

**Report on Drake P3 Validity**

**Jacob Hirsh, Ph.D.**

**May 14, 2018**

## **Executive Summary**

A series of analyses are presented to examine the validity of the Drake P3 questionnaires. Part 1 examines the internal validity of the questionnaire through a series of factor analyses and item reliability analyses using a sample of 82,242 respondents. Factor analyses indicate that the questionnaire has good internal structure and is well-described by the four key trait dimensions. Each item was identified as a valid indicator of its underlying trait domain. The questionnaires also demonstrated good reliability. Part 2 examines the predictive validity of the questionnaire, finding that it is indeed a significant and valid predictor of job performance outcomes. Part 3 examines gender differences in the questionnaire, finding only small to moderate differences between men and women that are consistent with the literature on gender and personality. No differences in the predictive validity of the questionnaire across genders were identified.

## Part 1: Factor Structure of the P3

In order for a questionnaire to be considered valid, it must first be established that all of the questions are valid indicators of their respective trait domains. An exploratory factor analysis was conducted on a global sample of 84,242 test-takers in order to examine the factor structure of the questionnaire. The primary and environmental scales were examined separately.

### Exploratory Factor Analysis of the Four Primary Scales

The results of the primary scale analyses, as presented below, indicate that there are indeed 4 factors underlying the covariance between the questionnaire items. The goodness-of-fit test indicates that a four factor solution provides a good fit to the model:

#### Goodness-of-fit Test

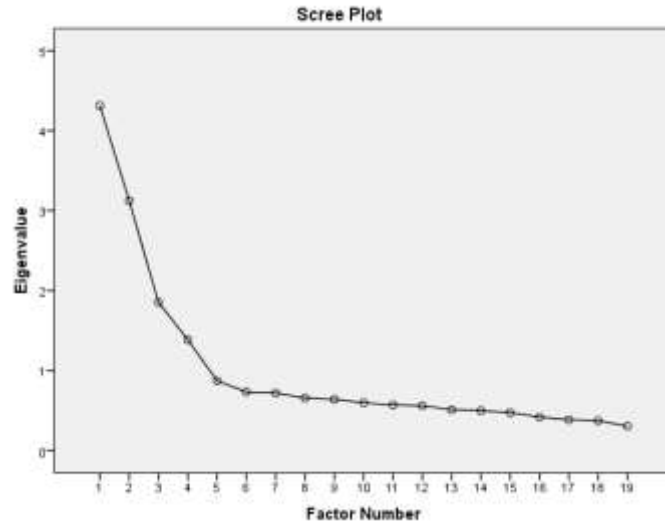
Chi-Square	df	Sig.
23693.950	101	.000

In the table below, the “Eigenvalues” describe the number of items worth of variance that is accounted for by each factor. The first factor is always the largest one, with subsequent factors getting smaller and smaller. Collectively, the first four factors account for 56.2% of the total questionnaire variance. Note that it is not necessary to achieve 100% variance accounted for, because some of the variance will be measurement error.

#### Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	4.313	22.699	22.699	3.785	19.919	19.919	3.299
2	3.124	16.444	39.143	2.597	13.670	33.589	2.759
3	1.854	9.758	48.901	1.291	6.794	40.383	1.448
4	1.385	7.291	56.192	.893	4.702	45.085	2.551
5	.872	4.591	60.783				
6	.733	3.859	64.642				
7	.720	3.788	68.430				
8	.657	3.459	71.889				
9	.641	3.372	75.261				
10	.598	3.148	78.409				
11	.571	3.003	81.412				
12	.560	2.949	84.361				
13	.512	2.696	87.057				
14	.500	2.630	89.686				
15	.475	2.499	92.185				
16	.419	2.206	94.391				
17	.386	2.032	96.424				
18	.374	1.966	98.390				
19	.306	1.610	100.000				

The scree plot below shows the decreasing eigenvalues for subsequent factors. Note that the first four factors are visibly distinguishable from the subsequent ones, which are small in size.



The structure matrix shows the extent to which each of the items is correlated with the 4 key factors. The ideal is to obtain a “simple structure”, where the items assessing each factor load strongly on their own trait dimension, but not on any of the others. A factor loading of 1.0 is the maximum, reflecting a perfect indicator of the trait dimension. A factor loading of -1.0 is the minimum, reflecting a perfect negative indicator of the trait dimension. Factor loadings between 0.4 and -0.4 are replaced with empty cells to facilitate presentation of the results. As can be seen below, each of the items loads on the correct factor, indicating good internal validity.

**Structure Matrix**

	Factor			
	Patience	Dominance	Extraversion	Conformity
S1A.Pat.Kindhearted	.812			
S1A.Pat.Compassionate	.756			
S1A.Pat.Pleasant	.747			
S1A.Pat.Cooperative	.635			
S1A.Pat.Agreeable	.597			
S1A.Pat.Tolerant	.518			
S1A.Dom.Commanding		.760		
S1A.Dom.Controlling		.714		
S1A.Dom.Authoritative		.670		
S1A.Dom.Powerful		.615		
S1A.Dom.Demanding		.584		
S1A.Dom.Aggressive		.490		
S1A.Ext.Shy			.743	
S1A.Ext.Reserved			.618	
S1A.Ext.Introverted			.577	
S1A.Conf.Detailed				-.780
S1A.Conf.Precise				-.750
S1A.Conf.Perfectionistic				-.549
S1A.Conf.Specific				-.489

Extraction Method: Maximum Likelihood. Rotation Method: Oblimin with Kaiser Normalization.

The table below indicates the correlations between the factors from this analysis (ranging from -1.0 to 1.0).

**Factor Correlation Matrix**

Factor	1	2	3	4
1	1.000	.000	-.103	-.394
2	.000	1.000	-.154	-.322
3	-.103	-.154	1.000	.038
4	-.394	-.322	.038	1.000

**Reliability Analysis of the Four Primary Scales:**

The reliability of a scale indicates the extent to which respondents are likely to get the same score every time they complete it. In principle, maximum reliability is 1.0, although this is rarely seen in commonly used scales because there is almost always some degree of measurement error. Reliability analyses were conducted on each of the four primary trait domains. Below is the Cronbach’s Alpha (a common reliability index) for each scale:

Trait Domain	Reliability
Dominance	.80
Patience	.83
Conformity	.72
Extraversion	.67
<b>Average</b>	<b>.76</b>

These analyses demonstrate that the average reliability of the four primary scales is comparable to the best-validated scales in personality psychology.

## Exploratory Factor Analysis of the Four Environmental Pressure Scales

The analyses conducted above were repeated for the four environmental pressure scales. As shown below, a four factor model fit the data well, accounting for 62.8% of the total variance.

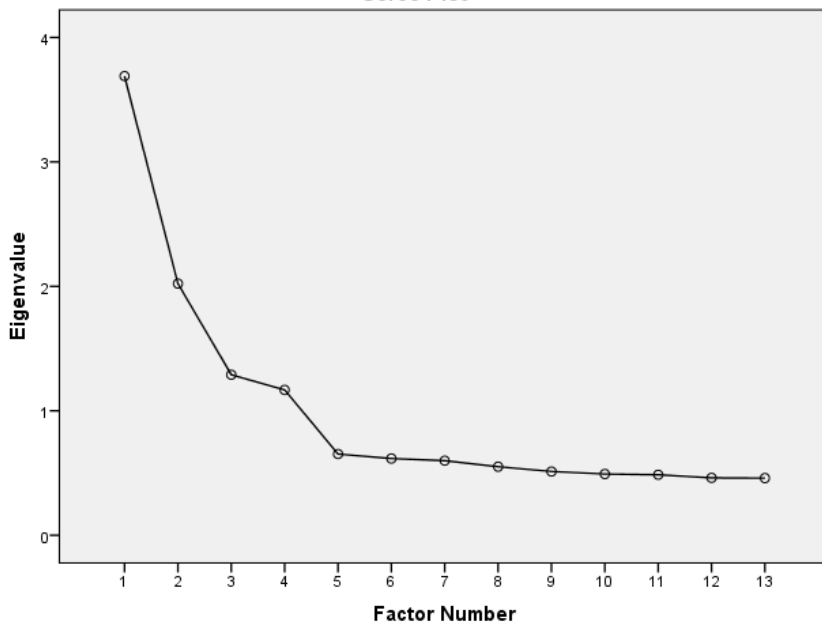
### Goodness-of-fit Test

Chi-Square	df	Sig.
2417.083	32	.000

### Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	3.690	28.381	28.381	3.163	24.327	24.327	2.116
2	2.022	15.554	43.936	1.489	11.457	35.784	1.814
3	1.290	9.921	53.857	.770	5.926	41.709	2.168
4	1.168	8.982	62.839	.634	4.878	46.587	2.104
5	.653	5.020	67.859				
6	.617	4.744	72.603				
7	.599	4.609	77.212				
8	.551	4.237	81.449				
9	.512	3.942	85.391				
10	.492	3.787	89.178				
11	.486	3.741	92.919				
12	.461	3.547	96.466				
13	.459	3.534	100.000				

Scree Plot



As can be seen in the table below, the environmental scales display a simple structure, with no problematic items that load on the wrong factor or fail to load on any factors.

**Structure Matrix**

	Factor			
	Extraversion	Patience	Conformity	Dominance
S1B.Ext.Social	.712			
S1B.Ext.Expressive	.711			
S1B.Ext.Communicative	.605			
S1B.Pat.Patient		.724		
S1B.Pat.Calm		.686		
S1B.Pat.Peaceful		.582		
S1B.Conf.Thorough			.722	
S1B.Conf.Systematic			.682	
S1B.Conf.Accurate			.606	
S1B.Dom.Strong				.679
S1B.Dom.Dominant				.674
S1B.Dom.Bold				.654
S1B.Dom.Competitive				.597

Extraction Method: Maximum Likelihood.

Rotation Method: Oblimin with Kaiser Normalization.

**Factor Correlation Matrix**

Factor	1	2	3	4
1	1.000	.310	.357	.341
2	.310	1.000	.400	.013
3	.357	.400	1.000	.290
4	.341	.013	.290	1.000

Extraction Method: Maximum Likelihood.

Rotation Method: Oblimin with Kaiser Normalization.

**Reliability Analysis of the Four Environmental Pressure Scales:**

Below are the Cronbach's Alpha reliabilities for each of the 4 key environmental pressure scales:

Trait Domain	Reliability
Dominance	.74
Patience	.70
Conformity	.70
Extraversion	.71
<b>Average</b>	<b>.71</b>

### **Additional Reliability Analyses**

The P3 also includes scales measuring conscientiousness, rational decision-making, and intuitive decision-making. Below are the reliability analyses for these additional scales.

<b>Trait Domain</b>	<b>Reliability</b>
Conscientiousness (Primary)	.85
Conscientiousness (Environmental)	.82
Rational (Primary)	.74
Rational (Environmental)	.69
Intuitive (Primary)	.83
Intuitive (Environmental)	.53



## Part 2: Predicting Performance Outcomes

An important component of questionnaire validation is the examination of predictive validity, or the extent to which scores on the questionnaire are able to predict performance outcomes of interest. Predictive validity was examined based on a sample of 228 employees from Drake International, spanning multiple different office locations across the globe. All respondents completed the questions from the P3, as well as a test of mental ability. Performance ratings were provided separately on a scale from 1 to 5.

A regression analysis was conducted, with the four key primary traits, conscientiousness, and mental ability entered into a model predicting employee performance. The results, presented below, indicate that higher performance was significantly predicted by 1) increased mental ability, 2) higher levels of conscientiousness, and 3) lower levels of extraversion. These results emerged simultaneously, indicating that they are independent of one another. The standardized beta weights in the table below are the equivalent to a correlation coefficient, ranging from -1.0 (perfect negative correlation) to 1.0 (perfect positive correlation). The R value in the model summary table is the equivalent of the cumulative correlation between all traits and the performance outcome (with 1.0 being perfect prediction). Although the specific trait domains that predict performance will vary by job context, these results are consistent with the finding that mental ability and conscientiousness usually predict performance outcomes across all job contexts. These results support the predictive validity of the questionnaire.

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.254 <sup>a</sup>	.064	.039	.921

### Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.195	.763		1.566	.119
	Extraversion	-.152	.071	-.150	-2.151	.033
	Conformity	-.053	.107	-.040	-.499	.618
	Patience	-.014	.125	-.009	-.113	.910
	Dominance	.040	.083	.035	.486	.627
	Conscientiousness	.361	.186	.176	1.940	.054
	Mental Abilities	.013	.005	.180	2.708	.007

### Part 3: Analysis of Gender Data

Although not all respondents reported their gender, there was a large enough sample to conduct an analysis of gender differences on the questionnaire. The analysis below compares the average scores of men and women. Below are the descriptive statistics for the standardized questionnaire scores, broken down by gender.

**Gender Differences in Trait Scores**

		N	Mean	Std. Deviation	Std. Error Mean
Extraversion	Women	8225	-1.5829	15.91327	.17547
	Men	7958	-2.1127	15.84599	.17763
Dominance	Women	8225	-4.3062	16.06704	.17716
	Men	7958	1.0545	15.15072	.16984
Conformity	Women	8225	-.6336	15.50990	.17102
	Men	7958	-1.1783	15.20228	.17041
Patience	Women	8225	1.7875	14.18233	.15638
	Men	7958	-3.4422	14.92828	.16734
Conscientiousness	Women	8225	60.0518	6.57330	.07248
	Men	7958	58.8249	7.34601	.08235

The table below indicates that there are significant differences between men and women on each of the primary trait dimensions from the questionnaire. The practical significance of these differences is rather small, however, as the effect sizes are only small to moderate in size. The Cohen's D value in the table below reflects the difference between genders in terms of standard deviations. The observed differences are consistent with existing research on aggregate gender differences in personality characteristics.

	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Cohen's D
Extraversion	2.122	16181	.034	.52979	.24970	.033
Dominance	-21.822	16181	.000	-5.36066	.24566	-.343
Conformity	2.255	16181	.024	.54469	.24151	.035
Patience	22.853	16181	.000	5.22976	.22884	.359
Conscientiousness	11.205	16181	.000	1.22692	.10950	.176

The analysis below examines the extent to which gender moderates the predictive validity of the questionnaire (i.e., whether trait scores are more or less effective at predicting performance outcomes for men or women). This analysis focuses specifically on conscientiousness, given that it emerged as the most important trait predictor of performance outcomes. No significant interaction between gender and conscientiousness was observed, indicating the equivalence of the questionnaire's predictive validity across genders.

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.167	.028	.014	.928

**Coefficients**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.113	.080		38.786	.000
	Consc./Gender Interaction	.030	.142	.018	.214	.831
	Gender	-.213	.139	-.108	-1.530	.127
	Conscientiousness	.090	.074	.100	1.218	.225

The magnitude of the relationship between Conscientiousness and performance also remains identical when examined separately by gender, underscoring the results reported above.

**Model Summary**

Gender.Coded	Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
.00	1	.110	.012	.005	.888
1.00	1	.111	.012	-.002	1.000

**Coefficients**

	Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
			B	Std. Error	Beta		
Women	1	(Constant)	2.367	.599		3.954	.000
		Conscientiousness	.012	.010	.110	1.272	.205
Men	1	(Constant)	1.904	1.053		1.807	.075
		Conscientiousness	.017	.018	.111	.926	.358

## **Summary and Conclusion:**

Part 1 of this report demonstrates that the primary and environmental questionnaires show good internal validity, with each item serving as a valid indicator of its underlying trait domain. Good reliability was also demonstrated for these scales. Part 2 demonstrates that the questionnaire is able to validly predict performance outcomes in an international sample of employees. Part 3 found statistically significant but small gender differences in each of the trait dimensions, as would be expected based on existing research in personality psychology. Importantly, there were no differences in predictive validity across genders.