

## Technical Report

### Global Diversity Survey®

This report contains information about the development of the Global Diversity Survey® and its psychometric properties and about the samples of respondents whose input aided in development.

### Development

Development proceeded in 3 stages.

**Stage 1 – item generation:** Authors used their expertise in global diversity to generate a pool of 72 statements (items), which were edited for clarity of meaning and critiqued for social desirability.

**Stage 2 – alpha test:** Edited items were placed online for access by individuals throughout the world, and responses were obtained anonymously. International organizations were contacted by authors to identify respondents and encourage them to log on to the survey site. A total of 105 surveys were completed in this manner, and the demographic characteristics of this sample are identified in Table 1.

**Table 1: Demographic characteristics of the alpha sample (N=106)**

<b>Gender:</b> (N=93)	male	28 (30%)	<b>Profession:</b> (N=92)	clerical/admin.	3 ( 3%)
	female	65 (70%)		supervisor/manager	8 ( 9%)
<b>Age:</b> (N=105)	<20	0	executive	12 (13%)	
	20-29	9 ( 9%)	skilled technician	0	
	30-39	17 (16%)	skilled laborer	1 ( 1%)	
	40-49	19 (18%)	artist/musician/author	1 ( 1%)	
	50-59	19 (18%)	self-employed	15 (16%)	
	60+	4 (4%)	business services	7 ( 8%)	
	no response	37 (35%)	health care	3 ( 3%)	
<b>Heritage:</b> (N=89)	African	6 (7%)	education/teacher	8 ( 9%)	
	Asian	4 (4.5%)	financial services	18 (20%)	
	European	61 (69%)	government service	2 ( 2%)	
	Hispanic/Latino	4 (4.5%)	customer service/retail	0	
	Middle Eastern	1 (1%)	shipping/transportation	0	
	Native Peoples	2 (2%)	not-for-profit	5 ( 5%)	
	Mixed	6 (7%)	other	9 (10%)	
	Other	5 (6%)	<b>Education:</b> (N=91)	primary	1 ( 1%)
		vocational/technical		4 ( 4%)	
		associate		2 ( 2%)	
		university (bacc.)		33 (36%)	
		advanced degree		31 (34%)	
		doctorate	15 (16.5%)		
		other	5 (5.5%)		

**Stage 3 – beta test:** Based on a statistical analysis of alpha test responses, 45 items were selected for further examination in a new sample. This group was a convenience sample acquired during diversity training with primarily human resource professionals in Seattle, Washington USA. The demographic characteristics of this group are summarized in Table 2.

**Table 2: Demographic characteristics of the beta sample (N=103)**

<b>Gender:</b> (N=101)	male	28 (28%)	<b>Profession:</b> (N=100)	clerical/admin.	6 ( 6%)
	female	73 (72%)		supervisor/manager	26 (26%)
				executive	5 ( 5%)
<b>Age:</b> (N=101)	<20	0		skilled technician	2 ( 2%)
	20-29	12 (12%)		artist/musician/author	1 ( 1%)
	30-39	38 (38%)		self-employed	2 ( 2%)
	40-49	35 (35%)		personnel/HR	40 (40%)
	50-59	14 (14%)		health care	0
	60+	2 ( 2%)		education/teacher	4 ( 4%)
	no response	2		financial services	0
				government service	0 ( 0%)
				customer service/retail	2 ( 2%)
				shipping/transportation	2 ( 2%)
				other	10 (10%)
<b>Heritage:</b> (N=103)	<b>First Language:</b>		<b>Education:</b> (N=99)	high school	7 ( 7%)
	English	86 (83%)		vocational/technical	4 ( 4%)
	Spanish	3 ( 3%)		associate	10 (10%)
	Vietnamese	3 ( 3%)		university (bacc.)	46 (45%)
	Romanian	1 ( 1%)		advanced degree	29 (29%)
	Tagalog	1 ( 1%)		doctorate	1 ( 1%)
	Turkish	1 ( 1%)		other	2 ( 2%)
	<b>Second Language:</b>				
	Spanish	6 ( 6%)			
	English	6 ( 6%)			
	French, English	2 ( 2%)			
	French, Spanish	3 ( 3%)			
	German	2 ( 2%)			
	Tagalog	1 ( 1%)			
	Hindi	1 ( 1%)			
	Japanese	1 ( 1%)			
	Korean	1 ( 1%)			
	Mandarin	1 ( 1%)			
	Russian	1 ( 1%)			
	None	77 (75%)			

A comparison of alpha and beta samples indicates the first (alpha) sample was more diverse, better educated, and more likely to be self employed or in finance than the second (beta) sample. The beta

sample was more often employed at a higher level; and while it represented some diversity of family background, it is a nonrepresentative sample limited by U.S. residence and a narrower range of employment.

### **Statistical Analysis**

This survey is designed primarily as a self-report tool, which means the following psychometric standards apply:

- Individual dimensions demonstrate reliabilities of at least .70 on a scale of  $-1.00$  to  $+1.00$ . This criterion establishes that items summed to produce a score are measuring the same thing.
- Intercorrelations between dimensions are lower than their respective reliability coefficients. This criterion establishes that statements attributed to a particular dimension actually measure that concept more closely than other concepts in the survey.
- Considerable independence among dimensions is not expected, since items in an instrument designed to measure one domain of interest – namely knowledge and attitudes about diversity – will be significantly correlated. The whole instrument can be regarded as a scale; individual dimensions, which are scored independently, are subscales.
- Given that reliability is expected to increase with more items in a subscale, the reliabilities for the 15-item dimensions of Self, Other, World and of Insight, Inclusion, and Adaptation are expected to be higher than the reliability coefficients for 5-item clusters at their intersection (Self-Insight, Self-Inclusion, Other-Inclusion, etc.)

### **Alpha Test Results (N=106)**

A set of 72 items was administered to a diverse sample of 106 respondents (alpha test). Analysis of this set of items was used to reduce the set of statements to 45 items that best met the criteria listed above.

Given the size of the sample, traditional factor analysis was not employed. Instead, items were subjected to several forms of nonlinear multidimensional analysis (hierarchical and nonhierarchical cluster analysis, multidimensional scaling) to determine which items grouped together. Internal consistency reliabilities were run on emergent item clusters. Item-scale correlations were used to determine which items contributed most to subscale reliability.

Inter-scale correlations were calculated among selected item clusters. When items are assigned to scales/subscales, the goal is to ensure each one is a better measure of the scale to which it is assigned than to another scale in the instrument. Thus, in a few instances, an item that was strongly associated with two or more subscales was replaced by one that was more uniquely associated with one subscale, even if it contributed slightly less to that subscale's internal consistency.

Using the 45 statements were selected by this approach, the following statistical results were achieved:

**Score Distribution:** Summed scores for the 6 major dimensions of *Self*, *Other*, *World* and *Insight*, *Inclusion*, and *Adaptation* were distributed approximately normally with the exception of *Self*, which demonstrated a modest amount of negative skew.

**Reliability:** Table 3 shows the internal consistency reliabilities for clusters of 5- and 15-item scales/subscales of the Global Diversity Survey, based on the alpha sample.

**Table 3: Internal consistency reliabilities (alpha) within cells and across dimensions (N=106)**

Dimension:	Insight	Inclusion	Adaptation	Overall Reliability (15 items)
<b>Self</b>	$r_{xx'} = .72$	.62	.65	.77
<b>Other</b>	.66	.75	.71	.82
<b>World</b>	.62	.71	.74	.82
<b>Overall Reliability (15 items)</b>	.76	.82	.86	

Fifteen-item dimensional scales clearly meet the reliability criterion of  $r_{xx'} = \geq .70$ . Five of 8 5-item subscales also meet the same criterion; the remaining subscales exceed .70 when a Spearman-Brown formula is applied. For example, a 5-item scale with an internal consistency coefficient of .62 has an estimated reliability of .77 with 10 similar items.

**Internal Validity:** The correlations between dimensions are shown in Table 4. Justification for using separate scales is found when inter-scale correlations are significantly lower than the internal reliability coefficients shown in the margins of Table 3 (Overall Reliability). In this sample, all dimensions except *Self* appear to offer unique contributions to understanding a person's response to diversity. *Self* ratings do not appear to offer a unique perspective apart from *Other*.

**Table 4: Inter-scale correlations among dimensions (N=106)**

Dimension:	Self	Other
<b>Other</b>	$r_{xy} = .78$	
<b>World</b>	.62	.71

Dimension:	Insight	Inclusion
<b>Inclusion</b>	$r_{xy} = .61$	
<b>Adaptation</b>	.65	.61

Correlations among 5-point subscales vary with the dimension. For example, coefficients among the 3 subscales measuring Insight (Self, Other, World) correlate .29 to .40; subscales measuring Inclusion correlate .34 to .49; and subscales measuring Adaptation correlate .53 to .66. Generally, even at the 5-item level, the unique contribution of each subscale is validated. This finding lends confidence to the strategy of examining each subcategory separately.

## Beta Test Results (N=103)

The set of 45 items extracted from alpha analysis was designed to maximize the reliability and validity of the set. Unavoidably, this process also fits the data to the sample used, by representing some of the unique characteristics of the sample. When the same set is analyzed using a different sample, the statistical properties of the first analysis lose some of their strength. Different and unique characteristics of the second sample replace unique characteristics of the first, producing a somewhat different set of responses and relationships.

**Score Distribution:** Summed scores for the 6 major dimensions of *Self*, *Other*, *World* and *Insight*, *Inclusion*, and *Adaptation* were distributed approximately normally with the exception of *Self*, which demonstrated a slight amount of negative skew.

**Reliability:** Table 5 shows the internal consistency reliabilities for clusters of 5- and 15-item scales/subscales of the Global Diversity Survey, based on the alpha sample.

**Table 5: Internal consistency reliabilities (beta) within cells and across dimensions (N=103)**

Dimension:	Insight	Inclusion	Adaptation	Overall Reliability (15 items)
<b>Self</b>	$r_{xx'} = .64$	.53	.61	.69
<b>Other</b>	.66	.50	.65	.70
<b>World</b>	.62	.37	.67	.61
<b>Overall Reliability (15 items)</b>	.69	.70	.84	Total Survey: .86

Fifteen-item dimensional scales meet or just about meet the reliability criterion of  $r_{xx'} = \geq .70$  with the exception of the dimension *World*, to which the 5 items labeled *World/Inclusion* are not well correlated. None of the 5-item subscales meet the criterion for self-report scales. When a Spearman-Brown formula is applied to offset the limited number of items in these subscales, all but *World/Inclusion* meet the criterion.

**Internal Validity:** The correlations between dimensions are shown in Table 6. Justification for using separate scales is found when inter-scale correlations are significantly lower than the internal reliability coefficients shown in the margins of Table 5 (Overall Reliability). The .10 to .20-point spread between reliability and validity coefficients indicates that all dimensions relate to a common construct, diversity, but each dimension also measures unique properties.

**Table 6: Inter-scale correlations among dimensions (N=103)**

Dimension:	Self	Other
<b>Other</b>	$r_{xy} = .70$	
<b>World</b>	.63	.64

Dimension:	Insight	Inclusion
------------	---------	-----------

<b>Inclusion</b>	$r_{xy} = .46$	
<b>Adaptation</b>	.45	.58

Correlations among 5-point subscales vary with the dimension. For example, coefficients among the 3 subscales measuring Insight (Self, Other, World) correlate -.03 to .32; subscales measuring Inclusion correlate .35 to .59; and subscales measuring Adaptation correlate .53 to .73. Subscales designed to measure Insight and Inclusion appear to make a more unique contribution than subscales designed to measure Adaptation.

Changes in item allocation to subscale were recommended to improve statistical results from the beta sample, but they were not implemented for two reasons: (1) The beta sample was considered a fairly small and non-representative sample on which to base changes in item selection; and (2) the authors were convinced of the content validity of their original item assignment to dimensions. As explained above, any set of development parameters can be expected to drop in value when applied to a new sample. It is assumed that as this survey is applied to more diverse samples, as that represented by the alpha test respondents, stronger reliability and validity will be demonstrated.

## Scoring

Distributions of scores in both samples were examined, and cut-off points were established at the 20 and 80 percentiles. The purpose was to distinguish the level of competency, attitude, and behavior that can be regarded as below, at, and above expectations for functioning effectively in a diverse global environment. Cutoff points were generally within 0-2 points of each other across samples on all dimensions. The middle band of 60% is slightly narrower than +/- 1 standard deviation of a distribution of scores that typically denotes “average,” but it is regarded as most appropriate for the purposes served by the Global Diversity Survey®.

## Summary

The Global Diversity Survey® is a conceptually-based, statistically validated self-report measure that is designed to serve as a guide to self understanding and development in a diverse global world. Because any decisions arising from its interpretation are made by the person responding, the level of psychometric validity is not expected to match the kinds of tests that are used by others to make decisions about a person. The respondent is able to verify using his or her own experience whether interpretation is appropriate to his or situation. On the other hand, there is sufficient psychometric integrity in the dimensions and subscales of this survey to avoid misleading a respondent. And the insights gained from taking the feedback and interpretation seriously, justify viewing the tool as a valid and reliable measure of one’s competence, attitudes, and behavior on the subject of diversity.