Base. Last Revised: 20 October 2006
- Vidoni, D.O. Factor analytic scales of the Adjective Check List (ACL) replicated across samples: Implications for validity. *Education and Psychological Measurement*, 1977, 37, 535-9.
- West, M. R., Kraft, M. A., Finn, A. S., Martin, R., Duckworth, A. L., Gabrieli, C. F., & Gabrieli, J. D. (2014). Promise and Paradox: Measuring Students' Non-cognitive Skills and the Impact of Schooling.
- Williams, J.E. & Best, D.L. *Measuring sex stereotypes: A multinational study*. 1990, Beverly Hills: Sage.
- Zimet, C.N., & Held, M.D. The development of views of specialties during four years of medical school, *Journal of Medical Education*, 1975, 50, 157-166.