The Crucible

THE PRIVATE EQUITY
HORSEPOWER ASSESSMENT®

Crucible Insights

Technical Manual

2023

INTRODUCTION

This document provides an overview of the intended purpose and interpretation of The Crucible. This document shall also function as a comprehensive tool for end-users of The Crucible assessment to understand the quality of The Crucible and evidence supporting its use in practice.

According to US legal requirements and standards for testing in the workplace, any documentation must provide clarity about the intended purpose, uses, and interpretations of scores on the assessment as well as provide any empirical evidence to support the intended use.[1]

PURPOSE

The Crucible is a performance enhancement research tool exclusively built for high-stakes environments. The Crucible is comprised of an intellectual component, a work styles analysis, and a self-reported behavioral component. Based on the results, The Crucible offers insights into how a person's profile translates into a private equity environment. To do this, we incorporate individual and team level analyses that can be compared to our established benchmarks or can be configured to a specific fund's formula for success. The Crucible was built by private equity experts upon reams of private equity-specific performance data to ensure that the most critical cognitive and behavioral qualities are considered.

INTENDED AUDIENCE

The Crucible is designed to provide insights for development at the individual and team level in high stakes environments. The tool is intended for use at diligence or mid-hold stages with executives and/or their teams, funds, operations groups, coaches and recruiters. Additionally, the Crucible assesses potential fit, and can be taken by individuals to assess their skill level and identify areas for improvement when entering the space.

KEY TAKEAWAYS

- ➔ The Crucible is a valid predictor of fit in a PE context. Regardless of previous experience, better scores on The Crucible were more likely to occur for individuals with successful PE exits.
- ➔ The Crucible research tool has demonstrable reliability and precision in the measurement of our conceptual model.
- ➔ There were no areas of potential bias or adverse impact identified in the Crucible assessment.
- ➔ Testing will be regularly reviewed and assessed to ensure that any potential bias or adverse impact will continue to be minimized, and that the tool remains valid and precise over time.

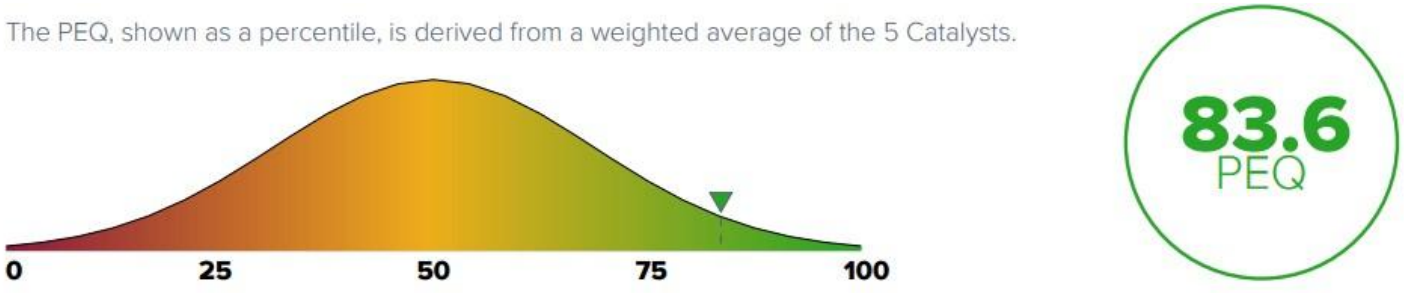
[1] https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf **(pgs. 23-31)**

THE CRUCIBLE MODEL

The Crucible provides information about an individual's standing on a variety of leader characteristics that are vital for success in private equity. The report includes an overall score, a proprietary archetype, and a granular breakdown of behavioral characteristics all delivered on a relative scale set by normative standards in a private equity context.

THE PEQ

The PEQ, or Private Equity Quotient, is an algorithmically generated score derived from a weighted average of the five Catalyst scores using the normative procedures outlined further in this document. The PEQ score is indicative of an individual's overall potential fit in the private equity environment.



The PEQ is shown as a percentile, from 0 to 100, where 0 represents a very poor fit to PE portfolio company leadership roles and 100 represents strong potential fit. As a percentile, the PEQ also provides relative information on where that individual stands compared to other leaders in private equity.

THE ARCHETYPE

The archetype is a complementary, qualitative characterization that provides more information regarding an individual's preferences and specific context for success based on two factors: transformation bias and leadership style. This archetype reflects a person's typical pace of action as well as their approach to communication, influence, and decision-making.

From left to right (columns), an individual's typical pace and rate of change in which they are most likely to succeed increases.

From bottom to top (rows), an individual's typical approach to leadership and decision making is likely to be more outgoing and gravitate toward instinct and speed over analysis and accuracy.

Any one individual archetype is not necessarily more or less desirable than another except in the context of a specific human capital challenge at hand. For this reason, the archetype is recommended to be used to garner information about one's potential formula for success and how they might interact with others in their group and for developmental purposes.

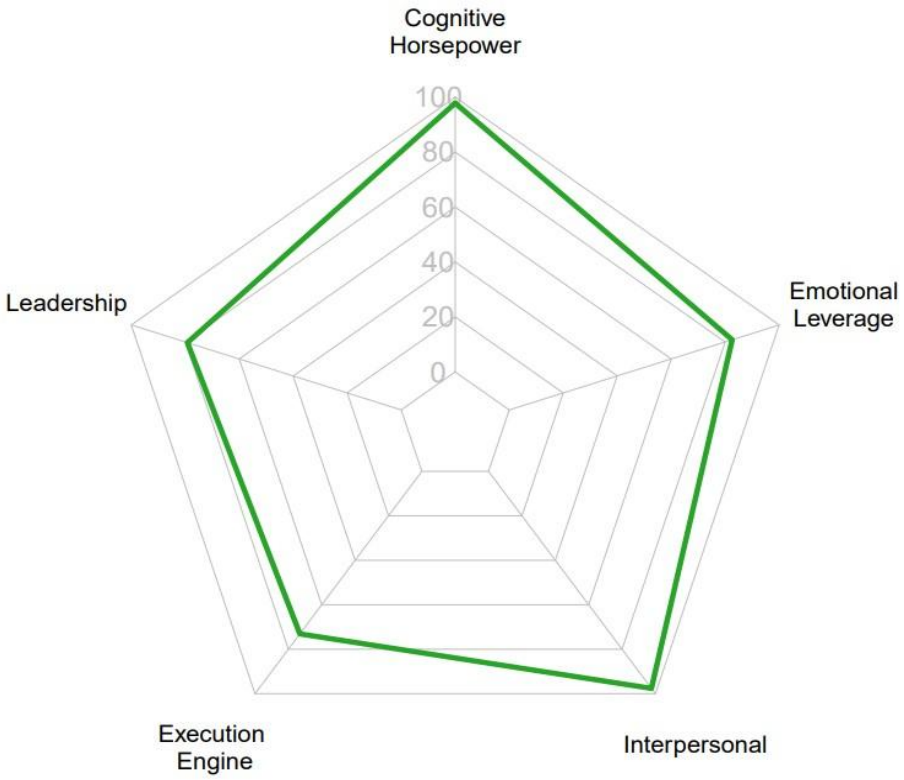
ACTOR	DIRECTOR	ACCELERATOR	DRIVER
Charismatic ACTOR	*Charismatic* DIRECTOR	*Charismatic* ACCELERATOR	*Charismatic* DRIVER
Expressive ACTOR	*Expressive* DIRECTOR	*Expressive* ACCELERATOR	*Expressive* DRIVER
Hybrid ACTOR	*Hybrid* DIRECTOR	*Hybrid* ACCELERATOR	*Hybrid* DRIVER
Deliberative ACTOR	*Deliberative* DIRECTOR	*Deliberative* ACCELERATOR	*Deliberative* DRIVER
Introspective ACTOR	*Introspective* DIRECTOR	*Introspective* ACCELERATOR	*Introspective* DRIVER

THE FIVE CATALYSTS

The Catalyst scores are derived from various Element level scores and reflect broader, and more complex categories of behavior and competencies that can decompose into smaller parts. Catalyst scores are a weighted algorithm, combining various elements which are then converted into percentile scores (from 0 to 100) using the norming procedures outlined further below.

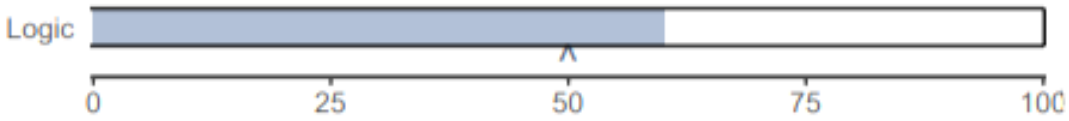
Each Catalyst score indicates the individual's strength on that Catalyst across the spectrum of behavioral performance. The table in the following section provides the conceptual framework and definitions for each Catalyst and the component Elements that make up each Catalyst score.

The five Catalysts include Execution Engine, Emotional Leverage, Leadership, Cognitive Horsepower, and Interpersonal. They are displayed together to provide a comprehensive understanding of an individual's strengths and areas for improvement, and are then individually analyzed further.



THE ELEMENTS

Element scores are derived from responses to cognitive or behavioral questions on The Crucible assessment. Element scores are converted into percentile scores (from 0 to 100). For each Element, higher percentile scores are generally positive, indicating a strength relative to other leaders in private equity. In some cases, marked in the table below, Element scores above a certain threshold represent a behavior that may not be ideal at the extreme end of the scale.



Each element is displayed on a percentile bar from 0 to 100 with a marker at the 50th percentile, also called the median. Individuals who score higher than 50 on an Element thus score higher than 50% of leaders assessed by The Crucible.

TABLE 1. The Crucible Model - Catalysts and Elements

Catalyst	Element	Element Definition
Execution Engine **The ability to act and respond quickly and accurately.**	Urgency	A powerful sense of tenacity and determination that aids in swift execution.
	Relentlessness	An ability to push forward even in cases of setbacks or obstacles.
	Impatience	A lack of tolerance for slow decision-making or slow action.
	Hands-on Approach	The desire to be directly involved in decisions about the way things should be done and in the development of others.
	Digs for Root Cause	The desire to search beneath the surface for a potential source of the underlying problem or issue.
	Detail Orientation	A strong focus on detail analysis and the capability to discern minute errors before they escalate.
	Systems & Processes Focused	An emphasis on following set systems & processes to instill discipline and repeatability.
	Metrics Focused	An orientation towards the use of key figures and indicators to leverage accountability.
	Intellectual Curiosity	The desire to learn more about the world, think abstractly, and find answers to deeper questions.
Emotional Leverage **A leader's ability to contagiously emote to drive results and change in order to positively impact an organization.**	Passion	The powerful sense of wanting to accomplish, achieve, and drive change.
	Intensity	Immense energy and concentration put into solving a problem or achieving a goal.
	Hunger/Drive	The drive to surpass "good enough" by always striving for success.

	Resiliency	The capacity to persevere through and recover from difficulties.
	Energy/Endurance	The capability to perform at a high level in a consistently intense, high-pressure environment.
	Disdain for Losing	A regard for losing as inferior or unworthy and the drive to avoid losing.
	Competitiveness	A strong orientation towards winning.
Leadership **The ability to drive MOIC through levers of psychology and performance.**	Sets High Standards	Consistently demands a high level of quality or achievement.
	Optimism	Hopefulness about the outcome of the future or situation.
	Inspires Confidence & Belief	The ability to instill values of ambition and dedication.
	Directness	Communicating authentically, constructively, and with candor.
	Strategic Perspective	Capable and driven to identify the long term or overall aims and the means of achieving them.
	Ensures Buy-in	Develops acceptance and mission affiliation among necessary stakeholders.
	Engages Deeply with Stakeholders	Builds deep and meaningful relationships and trust with people of interest.
	Drives Accountability	Instills a sense of personal responsibility and respect around achievement.
Cognitive Horsepower **The ability to problem solve and achieve higher levels of understanding.**	Logic	The ability to think rationally and make deductions.
	Math	Unrefined computational skills.
	Word Association	The ability to form semantic relationships between words.
	Pattern Recognition	The ability to identify nuances and make predictions given a series of events.
Interpersonal **The balance between emotional contagion and deeply rooted self-esteem to propel personal and team performance.**	Self-esteem	Uninhibited by self-doubt.
	Belief in Others	Firmly held confidence and trust in the abilities of others.
	Empathy	The ability to understand and share the feelings of individuals or feelings of the organization.
	Confidence	The belief in one's ability to succeed and overcome challenges.
	Authenticity	The state of being truthful to one's sense of self, personal values, and convictions and appearing truthful to others.
	EQ	The ability to recognize and understand one's emotions as well as the emotions of others.
	Collegiality	The quality of being able to engage with the hearts and minds of an organization.

THE CONTAMINANTS

In addition to providing an overall fit index and areas of strength for an individual, the Crucible provides Contaminant scores, six behavioral attributes that most likely limit an individual's upside or contribute to derailment. These scores are also converted into percentile scores (from 0 to 100).

Table 2. The Six Contaminants and Definitions

Contaminant	Contaminant Definition
Hesitancy	Lacking a bias for action or level of proactivity.
Lack of Awareness	Out of touch with the subtleties of team dynamics and perceptions and lacking self awareness.
Fragility	Lacking durability and resilience.
Fear of Failure	An elevated aversion to making mistakes.
Inability to Adapt	The lack of flexibility or the ability to adjust to changes.
Arrogance	Exaggerated levels of confidence or sense of importance.

The six Contaminants are recommended to be used in helping to identify developmental opportunities that can inform managerial tactics.

The Contaminant score is displayed underneath an orange arrow. The hashed area of the bar indicates the zone of elevated potential risk. When a score falls in this risk area, it is highlighted by an orange box and extra caution should be taken. Risk increases exponentially when an individual scores high on more than one Contaminant.

VALIDATION

Validation is the process that researchers use to assess the quality of a research tool by testing it against a variety of rigorous standards established by the scientific community and by ensuring the tool meets legal requirements. This process includes gathering a variety of evidence to support the use of the assessment.

The Crucible is a research tool used to measure abstract concepts like Urgency, Passion, and Authenticity. To test the quality of the tool, we investigate reliability, validity, and fairness. Validity is the idea that the assessment measures the intended concepts and appropriately fulfills its purpose. Simply put, if The Crucible is indeed able to measure an individual's potential fit in a private equity environment, it should be demonstrated that the assessment predicts meaningful outcomes in private equity and/or that the concepts are measured as precisely as possible.

The combination of cognitive, motivational, work styles, and behavioral components in assessments has been a valid and strong driver of leader performance based on 100 years of empirical research.[2] Particularly when combining measures of mental ability with integrity and conscientiousness (the tendency to be diligent, efficient, and self-driven). By leveraging these components the Crucible is at an advantage for demonstrating its effectiveness in predicting performance in private equity.

While all psychometric tools are subject to error and therefore cannot be perfect, an assessment is considered valid to the extent that data-driven evidence supports the intended use and interpretation of assessment scores. In this section we provide a variety of evidence to support the intended uses and score interpretations of The Crucible across various groups of leaders.

[2] Schmitt et al. (2016): The Validity and Utility of Selection Methods in Personnel Psychology.

NORMING PROCEDURE

Quantitative scores throughout The Crucible are expressed as percentiles to capture the market and to represent an individual's relative standing within the distribution of leaders in a similar context. This ensures relevant insights for fit within private equity and for individual development as desired.

SAMPLE AND DESCRIPTIVE STATISTICS

The sample used to create percentile scores comes from 1024 leaders across general management (n = 305) and private equity contexts (n = 719). This sample was used to capture the full range of performance on elements and catalysts within the broader population of individuals that may be interested in using The Crucible to assess fit within a private equity context.

Skewness and kurtosis help to further understand the distribution of scores by providing a sense of symmetry in the data. For these statistics, a value of 0 represents no deviation from normality, whereas values greater than 1.00 or less than -1.00 generally represent a notable deviation from normality.

A positive skewness means that the majority of scores are on the lower end of the scale, and a negative skewness means that the majority of scores are on the higher end of the scale.

A positive kurtosis means that scores peak together at one part of the scale, and a negative kurtosis means that scores are more broadly spread out along the entire scale.

Examining results from Table 1 shows that mean scores are slightly higher than the midpoint of the percentile range. Minimum and maximum values demonstrate that observed scores cover approximately the entire range of scores possible. Finally, skewness ranges from -0.44 to -0.79 and kurtosis ranges from -0.18 to -0.91, which does not indicate a large deviation from a normal score distribution.

TABLE 3. Descriptive Statistics for Norming Sample

Catalyst	Mean	SD	Min	Max	Skewness	Kurtosis
Execution Engine	61.0	26.5	1.4	100.0	-0.44	-0.91
Emotional Leverage	62.4	25.8	0.0	99.9	-0.54	-0.68
Leadership	65.4	23.4	1.1	99.9	-0.62	-0.33
Cognitive Horsepower	60.3	27.9	0.2	100.0	-0.47	-0.80
Interpersonal	62.7	26.0	0.1	100.0	-0.79	-0.18
PEQ	65.9	25.6	3.3	100.0	-0.63	-0.64

Note. N = 1024. SD = Standard deviation; Min = Minimum score; Max = Maximum score.

PSYCHOMETRIC PROPERTIES

SAMPLE

The results presented in this section are from 719 leaders in a private equity context who completed the Crucible electronically from 2021 to 2023. This sample is a small part of our dataset that reflects the most recent and up-to-date scores. This sample is appropriate to test the measurement precision of the Crucible as it reflects the target population for the assessment.

RELIABILITY

Internal consistency is a measure of reliability that captures the degree of correlation among related items or elements that measure the same concept. Each item that makes up an element, and each element that makes a catalyst should be related to one another.

A common metric used to test this is Cronbach's Alpha, a statistic that expresses the relationship between elements as a number ranging from 0.0 to 1.0. A value of 0.0 would indicate no relationship, while a value of 1.0 would indicate a perfect relationship. In practice, a 1.0 is impossible and undesirable because it translates to a lack of variety in the information gathered in the assessment. Thus, a commonly accepted gold standard is an Alpha value near .90, however values greater than .80 are considered very good, and values greater than .70 are considered satisfactory by the scientific community.

TABLE 4. Internal Reliability at Catalyst level

Catalyst:	Cronbach's Alpha:
Execution Engine	0.80
Emotional Leverage	0.91
Leadership	0.89
Cognitive Horsepower	0.71
Interpersonal	0.72

Alpha values for each Catalyst demonstrate satisfactory reliability (greater than .70), with the majority demonstrating very good or exceptional reliability.

INTERCORRELATIONS BETWEEN CATALYSTS

The following table shows the correlations between each of the five Catalysts from our most recent sample of private equity leaders. The correlation coefficient represents the relationship between two individual catalysts on the assessment.

Correlations between parts on an assessment should be unrelated, or be small to moderate in size. Any two Catalysts are expected to correlate more with each other to the extent that they are conceptually similar or co-occur in private equity leaders. Thus it is reasonable that some Catalysts, such as leadership,

have a moderate positive correlation (.32 - .59) with other Catalysts. Leadership is a broad concept that covers a range of behaviors and abilities, and should overlap with many other drivers of success.

Importantly, Cognitive Horsepower is unrelated to the other Catalysts, which are not intended to reflect the cognitive aspects of leaders. The lack of correlation with the other Catalysts demonstrates that they are each free of any contaminating influence of cognitive or intellectual ability and add substantial value.

TABLE 5. Correlations between The Crucible Catalysts

Catalyst	Execution Engine	Emotional Leverage	Leadership	Cognitive Horsepower	Interpersonal
Execution Engine	1.00				
Emotional Leverage	0.61*	1.00			
Leadership	0.52*	0.59*	1.00		
Cognitive Horsepower	0.07	0.05	-0.04	1.00	
Interpersonal	0.08	0.11*	0.32*	-0.05	1.00

N = 719. * = statistically significant at $p < .001$ level.

Intercorrelations between Elements within each of the five Catalysts can be found in Appendix A (pg. 21).

VALIDITY

SAMPLE

The results in this section rely on a portion of the overall research sample, including 302 leaders who completed the Crucible from December 2022 to March 2023. This subsample was identified based on the availability of performance data for these individuals.

DEFINING ALIGNMENT

The Crucible model should predict the potential of an individual in the private equity environment. In order to support using the tool in this manner, we compare assessment results from individuals in our sample to their performance on certain benchmarks of success in a private equity environment.

One benchmark of success in a private equity environment is referred to as an exit. When a company goes through a sale process, the executives involved, particularly those in the c-suite, are considered critical for a private equity firm to realize a return on investment. If an executive is in a c-suite role with a company during the hold period, but does not make it to the sale process due to being replaced or moving on to another opportunity, this is considered an indication of a failed exit.

As an additional test to demonstrate that The Crucible model predicts potential in a private equity environment and not just success, we also compare leaders who have private equity operating company leadership experience to those without, as well as those who have successful exits compared to those who are still in their first hold and thus have not succeeded or failed.

GROUP DIFFERENCES

One way to make a comparison between individuals with a history of success or failure in a private equity environment is to compute a t-test. This is a statistic that allows you to compare mean scores of two groups of individuals and detect statistically significant differences between the two groups. Mathematically it can be computed using the following formula[3]:

$$t = (X_1 - X_2) / s_p^2 \sqrt{(1/n_1) + (1/n_2)}$$

A significant difference is detected when the significance level (called a "p-value") is below 0.05. Thus, we use this value as a threshold that determines whether our assessment can predict potential success in private equity. A two-tailed t-test is generally more appropriate and has a stricter threshold for detecting significant differences, thus we use it to provide more convincing evidence in favor of the validity of the Crucible.

In the first test, we compare PEQ scores for individuals with PE exits to those without any PE exit, removing individuals in the first hold, as they have not yet had the opportunity to complete an exit.

The group of individuals with exit experience had a higher mean PEQ score. With a p-value less than .05, this indicates a significant difference, suggesting that individuals with exits score higher on The Crucible than those without any exits and that scores can be used to measure individual potential for PE.

[3] https://www.statology.org/welchs-t-test/

TABLE 6. t-test comparing individuals with and without PE Exit.

Group	Mean (X)	Variance (s)	n	t	p
PE Exit	.8316	.0193	127	3.227	0.0013
No Exit	.7767	.0229	175		

Note. There is only one t-value and p-value because these are calculated using the mean, variance, and group size from each group. Using the formula given above, X = mean, s_p^2 = pooled variance, n = group size.

In the second test, we compare PEQ scores for individuals with a successful exit with those who have failed exit experience. For this latter group, we included individuals with failed exits and no previously successful exits (n = 13), however similar results were obtained when including individuals with both failed and successful exits (n = 21).

The group of individuals with exit experience had a higher mean PEQ score. With a p-value less than .05, this suggests a significant difference, further confirming the Crucible's alignment with success in PE.

TABLE 7. t-test comparing individuals with exits and failed exits.

Group	Mean (X)	Variance (s)	n	t	p
PE Exit	.8316	.0193	127	2.550	0.0119
Failed Exit	.7306	.0105	13		

Note. There is only one t-value and p-value because these are calculated using the mean, variance, and group size from each group. Using the formula given above, X = mean, s_p^2 = pooled variance, n = group size.

To confirm that The Crucible predicts potential in private equity, it is important to demonstrate that there are no significant differences detected between individuals with and without general PE experience. If there was a significant difference, this might indicate bias favoring those with PE experience.

This analysis suggests that there is no significant difference between these two groups (p > .05), which helps rule out a bias in favor of those with PE experience. This means that The Crucible is a valid predictor of fit within a PE environment, regardless of the individual's personal background in PE.

TABLE 8. t-test comparing individuals with and without PE experience.

Group	Mean (X)	Variance (s)	n	t	p
PE Experience	.8106	.0195	200	0.5161	0.6061
No PE Experience	.8014	.0246	102		

Note. There is only one t-value and p-value because these are calculated using the mean, variance, and group size from each group. Using the formula given above, X = mean, s_p^2 = pooled variance, n = group size.

POPULATION VARIANCE TESTING

Population variance testing provides insight into how data from specific populations deviate, which can help detect and adjust for bias in a test or assessment. In order to have a valid assessment, it must function the same for all people. Bias refers to any systematic error where the outcome of the assessment or any individual part of the assessment is different, typically for a specific group of people.

The Crucible is designed specifically for private equity firms, their portfolio companies, and similar high-stakes environments. To accurately assess individuals and engage the best leaders, assessments must be free of any bias that would reduce diversity in the portfolio company. By testing for potential bias, this demonstrates that The Crucible provides a fair, and high quality evaluation for all leaders.

The Crucible is fully prepared to evaluate leaders of all backgrounds and identities for fit in a private equity environment.

SAMPLING PROCEDURE

To collect the data, The Crucible partnered with Qualtrics, the most trusted research software, to conduct a general population survey. Qualtrics is used by 85% of Fortune 100 companies and conducts over a billion surveys each year.[4]

An important part of defining the sample is determining the appropriate sample size. Four factors were evaluated to determine sample size: population size, margin of error, confidence level, and standard deviation. These factors are used to calculate a z-score, which is used to calculate sample size.[5]

$$\text{Calculated sample size} = \frac{Z^2 * SD * (1-SD)}{(margin\ of\ error)^2}$$

A sample size of 305 was determined with a pre-established confidence level of 95% which is commonly used with the analyses conducted in this testing effort. Additionally, quotas were added to tailor the sample towards those most likely to be taking The Crucible. Respondents were required to be 30 years of age or older, have obtained at least a Bachelor's degree, and have previously held a management position. Respondents were also removed if they did not spend a reasonable amount of time completing the assessment (speeding) or if their response patterns were clearly inattentive (selecting only the first answer for every question).

Guidelines for general US population age brackets, gender, and race quotas were added to ensure the groups were large enough to detect significant differences across all identities.

Gender (Quota)
Male (~30%) Female (~70%)

Race (Quota)	
Non-hispanic White (30%) Non-hispanic Black (23%)	Hispanic (~23%) Other Race (~23%)

[4] https://www.qualtrics.com/core-xm/survey-software/Qualtrics
[5] https://www.qualtrics.com/experience-management/research/determine-sample-size/

[6]

SAMPLE

The following tables define the sample of respondents used to complete population variance testing including the gender, race/ethnicity, and age of respondents.

TABLE 9. Sample of individuals for population variance testing

Race	N male	% of Male Sample	N female	% of Female sample	N total	% of Total Sample
Asian	9	10%	57	27%	66	21.64%
Black or African American	16	17%	63	30%	79	25.90%
Hispanic or LatinX	9	10%	56	26%	65	21.31%
Middle Eastern or North African	0	0%	3	1%	3	0.98%
Native American	0	0%	3	1%	3	0.98%
Native Hawaiian or Pacific Islander	0	0%	1	0%	1	0.33%
White	60	65%	43	20%	103	33.77%
Total	**92**		**213**		**305**	

Age	N male	% of Male Sample	N female	% of Female sample	N total	% of Total Sample
30-40	61	66%	103	48	164	53.77%
40-50	29	32%	81	38%	110	34.07%
50-60	1	1%	18	8%	19	6.23%
60-70	1	1%	11	5%	12	3.93%
Total	**92**		**213**		**305**	

Note. There were no respondents below the age of 30 or above the age of 70.

It is typical to first examine population variance at the broadest or highest level of assessment, and then if potential bias is found, one should dig deeper into the specific parts to locate the source of the bias.

Beginning at the element level and working up through the PEQ, there was no evidence of bias identified. Contaminants were also tested and analyzed.

[6] https://www.census.gov/quickfacts/fact/table/US/RHI125220; percentages used reflect approximations by the United States Census to capture results in accordance with the general population.

GROUP DIFFERENCES

A thorough statistical analysis of all scores at the element level, catalyst level, and finally the overall PEQ score, demonstrates no apparent bias in the assessment.

A preliminary way to determine potential areas of bias amongst the Catalysts is to use a t-test to compare group mean differences. This test is used to determine if there is a significant difference between the average scores of two different groups, determined by a p-value below 0.05. A two-tailed test was used as it does not assume any directionality in group differences and is thus more appropriate. Welch's t-test was used because it is more appropriate when group sizes and variances are unequal.[7]

$$t = (X_1 - X_2) / \sqrt{(s_1^2/n_1) + (s_2^2/n_2)}$$

A cohesive summary including the t-tests comparing mean scores across all parts of the assessment for subgroups of interest can be found in the appendix. Here we note any statistically significant effects that were detected.

At the Catalyst level, we identified one area for exploration. There were statistically significant differences detected for the Interpersonal Catalyst such that female respondents scored higher than male respondents (p = .036). Also for the Interpersonal Catalyst, Black respondents scored higher than White respondents (p = .048). The details of these tests can be found in the appendix.

To explore these further we looked at Element level comparisons within the Interpersonal Catalyst. For the comparison of female to male respondents, female respondents scored statistically significantly higher than male respondents on two Elements - Authenticity (p = .005) and Empathy (p = .003). For the comparison of Black and White respondents, Black respondents scored statistically significantly higher than White respondents on two different Elements - EQ (p = .009) and Collegiality (p = .050). The Crucible is not concerned about negative impact being introduced by these significant differences.

While these tests show differences in mean scores between the tested groups, mean differences are not indicative of potential biases or adverse impact.[8] When mean differences are detected, this can reflect true underlying differences in the overall populations, and this explanation is more likely when supported by other empirical research. The Crucible will continue to monitor and conduct follow-up studies.

It is a well-established finding that women score higher than men on integrity tests[9] and measures of emotional intelligence.[10] These are similar to what The Crucible captures in the Interpersonal Catalyst, specifically Authenticity and Empathy. According to the research women generally have more complex emotional knowledge than men and are better at perceiving emotional cues. Furthermore, women who have not yet made it to leadership positions are often disadvantaged due to societal gender roles, however women who have emerged in top leadership positions often have a leadership advantage because they are seen as successful agentic and communal leaders[11]

[7] https://www.statology.org/welchs-t-test/
[8] https://www.testingstandards.net/uploads/7/6/6/4/76643089/standards_2014edition.pdf **(Pg. 65, Standard 3.6)**
[9] **Berry et al. (2007)**; https://doi.org/10.1111/j.1744-6570.2007.00074.x
[10] **Newman et al. (2010)**; https://shrm.org/hr-today/news/hr-magazine/Documents/Joseph-Newman-2010.pdf
[11] https://msbfile03.usc.edu/digitalmeasures/tost/intellcont/Rosette%20and%20Tost%202010%20JAP-1.pdf

Thus, the difference we detected is likely to be a reflection of real differences between men and women in interpersonal abilities rather than a function of test bias.

Research has also supported the finding that minority racial groups perform better on self-reported emotional intelligence measures, suggesting a similar heightened interpersonal ability.[12]

Gender and racial differences can be further explained by the phenomena known as intersectionality, which refers to the way that a person's many social identities such as race, gender, sexuality, and socioeconomic status interact with one another to create unique experiences of privilege or oppression. Those who identify with more than one minority identity often face more interpersonal challenges and these experiences allow them to better connect with and empathize with others, especially in times of adversity[13], thus explaining higher scores in EQ and collegiality.

Mathematically, increased minority representation in the sample could be causing additional differences to be detected. That is, if women have a real difference in interpersonal ability, we expect to find a difference across racial groups as well because 80% of the Black identifying respondents in our sample also identify as female. The majority of female respondents in our sample identified as non-white, while 65% of the male respondents identified as white. Increased minority representation in this sample was intended to ensure that minority groups were not disadvantaged by the Crucible assessment.

Importantly, these differences do not cause concern for adverse impact. The EEOC defines adverse impact as "a substantially different rate of selection in hiring, promotion, or other employment decisions which work to the disadvantage of members of a race, sex, or ethnic group."[14] To test this, a user of The Crucible may adopt a practical measure taken directly from The Uniform Guidelines on Employee Selection Procedures known as the 4/5ths or 80% rule.[15] The rule states that adverse impact occurs when any group does not have a selection or passing rate equal to or greater than 80% of the group with the highest rate (most often white men). The Crucible, as a research tool, does not have a defined passing rate, so it cannot violate the 4/5ths rule absent the specific application made by a test user. Furthermore, because small mean differences found in this investigation are in favor of minority groups and the population of private equity leaders is disproportionately represented by the white male population,[16] it is very unlikely that any calculated selection ratio using the Crucible scores could violate the 4/5ths rule. The selection ratio would have to be higher for a minority group that is underrepresented in private equity to demonstrate adverse impact based on these differences. Thus, when used in a specific application, there is no potential for adverse impact to occur when using The Crucible.

[12] Newman et al. (2010); https://shrm.org/hr-today/news/hr-magazine/Documents/Joseph-Newman-2010.pdf
[13] Lim & DeSteno (2016); https://psycnet.apa.org/record/2016-01245-001
[14] https://www.govinfo.gov/content/pkg/CFR-2006-title29-vol4/xml/CFR-2006-title29-vol4-sec1607-16.xml
[15] https://www.govinfo.gov/content/pkg/CFR-2014-title29-vol4/xml/CFR-2014-title29-vol4-part1607.xml
[16] https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/the-state-of-diversity-in-us-private-equity

CONCLUSION

Based on the evidence in this document, The Crucible is a reliable, valid, and fair research tool. The tool shows satisfactory levels of internal reliability regarding the measurement of Elements, Catalysts, and the overall PEQ score.

Further analysis supports the validity of The Crucible model as a predictor of potential success in a private equity environment. Regardless of an individual's previous experience, better scores on The Crucible are more likely to occur for individuals with successful PE exits.

Finally, no areas of concern for adverse impact were identified in The Crucible assessment. While there were small group mean score differences detected, the differences favor populations that are grossly underrepresented in private equity and thus do not create a concern for adverse impact.

APPENDIX A: INTERCORRELATIONS BETWEEN ELEMENTS

TABLE A1. Correlations for Elements within Execution Engine Catalyst

Element	Urg.	Rel.	Imp.	Hands-on.	RC	DO	SPF	MF	IC
Urgency	1.00								
Relentlessness	0.41	1.00							
Impatience	0.59	0.27	1.00						
Hands-on Approach	0.19	0.38	-0.04	1.00					
Digs for Root Cause (RC)	0.08	0.20	-0.02	0.35	1.00				
Detail Orientation (DO)	-0.13	0.02	-0.27	0.46	0.33	1.00			
Systems & Processes Focused (SPF)	0.12	0.16	-0.03	0.34	0.34	0.39	1.00		
Metrics Focused (MF)	0.20	0.31	0.01	0.42	0.29	0.26	0.75	1.00	
Intellectual Curiosity (IC)	0.13	0.20	-0.02	0.21	0.37	-0.04	0.01	0.11	1.00

TABLE A2. Correlations for Elements within Emotional Leverage

Element	Passion	Intensity	Hunger/Drive	Resiliency	EE	DFL	Comp.
Passion	1.00						
Intensity	0.37	1.00					
Hunger/Drive	0.26	0.56	1.00				
Resiliency	0.22	0.32	0.34	1.00			
Energy/Endurance (EE)	0.46	0.66	0.49	0.60	1.00		
Disdain for Losing (DFL)	0.34	0.73	0.85	0.49	0.62	1.00	
Competitiveness	0.22	0.46	0.68	0.18	0.32	0.86	1.00

TABLE A3. Correlations for Elements within Leadership

Element	Inspires	SHS	DA	Engages	EB	Directness	Optimism	SP
Inspires confidence & belief	1.00							
Sets High Standards (SHS)	0.34	1.00						
Drives Accountability (DA)	0.29	0.31	1.00					
Engages Deeply with Stakeholders	0.31	0.20	0.15	1.00				
Ensures Buy-in (EB)	0.39	0.15	0.09	0.62	1.00			
Directness	0.48	0.42	0.62	0.27	0.23	1.00		
Optimism	0.40	0.06	0.06	0.16	0.19	0.07	1.00	
Strategic Perspective (SP)	0.40	0.31	0.31	0.32	0.29	0.36	0.09	1.00

TABLE A4. Correlations for Elements within Cognitive Horsepower Catalyst

Element	Logic	Word Association	Math	Pattern Recognition
Logic	1.00			
Word Association	0.47	1.00		
Math	0.61	0.52	1.00	
Pattern Recognition	0.26	0.33	0.20	1.00

TABLE A5. Correlations for Elements Within Interpersonal Catalyst

Element	Self Esteem	Belief in Others	Empathy	Confidence	Authenticity	EQ	Collegiality
Self Esteem	1.00						
Belief in Others	-0.05	1.00					
Empathy	-0.08	0.07	1.00				
Confidence	0.83	-0.06	-0.07	1.00			
Authenticity	0.08	-0.03	0.06	0.02	1.00		
EQ	0.30	-0.09	0.11	0.19	0.15	1.00	
Collegiality	0.02	0.09	0.11	0.00	0.02	0.02	1.00

APPENDIX B: T-TESTS FROM POPULATION VARIANCE TESTING

The following set of tables contain the t-tests (two-tailed) for The PEQ score and Catalyst level comparing female and male respondent group mean scores.

TABLE B1. Female vs. Male - PEQ

Group	Mean	Variance	n	t	p	Cohen's d
Female	33.2	269.9	213	-0.20	0.840	-0.02
Male	33.6	210.8	92			

TABLE B2. Female vs. Male - Execution Engine

Group	Mean	Variance	n	t	p	Cohen's d
Female	30.1	325.9	213	0.43	0.667	0.05
Male	29.2	251.4	92			

TABLE B3. Female vs. Male - Emotional Leverage

Group	Mean	Variance	n	t	p	Cohen's d
Female	30.5	318.8	213	-1.30	0.194	-0.16
Male	33.2	258.6	92			

TABLE B4. Female vs. Male - Leadership

Group	Mean	Variance	n	t	p	Cohen's d
Female	37.3	331.0	213	-0.67	0.50	-0.08
Male	38.5	192.5	92			

TABLE B5. Female vs. Male - Cognitive Horsepower

Group	Mean	Variance	n	t	p	Cohen's d
Female	38.5	731.3	213	-0.10	0.924	-0.01
Male	38.8	658.8	92			

TABLE B6. Female vs. Male - Interpersonal

Group	Mean	Variance	n	t	p	Cohen's d
Female	33.7	465.3	213	2.12	0.036*	0.26
Male	27.9	478.9	92			

* = statistically significant with $p < .05$.

The following set of tables contain the t-tests (two-tailed) for Elements within the Interpersonal Catalyst comparing female and male respondent group mean scores.

TABLE B7. Female vs. Male - Self Esteem

Group	Mean	Variance	n	t	p	Cohen's d
Female	38.5	409.5	213	0.10	0.920	0.01
Male	38.3	312.8	92			

TABLE B8. Female vs. Male - Belief in Others

Group	Mean	Variance	n	t	p	Cohen's d
Female	38.0	777.7	213	1.41	0.160	0.17
Male	33.3	681.1	92			

TABLE B9. Female vs. Male - Empathy

Group	Mean	Variance	n	t	p	Cohen's d
Female	38.4	614.7	213	2.98	0.003*	0.36
Male	29.8	496.2	92			

TABLE B10. Female vs. Male - Confidence

Group	Mean	Variance	n	t	p	Cohen's d
Female	34.9	403.0	213	0.31	0.757	0.04
Male	34.2	281.0	92			

TABLE B11. Female vs. Male - Authenticity

Group	Mean	Variance	n	t	p	Cohen's d
Female	30.6	681.4	213	2.85	0.005*	0.35
Male	21.7	596.4	92			

TABLE B12. Female vs. Male - EQ

Group	Mean	Variance	n	t	p	Cohen's d
Female	34.9	594.7	213	1.61	0.108	0.19
Male	30.5	430.4	92			

TABLE B13. Female vs. Male - Collegiality

Group	Mean	Variance	n	t	p	Cohen's d
Female	33.1	810.4	213	-0.24	0.810	-0.03
Male	34.0	784.9	92			

* = statistically significant with $p < .05$.

The following set of tables contain the t-tests (two-tailed) for Elements within the Interpersonal catalyst comparing Black and White respondent group mean scores.

TABLE B14. Black vs. White - PEQ

Group	Mean	Variance	n	t	p	Cohen's d
Black	34.6	278.0	77	1.18	0.239	0.18
White	31.8	217.0	100			

TABLE B15. Black vs. White - Execution Engine

Group	Mean	Variance	n	t	p	Cohen's d
Black	31.4	323.7	77	0.33	0.739	0.05
White	30.5	287.7	100			

TABLE B16. Black vs. White - Emotional Leverage

Group	Mean	Variance	n	t	p	Cohen's d
Black	33.1	285.6	77	0.20	0.837	0.03
White	33.0	272.9	100			

TABLE B17. Black vs. White - Leadership

Group	Mean	Variance	n	t	p	Cohen's d
Black	39.4	271.0	77	1.43	0.155	0.21
White	36.2	229.6	100			

TABLE B18. Black vs. White - Cognitive Horsepower

Group	Mean	Variance	n	t	p	Cohen's d
Black	35.6	644.9	77	0.71	0.479	0.10
White	32	597.5	100			

TABLE B19. Black vs. White - Interpersonal

Group	Mean	Variance	n	t	p	Cohen's d
Black	32.8	597.5	77	1.99	0.048*	0.30
White	26.1	422.5	100			

* = statistically significant with $p < .05$.

The following set of tables contain the t-tests (two-tailed) for Elements within the Interpersonal Catalyst comparing Black and White respondent group mean scores.

TABLE B20. Black vs. White - Self Esteem

Group	Mean	Variance	n	t	p	Cohen's d
Black	40.7	419.6	77	1.36	0.177	0.20
White	36.8	276.1	100			

TABLE B21. Black vs. White - Belief in Others

Group	Mean	Variance	n	t	p	Cohen's d
Black	34.5	885.2	77	0.46	0.648	0.07
White	32.6	597.5	100			

TABLE B22. Black vs. White - Empathy

Group	Mean	Variance	n	t	p	Cohen's d
Black	39.0	589.7	77	1.18	0.240	0.17
White	34.5	669.0	100			

TABLE B23. Black vs. White - Confidence

Group	Mean	Variance	n	t	p	Cohen's d
Black	35.4	398.9	77	0.81	0.418	0.12
White	33.1	284.8	100			

TABLE B24. Black vs. White - Authenticity

Group	Mean	Variance	n	t	p	Cohen's d
Black	24.6	627.0	77	1.23	0.219	0.19
White	20.0	578.0	100			

TABLE B25. Black vs. White - EQ

Group	Mean	Variance	n	t	p	Cohen's d
Black	36.8	705.1	77	2.65	0.009*	0.41
White	27.0	467.3	100			

TABLE B26. Black vs. White - Collegiality

Group	Mean	Variance	n	t	p	Cohen's d
Black	37.2	934.7	77	1.97	0.050*	0.30
White	28.5	713.1	100			

* = statistically significant with $p < .05$.