

**The MERIT Profile™ : Studies that Evaluate the
Measurement of Character and Behavior**

Mike Williamson, Ph.D.
Precision Human Development, Ltd.
VP and Director of Research and Development

**February 2008
San Diego Office,
San Diego, California 92106**

The MERIT Profile™ : Studies that Evaluate the Measurement of Character and Behavior

Mike Williamson, Ph.D.
Precision Human Development, Ltd.
VP and Director of Research and Development

Executive Summary

Precision Human Development (PHD) partnered with Future Achievement International (FAI) and A-Check America (ACA) to develop the MERIT Profile survey. The objectives provide an online, interactive survey that assesses four factors of behavior and ten factors of character as defined by concepts referred to as the MAXIMIZERS™, copyrighted by FAI.

The present report summarizes the findings of two research studies that were designed to investigate and evaluate the following important issues:

1. The reliability and validity of the Survey;
2. The effectiveness of the Survey to qualify “recruits” for specific job openings;
3. The value of the Survey for recommending coaching services to individuals with specific character development needs;
4. The efficiency of benchmarks for predicting character competency;
5. The potential effectiveness of benchmarks for predicting “success” on a wide range of client-defined criteria.

All of the above were based on metrics made available by responses to the MERIT Profile survey, Version 1.0, with the intention of using results to refine and improve Version 2.0.

The first study, the Job Applicant Study, was conducted retrospectively on a sample of 943 individuals who had taken the MERIT Profile Survey. Forty-four, long-tenured, Office Managers from a dental services group were assigned to a Control Group and 899 job applicants were assigned to an Experimental Group for the purposes of the Study. Seven Control and 243 Experimental Group surveys were determined to be invalid, based on proprietary results screening procedures, reducing the sample to 693 respondents.

Subsequently, separate factor analyses were performed on the Control and Experimental Group data. The factor loadings ranged from 0.45 to 0.88, with only three of 96 loadings falling below 0.50. Thus, for 11 of the 14 factors, the items were very appropriate for the factor and all of the items that loaded for one group also loaded for the second group. Items that loaded for three factors for character were inappropriate, therefore justifying replacements for those items on MERIT Profile Survey Version 2.0.

At that point in the investigation, a Character Index (CI) was derived to represent each individual's overall character competency. Benchmarks developed on four Control Groups, each developed on a different behavioral trait, validated the Survey's *intrinsic validity* by accurate, blind classifications of CI values for respondents in the four Experimental Groups. Accuracies in classifications ranged from 89.3 to 100 percent.

The second study, referred to as the Student Study, was prospectively conducted to examine the instrument's *extrinsic validity*. In that study, the Survey was administered to three distinctly different groups, 42 students attending Point Loma Nazarene University, San Diego, California, a private, sectarian university that attracts students from families with primarily traditional values, 48 students attending Douglas College in Vancouver B.C., Canada, a non-sectarian, public college, and 41 adults in a CRASH drug rehabilitation program at Bay View Hospital, also in San Diego.

For the Student Study, the mean percentile scores for the CI—a scale that ranged from 0 to 1000—were in the predicted order for the three groups, as follows: PLNU 801; Douglas 432; CRASH 321. The difference between mean values was very significant for every comparison.

The Survey's *extrinsic validity* was demonstrated by deriving benchmark equations on odd-numbered Control Groups and applying them to even-numbered Experimental Groups. Accuracies in blind classifications for Experimental Group members ranged from 97.7 to 100 percent.

In a separate analysis that pooled the entire sample, *extrinsic validity* increased further to 99.8 percent effectiveness in the accurate classification of individual's whose Character Indexes reached an acceptable threshold.

The investigation provided insights for improving Version 2.0. Innovations included:

1. *Adding options*--increasing the number of Likert scale options from five to seven to reduce the frequency of extreme option selections;
2. *Reverse scoring* to decrease response set;
3. Invalidating surveys with long strings of *single option responses*;
4. Invalidating surveys by *exaggerating discrepancies* between scores within the same factor;
5. Using interviews to *challenge respondents* whose responses were exceptionally high to reduce false-positive errors in decisions;
6. Using interviews to *challenge respondents* whose responses were exceptionally low to reduce false-negative errors in decisions.

The long-term vision for the MERIT Profile survey is to continue its refinement and to enlarge its application. The need for the Instrument is timely and irrefutable. The scope of its application is virtually unlimited. Its use of scientific methods and benchmarks in controlled studies for answering important questions provides respected strategies for testing and confirming its validity. Its impact on decisions with reference to selection, retention, and succession planning in the workplace could be very significant. Its use for research purposes needs to be explored and exploited as a tool for solving major problems, not only in the behavioral sciences, but in other professional disciplines as well. Examples of measurable results could include decreasing abuses in claims for Workman's Compensation by at least 20 percent, and reducing accidents among drivers of large trucks by at least 20 percent. The savings could be in the billions of dollars.

The MERIT Profile™ : Studies that Evaluate the Measurement of Character and Behavior

Mike Williamson, Ph.D.
Precision Human Development, Ltd.
VP and Director of Research and Development

The purpose of this report was to describe the strategies and procedures that were applied in the development of the MERIT Profile Survey and to examine the results of its initial application.

Objectives

The MERIT Profile Survey was developed to meet the need for an instrument that measures both behavior (personality) and character, and provides the metrics for creating both generic and customized benchmark equations used for making better selection and succession planning decisions in any context. This report documents the progress that has been made, to date, toward those objectives.

The Objectives were based on the following presupposition:

If research evidence supports the conclusion that the two constructs, behavior and character, can be accurately and appropriately measured by a single instrument and the metrics from that instrument can be translated into mathematical equations (benchmarks) that predict “success” on various important criteria, then organizations will have access to a decision-making tool that increases profits and provides many other valuable benefits from a wide-range of applications.

Precision Human Development (PHD) has therefore partnered with Future Achievement International (FAI) and its partner A-Check America, Inc. (ACA) to develop the MERIT Profile Survey for several important Internet applications, as follows:

1. To assess, describe and graphically display individuals' personality profiles with reference to four cornerstone behavioral traits; Dominance, Extroversion, Patience and Conformity;
2. To assess, describe and graphically display individuals' character profiles, delimited to qualities defined by FAI's ten MAXIMIZERS™ character competencies;
3. To qualify selected individuals as “recruits” for job openings;
4. To identify certain individuals for coaching services designed to help them improve in specific areas of need with reference to character development; thereby insuring their future employability.

5. To create a Character Index (CI) that provides a single metric that represents an accurate, percentile assessment of an individual's overall character.
6. To produce databases for use in generic or customized benchmark development that captures the mathematical relevance between the metrics provided by the Survey and important client-defined criteria. ***When properly applied, those benchmarks add significant profits and other benefits to the organization for the purpose for which they were prepared.*** Common client-defined criteria include quality of performance, productivity, sales, reduction of turnover, reduction of accident-prone behaviors, risk management, reduction of fraudulent Workman's Compensation claims, and others.

Early Development

The development of the MERIT Profile Survey began in earnest in January 2005. The first experimental version of the survey contained 120 items. That version was administered to three sets of 100 plus individuals, with revisions between each administration based on results from each previous administration.

By November of 2005 the MERIT Profile Survey, Version 1.0, had been reduced to 60 items that were purported to measure four cornerstone factors of behavior (Dominance, Extroversion, Patience and Conformity) and ten character competencies, referred to as MAXIMIZERS™.

"MAXIMIZERS" is an acrostic, copyrighted by Future Achievement International, in which each letter represents a quality of character, or more commonly referred to as a character competency. Particular sets of competencies are organized such that they identify Attitudes, Beliefs and Commitments, also referred to as the A, B, C's of character, as follows:

Attitudes	MTH	-	Make things happen
	APS	-	Achieve personal significance
	XON	-	"X" out the negatives
Beliefs	IRP	-	Internalize right principles
	MTM	-	March to a mission
Commitments	IAL	-	Integrate all of life
	ZOP	-	Zero in on caring for people
	EIN	-	Energize internally
	RIG	-	Realign rigorously
	STC	-	Stay the course

The application of Version 1.0 produced large amounts of data and provided additional opportunities to conduct controlled studies. Results from two studies revealed the need to program certain innovations into MERIT Profile Version 2.0, which has been completed. The remainder of this report describes the methods used and the findings obtained in the two studies.

The first study, the Job Applicant Study, was retrospective in that the Investigator had access to scores from previously-administered surveys. The second study, the Student Study, was prospectively conducted on three groups whose values, backgrounds and lifestyles were distinctly different, each from the others.

The Job Applicant Study

In August of 2007, a dental services group ordered a benchmark for a position entitled Office Manager that included 44 employees from various locations throughout the company. There were reasons to think that those individuals would be less inclined to bias their responses to the MERIT Profile Survey than would job applicants to whom the Survey is more frequently given—they appeared to represent mature, experienced employees who were secure in the position.

Based on that perception, Control and Experimental Groups were formed that were composed of the 44 Office Managers and 899 job applicants, respectively. The large amount of data provided laboratory-type conditions for a thorough examination of Version 1.0 in anticipation that the findings would help to increase the reliability and validity of Version 2.0.

For any psychometric instrument to be valid, it first must be consistent in the measurement of the constructs it purports to measure. For the MERIT Profile Survey a key element of consistency depended upon the *commonality* of responses within factors for the two constructs, behavior and character.

Thus, an examination of the *internal consistency* of the instrument was conducted that led to the development of screening tests for “invalid” surveys. Self-rated surveys are notoriously vulnerable to skewing toward option choices that the respondent perceives to be personally advantageous. Any significant reduction of those sources of intrinsic error would also increase the *validity* of the instrument for its intended *predictive* purposes.

Special algorithms, developed by the Investigator, were applied that clearly identified “invalid” surveys. That result caused those individuals to fail to qualify for *recruitment referral*; however, some of those failures qualified respondents for coaching services and *needs development*.

The specific flaws that invalidated certain surveys included the following:

1. Surveys that contained long strings of items for which the option selected was the same number—the selection of one option over several consecutive items could not be true because each string contained items that required opposite, or near opposite, option selections.
2. Surveys that contained excessive numbers of selections of Option 5, the highest, most positive response.

3. Surveys that contained large inconsistencies in option selections within factor—by definition *commonality* means that items that measure a given factor elicit the same, or nearly the same, option choice.

In the present study, seven of the 44 cases from the Control Group (15.9 percent) exceeded the acceptable threshold for valid surveys and 243 of the 899 cases from the Experimental Group (27.0 percent) failed to pass the screening tests. The relatively small percentage of invalid surveys by the Control Group supported the view that it served its purpose and provided a standard to be emulated.

Following the elimination of the 250 *invalid* surveys, a factor analysis was applied to the Control and Experimental Groups, separately, to answer the question: How similar were the sets of items that measured each factor for the two Groups? The answer to the question is contained in results reported in Tables 1 through 4.

Factor loadings of 0.4 or higher meet the conventional standard of acceptability. In the present analysis the factor loadings ranged from 0.45 to 0.88, with only three of 96 loadings falling below 0.50. Even more remarkable, however, was the fact that precisely the same items that defined a factor for the Control Group also defined the factor for the Experimental Group. That finding occurred for all four behavioral factors and seven of ten factors for the MAXIMIZERS™.

It is also noteworthy, that the items that defined a particular factor by both Groups were exceptionally appropriate for all of the reported items. However, of the five Commitment factors, only two are reported in Table 4. The explanation resides in the fact that only a few items that loaded for Integrate all of life (IAL), Energize internally (EIN) and Realign rigorously (RIG) were appropriate for those factors. But the failure was that of the Survey Developer (the Investigator). That finding pointed to the need for more diligence in crafting items that would measure those factors in Version 2.0.

It is common in psychometrics for “subtests” to be summarized in a single *score* that represents the overall construct. In the present context, metrics from all items that measured character were transformed into a continuous variable of percentile scores ranging from 0 to 1000 and given the label *Character Index (CI)*.

As a result of the elimination of the *invalid* Surveys, the internal consistency of the Survey was increased, preparing the way for testing the instrument’s *intrinsic* and *extrinsic validity*.

To test the Survey’s *intrinsic validity*, four benchmarks¹ were developed to predict the Character Index. The sample was sufficiently large to create a separate benchmark for subgroups whose members all had the same highest behavioral trait. The ability to control for behavioral style in that way eliminated a major source of intrinsic error.

¹ Appendix A contains the rationale for the theory and some of the details about how benchmarks are developed.

Table 1

**MERIT Profile Survey, Version 1.0, Factor Loadings for Four
Factors of Behavior with Item Description
CG: N=37; EG: N=655**

Item Description	DOM		EXT		PAT		CNF	
	CG	EG	CG	EG	CG	EG		
25. I am driven by vision, what could be	.79	.68						
14. I make things happen	.74	.51						
26. I am direct and to the point...	.64	.65						
50. I march to a mission	.61	.58						
39. I take charge, naturally	.61	.53						
28. I'm very outgoing			.78	.74				
02. I stir people up and get them excited			.74	.64				
57. I generate enthusiasm			.66	.72				
51. I'm a people person			.63	.66				
55. I am a motivator by nature			.62	.63				
42. I have lots of patience					.83	.63		
19. I'm sensitive to the feelings of others					.77	.71		
30. I'm driven by a need to help people					.76	.66		
18. I am gentle, warm and likeable					.67	.70		
10. I go out of my way to help people					.66	.63		
03. I keep my space neat							.87	.86
48. I am highly organized							.84	.87
31. I keep everything in its place							.82	.86
08. I am a perfectionist							.68	.57
21. I need to do it right, by the rules...							.47	.53

Note: Factor analysis was performed separately on a Control Group (CG) and an Experimental Group (EG) with the above results for: Dominance (DOM); Extroversion (EXT); Patience (PAT); Conformity (CNF). Screening tests for invalid Surveys reduced the original groups by 15.9 %, from CG =44 to CG =37, and 27.1 % from EG=899 to EG=655, respectively.

Table 2

MERIT Profile Survey, Version 1.0, Factor Loadings for THREE CHARACTER Competencies for Attitude With Item Description
CG: N=37; EG: N=655

Item Description	MTH		APS		XON	
	CG	EG	CG	EG	CG	EG
39. I take charge, naturally	.85	.77				
29. I am very bold and aggressive	.84	.73				
09. I take the initiative	.64	.61				
14. I make things happen	.62	.60				
37. I have high aspirations			.85	.73		
41. I work hard to overcome weaknesses			.80	.71		
53. I am destined for significance			.73	.65		
49. I keep focused, not easily distracted			.59	.54		
04. I make a positive impact on other people					.83	.62
44. I'm always positive, optimistic					.79	.72
20. I'm a happy person					.79	.77
60. I like myself					.77	.54

Note: Factor analysis was performed separately on a Control Group (CG) and an Experimental Group (EG) with the above results for: Make things happen (MTH); Achieve Personal Significance (APS); X-out the negatives (XON). Screening tests for invalid Surveys reduced the original groups by 15.9 %, from CG=44 to CG=37, and 27.1 % from EG=899 to EG=655, respectively.

Table 3

MERIT Profile Survey, Version 1.0, FACTOR LOADINGS FOR TWO Character Competencies for Beliefs With Item Description
CG: N=37; EG: N=655

Item Description	IRP		MTM	
	CG	EG	CG	EG
52. I am respected for my integrity	.85	.83		
40. I am respected for my good qualities of character	.84	.77		
01. People respect me for my honesty	.84	.75		
43. People respect me for my high moral values	.80	.74		
12. I am goal oriented			.75	.57
50. I march to a mission			.74	.81
15. I am devoted to a cause			.72	.70
22. I have a known purpose for living and I know what it is			.47	.55

Note: Factor analysis was performed separately on a Control Group (CG) and an Experimental Group (EG) with the above results for: Internalize right principles (IRP); March to a mission (MTM). Screening tests for invalid Surveys reduced the original groups by 15.9 %, from CG=44 to CG=37, and 27.1 % from EG=899 to EG=655, respectively.

Table 4

**The MERIT Profile Survey, Version 1.0, FACTOR LOADINGS FOR TWO Character Competencies for COMMITMENTS With Item Description
CG: N=37; EG: N=655**

Item Description	ZOP		STC	
	CG	EG	CG	EG
47. I build others up	.83	.68		
07. I'm generous with compliments	.82	.64		
10. I go out of my way to help people	.77	.75		
30. I'm driven by a need to help people	.72	.78		
13. I just keep on trying			.88	.71
05. I never give up			.87	.78
24. I'm tenacious; I don't quit			.76	.83
33. I persevere to the end			.45	.73
Note: Factor analysis was performed separately on a Control Group (CG) and an Experimental Group (EG) with the above results for: Zero in on caring for people (ZOP) Stay the course (STC) and. Screening tests for invalid Surveys reduced the original groups by 15.9 %, from CG N=44 to CG=37, and 27.1 % from EG=899 to EG=655, respectively.				

To achieve the goal in this investigation, the CI was ranked from the highest to the lowest value. Odd-numbered cases were assigned to a Control Group (CG) and even-numbered cases formed an Experimental Group (EG).

The Character Index was also subdivided to form High and Low Subgroups. The results reported in Chart 1 exclude all but the Predicted High CI classifications, since those categories provide the base for calculating predictive accuracies and are the only individuals who would be considered for selection in an empirical application.

The methods validated the effectiveness of the benchmarks. Although the benchmarks were created exclusively on the Control Groups, they were applied independently to both Groups. Thus, the *blind* classifications of members of the Experimental Groups simulated the efficiency by which the benchmarks would be expected to identify job applicants in real applications.

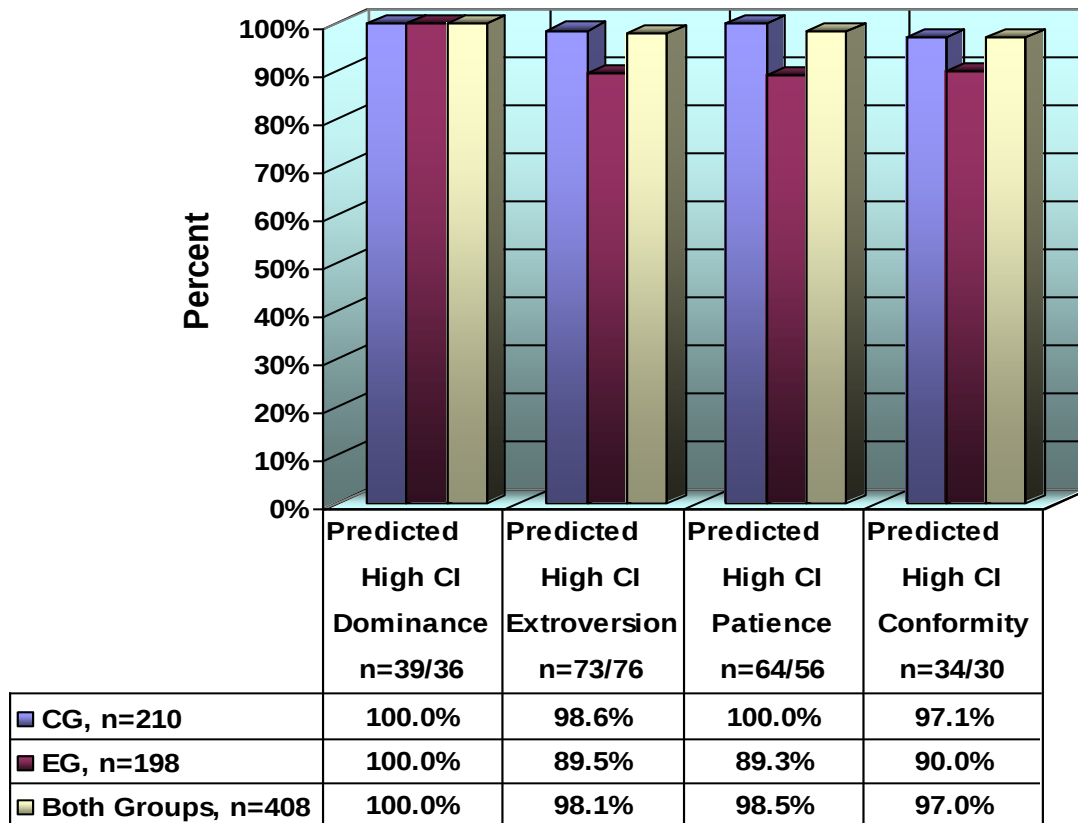
Predictive accuracies ranged from 97.1 to 100 percent in the Control Groups and 89.3 to 100 percent in the Experimental Groups. In contrast, if there had been no correlation between metrics provided by the Survey and the Character Index, the selection of High CI values would have been approximately 50 percent. Thus, it would have taken more than 10,000 similarly conducted studies to have produced these results once, if chance was the only "cause." ($p < 0.0001$)

Subsequently, four new benchmarks were derived that combined the Groups. That analysis, also reported in Chart 1, produced predictive accuracies from 97.0 to 100 percent.

While the separate benchmarks for the Control and Experimental Groups validated the benchmarks and provided evidence of their effectiveness, the benchmarks developed on the entire sample ultimately would be used, due to their greater stability.

Chart 1

The Merit Profile: Efficiency of Benchmarks to Predict the Character Index for Four Factors of Personality by Group

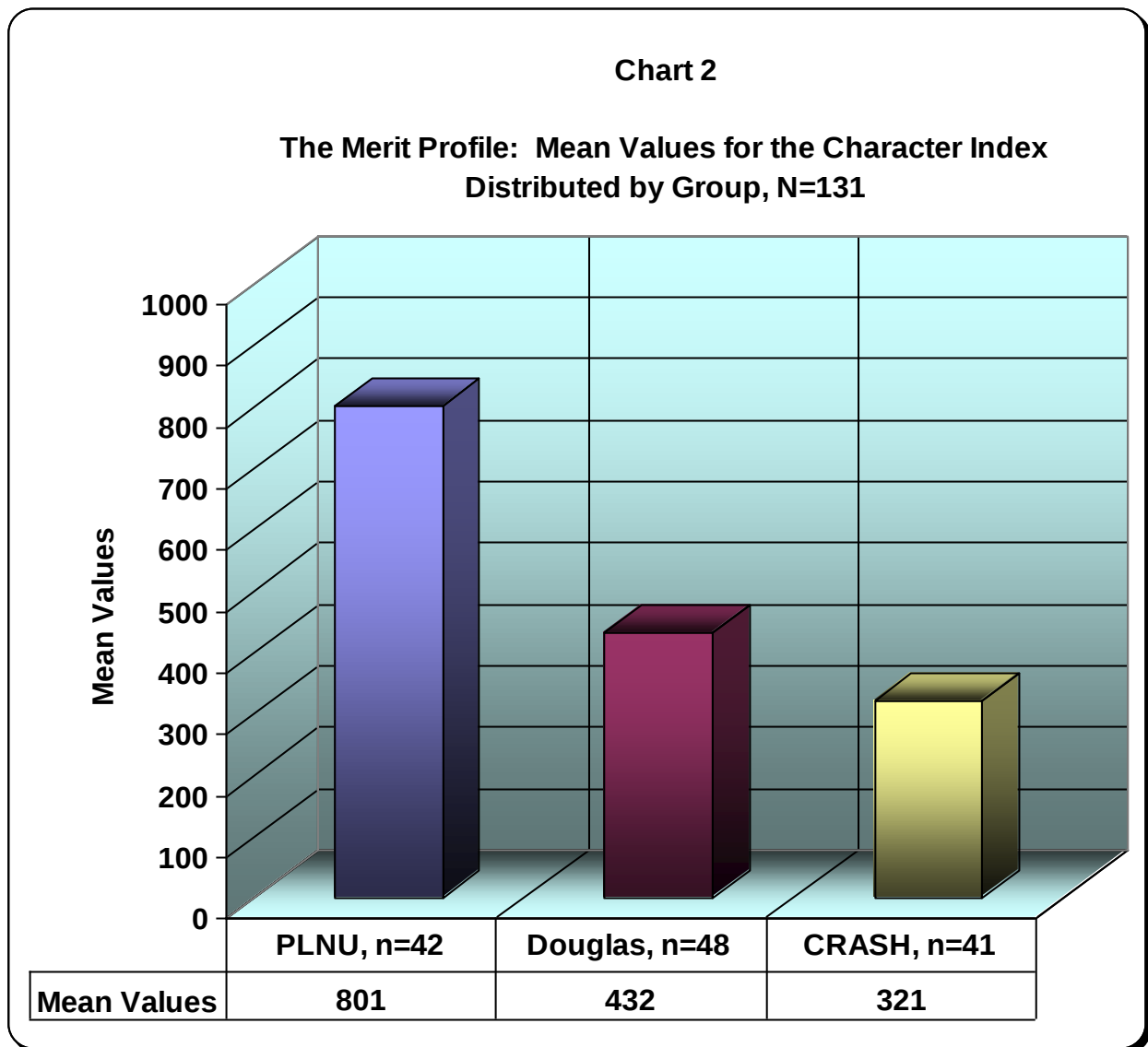


Note: Predicted Low CI Values were not included in the Chart, because the purpose of the benchmark was to "select" individuals who have above average character, as defined. The total sample size was 943; the sample for this analysis was 693. The Predicted High CI for Both Groups (408) represented 43.3 percent of the total sample and 58.9 percent of the above average sample.

The Student Study

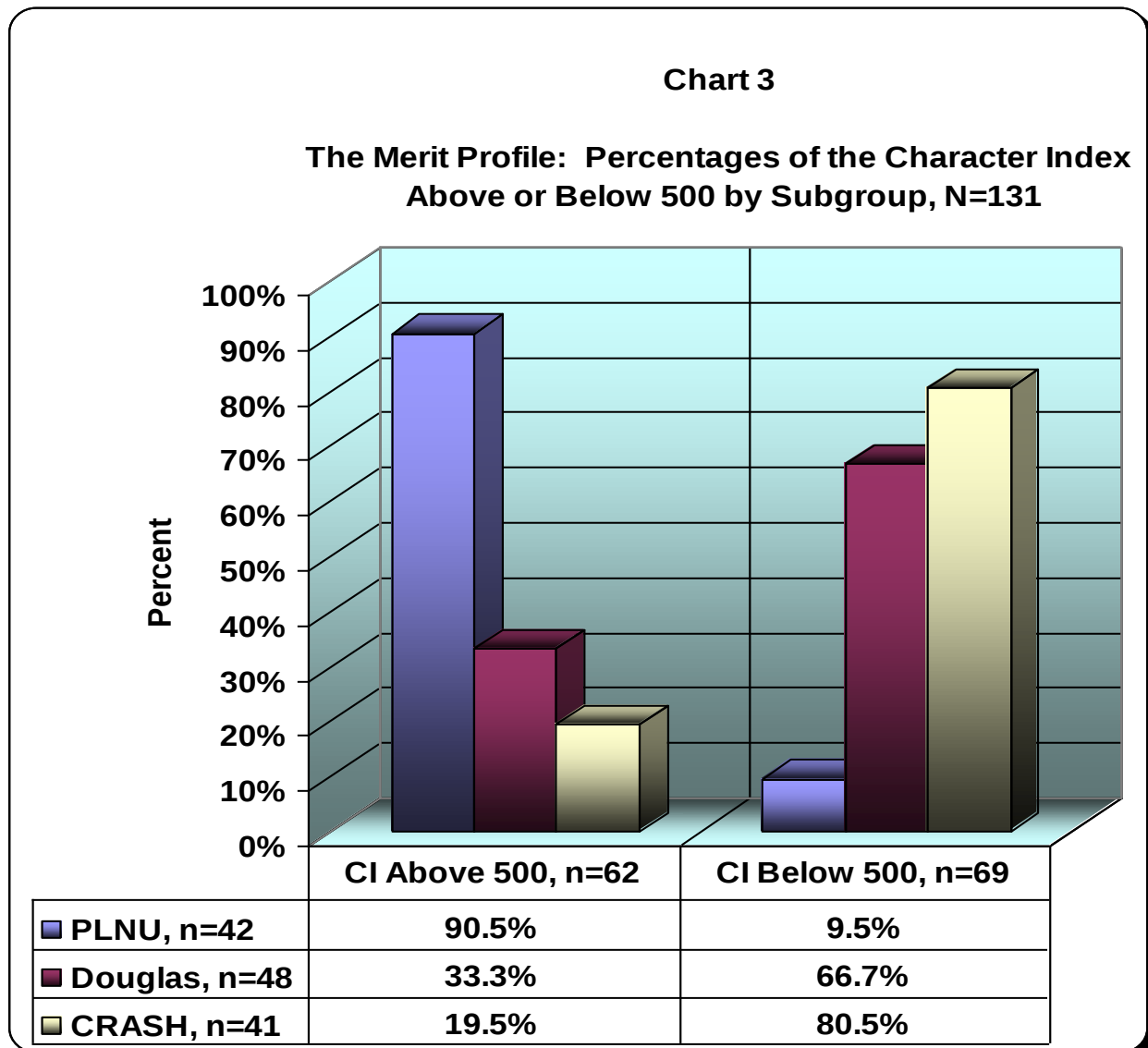
The second study, also referred to as the Student Study, was a prospectively conducted study that was specifically designed to maximize expected differences among the groups on their responses to the MERIT Profile Survey. Group 1 of three selected groups included 42 students at Point Loma Nazarene University (PLNU), a private university in San Diego, California, made up of students from mostly upper middle class families who represented conservative, traditional values. Group 2 contained 48 students from Douglas College (Douglas), a non-sectarian college in Vancouver B.C., Canada. Group 3 was composed of 41 adults in a drug rehabilitation program, referred to as CRASH, who were temporary residents at Bay View Hospital also in San Diego, many of whom had criminal records.

Statistical analysis of the data produced the results reported in Charts 2, 3 and 4. Chart 2 reports the CI mean values of each group and displays their relative differences. The CI mean for the PLNU students was 1.85 times larger than that of the students at Douglas College and 2.5 times larger than it was for the individuals in the CRASH program. The ranking of the means was in the expected order, affirming the Investigator's hypothesis and providing compelling initial evidence of the instrument's *extrinsic validity*.



Note: The Standard Errors for each of the groups was: 26.2; 38.6, and 33.6, respectively. The difference between the mean values for the PLNU and Douglas groups was significant at $p < 0.000$. The difference between the means for the Douglas and CRASH groups was significant at $p = 0.020$.

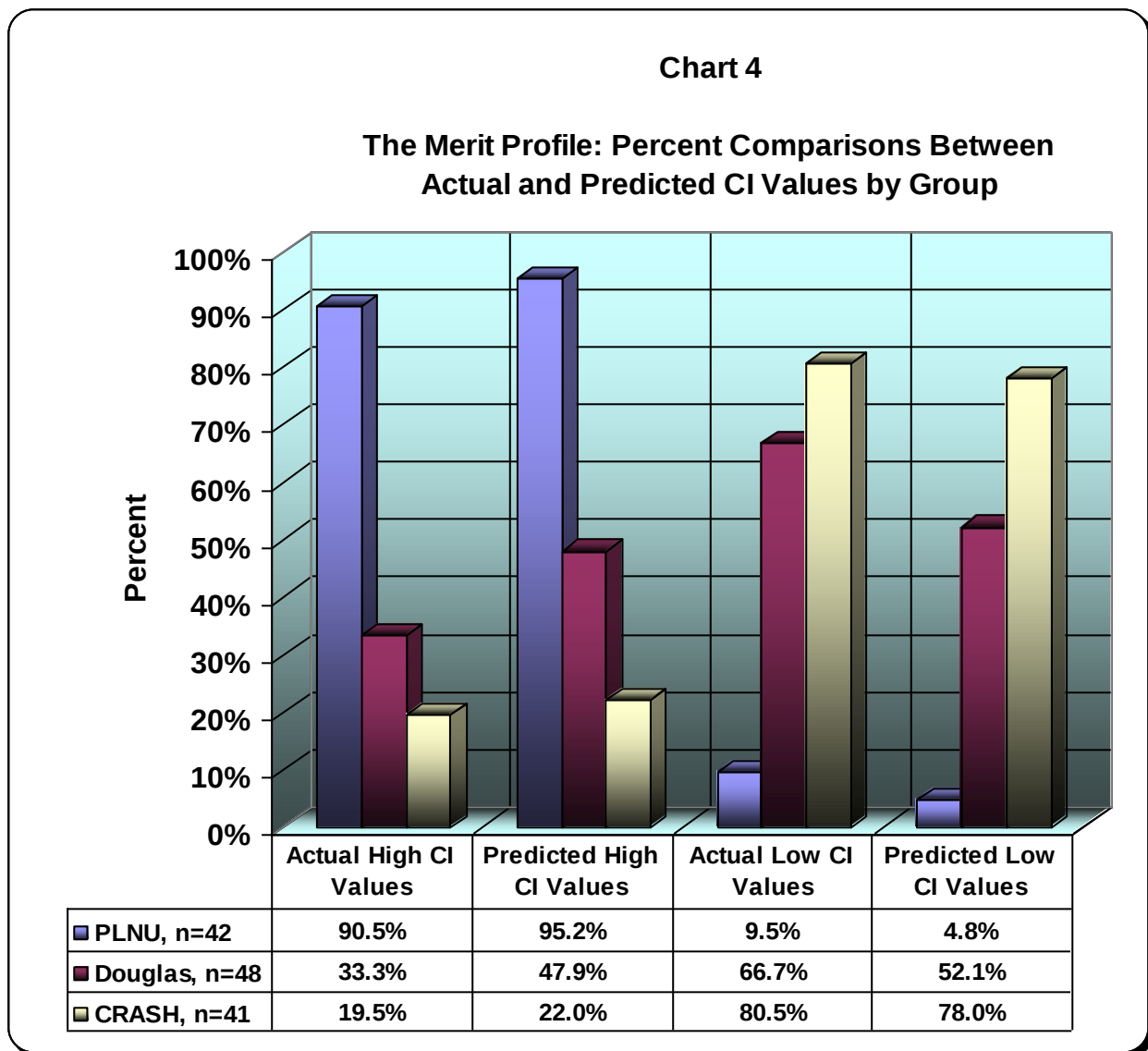
Chart 3 displays the percentages of survey respondents in the three groups whose CI values were above or below the normative mean value of 500. While the distributions were very different for the three groups, it is interesting to note that values for some portion of each group fell on both sides of the mean value (500). Thus, nearly 20 percent of the CRASH participants had above average CI values, and the scores for almost 10 percent of the PLNU students fell below that threshold.



Those findings suggested that a *generic benchmark* would identify “recruits” for job openings with acceptable competencies of character just as effectively as *customized benchmarks* now identify “success” on various criteria in the selection process. To investigate that possibility, a benchmark was developed that was designed to answer the key question: Are benchmarks effective predictors of character, as defined by the MAXIMIZERS™ and represented in the Character Index?

The procedure employed the following steps: Control and Experimental Groups were formed in the same manner as was used in the previous study. Similarly, a cut point on the Character Index was made to create High and Low Subgroups. A statistical analysis performed exclusively on the Control Group produced the equation for the benchmark that was also applied to members of the Experimental Group to simulate the selection process just as it occurs in typical applications by other benchmarks.

Chart 4 adds the benchmark predictions to the same data that were reported in Chart 3, such that, the Actual High CI Values and the Predicted High CI Values are compared by Group. The benchmark misclassified 2 PLNU students, 7 Douglas students, and 1 individual in the CRASH program for an overall predictive accuracy of 86.1 percent. The results provided a highly significant positive answer to the question, indicating that the benchmark in this test indeed was an effective predictor of character as represented in the Character Index.



Discussion

Current Development

Historically, character has been a difficult construct to accurately assess. The procedures used in the present studies represent another attempt to do that. While the results in this investigation are strong indicators of success in that regard, additional studies will need to be conducted to corroborate them.

To the extent that future studies show results similar in significance to those reported here, and there are many reasons for thinking that such goals are attainable, a new era in the assessment of character and behavior, together, will have begun, perpetuated by many valuable benefits that they produce.

Despite the work that remains to be done, it is safe to say that the objectives of this investigation were clearly met. Indeed, results provided realistic expectations that a sixty-item instrument that requires only a few minutes to complete has the ability to provide information that is highly significant, statistically, and highly profitable in many practical applications by end users.

After the removal of invalid Surveys, every test of the instrument's intrinsic and extrinsic validity was highly effective. The most notable findings were the accuracies of blind classifications of individuals in the Experimental Groups, because they provide a preview of what can be expected in actual marketplace applications.

As expected, results from the present investigation revealed the need for several key changes. Those innovations have been programmed into Version 2.0 and that Survey is currently being administered through the Website at www.acheckamerica.com.

Innovations

The changes that upgraded the MERIT Profile Survey, Version 2.0, included the following:

1. *Adding options*--increasing the number of Likert scale options from five to seven to reduce the frequency of extreme option selections;
2. *Reverse scoring* to decrease response set;
3. Invalidating surveys with long strings of *single option* responses;
4. Invalidating surveys by *exaggerating discrepancies* between scores within the same factor;
5. Using interviews to *challenge respondents* whose responses were exceptionally high to reduce false-positive errors in decisions;
6. Using interviews to *challenge respondents* whose responses were exceptionally low to reduce false-negative errors in decisions.

The Investigator understands that it would be presumptuous and inappropriate to impugn the character of a respondent based exclusively on his/her responses to the Survey. But, information produced by the two studies also revealed that it would be inappropriate to

©Copyright 2008, Precision Human Development, Ltd. All rights reserved.

refer job applicants for position openings whose character was masked by invalid responses to the Survey.

In fact, a practical solution to this dilemma has been devised and implemented. The scores earned on the Survey produce a Status Indicator, commonly referred to as "SI," that helps management make an unbiased decision about the disposition of each respondent. A description of the SI is in order at this point.

Status Indicator

Every respondent is automatically given a Status Indicator classification upon the completion, submission, and scoring of the Survey. The SI is represented by four categories, defined as follows:

- ❖ SI: 1 means: All screening tests for validity were passed and the respondent's Character Index fell between 159 and 840
- ❖ SI: 2 means: All screening tests for validity were passed and the respondent's Character Index fell above 840. These individuals either have impeccable character, or they altered their true responses in an attempt to give themselves an advantage. Respondents with this SI classification qualify for a face-to-face interview by a recruiter where additional questions may be asked before a final decision and recommendation is made.

The strategy to require an interview for these respondents was designed to reduce false-positive errors in decisions. Following the interview, individuals who do not qualify for recruitment referral automatically qualify for coaching services and the opportunity to improve in specific areas of needs development.

- ❖ SI: 3 means: All screening tests for validity were passed and the respondent's Character Index fell below 159. These individuals either truly have poor character by their own admission, or they are naturally modest and self-critical. Respondents with this SI classification also qualify for a face-to-face interview with a recruiter where additional questions may be asked.

This procedure was designed to reduce false-negative errors in decisions. Individual's who do not qualify for recruitment referral automatically qualify for coaching services and the opportunity to improve in specific areas of needs development.

- ❖ SI: 4 means: One or more screening tests for validity did not meet acceptable thresholds regardless of the respondent's Character Index. These individuals will not be recommended as recruitment referrals, and, typically, no further follow-up or action will be taken on their behalf.

The behavioral component of the MERIT Profile Survey adds a second important dimension of understanding of each individual. And, behavior and character together provide a unique combination of descriptive material as well as metrics that are the basic ingredients for benchmark development. Together the *package* provides a powerful tool for ensuring that the best available recruits are recommended for selection for position openings.

Strategic Vision

It is an unavoidable fact that the retirement of the Baby Boomers will soon be felt in the workplace due to the diminishing availability of qualified job applicants. Thus, it is a step in the right direction to try to salvage every person whose character flaws hinder their employability. That was the vision behind coaching services that are offered to individuals who have *development needs* and who want to improve them.

As the new era approaches, it introduces new challenges and opportunities. For example, with proper publication, it is conceivable that the CI could become information that individuals in general would seek to know about themselves to the same degree that they now seek to know their IQ (intelligence quotient) or EQ (emotional quotient). In those situations, the CI could become a catalyst for change—character is an expression of the heart and, as such, it is more susceptible to change than are personality, cognitive abilities and emotional intelligence. Thus, to the extent that society is made aware of the need for character improvement as a prerequisite for employment, and to the extent that services that affect those improvements are readily available, it is reasonable to think that usage would increase exponentially.

Without question, the long-term vision for the MERIT Profile Survey is to continue its refinement and to enlarge its application. The need for the Instrument is timely and irrefutable. The scope of its application is virtually unlimited. ***The metrics the Survey makes available, the benchmarks produced by them, and the use of scientific methods in controlled studies constitute a viable and respected strategy for confirming the validity of the instrument and the value of the procedure.***

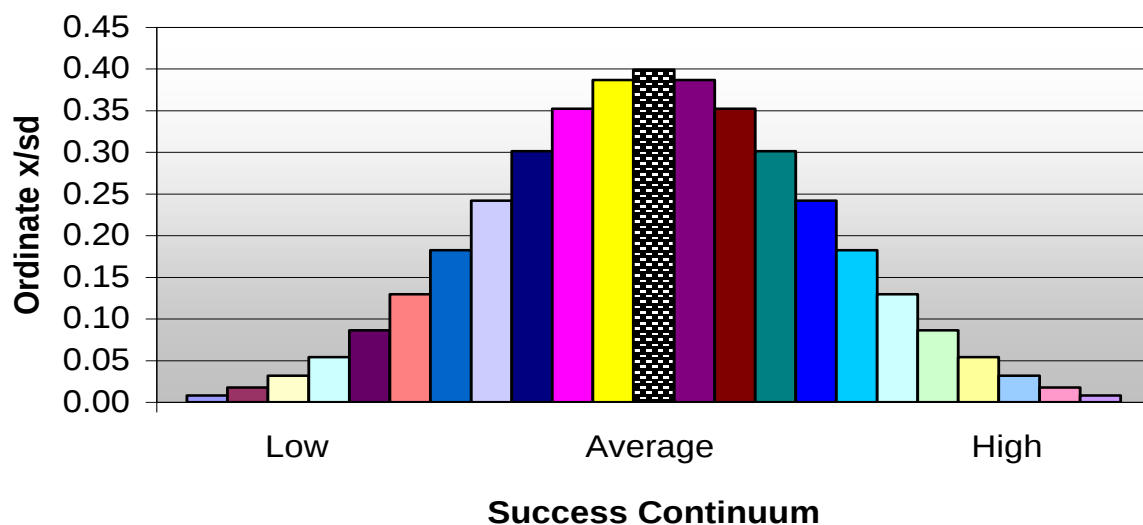
Appendix A

Benchmark Development

Many objective measures of success can be distributed along a continuum that forms a normal distribution curve, as displayed in Chart 5. One side of the curve will contain values that are preferred over their counterparts on the opposite side of the curve.

In general, individuals who are most successful one year will continue to perform at better than average levels on repeated evaluations, and the opposite is true for the less successful, as well. The consistency of performance over time indicates that "success," however it is defined by any objectively measured criterion, is not accidental or random. Indeed: ***Success is caused by attributes that are possessed by certain members of the group and possessed in lesser amounts, or not at all, by other members of the group.***

Chart 5
Normal Distribution Curve



When true causes of success can be identified, the predictive co-relationships between particular attributes and important outcomes can be reduced to mathematical equations. Those equations, in turn, can be valuable tools in the selection process in the workplace. Thus, if attributes "A," "B" and "C" are found to cause *above average* success and the absence of those attributes to cause *below average* success, the equation would be able to identify job candidates who are most likely to succeed **before** they are hired.

When selection continues to be based on the organization's historically-used procedures, the distribution will remain the same, such that, only 50 percent of the selections, by definition, will possess the attributes on the preferred side of the median value on the success continuum. When benchmarks are applied appropriately, the percentage of preferred selections will increase significantly from 50 percent to 70 percent or higher.

Appendix B

Significant Studies

The following references highlight the Investigator's research experiences and are representative of his 70 publications, most of which are in peer-review, scientific journals.

Google keywords and links to abstracts of the original documents and annotations that summarize the most salient features and findings are associated with the selected references that follow:

1. Williamson, M.L. and K. Hopkins. The Use of "None of These" Versus Homogenous Alternatives on Multiple-choice Tests. Journal of Educational Measurement, 4:2, Summer 1967.

Google keywords: None of these, Item Difficulty, Williamson, M.L. (~315,000 citations).

Note: This publication reports the Investigator's research for the Master's Degree that was awarded "Thesis of the Year" by the University of Southern California Alumni Association. The study identified four commonly used standardized achievement tests as being too difficult to be useful for below average students.

2. Williamson, M.L., R. Koch and R.A. Henderson. Phenylketonuria in School Age Retarded Children. American Journal of Mental Deficiency, 72:5, March 1968.

Google keywords: Phenylketonuria in School Age Retarded Children Williamson

Note: This was a Statewide project, directed by the Investigator, that screened over 30,000 children who were high-risk for phenylketonuria, PKU. The project was instrumental in passage of a State law for the routine screening of newborns for genetic disorders. That law has been in effect since January 1966 and has provided screening for 99 percent of all live births in California since it became effective.

3. Williamson, M.L., J. Dobson and R. Koch. Collaborative Study of Children Treated for Phenylketonuria: Study Design. Pediatrics, 60:6, pp. 822-827, December 1977.

Google keywords: Collaborative Study Phenylketonuria Williamson

Note: During his tenure as an Associate Clinical Professor at the University of Southern California, School of Medicine, the Investigator was the principal designer of the National Collaborative Study, cited above, then served as Chief Biostatistician and Co-Investigator to the Study for 18 years. The Study was conducted in 16 university treatment centers in 15 states. Seven professional disciplines were involved at each treatment center.

4. Acosta, P., E. Wenz and M.L. Williamson. Nutrient Intake of Treated PKU Infants. The Journal of Clinical Nutrition, Vol. 30, pp. 198-208, February 1977.

Google keywords: Acosta Wenz Williamson

Note: This article reports the results of a study, designed by the Investigator, which showed a positive correlation between amounts of aspartame ingested and blood phenylalanine levels in PKU children. The findings resulted in placement of a warning label on all products containing aspartame, including, NutraSweet and soft drinks.

5. Bessman, S., M.L. Williamson and R. Koch. Diet, Genetics and Mental Retardation Interaction Between Phenylketonuric Heterozygous Mother and Fetus to Produce Non-specific Diminution of IQ: Evidence in Support of the Justification Hypothesis. Proceedings of the National Academy of Sciences, USA, 75:3, pp. 1562-1566, March 1978.

Google keywords: Bessman Williamson Koch

Note: This study tested the Justification Hypothesis for evidence that diminished availability of non-essential amino acids during the gestational period could be the cause of 85 percent of familial-type mental retardation. The study found that in PKU families the mean IQ was 120 for homozygous normal siblings and 100 for their heterozygous siblings. The implications are that familial-type mental retardation, caused by genetic deficiencies in amino acid metabolism, could be eliminated by amino acid supplements to the diet of the mother during pregnancy.

6. Williamson, M.L., R. Koch, C. Azen and E.G. Friedman. Diet Termination: Matched Pair Control Between PKU Children and Their Non-PKU Siblings. In H. Bickel and U. Wachtel (Eds.), Inherited Diseases of Amino Acid Metabolism: Recent Progress in the Understanding, Recognition, and Management, International Symposium in Heidelberg, 1984. Stuttgart, New York: Georg Thieme Verlag, Inc. c. 1985.

Google keywords: Williamson, M.L., Diet Termination

Note: At the end of the 20 year study, the authors were able to conclude: "Proper dietary management is completely effective in protecting the normal development of the brain in children with phenylketonuria." The benefits of treatment resulted in an average IQ of 100 for the treated PKU children, comparable to their non-PKU siblings, versus an average IQ of 30 estimated for their untreated PKU siblings, a savings of 70 IQ points per child. The Study set the standards for treatment of PKU, worldwide.

White Papers

1. Williamson, M.L. and R. Larson. Final Report: Customer Satisfaction Survey, Smith Engineering and Environmental Corp., Ontario, California, © September 1995, pp. 1-140.
2. Williamson, M.L. Effects of Personality on the Performance of Truck Drivers, Precision Human Development, Ltd., San Diego, California, pp. 1-6, © 1998.
3. Williamson, M.L. The Effects of Behavioral Attributes on the Sale of Automobiles, Precision Human Development, Ltd., San Diego, California, pp. 1-22, © 2001.
4. Williamson, M.L. and J. Clark. The Use of Benchmarks to Reduce Early Terminations, Precision Human Development, Ltd., San Diego, California, pp. 1-26, © 2003.
5. Williamson, M.L. Attributes Versus Traits as Effective Predictors of Success in the Workplace. Precision Human Development, Ltd., San Diego, California, pp. 1-6, © 2004.

6. Williamson, M.L. and J. Clark. The Use of Benchmarks to Reduce Early Terminations: Follow-up Report, Precision Human Development, Ltd., San Diego, California, pp. 1-6, © 2005.
7. Williamson, M.L. and J. Clark. Study of the Effectiveness of Benchmarks on Early Terminations at H.Y. Louie Co., Marketplace IGA Stores. Precision Human Development, Ltd., San Diego, California, pp. 1-13. © 2006.
8. Williamson, M.L. The Merit Profile: Studies that Evaluate the Measurement of Character and Behavior. Precision Human Development, Ltd., San Diego, California, pp. 1-34, © 2008.

The Investigator's entire Curriculum Vitae are contained in Appendix C.